Author/s:
Dang, An Binh

Title:
Evaluating the Utility of Severity Ratings for Eating Disorders: A Comparison of DSM-5, ICD-11, and Cognitive-Based Indicators

Date:
2025

Persistent Link:
https://hdl.handle.net/11343/356743

Evaluating the Utility of Severity Ratings for Eating Disorders: A Comparison of DSM-5, ICD-11, and Cognitive-Based Indicators

An Dang

January 2025

OCRID Number: 0000-0001-7961-6913

Melbourne School of Psychological Sciences

The University of Melbourne Thesis submitted in partial fulfilment of the degree Master of Psychology (Clinical)/Doctor of Philosophy

Thesis summary

Eating disorders (EDs) are currently diagnosed using the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013), its text-revised version (DSM-5-TR; APA, 2022), or the 11th edition of the International Classification of Diseases (ICD-11; World Health Organization [WHO], 2019). Both the DSM-5 and ICD-11 use Body Mass Index (BMI) as a severity indicator for Anorexia Nervosa (AN), but only the DSM-5 includes severity indicators for Bulimia Nervosa (BN) and Binge Eating Disorder (BED), based on the weekly frequency of purging behaviours and binge eating episodes, respectively. This thesis by publication aimed to evaluate the current severity ratings for AN, BN, and BED as proposed by the DSM-5 and ICD-11, as well as alternative severity ratings suggested by researchers, which focus on common cognitive factors (e.g., drive for thinness [DT], and overvaluation of weight and shape (OWS) shared across EDs. By examining the strengths and limitations of these severity ratings, this thesis proposed recommendations for both clinical practice and research contexts.

Study 1 included a systematic review ($N = 22$) and a meta-analysis ($N = 19$) of studies evaluating both DSM-5 and alternative OWS severity indices for EDs. The meta-analysis revealed significant differences in ED psychopathology among most DSM-5 severity groups for BN, but not for AN or BED. For the alternative OWS rating, the meta-analysis found significant differences in ED psychopathology between individuals with low and high OWS in those with BED, thus providing support to further investigate OWS as alternative for EDs.

Given the primary purpose of these severity ratings is to assist in clinical practice, there is limited understanding of whether mental health clinicians within Australia (e.g., psychiatrists, psychologists) utilise these ratings in their practice. Additionally, little is known about clinicians' views on the usefulness of these severity ratings for treatment decisions. Therefore, Study 2 used qualitative and quantitative methods to assess whether clinicians use

these severity ratings in practice and their opinions regarding these ratings. Findings from 38 clinicians (mean age = 38.5, 97.37% female) showed that the majority (60%) reported not using these ratings in their clinical practice. Qualitative findings from clinical interviews of six clinicians, coupled with free-text entries from the remaining 32 clinicians, suggested that clinicians perceived the DSM-5 severity ratings as (1) invalidating, (2) pathology-reinforcing, (3) lacking clinical relevance, and (4) communication of ED severity to others. Specifically, for theme 4, while clinicians reported not communicating ED severity to treating clients, some noted that these ratings can aid in triage and communication among clinicians in tertiary settings.

Notwithstanding that AN is a psychiatric condition with severe, potentially life-threatening medical complications, research on severity ratings has often overlooked whether the severity ratings for AN accurately capture differences in both biological and psychological factors among treatment-seeking individuals with AN. Using a sample of 312 treatment-seeking patients with AN from St Vicent Hospital (mean age = 25.3, SD = 7.6; mean BMI = 16.8 kg/m², SD = 2.4, Studies 3a and 3b aimed to evaluate the clinical utility of the DSM-5, ICD-11 and alternative OWS severity ratings for AN on a wide range of biological and psychological variables.

The aim of Study 3a was to evaluate the DSM-5 and OWS severity ratings in relation to ED psychopathology (e.g., dietary restraint, eating concerns), general psychopathology (e.g., quality of life, disability), and biological variables (e.g., potassium, magnesium). As the study was completed before the publication of the ICD-11 in 2019, the published manuscript did not include the assessment of ICD-11 AN severity rating. To address this limitation, additional analyses were conducted to explore the ICD-11 BMI-based AN severity rating and its relationship with psychological and biological variables, with the results presented in section 5.3 (Chapter 5). Study 3b, on the other hand, aimed to determine psychiatric

comorbidity rates across the DSM-5, ICD-11, and OWS severity groups within the same treatment-seeking population.

Results from Study 3a showed no significant differences in biological variables (e.g., potassium, magnesium, phosphate) across the DSM-5, ICD-11, or OWS severity groups. However, results from Studies 3a and 3b revealed a paradoxical pattern in both the DSM-5 and ICD-11: higher BMI categories (classified as “mild” in DSM-5 and “significantly low BMI” in ICD-11) were associated with greater ED psychopathology (Study 3a) and higher rates of psychiatric comorbidities (Study 3b). Specifically, in DSM-5, individuals in the milder severity group reported more ED and general psychopathology (e.g., stress or anxiety symptoms) (Study 3a) and higher rates of comorbidities, such as panic disorder, social anxiety disorder, generalised anxiety disorder, and post-traumatic stress disorder (Study 3b). Similarly, in ICD-11, those in the "significantly low BMI" category had higher levels of reported depressive, stress, and anxiety symptoms, as well as eating, shape, and weight concerns (Study 3a), and higher rates of major depressive disorder and obsessive-compulsive disorder (Study 3b) compared to those in the "dangerously low BMI" category.

In contrast, the OWS severity ratings aligned with the expected linear progression of severity: individuals in the higher OWS (more severe) group exhibited greater ED psychopathology (Study 3a) and a higher prevalence of major depressive disorder (Study 3b) than those in the lower OWS group. Overall, the OWS severity ratings were more effective in capturing a linear relationship between severity and outcomes, with higher OWS scores associated with increased levels of psychiatric comorbidities and psychopathology. However, despite its relative superiority, the OWS measure was unable to effectively capture biological variables (Study 3a) and only accounted for a limited number of psychiatric comorbidities (Study 3b). These limitations suggest that, while the OWS severity ratings show promise, they are not yet suitable for clinical application.

While the relationships between risk factors and severity have been rigorously explored in other medical conditions (e.g., COVID-19, acute stroke), they remain underexplored in psychiatric disorders like AN. If the DSM-5 severity rating is valid, it should not only index ED psychopathology but also distinguish factors related to the development and maintenance of AN. Study 4 aimed to assess whether AN-related risk factors, assessed using the Oxford Risk Factor Interview, were related to DSM-5 BMI and DT alternative severity ratings in 153 pairs of individuals with AN and their healthy sisters (N = 306, mean age = 26.53) part of a multicentre European project. Multinomial logistic regressions were conducted and revealed that both DSM-5 and DT severity ratings identified similar risk factors associated with AN. Only childhood perfectionism was found to be more common in the extreme severe DSM-5 AN severity group compared to the DSM-5 AN severe group. Overall, the findings from Study 4 provide little evidence for an association between AN risk factor and either DSM-5 or DT severity ratings, warranting replication of these findings.

Despite DSM-5 AN and BN severity ratings being in place for over a decade, few studies have assessed their ability to predict treatment outcomes. Similarly, although the ICD-11, published in 2019, included a proposed severity rating for AN, no study has yet directly assessed its usefulness in aiding clinical practice. Study 5 addressed this gap by comparing treatment outcomes associated with DSM-5 AN and BN severity ratings, ICD-11 severity ratings for AN, and the alternative DT severity rating for AN and BN. A total of 628 female participants diagnosed with either AN (n = 266; mean age = 26.71) or BN (n = 362; mean age = 29.49) were recruited from an ED treatment unit in Spain. Participants were classified into DSM-5, ICD-11, and DT severity categories upon admission and underwent enhanced cognitive behaviour therapy programs tailored to the ED subtype. Significant associations were found only for ICD-11 AN severity groups, where "dangerously low BMI" was linked

to poorer treatment outcomes compared to "significantly low BMI." No significant associations were found for DSM-5 or DT AN and BN severity groups, again raising questions about the validity of these severity ratings.

Since all the studies in this PhD utilised clinical samples, there is limited understanding of how DSM-5, ICD-11, and alternative ED severity ratings function within community-based populations. Community-based research is important for enhancing the generalisability of findings. To address this gap, Study 6 used a community-based sample of 686 adults (75% female, mean age = 19.95) and employed ecological momentary assessment (EMA) to investigate whether BMI and weight suppression (another promising alternative severity rating, defined as the difference between highest and current weight) predicted state-based experiences of body dissatisfaction and disordered eating urges (e.g., dietary restriction, excessive exercise, binge eating). Participants completed six daily EMA surveys over seven days, reporting their levels of satisfaction with their appearance and urges to engage in disordered eating behaviours. Multilevel modelling analyses showed that individuals with a lower BMI and a lower weight suppression tend to report having a higher body dissatisfaction average. Furthermore, those with a higher BMI and a higher weight suppression were more likely to report having urges to engage in unhealthy eating. However, both BMI and weight suppression had limited predictive value for state-based experiences of body dissatisfaction and other disordered eating urges (i.e., dietary restriction, excessive exercise, and binge eating), highlighting the limitations of weight-based severity ratings in capturing ED-related state-based variables in a community sample.

To conclude, severity indicators for psychiatric disorders aim to guide prognosis and treatment planning. However, our six studies evaluating severity ratings for EDs revealed that these indicators failed to capture the complex biological and psychosocial profiles inherent to EDs, which can lead to inaccurate assessments of their impact. This, in turn, can hinder

effective treatment planning and limit access to appropriate care. Given the multifaceted nature of EDs, single-indicator measures (e.g., BMI as a severity rating for AN) appear to not adequately reflect the complexity of an ED and should therefore not be used in clinical practice. This underscores the need for a multivariate approach that considers how combinations of symptoms (e.g., number of purging methods) and clinical indicators (e.g., rate of weight loss, treatment history) predict critical outcomes, such as recovery and relapse rates.

Preface

The current thesis will be presented with publications. This thesis comprises six studies, of which four studies (i.e., Studies 1, 3a, 3b, 4, and 5) have been published. Studies 2 and 6 are currently under the second round of reviews in international journals. I contributed over 60% of the content within each publication – I was primarily responsible for writing and revising manuscripts, conducting data analyses, and implementing edits in line with feedback from co-authors and peer reviewers. The references for these papers are provided below, along with conference presentations resulting from this thesis are outlined below:

Papers:

Study 1 (Published). **Dang, A. B.,** Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the *DSM–5* severity ratings for eating disorders. *Clinical Psychology: Science and Practice, 29*(4), 325–344. https://doi.org/10.1037/cps0000078

Study 2 (Under Review). **Dang, A.,** Kirk, H., Kiropoulos, L., & Krug, I. Exploring Clinician Perspectives on the DSM-5 Eating Disorder Severity Ratings: A Qualitative Study

Study 3a (Published). **Dang, A. B**., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2023). Assessing severity in anorexia nervosa: Do the DSM‐5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?. *European Eating Disorders Review*, 31(4), 447-461. https://doi.org/10.1002/erv.2969

Study 3b (Published). **Dang, A. B.,** Kiropoulos, L., Castle, D., Jenkins, Z., Phillipou, A., Rossell, S., & Krug, I. (2024). Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings. *Eating Disorders*, 1-17. https://doi.org/10.1080/10640266.2024.2346001

Study 4 (Published). **Dang, A. B.,** Kiropoulos, L., Anderluh, M., Collier, D., Fernandez-Aranda, F., Karwautz, A., Treasure, J., Wagner, G., & Krug, I. (2024). Do risk factors differentiate DSM-5 and drive for thinness severity groups for anorexia nervosa?. *Journal of Eating Disorders*, 12(1), 5. https://doi.org/10.1186/s40337-024-00966-5

Study 5 (Published). **Dang, A.,** Granero, R., Sanchez, I., Solé, L.G., Toro, J.J.-d., Rosinska, M., Jimenez-Murcia, S., Krug, I. and Fernandez-Aranda, F. (2024), Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 Versus ICD–11 for Thinness. *European Eating Disorders Review*. https://doi.org/10.1002/erv.3156

Study 6 (Under review). **Dang, A.,** Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample?

Conference Presentations

Oral Communications

Dang. B. A., Krug. I., Granero. R., Aguera. Z., Sanchez., I., Riesco. N., Jimenez-Murcia. S., Menchon. J.S., Fernadez-Aranda. F. (October, 2020). *Drive for Thinness Provides an Alternative, More Meaningful, Severity Indicator than the DSM-5 Severity Indices for Eating Disorders.* Oral presentation at Eating Disorder Research Society.

Dang. B. A., Giles. S., Fernadez-Aranda. F., Kiropoulos. L., Fuller-Tyszkiewicz. M., Krug. I. (June, 2021). *A systematic review and meta-analysis on the DSM-5 and the alternative severity indices for Eating Disorders*. Oral presentation at International Conference on Eating Disorders.

Dang. B. A., Krug. I., De La Harpe. S., Rozenblat. V., Giles. S., Kiropoulos. L., Fuller-Tyszkiewicz, M (June, 2021). *Do women with differing levels of trait eating pathology*

experience daily stress and body dissatisfaction differently? Oral presentation at International Conference on Eating Disorders.

Dang. B. A., Krug. I., Granero. R., Riesco. N., Jimenez-Murcia. S., Menchon. J., Fernadez-Aranda. F. (June, 2021). *How to assess severity in males with Eating Disorders? The DSM-5 severity index versus severity based on Drive for Thinness*. Oral presentation at International Conference on Eating Disorders

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2022). *Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?* Oral presentation at Australia and New Zealand Conference for Eating Disorders

Dang, A., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample? (2023). Oral presentation at Australia and New Zealand Conference for Eating Disorders

Dang, A., Kirk, H., Kiropoulos, L., & Krug, I. Exploring Clinician Perspectives on the DSM-5 Eating Disorder Severity Ratings: A Qualitative Study (2024). Oral presentation at Australia and New Zealand Conference for Eating Disorders

Poster Communications

Dang. B. A., Giles. S., Fernadez-Aranda. F., Kiropoulos. L., Fuller-Tyszkiewicz. M., Krug. I. (October, 2020). *A systematic review and meta-analysis on the DSM-5 and the alternative severity indices for Eating Disorders*. Poster presented at Eating Disorder Research Society.

Dang. B. A., Giles. S., Fernadez-Aranda. F., Kiropoulos. L., Fuller-Tyszkiewicz. M., Krug. I. (April, 2021). *A systematic review and meta-analysis on the DSM-5 and the alternative*

severity indices for Eating Disorders. Poster presented at European Congress of Psychiatry.

Dang. B. A., Krug. I., De La Harpe. S., Rozenblat. V., Giles. S., Kiropoulos. L., Fuller-Tyszkiewicz (April, 2021). M. *Do women with differing levels of trait eating pathology experience daily stress and body dissatisfaction differently?* Poster presented at European Congress of Psychiatry.

Dang. B. A., Krug. I., Granero. R., Riesco. N., Jimenez-Murcia. S., Menchon. J., Fernadez-Aranda. F. (April, 2021). *How to assess severity in males with Eating Disorders? The DSM-5 severity index versus severity based on Drive for Thinness.* Poster presented at European Congress of Psychiatry.

Dang. B. A., Kiropoulos. L., Collier. D., Fernadez-Aranda. F., Karwautz.A., Wagnere. G., Krug.I. *The DSM-5 Severity Index for Anorexia Nervosa: Association Between the DSM-5 Anorexia Nervosa Severity Groups and Risk Factors* (August 2022). Presented at the Australia New Zealand Eating Disorder Conference.

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (August, 2022). *Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?*. Presented at the Australia New Zealand Eating Disorder Conference.

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (March, 2023). Assessing severity in anorexia nervosa: *Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?*. Presented at the International Eating Disorder Conference.

Invited Oral Talks

1.Delivered a seminar titled "Australia Awards Guest Lecture – Mental Health" on 20 October 2023, organized by the University of Melbourne for Mental Health Awareness Month. The presentation focused on the severity of eating disorders and their relevance across different cultural groups.

2. Presented findings from the current PhD thesis at the Monthly Seminar Series hosted by the Australian Eating Disorders Research & Translation Centre on Wednesday, 13 November 2024.

Grants and Awards Received During Candidature

2024	Peter Beumont Young Investigator Finalist, ANZAED Conference
2023	Peter Beumont Young Investigator Finalist, ANZAED Conference
2021	Academy of Eating Disorders, Early Career Investigator Fellowship, International Conference of Eating Disorders
2021	Australian Government Research Training Program (RTP) Scholarship

Other Publications Undertaken During PhD Candidature

Krug, I., **Dang, A**., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia,S., Menchón, J.M., & Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the DSM-5 severity indices for eating disorders. *European Eating Disorders Review*, *29*(3), 482-498. https://doi.org/10.1002/erv.2776

Krug, I., Arroyo, M. D., Giles, S., **Dang, A. B.,** Kiropoulos, L., De Paoli, T., ... & Fuller-Tyszkiewicz, M. (2021). A new integrative model for the co-occurrence of non-suicidal self-injury behaviours and eating disorder symptoms. *Journal of Eating Disorders*, *9*, 1-17. https://doi.org/10.1186/s40337-021-00508-3

Kaur, J., **Dang, A. B**., Gan, J., An, Z., & Krug, I. (2022). Night eating syndrome in patients with obesity and binge eating disorder: a systematic review. *Frontiers in psychology*, *12*, 766827. https://doi.org/10.3389/fpsyg.2021.766827

Krug, I., **Dang, A. B**., & Hughes, E. K. (2024). There is nothing as inconsistent as the OSFED diagnostic criteria. *Trends in Molecular Medicine*. 30(4), 403-415. https://doi.org/10.1016/j.molmed.2024.01.006

Dang, T. B., Hughes, E. K., **Dang, A. B**., Lai, H. Y., Lee, J., Liu, S., Portingale, J., Fuller-Tyszkiewicz, M., & Krug, I. (2024). Taking a Deeper Dive Into OSFED Subtypes: A Meta-Analysis and Systematic Review. *International Journal of Eating Disorders*, 57(10), 2006-2040. https://doi.org/10.1002/eat.24280

Krug, I., **Dang, A. B.,** Sánchez, I., Granero, R., Agüera, Z., Gaspar-Perez, A., ... & Fernandez-Aranda, F. (2024). How to assess eating disorder severity in males? The DSM-5 severity index versus severity based on drive for thinness. Eating Disorders, 32(1), 81-97. https://doi.org/10.1080/10640266.2023.2259682

Acknowledgements

I wouldn't have been able to complete my PhD (as well as my entire psychology degree) without the unwavering support of my family, friends, and supervisors.

First and foremost, I want to express my gratitude to my incredible supervisor, Associate Professor Isabel Krug. Thank you for being so approachable and supportive throughout this journey. It has been an absolute privilege to learn from such an exceptional researcher and teacher like you. Your encouragement and belief in me, especially during moments of doubt, have been invaluable, and I cannot thank you enough.

I would also like to extend my heartfelt thanks to Associate Professor Litza Kiropoulos for your guidance, encouragement, and support throughout my candidacy. Your mentorship and belief in my abilities have been instrumental in shaping this research and my development as a researcher.

To my PhD advisory committee – Professor Andy Perfors and Dr. Vanja Rozenblat – I am deeply grateful for your support and insights throughout my PhD journey. Your guidance has been instrumental in helping me navigate this process.

Feiven and Carmen, you both have been my buddies since the very beginning of my Master degree. It has been such a heartwarming journey to go through this degree with you both. I can't wait to learn more with you both, and more importantly discover more warm and comforting food spots together!

To my lab buddy and one of my closest friends, Jade, thank you for endlessly supporting and encouraging me throughout this degree. I'm so glad we got to travel and learn together, and I can't wait for more glasses of wine and adventures in the future! Thank you for always having my back.

Raina and Andrew, thank you for being there for me (literally almost every day) throughout this journey. Knowing you both has been one of the greatest blessings I've gained

over the course of this four-year degree. You guys, along with Jie, have truly become my family away from home. I can't wait for many more years of our weekly Saturday night dinners, board games, and movie nights together!

Jie, you have been my number one support from start to finish. Thank you for always taking care of me, supporting me, and believing in me throughout the years. You've probably learned more about eating disorders and severity ratings than you ever wished to, but your unwavering patience and encouragement have meant the world to me.

Lastly, I would like to extend my deepest gratitude to my family, to whom I dedicate this thesis. Con cảm ơn bố mẹ đã nuôi dạy và dành hết tất cả cho con. Mẹ à, con cảm ơn mẹ nhiều lắm. Con lúc nào cũng thấy mình thật may mắn khi là con của bố mẹ. Cảm ơn mẹ đã tin tưởng và yêu thương con. Bố ơi, con cảm bố đã luôn ở đó cho con khi con cần bố. Bây giờ con đã làm xong luận án của mình, và con hi vọng gia đình mình sẽ có nhiều thời gian bên nhau hơn.

Anh 2, em cảm ơn và Charlotte đã chăm sóc em trong nhưng năm qua, đã nấu ăn cho em và dẫn em ăn thật nhiều món ngon. Em hi vọng bố mẹ, anh em mình, Charlotte và Alexandra sẽ có nhiều thời gian cùng nhau hơn, anh nhé.

Table Contents

Thesis summary **2**
Declarations **8**
Preface **9**
Acknowledgements **15**
List of Tables **20**
List of Figures **22**
List of Abbreviations **23**
Chapter 1: Eating Disorders Diagnosis: Main Characteristics and Outcomes **24**
1.1 Eating Disorders Diagnosis, Prevalence, and Burden **25**
1.1.1 Anorexia Nervosa 26
1.1.2 Bulimia Nervosa 30
1.1.3 Binge Eating Disorder 33
1.1.4 Other Specified Feeding or Eating Disorders 36
1.2 Risk Factors for Eating Disorders **39**
1.2.1 Risk Factors for Anorexia Nervosa 39
1.2.2 Risk Factors for Bulimia Nervosa 41
1.2.3 Risk Factors for Binge Eating Disorder 43
1.3 Disordered Eating Spectrum **45**
1.4 Theoretical Models of Eating Disorders **47**
1.4.1 Transdiagnostic Maintenance Model of Eating Disorders 48
1.4.2 Dual-Pathway Model for Bulimia Nervosa 50
1.5 Course and Treatment of Eating Disorders **52**
1.5.1 Anorexia Nervosa 53
1.5.2 Bulimia Nervosa 55
1.5.3 Binge Eating Disorder 57
1.6 Severity Ratings for Eating Disorders **59**
1.6.1 The Concept of Severity in Psychiatry 59
1.6.2 The Introduction of the DSM-5 Severity Ratings for Eating Disorders 61
1.6.3 The ICD-11 Severity Ratings for AN 63
1.6.4 Alternative Severity Ratings for Eating Disorders 65
1.7 Chapter Summary **71**
Chapter 2: A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders. **73**
2.1 Chapter Introduction **73**

2.2 Study 1. ..... **75**
Chapter 3: Summary and Rationale for the Current Thesis **76**
3.1 Chapter Introduction **76**
3.2 Overall Objectives of the Current Thesis **79**
Chapter 4: Clinical Utility of DSM-5 Severity Ratings for Eating Disorders: Perspectives and Practices of Mental Health Clinicians **81**
4.1 Chapter Introduction **81**
4.2 Study 2 ..... **83**
Chapter 5: Comprehensive Comparison of DSM-5, ICD-11, and OWS Severity Ratings for Anorexia Nervosa: Psychological and Biological Correlates **104**
5.1 Chapter Introduction **104**
5.2 Study 3a. **107**
5.3 Additional analyses. **108**
5.4 Study 3b **118**
Chapter 6: Do Risk Factors Differentiate DSM-5and Drive for Thinness Severity Groups for Anorexia Nervosa? **119**
6.1 Chapter Introduction **119**
6.2 Study 4 ..... **121**
Chapter 7: Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 versus ICD–11 versus Drive For Thinness **122**
7.1 Chapter Introduction **122**
7.2 Study 5. ..... **124**
Chapter 8: Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample? **125**
8.1 Chapter Introduction **125**
8.2 Study 6. ..... **127**
Chapter 9: General Discussion **164**
9.1. Summary of Findings **164**
9.2. Discussion of Key Findings **168**
Objective 1: Provide a comprehensive and systematic synthesis of the current knowledge on ED severity ratings 168
Objective 2: Mental health clinicians' usage and perceptions about the usefulness of the proposed ED severity ratings in daily clinical practice. 170
Objective 3: Relationship between ED severity ratings (DSM-5, ICD-11 and alternative severity ratings) across a wide range of other ED-related features (e.g., risk factors, biological and psychological correlates) and treatment outcomes 172
Psychological variables. 173

Biological factors 177
Risk Factors 180
Treatment response 182
Objective 4: To evaluate DSM-5 and alternative severity ratings for EDs on a continuous scale within a community-based setting, aiming to enhance understanding of these ratings across the broader spectrum of disordered eating and EDs. 186
9.3. Implications 189
9.4. Strengths 194
9.5. Limitations and Future Directions 195
9.5.1 Methodology Limitations and Future Directions 195
9.5.2 Theoretical Limitations and Future Directions 198
9.6. Conclusions 201
References 202

List of Tables

Table 1 DSM-5 and ICD-11 Diagnostic Criteria for Anorexia Nervosa

Table 2 DSM-5 and ICD-11 Bulima Nervosa Diagnostic Criteria

Table 3 DSM-5 and ICD-11 Diagnostic Criteria for Binge Eating Disorder

Table 4 Summary of the current DSM-5, ICD-11 and Alternative Severity Ratings for EDs

Table 5 (Study 1; Table 1) Studies Investigated DSM-5 Severity Ratings for AN, BN and BED.

Table 6 (Study 1; Table 2) Alternative Classification Using Overvaluation of Weight and Shape

Table 7 (Study 1; Table 3) Measurements Used Across Studies

Table 8 (Study 1; Table 4) Downs and Black (1998) Checklist for Methodological Quality Adapted to Evaluate Studies Identified in the Systematic Review of the DSM-5 Severity Ratings for ED

Table 9 (Study 1; Table 5) Summary of Overall Effect Sizes, Heterogeneity, and Egger's Test P-value for Comparison Between DSM-5 AN, BN and BED Severity Groups

Table 10 (Study 2; Table 1) Summary of Participant Demographics

Table 11 (Study 2; Table 2) Demographic Characteristics of Interview Participants

Table 12 (Study 3a; Table 1) Demographic Information for the DSM-5 and Overvaluation of Weight/shape severity groups

Table 13 (Study 3a; Table 2) Clinical Profiles across the Anorexia Nervosa Severity Groups Based on the DSM-5 and Overvaluation of Weight/shape

Table 14 Demographic Information for the ICD-11 Anorexia Nervosa Severity Groups

Table 15 Clinical Profiles across the Anorexia Nervosa ICD-11 Severity Groups Based BMI

Table 16 (Study 3b; Table 1) Demographic Information for Each Severity Groups Within the DSM-5, ICD-11 and Overvaluation of Weight/Shape Severity Ratings for AN

Table 17 (Study 3b; Table 2) Zero-Inflated Poisson Regression for Severity Groups Contrasts Predicting Psychiatric Comorbidity

Table 18 (Study 3b; Table 3) Psychiatric Comorbidity Rate between DSM-5, ICD-11 and OWS Severity Groups.

Table 19 (Study 4; Table 1) Social and clinical related information of healthy sisters and individuals with AN

Table 20 (Study 4; Table 2) Distribution of risk factors for across the mild, the moderate, the severe and the extreme severe AN DSM-5/BMI severity groups, and high and low drive for thinness groups results using multinominal logistic regression, with healthy sisters as the reference group

Table 21 (Study 4; Table 3) Follow-up chi-square analyses on the risk factors that significantly contributed to the development of AN based on the DSM/BT and DT classification

Table 22 (Study 5; Table 1) Discriminative capacity for the severity groups at baseline: total sample

Table 23 (Study 5; Table 2) Discriminative capacity for the severity groups at baseline on the therapy outcome

Table 24 (Study 6; Table 1) Demographic characteristics of the total sample.

Table 25 (Study 6; Table 2) Effects of Order of Assessment, Time of Day, and Day of Week on Outcomes

Table 26 (Study 6; Table 3) Descriptive statistics for Level 1 and Level 2 variables

Table 27 (Study 6; Table 4) Fixed Effects for BMI and Weight Suppression, Predicting Outcomes and Moderation Effects of EAT-26 on the tested relationships.

Table 28 Summary of Findings on the relationship between Severity Ratings for AN, BN, and BED and Assessed Variables,

**Listed order according to the sequence in which they appear in the thesis.*

List of Figures

Figure 1 Continuum of Disordered Eating and Eating Disorders

Figure 2 Fairburn et al.'s (2003) Transdiagnostic model of Eating Disorders

Figure 3 Stice et al.'s (1998) Dual-Pathway Model for Bulimia Nervosa

Figure 4 (Study 1, Figure 1) PRISMA Flow Chart of Systematic Review and Meta-Analysis Selection Process

Figure 5 (Study 5, Figure 1) Prevalence of the severity groups in the study

*****_Listed order according to the sequence in which they appear in the thesis._

List of Abbreviations

AN	Anorexia Nervosa
ANOVA	Analysis of Variance
Atypical AN	Atypical Anorexia Nervosa
BED	Binge Eating Disorder
BMI	Body Mass Index
BN	Bulimia Nervosa
CBT	Cognitive Behavioural Therapy
CBT-E	Enhanced Cognitive Behavioural Therapy
DASS-21	Depression Anxiety and Stress Scale 21
DT	Drive For Thinness
DSM	Diagnostic and Statistical Manual of Mental Disorders
ED(s)	Eating Disorder(s)
EDE	Eating Disorders Interview
EDE-Q	Eating Disorders Examination Questionnaire
EDI-2	Eating Disorder Inventory-2
EMA	Ecological Momentary Assessment
FBT	Family Based Treatment
ICD	International Classification of Diseases
MINI	Mini International Neuropsychiatric Interview
NES	Night Eating Syndrome
NICE	National Institute of Health and Care Excellence
ORFI	Oxford Risk Factor Interview
OSFED	Other Specified Feeding or Eating Disorder
OWS	Overvaluation of weight and shape
PD	Purging Disorder
RANZCP	Royal Australian and New Zealand College of Psychiatrists
Sub-BED	Subthreshold Binge Eating Disorder
Sub-BN	Subthreshold Bulimia Nervosa
SEM	Structural Equation Modelling

Chapter 1: Eating Disorders Diagnosis: Main Characteristics and Outcomes

Chapter 1 provides a structured overview of anorexia nervosa (AN), bulimia nervosa (BN), and binge eating disorder (BED) – beginning with their definition, prevalence, onset, and associated burdens (Section 1.1). It then outlines the biological, psychological, and social risk factors associated with these eating disorders (EDs) (Section 1.2). Recognising that disordered eating exists on a continuum - from dieting behaviours to clinically diagnosed EDs - Section 1.3 focuses on defining disordered eating, highlighting its prevalence, and its consequences.

Two key theoretical models underpinning this thesis are introduced in Section 1.4: the transdiagnostic model of EDs (Fairburn et al., 2003) and the dual pathway model for BN (Stice et al., 1998; 2001). After discussing the course and treatment of AN, BN, and BED (Section 1.5), the concept of severity ratings is introduced in Section 1.6. Furthermore, in Section 1.6, the severity ratings outlined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013) and its text-revised version (DSM-5-TR; APA, 2022), and the 11th edition of the International Classification of Diseases (ICD-11; World Health Organization [WHO], 2019), as well as alternative disorder-specific measures (e.g., overvaluation of weight and shape) and transdiagnostic measures (e.g., drive for thinness), are outlined.

Finally, the chapter concludes by presenting the primary aim of this PhD thesis: to evaluate the DSM-5, ICD-11, and alternative severity ratings for EDs across a broad spectrum of variables, including psychological and biological correlates, risk factors, and treatment outcomes. The rationale for conducting a systematic review and meta-analysis on severity ratings (Chapter 2, Study 1) is also discussed, setting the stage for the research of this PhD to follow.

1.1 Eating Disorders Diagnosis, Prevalence, and Burden

EDs have a significant lifetime prevalence, affecting an estimated up to 8.4% of women and 2.2% of men globally (Galmiche et al., 2019; Schmidt et al., 2016). EDs involve severe disruptions in eating behaviours, which negatively impact both physical and psychological well-being (Schmidt et al., 2016). In Australia, recent data estimate that 10.46% of the population have experienced an ED at some point in their lives, equivalent to approximately 2,754,446 individuals (Deloitte Access Economics, 2024, p. 30).

In a recent study, research has found that EDs contributed 43.4 age-standardised disability-adjusted life years (DALYs) per 100,000 people globally, which represents the years of life lost due to premature death and the years lived with a disability (Hay et al., 2023). This marked a 9.4% increase in the disease burden from 2007 to 2017. In Australia, the economic burden was substantial, with $84 billion attributed to years of life lost from disability and death, alongside $1.65 billion in annual lost earnings (Hay et al., 2023).

Key features of EDs include behaviours that hinder maintaining a healthy minimum body weight, binge eating, compensatory actions (e.g., excessive laxative use or self-induced vomiting) to avoid weight gain and distortions in self-evaluation that are excessively influenced by weight, shape, and eating behaviours. Individuals with EDs have a significantly high mortality rate, with AN classified as the deadliest psychiatric disorder (Edakubo & Fushimi, 2020; Krug et al., 2024). Around 5% of those diagnosed with AN die within four years of receiving the diagnosis (Westmoreland et al., 2022). Additionally, relapse rates and chronic presentations of EDs are alarmingly high, with around 20% of individuals developing a chronic course of the illness despite treatment (Steinhausen et al., 2009). Recognising the consequences and chronicity of EDs, it is important to be able to identify and provide appropriate levels of interventions for affected individuals.

There are various types of EDs, as outlined in the DSM-5-TR (APA, 2022) and the ICD-11 including AN, BN, BED, Other Specified Feeding or Eating Disorders (OSFED), and Unspecified Feeding and Eating Disorders (UFED). Other disorders, including as Pica, Rumination Disorder, and Avoidant/Restrictive Food Intake Disorder, are also classified as Feeding and Eating Disorders within the DSM-5 and ICD-11, but are believed to have distinct underlying pathophysiology and patterns of psychopathology (e.g., no role of body weight or shape; APA, 2022) and are not the focus of the present thesis.

1.1.1 Anorexia Nervosa

Diagnostic, Prevalence and Burden

AN is a distinct and severe mental disorder that can impact individuals across all demographics, although adolescent girls and young adult women are especially vulnerable (Hoek et al., 2006; van Eeden et al., 2023). Both the DSM-5 and ICD-11 characterise AN by a refusal to maintain a minimal body weight, an intense fear of gaining weight, and distorted perceptions of one's body weight or shape (please see Table 1 for a direct comparison between ICD-11 and DSM-5 AN diagnostic criteria). Both diagnostic classifications also categorise AN into two subtypes: the AN restrictive subtype (AN-R), characterised by weight-control behaviours focused on calorie restriction and excessive exercise, and the AN binge-eating/purging subtype (AN-BP), which involves episodes of binge eating followed by purging to maintain a low body weight. The main difference between the DSM-5 and ICD-11 AN diagnostic criteria is that while both specify diagnostic criteria for significantly low body weight due to restricted energy intake, only the ICD-11 sets specific thresholds: Body mass index (BMI) below 18.5 kg/m² in adults and BMI-for-age under the 5th percentile in children and adolescents.

AN frequently begins during adolescence, with the lifetime prevalence estimated at 0.16%, ranging from 0.06% to 0.31% (Qian et al., 2021). However, the actual number of

cases is likely higher due to many individuals not seeking medical attention (Van Eeden et al., 2021). Within Australia, the estimated prevalence is around 1.8% (Deloitte Access Economics, 2024, p.30). The global disease burden for AN was 33.8 age-standardised DALYs per 100,000, reflecting a 6.1% increase from 2007 to 2017 (Hay et al., 2023). AN is significantly more prevalent in females than in males, with risk ratios in lifetime prevalence studies often exceeding 20:1 (Lindberg & Hjern, 2003; Nicholls & Viner, 2009). However, this gender difference may partly stem from lower detection rates of the disorder in males (Griffiths et al., 2015), as most ED research has largely focused on girls and women (Brown et al., 2023).

Table 1

DSM-5 and ICD-11 Diagnostic Criteria for Anorexia Nervosa

DSM-5 Anorexia Nervosa Criteria	**ICD-11 Anorexia Nervosa Criteria**
- Restriction of body weight leading to significantly low body weight,	- Significantly low body weight for the individual's height, age, developmental stage, or weight history. A commonly used threshold is body mass index (BMI) less than 18.5 kg/m2 in adults and BMI-for-age under 5th percentile in children and adolescents. Rapid weight loss (e.g., more than 20% of total body weight within 6 months) may replace the low body weight essential feature as long as other diagnostic requirements are met.
- Intense fear of gaining weight or persistent behaviour that interferes with body weight,	- Low body weight is not better accounted for by another medical condition or the unavailability of food.
- Disturbance in way ones shape or weight is experienced, or, undue influence of shape or weight on self-evaluation, or, lack of recognition of seriousness of current weight	- A persistent pattern of restrictive eating or behaviours to maintain an abnormally low body weight, often driven by an intense fear of weight gain. These behaviours may include fasting, consuming low-calorie foods, eating very slowly or minimally, spitting out food, self-induced vomiting, or using laxatives, diuretics, enemas, or skipping insulin doses (in diabetes). Additionally, individuals may engage in excessive exercise, motor hyperactivity, exposure to cold, or use stimulants, weight loss medications, or thyroid hormones to increase energy expenditure.
	- Excessive preoccupation with body weight or shape, often tied to self-evaluation or distorted perceptions of body size. This may manifest as frequent weight or body checks, calorie monitoring,

	seeking weight-loss information, or avoidant behaviours like avoiding mirrors, tight clothing, or knowledge of one's weight or clothing size.
DSM-5 Anorexia Nervosa Subtypes	**ICD-11 Anorexia Nervosa Subtypes**
Restricting: Individual has not engaged in recurrent binge-purge behaviours during the last three months **Binge eating/purging:** Individual has engaged in recurrent binge-purge behaviours during the last three months	**Restricting:** Individuals induce weight loss and maintain low body weight through restricted food intake or fasting alone or in combination with increased energy expenditure (but who do not engage in binge eating or purging behaviours. **Binge eating/purging:** Individuals who present with episodes of binge eating or purging behaviours aimed at getting rid of ingested food This type of AN also includes individuals who exhibit binge eating episodes but do not purge.

Note. Table adapted from the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (APA, 2013) and International Classification of Diseases, 11th edition (WHO, 2019).

Co-Occurring Medical, Psychiatric Conditions, and Consequences of AN

Restricted energy intake and weight-control practices in AN result in profound physical and psychological consequences (Meneguzzo et al., 2023). Physically, AN can affect the functioning of the entire body. Up to 42% of individuals with AN experience haematological and electrolyte disturbances due to dietary restrictions and purging (Baenas et al., 2024; Caregaro et al., 2005; Cass et al., 2020). These disturbances can have serious health implications, often requiring urgent medical care. For instance, reduced levels of white blood cells, haemoglobin, potassium, magnesium, and phosphate are linked to severe complications like increased infection rates, cardiac muscle weakness or paralysis, and cardiac arrhythmias (Caregaro et al., 2005; Puckett et al., 2021).

Individuals with AN face significant psychological challenges, including a high prevalence of psychiatric comorbidities, reduced quality of life, and substantial disability (Miskovic-Wheatley et al., 2023). These challenges often result in frequent absences from work and education. Psychiatric comorbidities are present across all stages of the illness and may even persist after recovery (Nagy et al., 2023). Research indicates that at least 80% of individuals with AN have at least one co-occurring psychiatric disorder (Andrés-Pepiñá et al.,

2020; Herzog et al., 2018; Nagy et al., 2023). Common co-occurring conditions include major depressive disorder, which affects up to 70% of individuals with AN (Calvo-Rivera, 2022), and anxiety disorders, which are observed in 21% to 71% of cases, including a high prevalence of obsessive-compulsive disorder (Riquin et al., 2021). Additionally, up to 16% experience substance use disorders (Devoe et al., 2021), while personality disorders, including borderline personality disorder, are reported in up to 30% of cases (Himmerich et al., 2019; Laczkovics et al., 2023; Miller et al., 2021; Riva et al., 2024). A link between AN and autism spectrum disorders has been proposed, with overlapping socioemotional difficulties observed in AN, though these are typically less severe than those seen in autism spectrum disorders (Boltri & Sapuppo, 2021). These comorbidities significantly influence AN's clinical presentation, prognosis, hospitalisation course, and treatment decisions (Catone et al., 2023).

Death is a significant risk in AN, which has the highest mortality rate among all mental health disorders–five times higher than in the general population (Arcelus et al., 2011; Patel et al., 2023; Krug et al., 2024). Deaths are caused by physical complications from severe low weight (Mehler et al., 2023; Millar et al., 2005; Westermair et al., 2024), psychiatric comorbidities (Soeby et al., 2024), and suicide (Tseng et al., 2023). Suicide is the second leading cause of death in individuals with AN, after cardiovascular complications, with those affected being over 31 times more likely to attempt suicide compared to the general population (Preti et al., 2011; Amiri et al., 2023; Smith et al., 2018).

Finally, AN also severely impacts interpersonal functioning. Isolation is common as individuals avoid food-related gatherings and withdraw from relationships, while conflicts over eating habits or treatment strain family, friendships, and romantic partnerships (Fletcher et al., 2021). Cognitive impairments and the demands of treatment further disrupt education and work, compounding the disorder's far-reaching social and functional consequences (Tan

et al., 2022). Overall, AN is a serious and debilitating mental health condition that profoundly affects not only individuals but also their families, with wide-ranging social and functional consequences.

1.1.2 Bulimia Nervosa

Diagnostic, Prevalence and Disorder Burden

BN is recognised in both the DSM-5 and ICD-11 (Table 2) as an ED associated with significant psychological distress and functional impairment (Milano et al., 2018; Tith et al., 2020; Steinhausen et al., 2020). It is characterised by at least one binge eating episode per week lasting three months or longer, during which individuals consume a large quantity of food, far exceeding what most people would eat in similar situations, accompanied by a loss of control over overeating. To counteract weight gain, individuals engage in compensatory behaviours such as self-induced vomiting, misuse of laxatives or diuretics, fasting, or excessive exercise. The main difference between the DSM-5 and ICD-11 diagnostic criteria for BN is that the DSM-5 requires binge eating and compensatory behaviours to occur at least once a week for three months, while the ICD-11 requires them only once a week for one month.

The lifetime prevalence of BN is up to 0.63%, with reported rates ranging from 0.33% to 1.02% (Qian et al., 2021). In Australia, the lifetime prevalence of BN is estimated to be 1.85% (Deloitte Access Economics, 2024, p.30). Like AN, BN commonly emerges in adolescence or young adulthood (Hudson et al., 2007). The global disease burden for BN was at 9.5 age-standardised DALYs per 100,000, reflecting a 10.3 % increase from 2007 to 2017. This highlights the significant prevalence and growing global burden of BN, underscoring its substantial impact on public health

Table 2

DSM-5 and ICD-11 Bulima Nervosa Diagnostic Criteria

DSM-5 Bulimia Nervosa Criteria	**ICD-11 Bulimia Nervosa Criteria**
- Recurrent episodes of binge eating, involving eating a very large amount of food for the given circumstance within a given time period, accompanied by loss of control over eating	- Frequent, recurrent episodes of binge eating (e.g., once a week or more over a period of at least 1 month). Binge eating is defined as a distinct period of time during which the individual experiences a loss of control over his or her eating behaviour. A binge eating episode is present when individual eats notably more and/or differently than usual and feels unable to stop eating or limit the type or amount of food eaten. Other characteristics of binge eating episodes may include eating alone because of embarrassment, eating foods that are not part of the individual's regular diet, eating large amounts of food despite not feeling hungry, and eating faster than usual.
- Recurrent inappropriate compensatory behaviours to avoid weight-gain	- Repeated inappropriate compensatory behaviours to prevent weight gain (e.g., once a week or more over a period of at least 1 month). The most common compensatory behaviour is self-induced vomiting, which typically occurs within an hour of binge eating. Other inappropriate compensatory behaviours include fasting or using diuretics to induce weight loss, using laxatives or enemas to reduce the absorption of food, omission of insulin doses in individuals with diabetes, and strenuous exercise to greatly increase energy expenditure.
- Binge eating and compensatory behaviours occur on average at least once per week for 3 months	- Excessive preoccupation with body weight and shape. When not explicitly stated, preoccupation with weight and shape may be manifested by behaviours such as repeatedly checking body weight using scales, checking one's body shape using tape measures or reflection in mirrors, constant monitoring of the calorie content of food and searching for information on how to lose weight or by extreme avoidant behaviours, such as refusal to have mirrors at home, avoidance of tight-fitting clothes, or refusal to know one's weight or purchase clothing with specified sizing.
- Self-evaluation is unduly influenced by body weight or shape	- There is marked distress about the pattern of binge eating and inappropriate compensatory behaviour or significant impairment in personal, family, social, educational, occupational or other important areas of functioning.
- The disturbance does not occur exclusively during episodes of anorexia nervosa	- The symptoms do not meet the definitional requirements for Anorexia Nervosa.

Note. Table adapted from the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (APA, 2013) and International Classification of Diseases, 11th edition (WHO, 2019).

Co-Occurring Medical, Psychiatric Conditions, and Consequences of BN

The health consequences of BN vary widely, ranging from minor biological and physical damage to severe and life-threatening conditions (Mehler & Rylander, 2015; Milano et al., 2018; Sagar et al., 2005). Specifically, BN can lead to significant health issues, including gastrointestinal tract injury, skin problems, hair loss, elevated serum amylase levels, metabolic acidosis, and hormonal imbalances that cause amenorrhea in women (Milano et al., 2018). It can also result in electrolyte imbalances, such as hypokalaemia, hyponatremia, and hypochloremia (Milano et al., 2018; Mehler & Rylander, 2015). These imbalances, due to compensatory behaviours, may lead to cardiac arrhythmias and potentially be life-threatening (Milano et al., 2018).

Psychiatric comorbidities are highly prevalent among individuals with BN, closely resembling the patterns seen in AN. Research shows that up to 95% of individuals with BN have at least one comorbid disorder, with 64% experiencing three or more (Howard et al., 2020; Keshishian et al., 2021). Anxiety disorders are the most common, affecting up to 65% of individuals (National Eating Disorders Association, 2022; Mashalpourfard et al., 2018; Jakovina et al., 2018), followed by major depressive disorder, which impacts up to 33% (National Eating Disorders Association, 2022; Mashalpourfard et al., 2018). Additionally, 30% to 50% of individuals with BN are affected by substance use disorders, with tobacco, caffeine, alcohol, sedatives, and cannabis being the most used substances (Bahji et al., 2019). The link between substance uses and BN is rooted in shared characteristics such as impulsivity, emotional dysregulation, and negative urgency (Schaefer et al., 2021). Both behaviours, substance use and disordered eating, function as maladaptive coping mechanisms, offering temporary relief from emotions like anxiety, guilt, or shame (Claudat et al., 2020). BN is also associated with a heightened risk of mortality, either from medical complications or suicide (Gibson et al., 2019; Krug et al., 2024; Lieberman et al., 2023).

Suicide risk is linked to impulsivity, feelings of hopelessness, and co-occurring psychiatric conditions (Lieberman et al., 2023; Van Eeden et al., 2021; Himmerich et al., 2019). BN also significantly impacts social functioning, as individuals frequently experience feelings of shame, guilt, and preoccupation with weight and body shape. These challenges often interfere with the ability to form healthy relationships, succeed academically, and maintain stable employment (Keel & Brown, 2010; Kessler et al., 2014). In summary, the interplay of psychological, social, and physical factors underscores the severe and multifaceted impact of BN.

1.1.3 Binge Eating Disorder

Diagnostic, Prevalence and Disorder Burden

BED was formally included as an official ED diagnosis in the DSM-5 and later in the ICD-11. BED is primarily characterised by recurrent episodes of binge eating, during which individuals consume unusually large amounts of food within a short period of time (such as within 2 hours) and feel a lack of control over their eating. For a DSM-5 BED diagnosis, these binge episodes must occur at least once a week for a minimum of three months and cause significant distress (see Table 3). The DSM-5 also specifies that binge-eating episodes must involve at least three of the following five behaviours: eating much faster than usual, eating until feeling uncomfortably full, eating when not physically hungry, eating alone due to embarrassment about how much is being consumed, and feeling negative emotion after a binge episode. BED shares the feature of regular binge-eating episodes with BN; however, BN includes the frequent use of inappropriate compensatory behaviours, like self-induced vomiting or fasting, to avoid weight gain – behaviour not commonly seen in BED. Additionally, while AN and BN both involve body image disturbances, such as an overemphasis on weight and shape, these are not required for a BED diagnosis. The ICD-11

criteria for BED closely align with those of the DSM-5 (see Table 3). However, the ICD-11 also includes broader criteria that encompass individuals with subjective binge eating.

The lifetime prevalence of BED is estimated to be 0.63%, ranging from 0.33% to 1.02%. (Qian et al., 2021). In Australia, BED has the second-highest lifetime prevalence among EDs at 2.2% (Deloitte Access Economics, 2024, p. 30). BED also has the latest average onset age of any ED, typically around 25 years old (Butterfly Foundation, 2012). In 2019, BED accounted for 0.8 million DALYs globally, equating to 9.8 age-standardised DALYs per 100,000 (Santomauro et al., 2021). These findings highlight the substantial public health impact of BED and the need for early detection and targeted interventions

Table 3

DSM-5 and ICD-11 Diagnostic Criteria for Binge Eating Disorder

DSM-5 Binge Eating Disorder Criteria	**ICD-11 Binge Eating Disorder Criteria**
- Recurrent episodes of binge eating, involving eating a very large amount of food for the given circumstance within a given time period, accompanied by loss of control over eating	- Frequent, recurrent episodes of binge eating (e.g., once a week or more over a period of 3 months). Binge eating is defined as a distinct period of time during which the individual experiences a loss of control over his or her eating behaviour. A binge eating episode is present when an individual eats notably more or differently than usual and feels unable to stop eating or limit the type or amount of food eaten. Other characteristics of binge eating episodes may include eating alone because of embarrassment or eating foods that are not part of the individual's regular diet.
- Binge eating episodes associated with three or more of the following: Eating much more than rapidly normal, eating until feeling uncomfortably full, eating large amounts when not physically hungry, eating alone due to embarrassment regarding how much one is eating, and feeling disgusted with oneself, depressed, or very guilty afterward .	- The binge eating episodes are not regularly accompanied by inappropriate compensatory behaviours aimed at preventing weight gain.
- Marked distress regarding binge eating is present	- The symptoms and behaviours are not better explained by another medical condition (e.g., Prader-Willi Syndrome) or another mental disorder (e.g., a depressive disorder) and are not due to the effect of a

	substance or medication on the central nervous system, including withdrawal effects.
- Binge eating occurs on average once per week for 3 months	- There is marked distress about the pattern of binge eating or significant impairment in personal, family, social, educational, occupational or other important areas of functioning.
- Binge eating is not associated with recurrent use of inappropriate compensatory behaviour	

Note. Table adapted from the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (APA, 2013) and International Classification of Diseases, 11th edition (WHO, 2019).

Co-Occurring Medical, Psychiatric Conditions, and Consequences of BED

BED poses significant physical and psychological health risks. Physically, the recurrent binge eating associated with BED increases the likelihood of obesity (BMI ≥ 30.0 kg/m²) and related conditions such as type 2 diabetes, hypertension, and cardiovascular disease (Keski-Rahkonen et al., 2021; Marzilli et al., 2018). Additionally, individuals with BED commonly experience gastrointestinal issues, including bloating and acid reflux, as well as a higher prevalence of sleep disorders like sleep apnoea (Giel et al., 2022; Keski-Rahkonen et al., 2021). Other associated conditions include asthma and chronic pain, further highlighting the substantial health burden of BED (Giel et al., 2022).

Psychologically, individuals with BED often experience feelings of shame, guilt, and low self-esteem, which can perpetuate the cycle of binge eating and emotional distress (Liu et al., 2023; Ratković et al., 2023). Comorbid psychiatric disorders are highly prevalent, with 94% of individuals meeting criteria for at least one additional psychiatric disorder (Udo and Grilo, 2019; Udo et al., 2019). Common comorbidities include mood disorders (70%), post-traumatic stress disorder (32%), and anxiety disorders (16%). BED is also frequently associated with disorders involving poor impulse control, such as borderline personality disorder (Udo and Grilo, 2019), alcohol use disorder (Udo and Grilo, 2019), gambling disorder (Giel et al., 2022; Mestre-Bach & Potenza, 2023), and attention-deficit/hyperactivity disorder (Hanson et al., 2020; Villa et al., 2023). Death is also risk for BED, either due to

health complications from being overweight or obese or psychiatric comorbidity (Giel et al., 2022). Suicidality is prevalent among individuals with BED, with suicidal ideation reported in 26.3% to 51.7% of cases and suicide attempts ranging from 2.3% to 34% (Mandelli et al., 2019; Bulik et al., 2021).

The psychological and physical consequences of BED often extend to social domains, leading to strained relationships and social isolation as individuals may avoid interactions due to fear of judgment or embarrassment about their eating behaviours (Marzilli et al, 2018). This social withdrawal, combined with the emotional toll and stigma of the disorder, frequently results in difficulties across work, education, and personal life (e.g., Hilbert et al., 2018; Pawaskar et al., 2017). Altogether, this highlights the far-reaching impact of BED.

1.1.4 Other Specified Feeding or Eating Disorders

Diagnostic, Prevalence and Disorder Burden

OSFED is diagnosed when individuals experience clinically significant distress or impairment related to eating and feeding behaviours but do not meet the full criteria for a specific EDs (APA, 2013). It is recognised as a formal diagnostic category in both the DSM-5 and ICD-11. While the ICD-11 does not specify subtypes for OSFED, the DSM-5 includes five subtypes under this category. These subtypes are atypical AN, subthreshold BN (Sub-BN), subthreshold BED (Sub-BED), purging disorder (PD), and night eating syndrome (NES). Atypical AN refers to individuals who meet all the criteria for AN except that their weight remains within or above the normal range. Sub-BN and sub-BED describe cases where the symptoms of BN and BED are present but occur at a lower frequency or over a shorter duration than outlined in the DSM-5. PD is characterised by recurrent purging behaviours, such as vomiting or laxative use, aimed at influencing weight or shape, without the presence of binge eating episodes. NES involves recurring episodes of eating at night that

are not culturally appropriate and cannot be fully explained by external factors or other eating disorders.

Given the majority of ED diagnoses in clinical practice fall under the atypical/subthreshold ED category, making OSFED the most prevalent ED (Santomauro et al., 2021). Studies have found that up to 12.1% of adolescents (Smink et al., 2014) and 7.4% of adults (Micali et al., 2017) meet the criteria for an OSFED diagnosis. Prevalence rates for specific OSFED subtypes are limited, but some studies report lifetime prevalence rates ranging from 1.1% to 5.3% for PD, 2.8% to 3.6% for atypical AN, 4.4% for subthreshold BN, and 1.5% to 5.7% for NES (Murray & Anderson, 2015; Hammerle et al., 2016). Within Australia, out of 1,102,977 Australians with an ED, 27% have suffered from OSFED (Deloitte Access Economics, 2024). The global disease burden of OSFED is estimated 38.5 standardised-age DALYs per 100,000 (Santomauro et al., 2021), reflecting the substantial impact these disorders have on both mortality and years lived with disability. Furthermore, the inconsistent definitions of OSFED subtypes contribute to a lack of clarity in both research and clinical practice, limiting our understanding of these subtypes (Dang et al., 2024). This uncertainty hampers the development of targeted interventions and results in the underrepresentation of OSFED in research, despite its high prevalence and significant impact on individuals' health and quality of life (Krug et al., 2024).

Consequences of OSFED

A plethora of research, including systematic reviews and meta-analyses (Dang et al., 2024; Walsh et al., 2023), has found that individuals with atypical AN or subthreshold BN may experience levels of ED psychopathology comparable to, or even greater than, those with full-threshold disorders like AN or BN (Chapa et al., 2018; Golden & Mehler, 2020). Although the physiological effects may be more severe in full-threshold BN due to higher levels of bulimic or binge eating behaviours and associated medical comorbidities,

individuals with OSFED often seek treatment only after a prolonged illness (Golden & Mehler, 2020; Wilkop et al., 2023). They are less likely to receive inpatient care, partly due to the lower recognition of the condition when individuals are within a normal weight range (Chapa et al., 2018). This delay in treatment can exacerbate impairment and medical complications for those with OSFED.

Individuals with OSFED commonly experience psychiatric comorbidities, with major depressive disorder (43%) and obsessive-compulsive disorder (21%) being the most prevalent (Hambleton et al., 2022). Other psychiatric symptoms, such as those associated with post-traumatic stress disorder, phobias, and psychosis, are also frequently reported (Hambleton et al., 2022; Milano et al., 2020). Although research on suicidal ideation and attempts in OSFED remains limited, the disorder poses a significant mortality risk (Arcelus et al., 2011). Studies suggest that individuals with OSFED face a comparable likelihood of dying from their EDs as those with AN, BN, and BED (Hambleton et al., 2022). A report by Deloitte Access Economics (2020) further found that approximately 33% of ED-related deaths, whether due to medical complications or suicide, were associated with OSFED. Similarly, a meta-analysis by Krug et al. (2005) revealed that mortality rates for OSFED are highly elevated, second only to those for AN – emphasising the severe impact of this disorder.

OSFED, like AN, BN and BED, frequently causes significant psychological and physical challenges that ripple into social aspects of life. Individuals with OSFED may distance themselves from others, avoiding social interactions out of fear of criticism or shame surrounding their eating behaviours (Zaccagnino et al., 2018). This tendency toward isolation, combined with the emotional strain and societal stigma tied to the disorder, often leads to challenges in maintaining healthy relationships and navigating professional, academic, and personal environments (e.g., Bhattacharya et al., 2020; Milano et al., 2020).

These impacts underscore that OSFED is just as severe as AN, BN and BED, warranting equal research and clinical attention.

1.2 Risk Factors for Eating Disorders

To effectively identify and provide effective interventions to address EDs, it is essential to understand the factors that increase vulnerability to these conditions. ED development is influenced by a combination of sociocultural, biological, genetic, and psychological factors (Striegel-Moore & Bulik, 2007). While the causes of EDs are complex and varied, research has identified common risk factors that predict their onset. Several studies have indicated that factors such as pressure to attain a thin body, striving for the thin ideal, body dissatisfaction, parental overweight, dieting behaviours, and negative emotional states are predictive of the future development of any ED (Barakat et al., 2023; Dakanalis et al., 2023; Fatt et al., 2024). Recognising these risk factors is key to identifying at-risk individuals and implementing effective screening, prevention, and targeted interventions (Barakat et al., 2023). The following paragraphs will detail the risk factors for AN, BN and BED – the three main EDs that are the focus of the current thesis.

1.2.1 Risk Factors for Anorexia Nervosa

Extensive reviews have highlighted various risk factors associated with AN. AN is influenced by genetic factors, with familial patterns commonly observed (Paolacci et al., 2020). Family members of individuals with AN are four times more likely to develop the disorder, and female relatives face an 11-fold increased risk compared to the general population (Zipfel et al., 2015). Genetic studies indicate that heredity contributes to 50-74% of the risk for developing AN (Moskowitz & Weiselberg, 2017), and monozygotic twins have a higher likelihood of developing AN than dizygotic twins (Yilmaz et al., 2015). This heritable risk is even more pronounced when strict diagnostic criteria are used to define AN, as opposed to broader definitions (Paolacci et al., 2020).

Several genome-wide association studies (GWAS) on AN have identified genetic associations, including a locus on chromosome 12 (rs4622308) linked to autoimmune disorders and diabetes, as well as eight chromosomal regions containing 120 genes (Watson et al., 2019; Paolacci et al., 2020; Yokokura et al., 2019). Key candidates include CADM1, MGMT, FOXP1, PTBP2, and loci linking AN risk to increased BMI, such as CTBP2, CCNE1, CARF, and NBEAL1 (Zeeland et al., 2014; Watson et al., 2019; Paolacci et al., 2020; Yokokura et al., 2019). Despite challenges like small sample sizes and cohort heterogeneity, these findings suggest that AN is a highly polygenic disorder (Donato et al., 2022).

In addition to a genetic predisposition, the interplay between genes and environmental factors has been widely studied, with particular focus on the serotonin-transporter-linked polymorphic region (5-HTTLPR) (Rozenblat et al., 2017). Karwautz et al. (2011) demonstrated that problematic parenting styles increased the risk of AN, but only in individuals carrying one or two copies of the 5-HTTLPR s-allele, while those homozygous for the l-allele did not exhibit the same risk. This highlights the critical role of gene-environment interactions in AN development.

Environmental factors can also lead to epigenetic changes in gene expression (Käver et al., 2024). Although research on epigenetics in AN is still in its early stages, preliminary findings suggest that significant alterations occur during acute AN, potentially influencing the disease's maintenance and chronicity. These changes appear to be highly reversible upon recovery (Hübel et al., 2019; Käver et al., 2024). However, larger international studies are needed to validate these findings and account for various influencing factors (Käver et al., 2024; Hirtz et al., 2020).

In addition to genetic predispositions, certain environmental and psychological factors increase the risk of AN. Longitudinal research has identified key predictors of AN onset in

high-risk individuals, including a low BMI, weight suppression, an overemphasis on weight and body shape, intense fear of weight gain, negative emotional states, and psychosocial difficulties (Stice et al., 2017; 2020; 2021). Several systematic reviews and meta-analyses suggest that individuals with AN exhibit greater severity and higher rates of perfectionism (Dahlenburg et al., 2019) and family discord (Del Casale et al., 2023) than those with other psychiatric disorders. Additional risk factors include limited socialisation during childhood (Krug et al., 2012), exposure to adverse childhood events, early childhood eating problems, and increased parental demands (Bulik et al., 2006; Jacobi et al., 2004; Lavinia-Maria et al., 2020). Furthermore, women with autism are at a notably higher risk of developing AN, with research showing that 20–35% of women with AN meet diagnostic criteria for autism (Boltri & Sapuppo, 2021; Carpita et al., 2022; Kerr-Gaffney et al., 2021).

Social factors also play a crucial role in the development of AN. In the case of AN, cultural ideals of "thinness" and the normalisation of dieting drive the focus on physical appearance. Falling short of these rigid standards breeds shame, often triggering the onset of ED (Mitchell et al., 2024). Media influences with peer diet, peer criticism, and poor social support have also been found to further fuel these ideals, reinforcing societal pressure to conform to the "perfect" body image (Del Casale et al., 2023; Haynos et al., 2016). Altogether, this highlights various risk factors that contribute to the susceptibility to developing AN.

1.2.2 Risk Factors for Bulimia Nervosa

Genetic, psychological and social influences have also been found to be significant risk factors in the development of BN. In terms of the genetic contribution, familial and twin population-based studies have shown that approximately 50% of the genetic risk factors for BN overlap with those for AN, while the remaining factors are unique to each disorder (Barakat et al., 2023; Bulik et al., 2010; Yao et al., 2021). Individuals with a family member

diagnosed with BN are 9.6 times more likely to develop the disorder themselves (Thornton et al., 2011). Similar to AN, 5-HTTLPR has also been observed to contribute to compulsive personality traits and the development of BN (Calati et al., 2011; Steiger et al., 2012). Additionally, a polymorphism in the neurotransmitter inhibition gene (HTR1B) has been linked to an increased risk of developing BN in a family-based study (Hernández et al., 2016). Notably, no GWAS have been conducted on BN to date.

The interplay between genes and environmental factors has also been highlighted in BN. For example, research has shown that childhood trauma may contribute to BN risk, particularly through genetic variations in the glucocorticoid receptor pathway, which are associated with binge eating and a loss of control overeating (Barakat et al., 2023; Sardahaee et al., 2017; Steiger et al., 2012). Monteleone et al. (2017) further supported this finding, showing that individuals with a history of childhood maltreatment and BN had lower cortisol levels compared to both healthy controls and those with BN but no history of trauma.

Shifting focus to psychological and environmental factors, only a limited number of prospective longitudinal studies have been conducted to identify risk factors BN. These studies have identified risk factors such as body dissatisfaction and negative affect (Stice et al., 2021b; Yamamiya & Stice, 2024), dieting and fasting behaviours (Stice et al., 2008; Yamamiya & Stice, 2022; 2024), concerns about weight and eating (Allen et al., 2016), and low parental support (Yamamiya & Stice, 2024). Cross-sectional and retrospective studies, on the other hand, highlight common correlations to the development of BN, including teasing or critical remarks about weight or shape (Day et al., 2022; Dahill et al., 2021), obstetric complications (e.g., low birth weight; Larsen et al., 2021; Milembamane et al., 2023; Raevuori et al., 2014) and experiences of sexual abuse (Monteleone et al., 2022; Hilbert et al., 2014).

Sociocultural factors also play a critical role in the development of BN. Similar to AN, pressure to conform to societal ideals of thinness drives body dissatisfaction, which often leads to dieting – a behaviour that increases the risk of bingeing and purging (Barakat et al., 2023; Weissman et al., 2019). This dissatisfaction is further intensified through social comparison with peers or celebrities, while family dynamics that focus on appearance or criticise weight add to the risk (Diedrichs et al., 2017; Jaeger et al., 2002). Cultural attitudes that normalise dieting and weight control behaviours also reinforce these pressures. When combined with personal vulnerabilities, these sociocultural influences significantly heighten the likelihood of developing BN (Weissman et al., 2019).

Overall, the development of BN has been linked to a combination of biological psychological, social and environmental factors. Notably, many of the risk factors and traits associated with BN such as concerns about weight and eating, dieting, exposure to critical comments about body shape and negative affect have also been identified in AN (Jacobi et al., 2004), suggesting a common underlying pathology between the two disorders. A key difference between AN and BN, however, is that BN often develops following a pattern of binge eating, whereas AN is more commonly preceded by restrictive dieting behaviours (Hilbert et al., 2014).

1.2.3 Risk Factors for Binge Eating Disorder

BED, like AN and BN, arises from a complex interplay of biological, psychological, environmental and sociocultural factors. Twin studies estimate the heritability of BED to be between 41% and 57%, independent of obesity status (Hilbert et al., 2019; Thornton et al., 2011; Himmerich et al., 2019). Although research on candidate genes related to reward processing, homeostasis, and mood regulation is limited, some evidence highlights the potential role of dopamine (e.g., DRD2) and μ-opioid receptor (e.g., OPRM1) genes in BED (Hilbert, 2019). Specifically, researchers suggest that the expression of certain alleles in the

dopamine D2 receptor may heighten sensitivity to reward, a factor that could be critical in the development of BED (Hilbert et al., 2019).

Research on gene-environment interactions in BED is scarce, with most findings proving nonsignificant (Barakat et al., 2023). For example, Racine et al. (2009) found no evidence that 5-HTTLPR or the HTR2A (T102C polymorphism) interact with traits like impulsivity or dietary restraint to predict binge or emotional eating symptoms. To our best knowledge, to date, Veldic et al. (2017) conducted the only epigenetic study on BED, examining promoter methylation of SLC1A2, a gene involved in glutamate clearance, in individuals with bipolar disorder. It found reduced DNA methylation in individuals with bipolar disorder and BED compared to those with bipolar disorder alone, though the inclusion of both BN and BED cases complicates the interpretation. Additionally, no GWAS or broader epigenetic studies on BED have been published to date.

Since BED was first introduced in the DSM-5, research on its psychological risk factors has expanded, though prospective longitudinal studies remain scarce (e.g., Braun et al., 2019; Coffino et al., 2019). Existing longitudinal studies suggest that pressure to be thin, thin-ideal internalisation, negative affect, and low parental support are key contributors to BED risk (Stice et al., 2011; 2021; Yamamiya & Stice, 2024). Cross-sectional and retrospective studies have further linked BED risk to feelings of inadequacy stemming from childhood experiences, such as weight-related teasing or family concerns about eating and weight (Monteleone et al., 2022; Rubin et al., 2021). Other contributing factors include conflict avoidance, overvaluation of weight and shape, low self-esteem, and impulsive or obsessive tendencies (Braun et al., 2019; Coffino et al., 2019; Kukk & Akkermann, 2022).

Social and environmental factors further contribute to BED risk. Cultural ideals equating thinness with attractiveness and health, along with the normalisation of dieting, heighten vulnerability (Weissman et al., 2019). Food insecurity plays a distinct role,

triggering survival instincts that prompt individuals to binge or hoard food when it becomes available (Stice et al., 2017). Such behaviours are especially common among children from unstable environments who, after being placed in care, may binge or hide food as a coping mechanism for past scarcity (Borg et al., 2022). These survival-driven patterns often persist even after achieving food security, with individuals who have experienced trauma, homelessness, or financial hardship remaining at elevated risk for BED due to the long-lasting psychological impact of food instability (Hazzard et al., 2020).

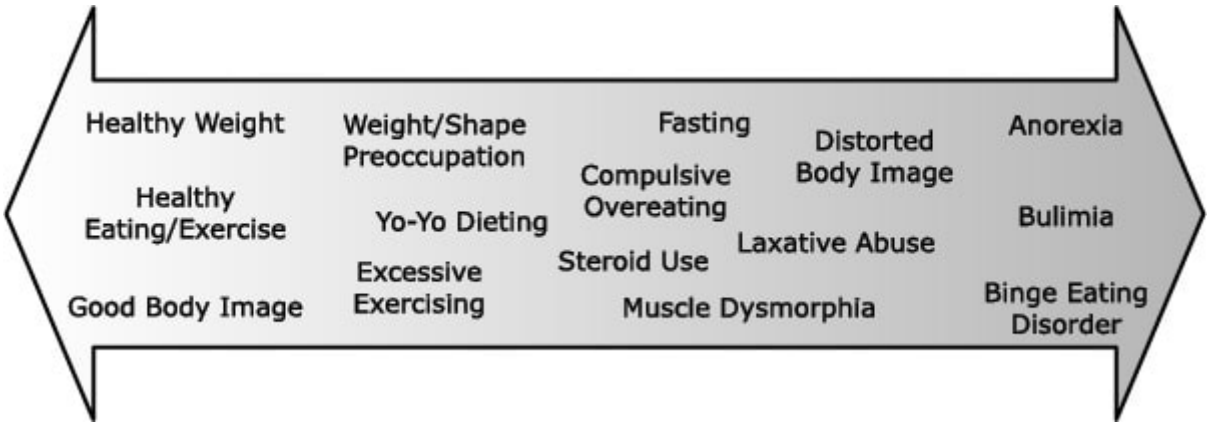
The significant overlap in aetiological factors among EDs, particularly between AN and BN and between BN and BED, highlights their interconnected nature (Hilbert et al., 2014). AN and BN share disruptions in eating behaviours rooted in body image concerns, while BN and BED are linked by binge eating accompanied by a loss of control. Diagnostic cross-over is common, with most transitions occurring from AN to BN (23-27%), followed by BN to BED (8-11%), BN to AN (8-9%), and BED to BN (7-8%) (Castellini et al., 2011; Miskovic-Wheatley et al., 2023). This overlap complicates the identification of risk factors unique to each DSM-5 diagnosis. Instead, focusing on broader disordered eating behaviours, such as a drive for thinness or binge-purge tendencies, may provide greater insight into the shared and distinct mechanisms underlying these disorders (Tozzi et al., 2005).

1.3 Disordered Eating Spectrum

A substantial number of individuals may not meet the formal diagnostic criteria for any of the EDs mentioned but still experience clinically significant symptoms that impair both mental and physical health (Romano et al., 2020). Disordered eating exists on a continuum (Figure 1) between normal eating patterns and full-blown EDs, featuring similar behaviours and symptoms, but at a reduced frequency or severity. This can include, but not limited to, restrictive eating, compulsive eating, or irregular and rigid eating habits.

Figure 1

Continuum of Disordered Eating and Eating Disorders



Note. Adapted from Krug et al. (2017)

Disordered eating is highly prevalent, with an estimated prevalence of 22-25% of individuals engaging in unhealthy weight control practices such as vomiting, fasting, excessive exercise, or using laxatives or diuretics (Alfalahi et al., 2022; Nagata et al., 2018). Up to 10% may also engage in binge eating (Potterton et al., 2020). These prevalence rates are concerning given the association between disordered eating and various negative outcomes, including low self-esteem, anxiety, distress, and suicidal ideation (Brausch & Gutierrez, 2009; Kärkkäinen et al., 2018; Sala & Levinson, 2016; Wade et al., 2012). Additionally, disordered eating is linked to comparable levels of emotional distress and functional impairment as full-threshold EDs (Krug et al., 2008; Stice et al., 2009).

Disordered eating patterns in childhood or early adolescence often foreshadow the development of clinical-level EDs later in life (Linardon et al., 2019; Haynos et al., 2018; Stice et al., 2010). Longitudinal research indicates that adolescents engaging in extreme dieting or unhealthy weight control behaviours are more likely than their peers to exhibit persistent disordered eating, increased binge eating, and compensatory behaviours such as

self-induced vomiting or misuse of diet pills and laxatives in later adolescence and early adulthood (Haynos et al., 2018; Loth et al., 2014; Yoon et al., 2020).

Severe dieting has been strongly implicated to the onset of EDs, with one prospective study reporting that females who engaged in extreme dieting faced an 18-fold higher risk of developing an ED, while moderate dieting was associated with a five-fold increase in risk compared to non-dieters (Jebeile et al., 2021; Patton et al., 1999). These findings underscore the role of disordered eating as a significant precursor to the onset of clinical EDs.

1.4 Theoretical Models of Eating Disorders

As highlighted in previous sections, EDs are severe psychiatric conditions with significant consequences, including heightened mortality and suicide rates (Iwajomo et al., 2021; Forrest et al., 2021) and substantial impairments in health and quality of life (Kärkkäinen et al., 2018; Thomas et al., 2009; Wade et al., 2012). These impacts emphasise the need for effective treatment development. Developing such interventions requires the rigorous testing and refinement of theoretical models (Campbell et al., 2000).

A systematic review by Pennesi and Wade (2016) identified 54 theoretical frameworks explaining the development of disordered eating and EDs. However, the next subsection focuses on two key models – the transdiagnostic model of EDs (Fairburn et al., 2003) and the dual pathway model for bulimic symptoms (Stice et al., 1998; 2001). These two models were selected because, according to Pennesi and Wade (2016), they were the only ones among the 54 reviewed models to satisfy both efficacy criteria (tested in randomised controlled trials) and effectiveness criteria (implemented or disseminated in real-world settings). Additionally, these two models are central to exploring alternative severity ratings for EDs, an area of focus of this thesis. For an extensive review of all theoretical models, see Pennesi and Wade (2016).

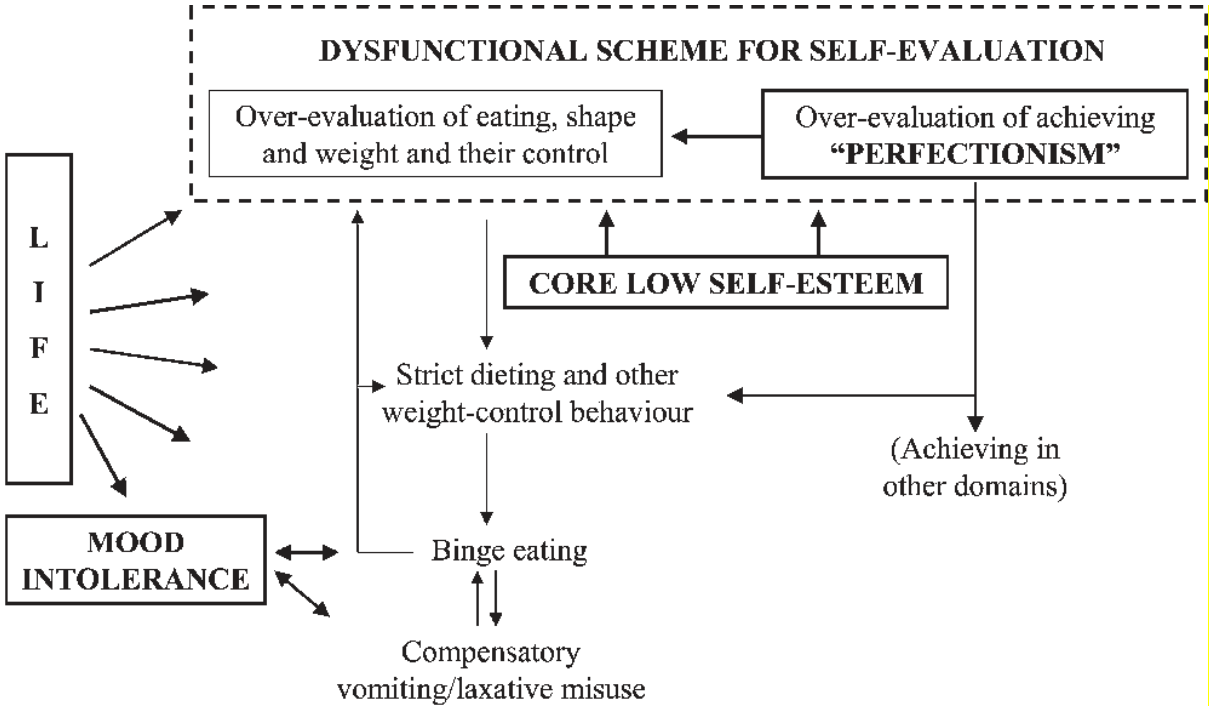
1.4.1 Transdiagnostic Maintenance Model of Eating Disorders

Transdiagnostic maintenance model of EDs, hereby transdiagnostic model, aims to identify common processes or maintenance factors across different EDs to improve psychological treatments (Mansell et al., 2009). Fairburn et al. (2003) developed the transdiagnostic model (Figure 2) that highlights shared maintenance factors in disorders such as AN, BN and BED. Central to Fairburn et al.'s model is the idea that a dysfunctional self-evaluation system plays a key role in the development and maintenance of EDs. According to the model, individuals' self-worth becomes overly dependent on controlling eating, weight, or shape, leading to an overemphasis on dietary restraint. However, strict dietary rules are hard to maintain, resulting in binge eating when cognitive control weakens.

In addition to this core concept, Fairburn et al. (2003) identified four other mechanisms that may perpetuate EDs in some individuals. First, low self-esteem can drive a focus on weight and shape as a means of boosting self-worth, reinforcing the overvaluation of these factors. Second, clinical perfectionism may cause individuals to assess their value based on their weight, shape, and eating, further amplifying dietary restraint. Third, interpersonal difficulties may increase efforts to control weight and shape in pursuit of a socially desirable ideal. Finally, mood intolerance might trigger binge eating or purging as ways of coping with difficult emotions. Evidence supporting the transdiagnostic model includes the observation that symptoms often shift from one ED to another during treatment (e.g., AN symptoms transitioning to BN symptoms), although they rarely transform into non-ED mental illnesses (Pennesi and Wade, 2016).

Figure 2

Fairburn et al.'s (2003) Transdiagnostic model of Eating Disorders



Note: From *Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatmen*t (Fairburn et al., 2003, *Behaviour Research and Therapy*, 41(5), 509–528). Copyright 2003 by Elsevier

This theory has led to the development of an "enhanced" cognitive behavioural therapy ([CBT-E]; Fairburn, 2008), designed to target the key maintenance mechanisms through collaborative case formulation. CBT-E has been validated across diverse populations, including adults (e.g., Mares et al., 2022), adolescents (Jones et al., 2020); mixed sample with both male and female samples (e.g., Dakanalis et al., 2015; Lampard et al., 2013; Mares et al., 2022). It is widely considered the treatment of choice for EDs in individuals with a BMI above 17.5 (Fairburn et al., 2015; National Institute of Health and Care Excellence [NICE]; 2014; Poulsen et al., 2014). Furthermore, recent studies, including longitudinal research (Calugi et al., 2017) and systematic reviews (Atwood et al., 2019), have demonstrated its efficacy for patients with AN, even in severe and enduring cases.

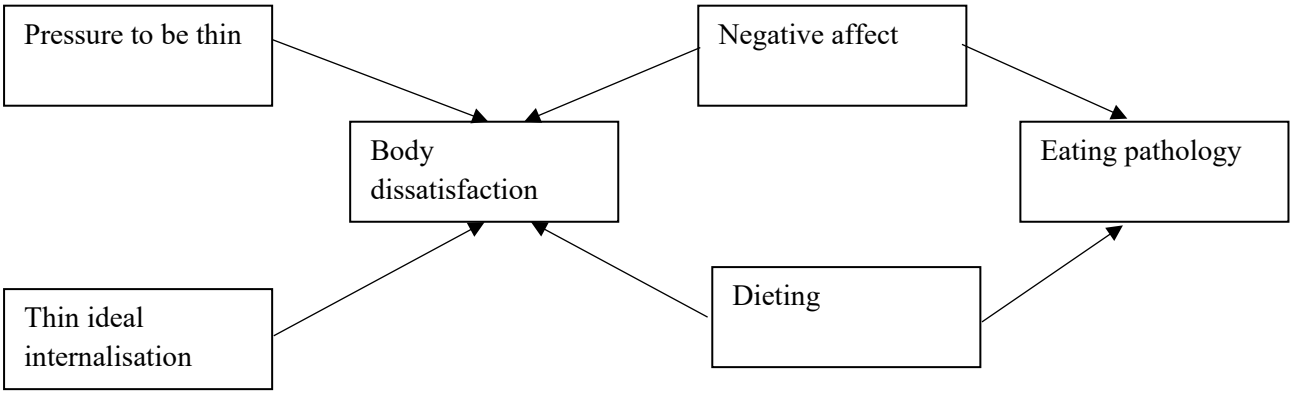
Despite its strengths, the current research on CBT-E has limitations. Most studies involve predominantly White female samples, reflecting the higher prevalence of EDs among women (Atwood et al., 2020). However, some studies have explicitly excluded men (Thompson-Brenner et al., 2016; Wade et al., 2017; Zipfel et al., 2014), introducing potential selection bias and limiting understanding of sex and cultural differences in treatment efficacy.

1.4.2 Dual-Pathway Model for Bulimia Nervosa

The dual pathway model for BN (Figure 3) is a socio-cultural theory that explains the development of BN in females by linking sociocultural pressures, body dissatisfaction, dieting and negative affect (Stice et al., 1998; 2001). This framework suggests that perceived sociocultural pressures to be thin foster both the internalisation of the thin ideal, defined as the adoption of societal standards for thinness to the extent of altering behaviours to meet these ideals (Thompson & Stice, 2001), and body dissatisfaction. Body dissatisfaction, in turn, contributes to bulimic behaviours through two primary pathways: dieting and negative affect. Dieting can lead to binge eating, either as a compensatory response to caloric deprivation or as a temporary coping mechanism for alleviating negative affect. Additionally, dietary restraint may exacerbate negative affect, either when the restraint is disrupted or due to the emotional impact of prolonged caloric restriction (Stice et al., 2002). The dual pathway model outlines specific causal and directional relationships among these factors, providing a framework for understanding the mechanisms underlying the development of bulimic symptoms.

Figure 3

Stice et al.'s (1998) Dual-Pathway Model for Bulimia Nervosa



Note. Adapted from Stice's dual-pathway model of Bulimic Symptoms (Stice et al., 1988; 2001)

Although the dual pathway model has received empirical support, only a limited number of studies have examined it in its entirety (Krug et al., 2024). To our best knowledge, to date, five cross-sectional studies (Anbari et al. 2020; Duemm et al. 2003; Hutchinson et al. 2010; Maraldo et al. 2016; Stice et al. 1996) and five longitudinal studies (Allen et al., 2012; Puccio et al., 2016b, 2019; Salafia & Gondoli, 2010; Stice et al., 1998) have tested the full model using structural equation modelling (SEM) or path analysis. Most of these studies have focused on female participants, either adults (Anbari et al., 2020) or adolescents (Anbari et al., 2020), from community populations (Maraldo et al., 2016), and none have been conducted in clinical populations. Despite these limitations, the findings from these studies have explained between 33% (Stice et al., 1998) and 54% (Allen et al., 2012) of the variance in bulimic symptoms, contributing valuable insights into key factors that can be targeted to prevent the development of these symptoms.

Interventions based on the dual pathway model are informed by cognitive dissonance theory, which suggests that the psychological discomfort caused by holding inconsistent beliefs motivates individuals to change their attitudes or behaviours to reduce this

inconsistency (Pennesi & Wade, 2016; Stice et al., 2000; 2001a; 2008; 2012a; 2009). These interventions specifically aim to reduce thin-ideal internalisation by encouraging individuals to critically evaluate and challenge this ideal. Research has shown that reductions in thin-ideal internalisation fully mediate the program's effects on body dissatisfaction and partially mediate changes in bulimic symptoms (Stice et al., 2011).

Despite the theoretical and clinical significance of the dual pathway model, as mentioned previously, only a limited amount of research has explored it in its entirety (Krug et al., 2024; Puccio et al., 2019). Moreover, while prevention programs based on the model have shown effectiveness in reducing bulimic behaviours, these interventions have been predominantly tested in younger age groups and primarily among females (Krug et al., 2024). This underscores the need for further research to validate the model across more diverse populations and to assess its applicability to broader demographic groups, thereby enhancing its relevance and effectiveness in addressing bulimic behaviours comprehensively.

1.5 Course and Treatment of Eating Disorders

As highlighted in earlier sections, EDs are among the most severe mental health conditions, marked by some of the highest mortality rates and a profound impact on individuals' physical health, emotional well-being, and overall quality of life (Miskovic-Wheatley et al., 2023). The complexity of these disorders is further illustrated by the range of genetic, psychological, and sociocultural risk factors that contribute to their development, as discussed in Section 1.2. Building on this understanding, in this section, I will first examine the general course and treatment approaches for AN, BN and finally BED. For each ED subtype, I will discuss general recovery outcomes, key predictors influencing recovery trajectories, and evidence-based interventions, assessing the strength of evidence supporting their efficacy in improving outcomes and promoting long-term recovery.

1.5.1 Anorexia Nervosa

Course and Recovery Outcomes of Anorexia Nervosa

Recovery outcomes for EDs vary significantly, but recent findings provide a more hopeful perspective on the core symptoms of AN (Dobrescu et al., 2020). For instance, a meta-analysis of 119 studies revealed that 59.6% of AN patients achieved weight normalisation, 57% resumed regular menstruation, and 46.8% demonstrated normalised eating behaviours (Steinhausen et al., 2002). These findings are echoed in a six-year follow-up study by Fichter and Quadflieg (1999), which found that 55.4% of women with AN fully recovered following inpatient care.

Gender comparisons in AN recovery outcomes offer additional insights. Strobel et al. (2019) found that men with AN not only gained more weight during treatment but also exhibited lower levels of ED-specific psychopathology compared to women. Despite this, remission rates were similar for both genders, remaining around 40–41% (Strobel et al., 2019).

Long-term studies further highlight the variability in recovery trajectories. Eddy et al. (2017) found that only two-thirds of females with AN achieved full recovery after a 22-year follow-up period, emphasising the prolonged course of the illness for many patients. The duration of AN has been reported to range widely, from an average of 6.5 years (Zerwas et al., 2013) to 14 years (Fernández-Aranda et al., 2021), underscoring the chronic nature of the disorder in some cases.

Several factors have been identified as critical predictors of recovery for AN. A higher starting weight consistently emerged as the strongest predictor of achieving a higher BMI and full recovery one-year post-therapy (Franko et al., 2018). In contrast, empirical studies (e.g., Gregertsen et al., 2019; Wild et al., 2016) as well as systematic reviews and meta-analyses (e.g., De Rijk et al., 2024; Sala et al., 2023) highlighted negative predictors,

including a duration of illness longer than six years, pre-treatment depression, and the presence of comorbid psychiatric conditions. These findings underscore the importance of early intervention and comprehensive treatment approaches in AN tailored to individual risk factors

Treatments for AN and Treatment Specific Outcomes.

According to the guidelines of NICE (2020), treatment for AN needs to be multidisciplinary, including medical care, weight monitoring, and psychological interventions. Due to medical complications, high relapse and mortality rates, as well as its long duration, treatment for AN is often challenging (Baenas et al., 2024; Van Eeden et al., 2021).

For adolescents with AN, Family-Based Therapy (FBT) is widely recognised as the first-line psychological intervention, supported by a recent meta-analysis (Zeeck et al., 2018). FBT has been shown to significantly improve ED psychopathology, such as dietary restraint, and promote weight restoration (Le Grange et al., 2021; Stewart et al., 2024; Zeeck et al., 2018). However, limitations in clinical trials of FBT, including insufficient follow-up periods and inconsistent outcomes, highlight areas for improvement. Additionally, long-term outcomes reveal challenges, with only 41.8% of participants achieving full recovery by the end of treatment and 26.6% continuing to meet the diagnostic criteria for AN two to four years later (Le Grange et al., 2014).

For adults with AN, the effectiveness of psychological treatments remains inconclusive, with no single approach showing clear superiority over others (Solmi et al., 2021; Watson & Bulik, 2013). International guidelines typically recommend standalone outpatient psychological therapies as the first-line intervention over pharmacological treatments for adults with AN (Solmi et al., 2021). However, the specific recommendations for treatment modalities vary across regions. Australian guidelines generally endorse any

specialist-led, manual-based therapy, ideally delivered within a multidisciplinary framework (Hay et al., 2014). In contrast, guidelines in the Netherlands and the United Kingdom (UK) provide more detailed recommendations, outlining first- and second-line psychological treatments (Hilbert et al., 2017). For example, the 2017 UK National Institute for Health and Care Excellence (NICE) guidelines recommend Specialist Supportive Clinical Management (McIntosh et al., 2006), Maudsley Anorexia Treatment for Adults (Wade et al., 2011), and CBT-E (Fairburn et al., 2008) as first-line treatments, while reserving focal psychodynamic therapy as a second-line option. A meta-analysis by Solmi et al. (2021) examined 16 randomised controlled trials investigating psychological interventions for adults with AN, focusing on weight restoration, ED symptoms, and all-cause dropout rates. The findings revealed that no psychological intervention consistently outperformed treatment as usual for the primary outcomes. However, CBT-E had a lower dropout rate than psychodynamic therapies, with an odds ratio of 0.54, indicating a 46% lower likelihood of treatment discontinuation. Altogether, these results highlight the pressing need for more effective psychological treatments for both adolescents and adults with AN.

1.5.2 Bulimia Nervosa

Course and Recovery Outcomes of Bulimia Nervosa

Recent research presents a more optimistic outlook for BN treatment, showing significant improvements in remission rates over a decade post-treatment despite frequent relapses. Around two-thirds of individuals with BN achieve substantial improvement or full recovery in the long term, while 15–20% may develop a chronic course of the disorder (Quadflieg et al., 2024). However, research has shown that recovery likelihood decreases significantly after ten years, with a small proportion transitioning to another ED diagnosis, with most individuals with BN develop AN (up to 14%), and 1–3% experiencing mortality (Quadflieg et al., 2024).

Regarding predictors of BN prognosis, empirical studies (e.g., Matherne et al., 2022) and comprehensive reviews of studies (Gorrell et al., 2022; Steinhausen & Weber, 2009) have provided limited clarity, with few consistent predictors identified. A study by Quadflieg and Fichter (2019), which analysed baseline predictors in inpatient BN groups and assessed outcomes over an 11-year follow-up, found a limited number of significant predictors. These included age at onset, age at treatment, marital status, and symptom severity (both ED-specific, such as binge eating, vomiting, or excessive exercise, and general psychopathology). Their findings indicated that older age at treatment, more severe symptoms, and shorter follow-up periods were associated with poorer long-term outcomes for those with BN.

Treatment for BN and Treatment Specific Outcomes

Research on treatment modalities for BN has highlighted both advancements and significant gaps in treatment efficacy. Randomised controlled trials (e.g., Barakat et al., 2023; Hartmann et al., 2024) and meta-analyses (e.g., Rozakou-Soumalia et al., 2021) demonstrate that several psychological interventions are more effective than wait-list controls, pill placebos, and antidepressants in reducing BN symptoms. These treatments include CBT (Poulsen et al., 2014; Hartmann et al., 2024), interpersonal psychotherapy (Zhang et al., 2024), dialectical behaviour therapy (Vogel et al., 2021; Rozakou-Soumalia et al., 2021), and adolescent family-based therapy (Couturier et al., 2013; Linardon et al., 2017).

Despite several psychological interventions having been developed for BN, achieving full symptom abstinence remains a critical challenge. A meta-analysis by Linardon and Wade (2018), synthesising 45 randomised controlled trials, found that self-help interventions for BN yielded lower rates of post-treatment symptom abstinence compared to clinician-led treatments. Interestingly, these differences disappeared at follow-up (3 to 12 months), with self-help outcomes becoming comparable to those of clinician-led approaches. Despite this, the review revealed that more than 60% of patients failed to achieve complete abstinence

from core BN symptoms, even after receiving empirically supported interventions (e.g., CBT). These findings highlight the pressing need to improve the effectiveness of psychological treatments for BN, with a particular focus on addressing persistent symptoms and promoting sustained recovery.

1.5.3 Binge Eating Disorder

Course and Recovery Outcomes of Binge Eating Disorder

Compared to AN and BN, individuals with BED generally experience more favourable long-term outcomes and higher treatment success rates (Miskovic-Wheatley et al., 2023). A recent meta-analysis emphasises that psychotherapy and behavioural interventions are particularly effective in reducing BED symptoms and supporting weight reduction within this population (Hilbert et al., 2020). With appropriate specialised treatment, post-treatment symptom abstinence rates for BED can improve by as much as 63% (Anderson et al., 2020; Linardon, 2018). However, without proper treatment, BED often leads to chronic or persistent issues (Solmi et al., 2024). Nevertheless, effective interventions and positive patient responses significantly enhance prognosis, offering better outcomes compared to those seen in AN and BN (Hilbert et al., 2018).

Predictors of treatment outcomes in BED share notable similarities with those in other mental health conditions. Milder cases and lower levels of psychopathology are generally associated with faster remission and better outcomes (Anderson et al., 2020; Fichter et al., 2008). However, the severity of ED symptoms remains a critical determinant of treatment success. Factors such as frequent binge-eating episodes, emotional eating, and heightened concerns about weight and body image are consistently linked to poorer outcomes (Castellini et al., 2011; Fischer et al., 2014). Overall, while BED often poses significant challenges, targeted and timely interventions can greatly improve treatment success and patient outcomes.

Treatment for BED and Treatment Specific Outcomes

International guidelines (e.g., NICE) recommend evidence-based psychological therapy as the first-line treatment for BED. These psychological interventions are typically delivered in outpatient settings but may also be integrated into comprehensive hospital programs with outpatient follow-up care (Hay et al., 2019). The primary therapies evaluated for BED include CBT (Fairburn et al., 2008), interpersonal psychotherapy (Wilfley et al., 2008), and dialectical behaviour therapy (Safer et al., 2009), all of which have been implemented in both group and individual formats (Miskovic-Wheatley et al., 2023). Interpersonal psychotherapy stands out with the highest abstinence rates for binge eating, with 63% of individuals achieving symptom abstinence post-treatment, according to a meta-analysis by Linardon (2018). In contrast, mindfulness-based approaches had the lowest abstinence rates, with only 22% of individuals achieving symptom abstinence. A comparative study by Hilbert et al. (2012) further highlighted the superiority of interpersonal psychotherapy in achieving long-term outcomes compared CBT. While individuals treated with CBT experienced an increase in EDs symptoms between the end of treatment and a four-year follow-up, those who received interpersonal psychotherapy showed continued improvements during this period. Notably, remission rates at the four-year follow-up were significantly higher for interpersonal psychotherapy (76.7%) compared to CBT (52.0%) (Hilbert et al., 2012).

These findings suggest that while treatment outcomes for BED are generally more favourable compared to AN and BN, there is still substantial room to enhance the effectiveness of psychological interventions. Future efforts should focus on optimising long-term remission rates and reducing the risk of relapse to maximise treatment success for BED.

1.6 Severity Ratings for Eating Disorders

Section 1.5 demonstrated that, despite significant efforts to develop effective treatments, treatment outcomes remain highly variable. This variability underscores the need for personalised approaches to care for the different ED subtypes, with severity ratings potentially playing a crucial role in guiding these interventions.

Severity ratings provide a structured method to evaluate symptom intensity, functional impairments, and medical risks. These ratings are crucial for tailoring treatment to individual needs and prioritising the most severe cases to optimise resource allocation (Zimmerman et al., 2018). In research, severity ratings can assist in standardising assessments across studies, supporting the development of evidence-based treatments. The following sections explore the concept of severity in psychiatry, with a focus on the introduction of severity ratings for EDs, including those in the DSM-5, ICD-11, and alternative severity ratings – key components of this PhD thesis.

1.6.1 The Concept of Severity in Psychiatry

The determination of illness severity has significant clinical implications, affecting when to seek care, the type and intensity of treatment, and decisions about continuing or discontinuing therapy. In psychiatry, the term "severe illness" can refer to various factors, such as the number, intensity, and frequency of symptoms, as well as their persistence and impact on functioning or quality of life (Lucas et al., 2022; Zimmerman et al., 2018). It may also imply a higher risk of permanent disability or mortality. Unlike medical conditions, psychiatric disorders lack biomarkers (e.g., structural damage) to measure severity objectively, so clinicians often rely on observable symptoms to conduct severity assessments (Zimmerman et al., 2018).

Despite the importance of defining severity, there is no clear consensus on how to do so, especially when it comes to resource allocation for patients with severe mental illness in

public health (Rugger et al., 2000). Current classification systems use a polythetic categorical approach, where diagnoses are based on meeting a certain number of predefined symptoms. This approach is widely used for justifying treatment, communicating diagnoses, and collecting epidemiological data (Widiger et al., 2018). However, these classifications have limitations, such as high within-category variability, frequent comorbidity, and challenges in capturing the spectrum and severity of symptoms (First et al., 2024).

To address these limitations, in the latest update, both the DSM-5 and ICD-11 have introduced a dimensional approach to measuring severity. The DSM-5 categorises severity into four levels: mild, moderate, severe, and extreme severe for a wide range of psychiatric disorders, including major depressive disorder, and obsessive and compulsive disorder (Zimmerman et al., 2018). In the ICD-11, severity classification levels differ depending on the disorder. For example, personality disorders are classified as mild, moderate, or severe, while AN is described in terms of "significantly low BMI" or "dangerously low BMI" severity.

These classifications consider factors such as the number of criteria met or the frequency of core symptoms. Unlike many physical illnesses, assessments of psychiatric severity typically do not factor in mortality risk. Instead, they concentrate on symptoms and their effects on distress and functioning (Zimmerman et al., 2018). The introduction of graded severity ratings in both DSM-5 and ICD-11 was intended to enhance diagnostic accuracy and aid clinical decision-making. This approach aims to communicate key aspects of a disorder, improve clinical practice, increase access to care, and raise public and professional awareness and understanding of these conditions (Aouad et al., 2022)

Illness severity can be indicated by several factors, including comorbidities, general psychopathological symptoms, and related cognitive impairments (Wilkop et al., 2023). For clinical usefulness, treatment outcomes should be a primary focus, as more severe cases

typically respond less favourably than milder ones (Lucas et al., 2022). Severity is also reflected in prognosis, with high-severity cases often showing greater treatment resistance, which can lead to chronic symptoms (Dieffenbacher et al., 2024). Overall, the severity of illness is multifaceted and should be evaluated in terms of various factors, rather than a single dimension. In the following three subsections (1.6.2, 1.6.3, and 1.6.4), I will provide a more in-depth introduction to the DSM-5, ICD-11, and alternative approaches to severity ratings for EDs.

1.6.2 The Introduction of the DSM-5 Severity Ratings for Eating Disorders

As the 5th edition of the DSM-5 (APA, 2013) was being reviewed, the ED Work Group presented new ED severity ratings. These severity ratings for EDs were an attempt to overcome the limitations of the dichotomous diagnostic approach used in the former DSM-IV-R (APA, 2000), and to provide clinicians with ED-specific markers of disorder severity (APA, 2013). This was accomplished by establishing gradients for disorder severity by classifying individuals with an ED into "mild", "moderate", "severe", and "extreme" categories. Specifically, the DSM-5 categorises individuals with AN into mild (≥17.0 kg/m^2), moderate (16–16.99 kg/m^2), severe (15–15.99 kg/m^2), and extreme (<15 kg/m^2) severity groups based on BMI. The severity for BN is determined by the frequency of inappropriate compensatory behaviours, and the severity of BED is assessed by the frequency of binge eating episodes. The severity categories for BN and BED are mild (1–3 episodes/week), moderate (4–7 episodes/week), severe (8–13 episodes/week), and extreme severe (14 or more episodes/week). For a comprehensive overview of all current severity ratings (e.g., DSM-5 and alternative ratings), please refer to Table 4.

As mentioned previously, BMI, a simple index of weight-for-height, is used as a severity indicator for AN. The DSM-5 likely uses BMI to index AN severity to reflect the centrality of body weight as a key feature of AN (Criterion A; APA, 2013). However, the

current literature shows inconsistent support for the use of BMI to index AN severity. Specifically, some studies report that a low BMI is associated with increased medical (e.g., cardiac problems) and psychological complications (e.g., comorbidity with anxiety), and a poorer prognosis (Berner et al., 2014; Maurel et al., 2024; Wolfi & Treasure, 2011), while other studies did not find such associations (e.g., Hebebrand et al., 1996; Toppino et al., 2024). These inconsistent findings may partially be due to the severity indices overshadowing important physical (e.g., gender, weight history, bone density), psychological or cognitive factors (e.g., drive for thinness), which might reflect the individuals' levels of impairments (Grilo et al., 2015a; Hirst et al., 2017). Therefore, it is unclear whether BMI is useful in indexing AN severity.

As articulated before, the DSM-5 ratings for BN are based on the frequency of inappropriate compensatory behaviours, and for BED these ratings are based on the frequency of binge eating episodes. Some studies have found support for the utility of these severity ratings with the frequency of compensatory and binge eating episodes being positively correlated with general and ED psychopathology for individuals with BN (e.g., Gibson et al., 2019; Peñas-Lledó et al., 2002; Weigel et al., 2019) and BED (Shroff et al., 2006; Wilfley et al., 2016).

However, despite these positive associations between them, the severity indices for BN and BED have been criticised due the arbitrariness of the frequency classifications (Lucas et al., 2022). Specifically, it is unclear why three episodes per week is considered mild while four episodes per week is classified as moderate, and whether these cut-off scores reflect meaningful differences in terms of psychopathology, distress, and prognosis (Toppino et al., 2022). As such, clarification is needed regarding the utility of the severity indicators for BN and BED in the DSM-5 and upcoming versions of the DSM. In the next chapter of the thesis (Chapter 2), I will introduce a systematic review and meta-analysis which will provide a

comprehensive overview of all studies that have assessed the clinical utility of the DSM-5 severity ratings for EDs.

1.6.3 The ICD-11 Severity Ratings for AN

Similar to the DSM-5, the ICD-11 incorporates BMI-based specifiers to classify the severity of underweight in individuals with AN. Specifically, the ICD-11 distinguishes between adults with a BMI of 14 kg/m² or higher and those with a BMI below 14 kg/m², defining these categories as "significantly low BMI" and "dangerously low BMI" respectively. This BMI threshold of 14 kg/m² is intended to enhance diagnostic precision by recognizing the impact of extreme underweight on AN prognosis and the level of treatment intensity required (Uher & Rutter, 2013). According to the Royal Australian and New Zealand College of Psychiatrists' clinical practice guidelines for ED treatment [RANZCP] (Hay et al., 2014), a BMI of 14 kg/m² typically indicates the need for psychiatric admission and medical monitoring, while a BMI below 12 kg/m² is associated with high medical risk, necessitating admission to a medical setting with psychiatric consultation. Thus, unlike the DSM-5 severity ratings, the ICD-11's threshold of 14 kg/m² is supported by clinical guidelines and research-based rationale, underscoring its role in guiding treatment decisions.

To date, only one study by Engelhardt et al. (2020) has explored the relationship between the ICD-11 AN severity rating and prognosis. Among 469 children and adolescents with AN, they found that those in the "dangerously low BMI" group had longer treatment durations and lower BMIs at discharge compared to the "significantly low BMI" group. However, these differences were not significant enough to aid in clinical risk assessment or treatment stratification.

Table 4

Summary of the current DSM-5, ICD-11 and Alternative Severity Ratings for EDs

Severity Classification Systems **DSM-5**	**Severity Groups**			
	Mild	Moderate	Severe	Extreme
DSM-5 AN BMI (kg/m^2)	>17	>16-16.99	15-15.99	<15
DSM-5 BN Compensatory behaviours	1-3	4-7	8-13	>14
DSM-5 BED Binge Eating Episodes	1-3	4-7	8-13	>14
ICD-11				
ICD-11 AN BMI (kg/m^2)	Significantly Low BMI ≥14		Dangerously Low BMI <14	
Disordered-specific Alternative Severity Ratings				
Overvaluation of weight and shape (OWS) for BED *	Low OWS (Score < 4 on either of the two OWS items on EDE or EDE-Q)		High OWS (Score ≥4 on either of the two OWS items on EDE or EDE-Q)	
Number of Purging Methods for BN **	Single Purging Method		Multiple Purging Method	
Transdiagnostic Alternative Severity Ratings				
Overvaluation of weight and shape *	Low OWS (as above)		High OWS (as above)	
Drive for thinness (DT)***	Low DT (of ≤ 14 for the EDI DT sub-scale)		High DT (score >14 for the EDI DT sub-scale)	
Duration of illness ****	Short duration of illness AN ≤ 12 years BN ≤ 13 years BED ≤ 21 years		Long duration of illness AN > 12 years BN > 13 years BED > 21 years	
Weight suppression	No cut-off has been determined			

Note.
OWS severity ratings cut-off for BED and transdiagnostically were based on study of Grilo et al., (2015a; 2015b) and Gianini et al. (2017)
*OWS were assessed based on the two OWS items from either the EDE (Cooper, 1987) and the EDE-Q (Fairburn, 2008): "Over the past four weeks, has your shape influenced how you feel about (judge, think, evaluate) yourself as a person?" and "Over the past four weeks, has your weight influenced how you feel about (judge, think, evaluate) yourself as a person?"
**Number of purging methods cut-off was based on Gianini et al., (2017)
***Drive for thinness cut-off was based on Garner et al.'s (1983) recommendation and has been assessed by Krug et al. (2021)
****Duration of illness cut-off was based on Fernández-Aranda et al.'s (2021) recommendations.

1.6.4 Alternative Severity Ratings for Eating Disorders

Researchers have questioned the utility of the DSM-5 and ICD-11 severity ratings for EDs, prompting the exploration of alternative single-indicator measures of severity (Gianini et al., 2017). Two main approaches to alternative severity ratings using a single indicator are considered: disorder-specific severity markers and a transdiagnostic approach to severity rating (see Table 4). In the following three subsections, I will present alternative disorder-specific and transdiagnostic severity ratings that have been discussed in the ED literature.

1.6.4.1 Disordered-specific Alternative Severity Rating

Taking a similar approach to the DSM-5 and ICD-11, the majority, if not all, alternative disorder-specific severity indictors for EDs are based on the single core feature of the disorder. Table 4 presents an overview of all the alternative severity ratings that are discussed here.

Overvaluation of weight and shape for BED. Overvaluation of weight and shape **(**OWS), the disproportionate influence of body evaluation on self-worth, a central feature of EDs and is emphasised in the transdiagnostic model of EDs (Fairburn et al., 2003). Given that OWS is a core feature of BED and differentiates levels of psychopathology (Coffino et al., 2019; Kenny & Carter, 2018), researchers have built upon this understanding to explore OWS as a single-indicator severity rating for BED.

Recent studies have examined OWS as an alternative indicator of BED severity, comparing its clinical and predictive utility against the DSM-5 severity ratings based on weekly binge frequency (e.g., Gianini et al., 2017). OWS in these studies was assessed using two items from the Eating Disorder Examination ([EDE; Cooper & Fairburn, 1987] and the Eating Disorder Questionnaire ([EDE-Q]; (Fairburn & Beglin, 1994): “Over the past four weeks, has your shape influenced how you feel about (judge, think, evaluate) yourself as a person?” and “Over the past four weeks, has your weight influenced how you feel about

(judge, think, evaluate) yourself as a person?" Individuals were categorised into low OWS (score < 4) or high OWS (score ≥ 4) based on their responses on either of these items. Using traditional statistical methods (e.g., Analysis of Variance [ANOVA]) and exploratory data mining approaches (e.g., SEM), findings indicate that OWS more effectively captures both ED-specific and general psychopathology in clinical (Forrest et al., 2022; Gianini et al., 2017; Grilo et al., 2015a) and community samples (Grilo et al., 2015b). These results underscore the potential of OWS as a more meaningful indicator of BED severity. To our knowledge, OWS has not been assessed as an alternative disorder-specific severity for either AN or BN.

Number of purging methods for BN. The number of purging methods used has been suggested as an alternative severity marker for BN (Gianini et al., 2017). Research has shown that a higher number of purging methods is linked to more severe general and ED-related psychopathology (e.g., Edler et al., 2009). Specifically, individuals who engage in multiple purging behaviours tend to experience greater levels of self-injury, psychiatric comorbidities, and ED psychopathology compared to those using only one purging method (Eddy et al., 2009; Edler et al., 2007). Two studies (Dieffenbacher et al., 2024; Gianini et al., 2017) have examined the number of purging methods as an alternative severity marker for BN, comparing it to the DSM-5 severity rating. Both found that using the number of purging methods offers a more robust measure of psychopathology, with individuals who engage in multiple purging methods showing higher levels of both general and ED-specific psychopathology than those who use only a single method.

1.6.4.2 Transdiagnostic severity ratings for ED

The transdiagnostic approach, which identifies shared factors driving the development and maintenance of ED, has gained traction in the field. Cognitive and behavioural traits like restrictive eating, binge eating, compensatory behaviours, and OWS are shared across AN, BN and BED, with diagnostic crossover frequently observed (Castellini et al., 2011).

Therefore, using transdiagnostic markers (e.g., OWS) as indicators of severity may offer deeper insights into symptom severity across and within diagnostic ED groups (Fairburn et al., 2003). If these markers reliably differentiate levels of impairment and psychopathology, they could enhance treatment by focusing on core features across EDs. Addressing this shared feature in treatment might improve outcomes across the spectrum of ED presentations and reducing diagnostic crossover.

Overvaluation of weight and shape. OWS, as previously defined, has not only been explored as an alternative severity indicator for BED but has also been proposed as a transdiagnostic severity measure across EDs. Gianini et al. (2017) pioneered this concept, investigating the clinical utility of OWS, the most prominent transdiagnostic component outlined in Fairburn et al.'s model, as an alternative severity rating across AN, BN, and BED. Gianini et al. categorised participants with AN, BN and BED into "low" and "high" OWS groups based on their scores (i.e., score ≥ 4) on the EDE (Cooper & Fairburn, 1987). Their findings demonstrated that OWS provided greater clinical utility than DSM-5 severity ratings for AN, BN, and BED, particularly in differentiating levels of ED symptomatology and general psychopathology across severity groups and diagnostic categories.

Drive for thinness. Drive for thinness (DT), defined as an extreme fear of weight gain, resulting in a disordered eating pattern, has been considered as an alternative transdiagnostic severity specifier across all EDs (e.g., Krug et al., 2021). The role of DT is aligned with what was conceptualised within the dual pathway model to ED. Specifically, DT has been found to triggers body dissatisfaction. This body dissatisfaction, in turn, leads to increased dieting and negative affect, both of which are linked to bulimic pathology (Stice et al., 1998; 2001).

While the dual pathway model was developed for BN, researchers have assessed individual components of the model, including DT across EDs (Krug et al., 2024). DT has

primarily been assessed using different versions of the Eating Disorder Inventory (EDI; Garner et al., 1983), including EDI-2 (Garner, 1991) and EDI-3 (Garner, 2004). Studies have found that individuals with AN and BN that experienced a lower DT experienced lower ED symptoms, lower psychiatric comorbidity and poorer prognosis than those with higher DT (Abbate-Daga et al., 2007; Keski-Rahkonen et al., 2013; Peñas-Lledó et al., 2015; Prost-Lehmann et al., 2018). Although the relationship between DT and BED is not widely investigated, some research has indicated that individuals with BED were found to have an excessive concern with shape and thinness and were more likely to have higher DT scores compared to non-bingers (Kuehnel & Wadden, 1994).

To date, three clinical studies (Krug et al., 20212024; Mora-Maltas et al., 2023) have evaluated the use of DT as an alternative transdiagnostic severity rating EDs. These studies compared the clinical utility of DT and DSM-5 severity ratings in indexing ED-specific psychopathology (e.g., bulimic behaviours) and general psychopathology (e.g., cognitive inflexibility), within treatment-seeking individuals with EDs. DT was assessed on EDI-2 (Garner, 1991), and a cut-off of >14 for the EDI DT sub-scale was used to differentiate those with low and high DT. These studies, which included samples of all males (Krug et al., 2024), all females (Krug et al., 2021), and mixed-gender participants (Mora-Maltas et al., 2023), consistently demonstrated that DT categories (low vs. high DT) provided greater differentiation in both ED-specific and general psychopathology compared to DSM-5 severity ratings. Overall, these findings provide promising research-based support for DT as an alternative transdiagnostic factors for EDs.

Duration of illness. Duration of illness has also been proposed as a transdiagnostic indicator of ED severity. Research has showed that shorter illness duration is associated with more favourable treatment outcomes across EDs (Errichiello et al., 2016; Fernández-Aranda et al., 2021; Reas et al., 2000; Vall & Wade, 2015). Specifically, Fernández-Aranda et al.

(2021) found that longer illness duration was linked to poorer treatment responses, supporting its potential as a marker of severity. Fernández-Aranda et al. (2021) identified a specific duration cut-off points that increase the risk of poor outcomes: 12 years for patients with AN, 13 years for those with BN, and 21 years for individuals with BED. However, the findings regarding the prognostic value of illness duration are inconsistent. Some studies have found no relationship between illness duration and treatment outcomes or ED psychopathology, whether across all EDs (Radunz et al., 2020), specific EDs (e.g., AN; Nordgaard et al., 2020), or specific treatments like CBT (Raykos et al., 2018). In contrast, other research has found that longer illness duration is associated with poorer treatment outcomes (Austin et al., 2021; Wild et al., 2016).

Regardless of the mixed findings, a recent study by Mora-Maltas et al. (2023) evaluated illness duration as a transdiagnostic severity indicator by comparing it to both the DSM-5 severity rating and DT as an alternative severity indictor in indexing cognitive flexibility. In their study, with a sample of 161 individuals with AN, BN, and BED, illness duration demonstrated greater utility in differentiating cognitive flexibility across diagnostic groups than the DSM-5 and DT severity ratings. However, it is important to note that this study focused exclusively on cognitive flexibility as the outcome variable, limiting its generalisability to other domains of functioning or treatment outcomes.

Weight suppression. Weight suppression, defined as the difference between an individual's highest and current weight (excluding pregnancy; Lowe, 1984; 1986), has emerged as a significant concept in ED research. It has been proposed as a transdiagnostic indicator of ED development, the severity of disordered eating, and psychopathology across various EDs, including AN and BN (Frank et al., 2020; Keel et al., 2010; Lowe, Haller, et al., 2020; Lowe, Piers, et al., 2018; Stice et al., 2020).

Research on weight suppression has been synthesised in both systematic (Jenkins et al., 2018) and narrative (Gorrell et al., 2019) reviews, highlighting its role in predicting psychopathology and treatment outcomes across EDs. A systematic review by Jenkins et al. (2018) analysed 12 studies involving a total of 1,566 individuals with AN, BN, BED or binge eating behaviours. This review found that greater weight suppression was associated with increased total weight gain at post-treatment but was not linked to other treatment outcomes, such as treatment completion, abstinence from ED behaviours, or symptom reduction.

Similarly, a narrative review by Gorrell et al. (2019) explored the relationship between weight suppression, ED psychopathology, and treatment outcomes. The review concluded that higher levels of weight suppression were associated with greater ED psychopathology in individuals with AN and BN. In terms of treatment outcomes, Gorrell et al. found that weight suppression was positively correlated with weight gain following treatment in AN and BN, as well as with treatment dropout in BN and BED, aligning with the findings of Jenkins et al. (2018).

Due to its relatively recent recognition as an independent ED in the DSM-5 (2013), the role of weight suppression in BED has been less studied. Stice et al. (2021) found that weight suppression predicted the development of AN and BN but not BED. However, Van Son et al. (2013) reported that weight suppression was associated with binge eating and loss of control in a community sample of 3,512 participants.

Given the role of weight suppression in predicting ED psychopathology and treatment outcomes across EDs, it is not surprising that studies have suggested the role of weight suppression as a severity indicator to predict ED symptoms among clinical samples with EDs (Brewerton et al., 2000; Hagan et al., 2017; Khalil et al., 2024; Lowe et al., 2011). Overall, these findings have spurred growing interest in weight suppression as a transdiagnostic indicator of ED severity. Emerging evidence suggests that weight suppression could provide

a more nuanced approach to differentiating severity levels in ED-related symptomatology compared to the DSM-5 framework, which is primarily based on a single disorder-specific core feature (Khalil et al., 2024). Despite its potential, no studies to date have directly compared the clinical or predictive utility of DSM-5 severity ratings with weight suppression as a transdiagnostic severity measure, highlighting an important area for future research.

1.7 Chapter Summary

The first chapter of this thesis provided an overview of the prevalence, diagnostic criteria (as outlined by DSM-5 and ICD-11), and the consequences of ED on physical health, psychological well-being, and economic burden. Sections 1.2 and 1.4 summarised the risk factors for the development of EDs and provided an overview of two theoretical models (i.e., transdiagnostic maintenance model of EDs and the dual pathway model for BN) that explain the onset and maintenance of these disorders. These sections emphasised the extensive efforts made to identify factors contributing to ED development and to develop theoretical frameworks that guide treatment approaches. However, as detailed in Section 1.5, recovery outcomes for EDs can vary widely. Factors such as BMI, illness duration, and psychiatric comorbidities have been identified as key influences on recovery trajectories. To improve clinical outcomes, researchers, along with the DSM-5 and ICD-11 working groups, have proposed severity ratings for EDs. These ratings are either disorder-specific, such as BMI for AN, or take a transdiagnostic approach, like using DT as a severity measure across all EDs.

Despite these efforts, little is known about the predictive and clinical utility of these severity ratings, whether based on the DSM-5, ICD-11, or the alternative approaches. Addressing this gap is the central aim of this PhD thesis, which sought to enhance the understanding of the clinical utility of the DSM-5 and ICD-11 severity ratings for EDs. Furthermore, the thesis explored the alternative severity measures that incorporate shared physical (e.g., weight suppression) and psychological (e.g., DT) factors.

The first aim of this thesis is to provide a comprehensive and systematic synthesis of the current knowledge on ED severity ratings. To achieve this, the project began with a systematic review and meta-analysis that critically evaluated and synthesised existing evidence. This first study (Study 1) not only provides an overview of the current landscape but also identifies key gaps in the literature, laying the foundation for the subsequent studies within the thesis. The remaining objectives of the thesis are presented in **Chapter 3,** following Study 1 **(Chapter 2)**, reflecting the chronological development of the research ideas and their progression throughout this PhD project.

Chapter 2: A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders.

2.1 Chapter Introduction

As outlined in Chapter 1, EDs are severe psychiatric conditions that profoundly impact individuals' physical, psychological, and social well-being (Peckmezian and Hay, 2017). Although several psychological treatment approaches have been developed and shown to effectively target core symptoms across EDs, a substantial number of individuals remain unwell (Solmi et al., 2024). To enhance clinical practice and improve outcomes, researchers, along with the DSM-5 and ICD-11 working groups, have proposed severity ratings for EDs. These ratings may be disorder-specific, such as using BMI for AN or transdiagnostic, such as employing DT to assess severity across all EDs.

Despite these advancements, the predictive and clinical utility of these severity ratings – whether derived from DSM-5, ICD-11, or alternative frameworks – remains poorly understood. To address this critical gap, **Chapter 2** introduces **Study 1,** which presents a systematic review and meta-analysis designed to rigorously evaluate and synthesise the evidence on the clinical utility of DSM-5 and alternative severity ratings for EDs. Systematic reviews and meta-analyses are recognised as the highest standard of evidence, providing a comprehensive and reliable summary of existing research. This approach was selected as the foundation of this PhD to ensure a methodologically rigorous and accurate evaluation of the literature.

It is important to note that this study did not assess the clinical utility of ICD-11 severity ratings, as Study 1 was conducted before the publication of the ICD-11 severity criteria for AN in 2019. Chapter 2 systematically reviews the study designs, methodologies, and potential sources of bias within the existing literature to assess the quality of the evidence. A meta-analysis was then performed to synthesise data from the included studies,

providing a clearer understanding of the current landscape and identifying key limitations and knowledge gaps in the field of ED severity ratings. The insights and gaps identified in Chapter 2 form the basis for the subsequent chapters of this thesis, which aim to address these unresolved issues and advance the understanding of severity ratings in EDs.

2.2 Study 1. A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders.

This chapter (Chapter 2) constitutes a manuscript published in Clinical Psychology: Science and Practice. (31st March 2022)

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022, March 31). A Systematic Review and Meta Analysis on the DSM–5 Severity Ratings for Eating Disorders. *Clinical Psychology: Science and Practice*, *29* (4), 325. https://doi.org/10.1037/cps00

Contribution to the Study

Author	Task contribution
An Dang	Conceptualisation; manuscript writing; editing; data collection; data analyses
Sarah Giles	Data collection, editing
Matthew Fuller-Tyszkiewicz	Data analyses, editing
Litza Kiropoulos	Editing
Isabel Krug	Conceptualisation; manuscript – reviewing and editing; supervision

Clinical Psychology: Science and Practice

ISSN: 0969-5893

2022, Vol. 29, No. 4, 325–344
https://doi.org/10.1037/cps0000078

A Systematic Review and Meta-Analysis on the *DSM–5* Severity Ratings for Eating Disorders

An Binh Dang[1], Sarah Giles[1], Matthew Fuller-Tyszkiewicz[2, 3], Litza Kiropoulos[1], and Isabel Krug[1]

[1] Melbourne School of Psychological Sciences, The University of Melbourne

[2] School of Psychology, Deakin University

[3] Centre for Social and Early Emotional Development, Deakin University

The current review assessed the clinical utility of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–5*) and alternative severity ratings for eating disorders (EDs) in their ability to distinguish ED psychopathology between the severity groups and frequency distribution of these severity ratings. EMBASE, MEDLINE, PsycARTICLES, PsycINFO, and ProQuest were used to identify relevant academic and grey literature published between 2013 and June 2021. Twenty-two studies were retrieved for the systematic review, and 19 studies qualified for the meta-analyses. We found that individuals with anorexia nervosa (AN) and bulimia nervosa (BN) were close to evenly represented across the *DSM–5* severity groups. Conversely, approximately 82% of those with binge eating disorder (BED) were categorized into the mild and the moderate severity groups. In terms of differences across the *DSM–5* severity groups for ED symptomatology, we found various significant differences for BN (g = .23–1.23), but only a few significant differences for AN (g = .002–.30) and BED (g = .13–1.55). Finally, using the alternative ED severity rating overvaluation of weight and shape, significant differences were observed in ED psychopathology between the low and high weight/shape overvaluation BED severity groups (g = .13–1.23). The findings of this review provide support for the clinical utility of the *DSM–5* severity ratings in BN, but not in AN or BED, and suggest an alternative severity classification for BED based on overvaluation of weight and shape that may have utility as a severity indicator.

Public Health Significance Statement

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–5*; American Psychiatric Association [APA], 2013) introduced severity ratings for anorexia nervosa (AN), bulimia nervosa (BN) and binge eating disorder (BED) to assist clinical practice. Our review concludes that except for the *DSM–5* severity ratings for BN, limited evidence supports the clinical utility of the *DSM–5* severity ratings for AN and BED as these severity ratings do not adequately capture the spectrum of ED psychopathology that patients experience. Therefore, there is a need to develop alternate, clinically useful, severity ratings for AN and BED.

Keywords: eating disorders, severity ratings, *DSM–5*, overvaluation of weight and shape, meta-analysis

Supplemental materials: https://doi.org/10.1037/cps0000078.supp

Eating disorders (EDs) are associated with numerous acute complications, from long-term to permanent medical, psychological, and social consequences (Wolfi & Treasure, 2011). As the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–5*; American Psychiatric Association [APA], 2013) was being reviewed, the ED Work Group presented new ED severity ratings. These severity ratings for EDs were an attempt to overcome the limitations of the dichotomous diagnostic approach used in the former *DSM–IV-R* (APA, 2000), and to provide clinicians with ED-specific makers of disorder severity and clinical prognosis (APA, 2013). This was accomplished by establishing gradients for disorder severity by classifying individuals with an ED into "mild," "moderate," "severe," and "extreme" categories. Specifically, the *DSM–5* categorizes individuals with anorexia nervosa (AN) into mild (≥17.0 kg/m^2), moderate (16–16.99 kg/m^2), severe (15–15.99 kg/m^2), and extreme (<15 kg/m^2) severity groups using body mass index (BMI). The severity for bulimia nervosa (BN) is determined by the frequency of inappropriate compensatory behaviors, whereas the severity of binge eating disorder (BED) is assessed by the frequency of binge eating episodes. The severity categories for BN and BED are mild (1–3 episodes/week), moderate (4–7 episodes/week), severe (8–13 episodes/week), and extreme (14 or more episodes/

This article was published Online First March 31, 2022.

An Binh Dang https://orcid.org/0000-0001-7961-6913
Sarah Giles https://orcid.org/0000-0001-5228-121X
Matthew Fuller-Tyszkiewicz https://orcid.org/0000-0003-1145-6057
Litza Kiropoulos https://orcid.org/0000-0002-1921-5904
Isabel Krug https://orcid.org/0000-0002-5275-3595

Correspondence concerning this article should be addressed to Isabel Krug, Melbourne School of Psychological Sciences, The University of Melbourne, Redmond Barry, Level 7, North Melbourne, VIC 3051, Australia. Email: isabel.krug@unimelb.edu.au

week). Other specified feeding or eating disorder (OSFED) and unspecified feeding or eating disorder (UFED) do not have severity ratings in the *DSM–5*.

The introduction of the *DSM–5* ED severity ratings sparked significant interest amongst ED researchers as these severity indices were not empirically identified prior to their inclusion in *DSM–5*, and evidence regarding their clinical utility is inconsistent (Lydecker et al., 2020). Therefore, there is a need to summarize and synthesize the existing literature examining the *DSM–5* ED severity indices to determine the clinical utility of markers of disorder severity. The insights gained from such undertaking may help to inform clinical practice (e.g., clinical supervision, or risk of complications) and policy (e.g., treatment access or funding) and ensure that individuals with EDs receive an appropriate level of care.

DSM–5 Severity Ratings for Eating Disorders

With respect to the severity indicators for AN, some studies report that a low BMI is associated with increased medical (e.g., cardiac problems) and psychological complications (e.g., comorbidity with anxiety), and a poorer prognosis (Button et al., 2010; Wolfi & Treasure, 2011), while other studies did not find such associations (e.g., Hebebrand et al., 1996). These inconsistent findings may partially be due to the severity indices overshadowing important physical (e.g., weight history), psychological or cognitive factors (e.g., drive for thinness), which might reflect the individuals' levels of impairment (Grilo et al., 2015a; Hirst et al., 2017). Therefore, it is unclear whether BMI is useful in indexing AN severity.

The *DSM–5* ratings for BN and BED are based on the frequency of inappropriate compensatory behaviors for BN and binge eating episodes for BED. Some studies have found support for the utility of these severity ratings with the frequency of compensatory and binge eating episodes being positively correlated with general psychopathology and ED symptoms for individuals with BN and BED, respectively (Peñas-Lledó et al., 2002; Shroff et al., 2006). However, despite being these positive associations between, the severity indices for BN and BED have been criticized due to the arbitrariness of the frequency classifications. Specifically, it is unclear why three episodes per week is considered mild while four episodes per week is classified as moderate, and whether these cut-off scores reflect meaningful differences in terms of psychopathology, distress, and prognosis. As such, clarification is needed regarding the utility of the severity indicators for BN and BED in *DSM–5*.

Alternative Severity Indices for Eating Disorders

Given the ambiguity surrounding the utility of the *DSM–5* ED severity ratings, researchers have sought alternative severity classifications for each ED subtype, including transdiagnostic indices (e.g., overvaluation of weight and shape; Gianini et al., 2017). While the *DSM–5* approach argues for the use of distinct severity markers (e.g., BMI) for each ED diagnosis, the transdiagnostic approach, defined as common factors that contribute to the development and maintenance across multiple ED subtypes, may provide greater insights into symptom severity both within and between diagnosis groups (Fairburn et al., 2003).

Overvaluation of weight and shape, defined as the excessive influence of one's body on self-worth, is one of the cardinal features of EDs which motivates and maintains disordered eating (Fairburn et al., 2003). Considerable research supports the view that the cognitive components of EDs, such as overvaluation of weight and shape, more so than behavioral ED symptoms (e.g., binge eating frequency), are core maintaining processes of EDs (e.g., Hoiles et al., 2012). Furthermore, overvaluation of weight and shape is also a prominent cognitive feature for individuals with BED, and reliably predicts elevated levels of psychosocial impairment (Hrabosky et al., 2007). Studies have found that individuals with BED who overvalue their shape and weight endorse significantly higher levels of ED psychopathology, functional impairments, distress, and demonstrate poorer prognosis compared with those with BED and low weight and shape overvaluation (e.g., Goldschmidt et al., 2011). Overall, overvaluation of weight and shape has been shown to be useful as either a distinctive alternative severity classification for BED or a transdiagnostic severity rating across AN, BN, and BED.

Drive for thinness, defined as an extreme fear of weight gain resulting in disordered eating patterns, mainly restrictive eating, has also been proposed as an alternative transdiagnostic severity marker for AN, BN, and BED (Peñas-Lledó et al., 2015). Research has shown that patients with AN and BN with a higher drive for thinness experienced more co-occurring psychiatric conditions and ED symptomatology than those with a lower drive for thinness (Abbate-Daga et al., 2007). Furthermore, using the Eating Disorder Inventory-2 (EDI-2; Garner et al., 1983), individuals with BED were found to have more excessive concerns with shape and thinness compared with healthy controls (Kuehnel & Wadden, 1994). Therefore, drive for thinness appears to be an important transdiagnostic cognitive process in EDs, which may represent a clinically useful severity index across EDs.

Finally, the number of idiosyncratic purging methods has been proposed as a severity classification (Gianini et al., 2017). The number of purging methods has been associated with ED related psychopathology (Edler et al., 2007), with studies demonstrating that patients who engaged in multiple purging behaviors exhibited significantly more self-injury behaviors, psychiatric comorbidities, and ED psychopathology in comparison to individuals with a single purging method (Eddy et al., 2009; Edler et al., 2007). Thus, the number of purging methods may provide valuable clinical information in differentiating severity within BN.

The Current Review

Research findings regarding the clinical utility of the *DSM–5* ED severity categories are mixed, creating confusion as to whether researchers and clinicians should use these ratings and how to best conceptualize ED severity. Therefore, systematically reviewing and synthesizing the existing literature may help to elucidate the clinical usefulness of both the *DSM–5* and the alternative severity indicators for AN, BN and BED.

Understanding the clinical utility of the *DSM–5* and alternative severity ratings for EDs can assist the development of ED severity ratings in the upcoming edition of the *DSM*, which in turn can inform clinical practice. Hence, the aims of this study were threefold. First, we systematically reviewed all the studies to date assessing the *DSM–5* and the alternative severity indicators (e.g.,

weight and shape concerns and drive for thinness) for AN, BN and BED to obtain an overall picture of the clinical utility of these approaches in indexing ED severity. Second, we conducted a quality appraisal of the studies retrieved. Third, using meta-analytic techniques, we synthesized the available data of the studies included in the review regarding the clinical utility of the *DSM–5* and alternative ED severity ratings.

Method

Systematic Review

Search Strategy

The aims and methods of the current review were pre-registered in the PROSPERO database (ID CRD42020171211). A search for relevant academic published or grey literature papers was carried out utilizing four international databases (EMBASE, PsycINFO, PsycARTICLES, MEDLINE) for papers written in English between January 1, 2013 (the year when the *DSM–5* was published) and June 5, 2021.

For each database, the following keywords in the title and or abstract were used: "severity rating*" and related terms (e.g., "severity indices" or "severity specifiers"). They were then cross-referenced using the conjunction "AND" with each of the following search terms (a) "eating disorder*" OR (b) anorex*, OR (c) bulimi*, OR (d) "binge eating disorder", OR (e) "inappropriate compensatory behav*" (Supplementary 1).

A complementary search of the ProQuest Dissertations & Theses database and the International Conference on Eating Disorder (ICED) hosted by the Academy for Eating Disorders—the largest international ED conference was conducted. The same parameters as the academic database search were used to search for unpublished papers in the ProQuest database. We hand-searched relevant abstracts which referred to ED severity ratings in ICED 2013 to 2020 conferences. To further minimize potential publication bias, two major listservs, one international (1,045 members) and one Australia-based, were sent requests for published or unpublished manuscripts meeting the review eligibility criteria. Forward and backward referencing was conducted on articles that met the inclusion criteria of the current review.

Eligibility Criteria

The following inclusion criteria were applied to studies that: (a) used the *DSM–5* severity ratings for either adolescents or adults with AN, BN, or BED, and (b) utilized at least one variable of interest, which were ED or general psychopathology, psychiatric comorbidity, putative maintenance factors, psychological distress, impairment, prognosis, or quality of life.

The following exclusion criteria were applied to studies that (a) did not use the *DSM–5* severity indices for AN, BN, or BED; (b) did not report at least one variable that reflected ED severity (e.g., ED psychopathology); (c) used ecological momentary assessment as this review did not aim to assess how these severity ratings changed in real-time; (d) were literature reviews or case reports; and (e) were published in a language other than English.

Initially, this review aimed to assess the clinical usefulness of the *DSM–5* severity ratings for AN, BN, and BED. While screening for these studies, we found studies that assessed the clinical utility of alternative severity ratings against the *DSM–5* severity ratings. We therefore decided to extract relevant information on the clinical utility of these alternative severity ratings in addition to the *DSM–5* severity ratings.

Study Selection and Data Extraction

Search outputs were cross-referenced, and duplicates removed. Titles and abstracts were assessed by the first author AD against the inclusion and exclusion criteria. The full text of all remaining articles was examined by authors AD and SG to determine their eligibility for inclusion in the systematic review. When there was uncertainty, the authors consulted with a senior author (IK). The agreement between the two raters was 85% at the full-text screen, and the Cohen kappa value was .70, which indicates a substantial level of agreement.

Once the study was included, the following data were extracted for each article: (a) bibliographical data (author, publication year, country); (b) study setting; (c) the number of individuals in each severity group; (d) sample characteristics (e.g., age, BMI); (e) ED and general psychopathology diagnostic tools; and (f) level of ED and general psychopathology.

Quality Appraisal

The quality and risk of bias of each included study was appraised independently by two coders (AD and SG) using a revised version of the Downs and Black (1998) criteria, which has recently been adapted to be used in the ED literature (Rozenblat et al., 2017). Based on the revised criteria, the quality of a study was appraised on whether they (a) described the aims, objectives, hypothesis; (b) outlined main outcomes in the introduction or method section; (c) specified participant characteristics that are appropriate for ED severity indices research; (d) described main findings; (e) stated characteristics of participants lost to follow-up (if relevant); (f) had a representative sample of the population examined (including clinical but not convenience samples); (g) included exact probability values reported (or confidence intervals included); (h) noted any "data-dredging" explicitly; (i) used appropriate statistical tests; (j) employed valid and reliable outcome measures; (k) recruited participants in different groups contemporaneously; (l) conducted adequate adjustment for confounding variables; and (m) had sufficient power. Any discrepancies amongst these criteria between the two coders were resolved by discussion and consensus with a senior author (IK).

Meta-Analysis

Inclusion and Exclusion Criteria

For the meta-analysis, ED and general psychopathology were chosen as outcome variables as they were consistently reported across studies. Specifically, to be included in the meta-analysis, studies had to report sufficient statistical information (e.g., mean, standard deviations, and/or raw data) to calculate effect sizes for our outcome variables (i.e., general and ED psychopathology) for each of the *DSM–5* severity categories for AN, BN, and BED. If this data were available for the alternative severity ratings (e.g., drive for thinness), this information was also extracted to calculate effect sizes for the alternative severity ratings. Further

information was requested from the authors if the paper had insufficient information.

Statistical Analysis

R package metafor (Viechtbauer, 2010) was used to analyze the data. To establish the proportion of each *DSM–5* severity index for AN, BN, and BED, the data were pooled using random-effects model based on the DerSimonian-Laird method. To compare the differences in eating psychopathology between the severity categories using either the *DSM–5* or the alternative severity classification systems, summarized mean differences (SMD) were calculated when comparing two severity groups successively within each ED diagnosis. Hedge's g was chosen as a measure of effect size, and was interpreted according to Cohen (1988), which specified 0.2 as small, 0.5 as medium, and 0.8 as a large effect size.

The Eggers' test was used to examine publication bias. If publication bias was indicated by the Eggers' test, a Duval and Tweedie's "trim-and-fill" analysis (Duval & Tweedie, 2000) was applied to re-calculate the pooled SMD for any studies that may introduce publication bias. For all analyses, Cochrane's Q with $p < .05$ was chosen to evidence statistical heterogeneity between studies, and I^2 was used to quantify the degree of heterogeneity; for I^2 a value of 25%, 50%, and 75% implies low, moderate, and high heterogeneity, respectively (Higgins et al., 2003).

Results

Systematic Review

The systematic search resulted in 3,766 studies. After removing the duplicates, 3,094 studies remained; of which, we screened the titles and abstracts. No additional articles were identified through forward and backward search. Thirty-three full-text records, including three unpublished studies, were assessed for eligibility to be included in the review. We excluded 11 studies as they did not meet the inclusion criteria. Figure 1 presents the selection process and the reasons for exclusion.

Tables 1 and 2 summarize the 22 studies, including one unpublished study, that met the inclusion criteria for the systematic review. Eight studies (36%) were conducted in the U.S.; eight studies (36%) in Italy; two studies (9%) in Spain; and the remaining four studies (19%) in Portugal, Japan, the Netherlands, and the United Kingdom. A total of 9,196 individuals (6,990 females) participated in the studies reviewed. Thirteen studies used clinical samples with predominantly adult females (i.e., 82% females; Dakanalis, Bartoli, et al., 2017; Dakanalis & Clerici, 2017; Dakanalis, Colmegna, et al., 2017; Dakanalis, Riva, et al., 2017; Dakanalis, Timko, et al., 2018; Forrest et al., 2020; Gianini et al., 2017; Gorrell et al., 2019; Grilo et al., 2015a; Jenkins et al., 2016; Machado et al., 2017; Nakai et al., 2017; Zayas et al., 2018). Two studies comprised mainly adult females from the community (i.e., 70% females; Grilo et al., 2015b, 2015c). Two studies utilized adolescent participants from clinical settings (Dakanalis, Colmegna, et al., 2018; Dakanalis, Zanetti, et al., 2018), while one study used community-based adolescent participants (Smink et al., 2014). Three studies employed only females from clinical settings (Dalle Grave et al., 2018; Krug et al., 2021; Smith et al., 2017), and one study (Krug et al., 2022) used an all-male clinical sample.

Study Measurements

Table 3 summarizes measurements that were used across the studies to measure ED psychopathology, putative factors (e.g., perfectionism), psychiatric comorbidities (e.g., depression), and psychological impairments. Twelve (54.5%) studies used the Eating Disorder Examination (EDE; Fairburn et al., 1993), five (22.7%) used Eating Disorder Questionnaire (EDE-Q; Fairburn & Beglin, 1994); and two (9.0%) used the EDI-2 to measure ED psychopathology. One study (4.5%) used both the EDE and EDE-Q, and one study (4.5%) used both the EDE and EDI to measure ED psychopathology. One study (4.5.%; Smink et al., 2014) did not measure ED psychopathology.

Transdiagnostic Alternative Severity Indicators

To assess severity based on the level of overvaluation of weight and shape, items on the EDE (Fairburn et al., 1993) or the EDE-Q (Fairburn & Beglin, 1994) were used. The overvaluation of shape/weight was assessed by the following same two items on either the EDE and EDE-Q: "Over the past 4 weeks, has your *shape* influenced how you feel about (judge, think, evaluate) yourself as a person?" and "Over the past 4 weeks has your *weight* influenced how you feel about (judge, think, evaluate) yourself as a person?" Three studies (Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015b) used the EDE with a cut-off score of 4 to differentiate clinical (i.e., score ≥ 4 on either of the two overvaluation items mentioned above) from subclinical levels of overvaluation of weight and shape. Only Grilo et al. (2015c) used a cut-off score of 5 on the EDE-Q to differentiate individuals with clinical from subclinical levels of overvaluation of weight and shape.

To assess drive for thinness, the EDI-2 (Garner et al., 1983) was used in Krug et al. (2021) and Krug et al. (2022), with both studies using a cut-off score of 14, according to the recommendations of Garner et al. (1984), to distinguish clinical from subclinical levels of drive for thinness.

DSM–5 Severity Indices for AN

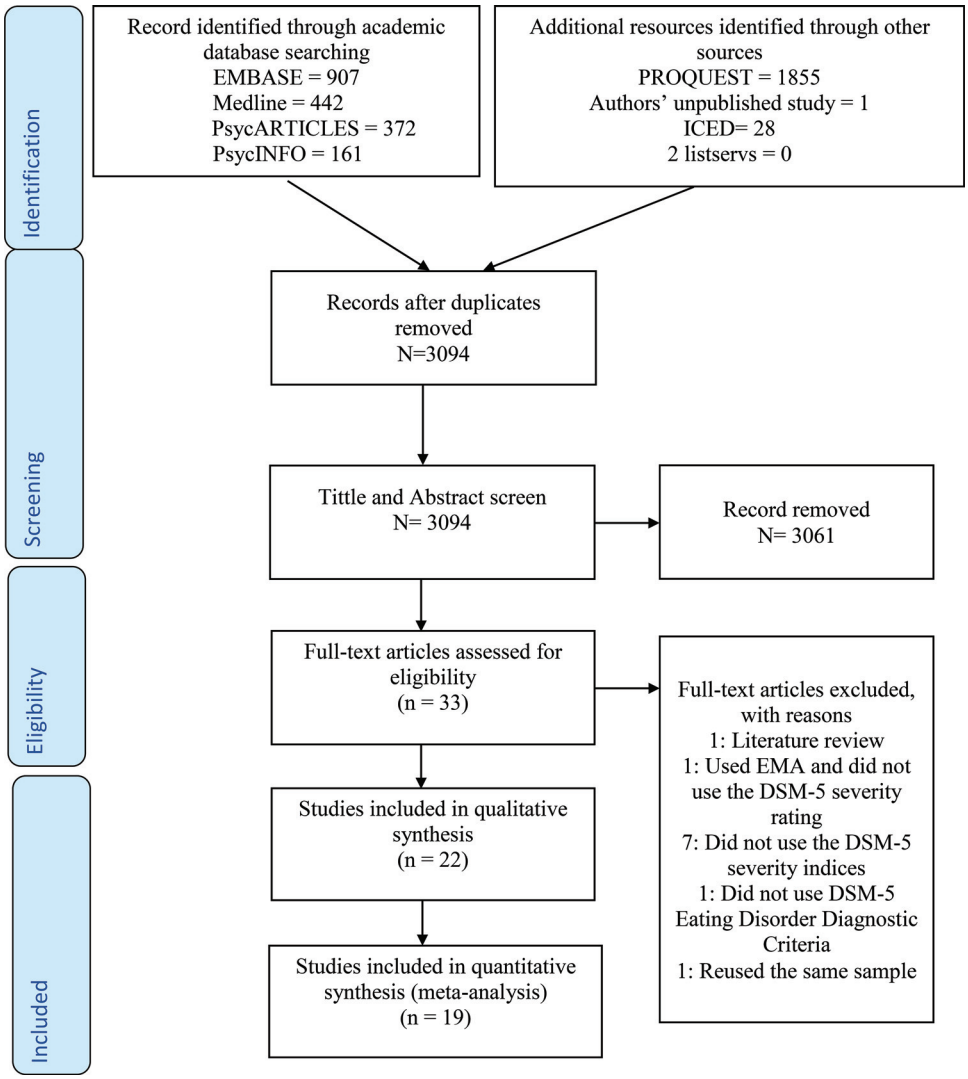
Ten papers examined the AN severity index using BMI, including nine clinical (e.g., Dakanalis, Timko, et al., 2018) and one community study (Smink et al., 2014). There was a total sample of 2,126 individuals with AN, with the sample sizes outlined in the studies ranging between 16 to 741 participants. Individuals with AN had mean ages ranging from 21.1 to 26.2 years of age, and BMIs varying from 15.43 to 17.16 kg/m^2.

Individuals with AN were quite evenly represented across the *DSM–5* severity groups (ranging from 16% to 28% in each of the four severity groups; Dakanalis, Timko, et al., 2018; Gianini et al., 2017). Only two studies (Krug et al., 2021; Nakai et al., 2017) provided the frequency distribution for the AN-restrictive (AN-R) and the AN-binge purging (AN-BP) diagnostic groups separately using the *DSM–5* severity rating for AN. Both studies revealed that more patients with AN-R (i.e., up to 70%) were categorized in the severe and extreme groups compared to the individuals with AN-BP (i.e., up to 42.5%).

A positive association was found between the number of hospitalizations and the *DSM–5* severity rating for AN (Gianini et al., 2017). However, most studies found no significant differences

Figure 1

PRISMA Flow Chart of Systematic Review and Meta-Analysis Selection Process



Note. See the online article for the color version of this figure.

between the clinical presentation associated with AN across the *DSM–5* severity categories (e.g., age of ED onset, distress, impairment, and ED psychopathology; Dalle Grave et al., 2018; Machado et al., 2017). Notably, while individuals with a *DSM–5* AN severity rating of severe and extreme presented with higher ED symptoms overall (Dakanalis, Timko, et al., 2018; Nakai et al., 2017), those in the mild and moderate categories exhibited more bulimic behaviors compared with the more severe categories (Dalle Grave et al., 2018). Even most studies examining the utility of the *DSM–5* AN severity ratings comprised an adequate sample size almost all studies included predominantly female participants. Only Zayas et al. (2018) assessed whether the *DSM–5* AN severity rating functioned differently between males and females, and found that BMI poorly capture AN severity in both males and females. Such findings were replicated by Krug et al. (2022), which used an all-male clinical sample.

Regarding AN subtypes, most studies did not examine whether the AN severity index differed between the AN-R and the AN-BP subtypes (e.g., Zayas et al., 2018) despite previous research having shown distinctive diagnostic characteristics (e.g., elevated emotion dysregulation in the AN-BP subtype) between these AN subtypes (Racine & Wildes, 2013). Only Krug et al. (2021) examined whether the AN severity index differed between the AN-R and the AN-BP subtypes. Results showed that the *DSM–5* severity ratings for AN provided more meaningful differentiation for AN-R in clinical information, self-harming behaviors, suicide attempts, and drug and alcohol abuse between the severity categories compared to AN-BP.

DSM–5 Severity Indices for BN

Regarding the *DSM–5* severity categories for BN, a total of 13 studies comprising 11 clinical samples (e.g., Dakanalis, Clerici, et

Table 1
Studies Investigated DSM–5 Severity Ratings for AN, BN, and BED

ED subtypes	Author(s)	Country and study setting	Sample size (female/total)	Sample of each severity index (1 = mild, 2 = moderate, 3 = serve, 4 = extreme serve)	Mean age (*SD*)	Measures	Relevant results	Conclusion
Anorexia nervosa	Dakanalis, Timko, et al. (2018)	Italy, Clinical	261/273	1 = 63, 2 = 71, 3 = 73, 4 = 66	25.26 (7.15)	EDE, SCID/ Patient Edition, BSI, EDI-3, and CIA-3.0	ED psychopathology and putative factors: 1 < 2<3 < 4	Support
	Dalle Grave et al. (2018)	Italy, Clinical	128/128	1 = 41, 2 = 20, 3 = 18, 4 = 49	25.5 (5.9)	EDE-Q and BSI	ED psychopathology: N.S.	Limited support
	Gianini et al. (2017)	U.S., Clinical	128/136	1 = 39, 2 = 38,3 = 27, 4 = 32	25.42 (7.76)	EDE, EDE-Q, CIA, BDI	ED psychopathology: N.S.	Limited support
	Krug et al. (2021)	Spain, Clinical	340/340	1 = 320, 2 = 210, 3 = 105, 4 = 116	25.37 (7.92)	EDI-2, SCL-90, TCI-R	ED and general psychopathology: N.S.	Limited support
	Krug et al. (2022)	Spain, Clinical	0/60	1 = 35, 2 = 12, 3 = 8; 4 = 5	25.55 (8)	EDI-2, SCL-90, TCI-R	ED and general psychopathology: N.S.	Limited support
	Machado et al. (2017)	Portugal, Clinical	193/201	1 = 73, 2 = 40, 3 = 30, 4 = 58	22.4 (9.5)	EDE, EDE Q	ED psychopathology: N.S.	Limited support
	Nakai et al. (2017)	Japan, Clinical	unavailable / 79	1 = 26, 2 = 16, 3 = 19, 4 = 46	24.32 (8.4)	EDE-Q	ED psychopathology: 1 & 2 > 3 & 4	Support
	Smink et al. (2014)	Netherlands, Community	15/16	1 = 5, 2 = 4, 3 = 4, 4 = 3	U/A	EDE	ED psychopathology: N.S.	Limited support
	Smith et al. (2017)	U.S., Mixed	109/109	1 = 70, 2 = 26, 3 = 9, 4 = 4	25.04 (8.48)	EDE, BDI, STAI, EDQOL	ED and general psychopathology: N.S.	Limited support
	Zayas et al. (2018)	U.S., Clinical	275/355	1 = 136, 2 = 90, 3 = 75, 4 = 54	23.31 (9.64)	EDE-Q, EDI-3	ED psychopathology: N.S in females and males	Limited support
Bulimina nervosa	Dakanalis, Bartoli, et al. (2017)	Italy, Clinical	261/281	1 = 54, 2 = 83, 3 = 85, 4 = 59	26.55 (8.04)	EDE, SCID/ Patient Edition, YBC-EDS, BSI, CIA, SAS-SR, CPQ	ED psychopathology: 1 < 2<3 < 4	Support
	Dakanalis, Clerici, et al. (2017)	Italy, Clinical	321/345	1 = 94, 2 = 90, 3 = 86, 4 = 75	27.05 (7.45)	EDE, SCID/ Patient Edition, EDI-3, K6, SASR	ED and general psychopathology: 1 < 2<3 < 4	Support
	Dakanalis, Colmegna, et al. (2018)	Italy, Clinical	256/272	1 = 79, 2 = 77, 3 = 64, 4 = 52	15.3 (1.7)	EDE, EDI-3, K-SADS, Peds QOL, SAAS, OBCS, SAS-SR	ED and general psychopathology: 1 < 2<3 < 4	Support
	Gianini et al. (2017)	U.S., Clinical	92/93	1 & 2 = 30; 3 = 35, 4 = 38**	23.03 (4.67)	EDE-Q, BDI, CIA, RSES.	Eating concern and clinical impairment: 1 < 2<3 < 4	Limited support

(*table continues*)

Table 1 (*continued*)

ED subtypes	Author(s)	Country and study setting	Sample size (female/total)	Sample of each severity index (1 = mild, 2 = moderate, 3 = serve, 4 = extreme serve)	Mean age (*SD*)	Measures	Relevant results	Conclusion
	Gorrell et al. (2019)	U.S., Clinical	110/185	1 = 17, 2 = 32, 3 = 37, 4 = 24	U/A	EDE	ED and general psychopathology at end of treatment: 1 < 2<3 < 4; at follow up: N.S	Moderate support
	Grilo et al. (2015b)	U.S., Community	103/199	1 = 77, 2 = 68, 3 = 32, 4 = 22	32.77 (10.05)	EDE-Q, BDI	ED and general psychopathology: N.S.	Limited support
	Jenkins et al. (2016)	United Kingdom, Clinical	204/214	1 = 94, 2 = 70, 3 = 32, 4 = 8	27.8 (9.5)	EDE-Q, CIA, CORE-OM	ED and general psychopathology: N.S.	Limited support
	Krug et al. (2021)	Spain, Clinical	1364/1364	1 = 640, 2 = 358, 3 = 63, 4 = 303	27.25 (8.22)	EDI-2, and SCL	ED and general psychopathology: N.S	Limited support
	Krug et al. (2022)	Spain, Clinical	0/83	1 = 48, 2 = 14, 3 = 3; 4 = 18	28.16 (90.8)	EDI-2, SCL-90, TCI-R	ED psychopathology: N.S.	Limited support
	Nakai et al. (2017)	Japan, Clinical	unavailable/125	1 = 54, 2 = 253, 3 = 21, 4 = 25	28.08 (7.05)	EDE-Q	Eating concern: 1 & 2 < 3 & 4	Limited Support
	Smink et al. (2014)	Netherland, Community	7/8	1 = 5, 2 = 1, 3 = 1, 4 = 1	U/A	EDE	Distribution: 1 & 2 < 3 & 4	Moderate Support
	Smith et al. (2017)	U.S., Mixed	68/76	1 = 17, 2 = 23, 3 = 16, 4 = 20	28.61 (10.8)	EDE, BDI, MOS SF-6,	Bulimic symptoms: 1 & 2 < 3 & 4	Limited support
	Zayas et al. (2018)	U.S., Clinical	134/177	1 = 33, 2 = 61, 3 = 37, 4 = 46	24.42 (8.67)	EDE-Q, EDI-3	Provided support for BN severity indices for females only	Limited support
Binge eating disorder	Dakanalis, Colmegna, el al. (2017)	Italy, Clinical	168/195	1 = 65, 2 = 69, 3 = 31, 4 = 30	37.25 (9.32)	EDE, BES SCID-I/Patient Edition, BSI, SAR-SR, EDI-3	ED and general psychopathology: 1 < 2 < 3 < 4	Support
	Dakanalis, Riva, et al. (2017)	Italy, Clinical	168/189	1 = 59, 2 = 63, 3 = 34, 4 = 33.	37.53 (8.49)	EDE, BES, OBCQ, YBOCS-BE, SCID-I/Patient Edition, K-6, SAS-SR, TOMS, IIP.	ED and general psychopathology: 1 < 2< 3 < 4	Support
	Dakanalis, Zanetti, et al. (2018)	Italy, Clinical	192/223	1 = 80, 2 = 85, 3 = 30, 4 = 28	15.30 (1.7)	EDE, PedsQOLs, K-SADS, EDI-3	ED and general psychopathology: 1 < 2< 3 < 4	Support
	Forrest et al. (2020)	U.S., Clinical	585/788	1 = 415, 2 = 284, 3 = 66, 4 = 23	45.93 (9.87)	EDE, SCID-I/ Patient Edition,	ED and general psychopathology: 1 < 2< 3 < 4	Support

(*table continues*)

Table 1 (*continued*)

ED subtypes	Author(s)	Country and study setting	Sample size (female/total)	Sample of each severity index (1 = mild, 2 = moderate, 3 = serve, 4 = extreme serve)	Mean age (*SD*)	Measures	Relevant results	Conclusion
	Gianini, et al. (2017)	U.S., Clinical	247/343	1 = 157, 2 = 142, 3 = 27, 4 = 17	46.71 (30.35)	EDE, BDI, RSES, MOS SF-36	ED and general psychopathology: 1 < 2< 3 < 4	Support
	Grilo et al. (2015a)	U.S., Community	299/338	1 = 264, 2 = 67, 3 = 6, 4 = 1	36.06 (11.86)	QEWP-R, EDE-Q, BDI	ED psychopathology: 1 < 2< 3 < 4. Depressive symptoms: N.S.	Moderate support
	Grilo et al. (2015c)	U.S., Clinical	623/834	1 = 331, 2 = 395, 3 = 83, 4 = 25	45.43 (10.02)	EDE, BDI	ED and general psychopathology: 1 < 2< 3 < 4	Moderate support
	Krug et al. (2021)	Spain, Clinical	329/329	1 = 138, 2 = 147, 3 = 12, 4 = 32	36 (11.11)	EDI-2, and SCL	ED and general psychopathology: N.S	Limited support
	Krug et al. (2022)	Spain, Clinical	0/35	1 = 16, 2 = 17, 3 = 1; 4 = 1	37.49 (12.71)	EDI-2, SCL-90, TCI-R	ED and general psychopathology: N.S.	Limited support
	Nakai el al. (2017)	Japan, Clinical	unavailable/41	1 = 9, 2 = 21, 3 = 11***	24.96 (7.07)	EDE-Q	ED and general psychopathology: N.S.	Limited support
	Smink et al. (2014)	Netherlands, Community	20/25	1 = 12, 2 = 10, 3 = 2, 4 = 1	U/A	EDE	ED and general psychopathology: 1 < 2< 3 < 4	Support
	Smith et al. (2017)	U.S., Mixed	191/216	1 = 79, 2 = 85, 3 = 44, 4 = 8	46.77 (10.35)	EDE, BDI, MOS SF-36	ED and general psychopathology: N.S.	Limited support

Note. BDI = Beck Depression Inventory; BCQ = Body Check Questionnaire; BES = Binge-Eating Scale; BSI = Brief Symptom Inventory; CIA = Clinical Impairment Assessment; CORE-OM = Clinical Outcomes in Routine Evaluation—Outcome Measure; CPQ = Clinical Perfectionism Questionnaire; EDE = Eating Disorder Examination—Interview; EDE-Q = Eating Disorder Questionnaire; EDI-2 = Eating Disorder Inventory (2.0); EDI-3 = Eating Disorder Inventory (3.0); EDQOL = Eating Disorder Quality of Life Questionnaire; IIP = Inventory of Interpersonal Problem; K6 = Kessler Screening Scale for Psychological Distress – 6 items; K-SADS = Schedule for Affective Disorder and Schizophrenia for School-Age Children; MOS SF-36 = Medical Outcomes Study Health Status Survey Physical Health Component; OBCS = Objectified Body Consciousness Scale; PedsQOL = Pediatric Quality of Life Inventory; RSES = Rosenberg Self-Esteem Scale; TOM = general subscale of the Tolerance of Mood States Scale; SAAS = Social Appearance Anxiety Scale; SAS-SR = Social Adjustment Scale-Self-Report; SCID/Patient Edition = Structured Clinical Interview *DSM–IV* Axis I Disorder Patient Edition; SCL-9 = Symptom Checklist Revised; STAI = State-Trait Anxiety Inventory; TCI-R = Temperament and Character Inventory-Revised; YBC-EDS = Yale-Brown Cornell Scale Eating Disorder Scale; N.S. = no significant differences in any variable across severity groups. Limited support = no significant differences in any variable between severity groups or only differences in one or two variables with small effect sizes between any groups. Moderate support = significant differences between severity groups across variables with medium effect sizes. Support = significant differences between all severity groups across a wide range of variables with large effect sizes.

** Severity groups were combined to have sufficient power.

*** There were no participants were categorized into “extreme” group.

Table 2
Alternative Classification Using Overvaluation of Weight and Shape

Alternative classifications	Author(s)	Country and settings	Sample of each severity index	Measure	Cut-off score	Findings	Conclusion
OWS	Forrest et al. (2020)	U.S., Clinical	Low OWS = 349, High OWS = 439	EDE	4	ED and general psychopathology: 1 < 2	Support
	Gianini et al. (2017)	U.S., Clinical	BED: Low OWS = 104, High OWS = 239; Transdiagnostic: Low OWS = 128, High OWS = 424	EDE	4	ED and general psychopathology, quality of life, self-esteem, psychological impairments: 1 < 2	Support.
	Grilo et al. (2015a)	U.S., Community	Low OWS = 130, High OWS = 196	EDE	4	ED and general psychopathology: 1 < 2	Support
	Grilo et al. (2015c)	U.S., Clinical	Low OWS = 384, High OWS = 449	EDE-Q	5	ED and general psychopathology: 1 < 2.	Support.
Number of purging methods	Gianini et al. (2017)	U.S., Clinical	Single method = 65, Multiple method = 28			ED and psychopathology: single method < multiple methods.	Moderate support
Drive for thinness	Krug et al. (2021)	Spain, Clinical	AN: Low DT = 590, High DT = 151; BN: Low DT = 166, High DT = 1198; BED: Low DT = 180, High DT = 149	EDI-2, SCL-90, TCI-R	14	ED and general psychopathology: Low DT <High DT.	Support
	Krug et al. (2022)	Spain, Clinical	AN: Low DT = 53, High DT = 7; BN: Low DT = 59, High DT = 24; BED: Low DT = 11, High DT = 24	EDI-2, SCL-90, TCI-R	14	ED and general psychopathology for AN and BED: Low DT <High DT. N.S for BED	Moderate support

Note. AN = anorexia nervosa; BN = bulimia nervosa; BED = binge eating disorder; DT = drive for thinness; OWS = overvaluation of weight and shape; EDE = Eating Disorder Examination—Interview; EDE-Q = Eating Disorder Questionnaire; EDI-2 = Eating Disorder Inventory (2.0); SCL-9 = Symptom Checklist Revised; TCI-R = Temperament and Character Inventory-Revised. Moderate support = significant differences between severity groups across variables with medium effect sizes. Support = significant differences between all severity groups across a wide range of variables with large effect sizes.

al., 2017) and two community samples (Grilo et al., 2015c; Smink et al., 2014) were included in the review. A total of 3,347 individuals with BN participated in these studies with the sample sizes ranging from eight to 1,365 participants, with mean ages ranging from 15.3 to 32.77, and mean BMI varying from 21 to 29.3 kg/m^2.

Regarding the frequency distribution of the BN severity groups, six studies (e.g., Dakanalis, Clerici, et al., 2017; Gorrell et al., 2019) found close to evenly distributed cases across the four severity groups. In contrast, six other clinical and community-based studies revealed that up to 76% of individuals with BN fell into the mild or the moderate categories (e.g., Grilo et al., 2015c; Krug et al., 2021). Specifically, Grilo et al. (2015c) and Jenkins et al. (2016) were only able to classify a small number of individuals into the severe and extreme *DSM–5* BN severity groups. These authors therefore collapsed the severe and extreme groups into one "severe-extreme" group, leaving the analyses to only three BN severity groups. Contrastingly, Gianini et al. (2017) had to combine the mild and moderate BN severity groups into one "mild-moderate" group for further analyses, as they only had seven participants in the mild category. Such inconsistencies between studies regarding the frequency distribution of BN severity groups may be due to the small sample sizes in these categories or differences in study settings (i.e., clinical versus community). Particularly, Jenkins et al.'s (2016) and Gianini et al. (2017) overall sample sizes were small, thus limiting the extent to which their sample represented the spectrum of severity within BN. Meanwhile, Grilo et al. (2015c) assessed their BN severity groups in a community setting, where, as expected, only a small number of individuals fell into the severe and extreme groups.

Concerning the utility of BN severity ratings, some studies found significant differences in BN severity groups for psychopathology, psychiatric comorbidity, and distress (e.g., Dakanalis, Clerici, et al., 2017), while other studies did not find significant differences in the aforementioned variables (e.g., Gianini et al., 2017; Jenkins et al., 2016). Only Zayas et al. (2018) and Krug et al. (2022) assessed how the *DSM–5* severity ratings for BN functioned in males. Both studies found that the *DSM–5* severity ratings for BN were able to differentiate ED psychopathology in females but not males.

DSM–5 Severity Indices for BED

Twelve papers, comprising 10 clinical-based studies (e.g., Dakanalis, Colmegna, et al., 2017) and two community-based studies (Grilo et al., 2015b; Smink et al., 2014) investigated the clinical utility of *DSM–5* BED severity ratings. Out of these 12 papers, only Dakanalis, Zanetti, et al. (2018) used an adolescent only sample. The studies included a total of 3,356 participants with BED. Sample sizes ranged between 25 to 788 participants, with mean ages ranging

Table 3
Measurements Used Across Studies

Variables	Measure	Studies	Studies
ED psychopathology	Eating Disorder Examination—Interview (EDE; Fairburn et al., 1993)	12	AN (n = 1): Dakanalis, Timko, et al. (2018). BN (n = 5): Dakanalis, Bartoli, et al. (2017); Dakanalis, Clerici, et al. (2017); Dakanalis, Colmegna, et al. (2018); Gorrell et al. (2019); Grilo et al. (2015c). BED (n = 6): Dakanalis, Colmegna, et al. (2017); Dakanalis, Riva, et al. (2017); Dakanalis, Zanetti, et al. (2018); Grilo et al. (2015b); Forrest et al. (2020); Grilo et al. (2015a). AN, BN, and BED (n = 1): Smith et al. (2017).
	Eating Disorder Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994).	5	AN (n = 2): Dalle Grave et al. (2018); Machado et al. (2017). BN (n = 2): Grilo et al. (2015c; Jenkins et al., 2016). AN, BN, and BED (n = 1): Nakai et al. (2017).
	Eating Disorder Inventory-2 (EDI; Garner & Olmsted, 1991)	2	AN, BN, and BED (n = 2): Krug et al. (2021); Krug et al. (2022).
	EDE and EDE-Q	1	AN, BN, and BED (n = 1): Gianini et al. (2017).
	EDE and EDI	1	AN and BN (n = 1): Zayas et al. (2018).
Putative factor for ED	Eating Disorder Inventory-3 (EDI-3; Garner, 2004)	6	AN (n = 1): Dakanalis, Timko, et al. (2018). BN (n = 2): Dakanalis, Clerici, et al. (2017); Dakanalis, Colmegna, et al. (2018). BED (n = 3): Dakanalis, Colmegna, et al. (2017); Dakanalis, Riva, et al. (2017); Dakanalis, Zanetti, et al. (2018).
General psychological impairment	Brief Symptom Inventory-3 (De Leo et al.,1993)	4	AN (n = 2): Dalle Grave et al. (2018); Dakanalis, Timko, et al. (2018). BN (n = 1): Dakanalis, Bartoli, et al. (2017). BED (n = 1): Dakanalis, Colmegna, et al. (2017).
	Clinical Impairment Assessment (Bohn & Fairburn, 2008	3	AN (n = 1): Dakanalis, Timko, et al. (2018). BN (n = 1): Jenkins et al. (2016). AN, BN, and BED (n = 1): Gianini et al. (2017).
	Kessler 6-item of Psychological Distress Scale (Furukawa et al., 2003)	2	BED (n = 2): Dakanalis, Colmegna, et al. (2017); Dakanalis, Riva, et al. (2017)
	Temperament and Character Inventory-Revised (TCI-R; Spanish version; Gutiérrez-Zotes et al., 2004)	2	AN, BN, and BED (n = 1): Krug et al. (2021, 2022)
	Clinical Outcome in Routine Evaluation—Outcome Measure (Barkham et al., 2001).	1	BN (n = 1): Jenkins et al. (2016)
Psychiatric comorbidity	Structured Clinical Interview for *DSM–IV* Axis I Disorders (SCID-I/P; First et al., 1996)	4	AN (n = 1): Dakanalis, Timko, et al. (2018) BN (n = 1): Dakanalis, Clerici, et al. (2017). BED (n = 2): Dakanalis, Colmegna, et al. (2017); Dakanalis, Riva, et al. (2017).
	Beck Depression Inventory (Beck et al., 1987)	4	BN (n = 1): Grilo et al. (2015c). BED (n = 1): Grilo et al. (2015b). AN, BN, and BED (n = 2): Gianini et al. (2017); Smith et al. (2017).
	Schedule for Affective Disorder and Schizophrenia for School-Age Children (K-SADS; Kaufman, 1997)	2	BN (n = 1): Dakanalis, Colmegna, et al. (2018). BED (n = 1): Dakanalis, Zanetti, et al. (2018).

from 15.3 to 46.77 years of age, and mean BMI from 28.98 to 38.07 kg/m$^{2.}$

These twelve studies found that up to 80% of those with BED were categorized into the mild and moderate groups. Despite only 20% of patients with BED being classified into the severe and extreme groups, this was still sufficient to warrant comparisons across categories.

Regarding the clinical utility of the *DSM–5* severity ratings for BED, some clinical-based studies found no differences between the BED severity groups in terms of ED related pathology, quality of life, and co-occurring depressive symptoms (e.g., Nakai et al., 2017; Smith et al., 2017), while other studies found that more severe (i.e., severe and extreme severity specifiers) binge-eating categories reported worse ED symptomatology, and higher levels of psychiatric comorbidity (e.g., Gianini et al., 2017; Grilo et al., 2015a, 2015b). Grilo et al.'s (2015b) community-based study found that individuals in the severe and extreme BED categories experienced significantly higher ED-related psychopathology. Relatedly, in a clinical-adolescent-sample, Dakanalis, Zanetti, et al. (2018) found that the *DSM–5* BED severity ratings significantly differentiated individuals with BED in terms of BMI as well as ED and general psychopathology. Overall, these findings provide

mixed support regarding the clinical utility of the current *DSM–5* severity index for BED. These inconsistent findings may be attributable to the heterogeneity of the BED participants included within the studies (e.g., comorbid obesity), which may have impacted the findings.

Alternative Classifications System for BED Using Shape and Weight Overvaluation

For the alternative severity classification system using overvaluation of weight and shape, three clinical-based studies (Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015a) and one community-based study (Grilo et al., 2015b) were identified. There were a total of 2,588 participants included in these studies. Participants age ranged from 36.06 to 45.93 years of age, and their BMI ranged from 21 to 29.3 kg/m^2.

Both clinical- and community-based studies found that individuals with BED were almost evenly represented in the low and high overvaluation of weight and shape categories (i.e., 42.23% versus 57.77%). All studies found weight and shape overvaluation to be superior to the *DSM–5* severity index for BED, as it offered greater distinctions between the BED severity groups in clinical information, ED psychopathology, putative maintenance factors (e.g., self-esteem), and psychiatric comorbidity (e.g., Grilo et al., 2015a, 2015b).

Alternative Transdiagnostic Severity Classification Systems

Three studies (Gianini et al., 2017; Krug et al., 2021; Krug et al., 2022) examined the discriminative capacity of the common transdiagnostic cognitive vulnerabilities across AN, BN, and BED in indexing ED severity compared with the *DSM–5* approach. Specifically, Gianini et al. (2017) investigated overvaluation of weight and shape, while Krug et al. (2021) and Krug et al. (2022) examined drive for thinness as an alternative transdiagnostic severity classification system across AN, BN, and BED.

Regarding the use of overvaluation of weight and shape as a transdiagnostic severity indicator across AN, BN, and BED, 552 individuals with AN, BN, and BED (mean age = 35.36 years, mean BMI = 31.5 kg/m^2) participated in Gianini et al.'s (2017) study. Using the clinical suggested cut-off score of 4 on the EDE (Fairburn et al., 1993) to differentiate clinical from subclinical levels of overvaluation of weight and shape, 30% and 70% of patients with EDs were categorized in the low and high overvaluation of weight and shape groups, respectively. Gianini et al. found significant differences between subclinical and clinical overvaluation of weight and shape regarding general and ED psychopathology, quality of life, depressive symptoms, and psychological impairments. However, Gianini et al.'s findings might have been confounded as they did not control for comorbid obesity in those with BED nor did they consider differences in treatment settings (e.g., inpatient, outpatient, and residential care).

Unlike Gianini et al. (2017), Krug et al. (2021) assessed drive for thinness as an alternative severity classification system for each ED subtype separately. The sample included 2,410 women with a threshold ED diagnosis (i.e., AN: n = 741, BN: n = 1,340, BED: n = 329), with a mean age of 27.6 years and a mean BMI of 24.01 kg/m^2. Using the EDI-2, Krug et al. (2021) found that 30% of individuals with AN, 95% of individuals with BN, and 45% individuals with BED were classified into the high drive for thinness group. Furthermore, Krug et al. (2021) supported the use of drive for thinness as an alternative transdiagnostic classification system across AN, BN and BED over the *DSM–5*. Specifically, they found significant differences between the low and high drive for thinness groups in impulsivity behaviors, general and ED psychopathology across ED diagnoses.

Using an all-male clinical sample (N = 178; AN: n = 60, BN: n = 83, BED: n = 35) with a mean age of 21.1 years and mean BMI of 25.34 kg/m^2 (SD = 8.84), Krug et al. (2022) assessed the clinical usefulness of: (a) the *DSM–5*, (b) drive for thinness, and (c) a combined *DSM–5*/drive for thinness severity categories for AN, BN, and BED. There was a low frequency of participants within the *DSM–5* severity categories moderate, severe, and extreme across all ED diagnoses. As such, the authors reclassified these groups into one moderate-to-extreme group. Krug et al. (2022) found that 11.7% of those with AN, 28.9% of those with BN, and 57.1% of the male individuals with BED were classified into the high drive for thinness group. Conversely, 56.7%, 44.2%, and 74.3% of male individuals with AN, BN, and BED were classified into the moderate to extreme/high combined *DSM–5*/drive for thinness severity group, respectively. This study revealed that the drive for thinness severity approach was better at differentiating between AN and BN severity groups in ED and general psychopathology in males with EDs than the *DSM–5* and the combined *DSM–5*/drive for thinness approach. Conversely, for patients with BED neither of the three severity classification approaches seemed clinically useful.

Alternative Classifications for BN Using Number of Purging Methods

Only Gianini et al. (2017) compared the *DSM–5* BN severity index with the number of purging methods as an alternative BN severity classification system using a sample of 93 patients with BN (mean age = 23.03 years, mean BMI = 22.08 kg/m2). Using the number of purging methods (i.e., self-induced vomiting, laxative, and diuretic misuse) utilized in the last 3 months as the purging method index, 65 (67%) individuals with BN were classified as "single purging method" users, while 28 (43%) individuals with BN were labelled as "multiple purging method" users. Gianini et al. found that patients with a single purging method had a significantly lower level of eating and weight concerns than those in the multiple purging methods group.

Findings of the Quality Appraisal

The results of the quality appraisal of included studies are presented in Table 4. The agreement between the coders was 85%. The quality and risk appraisal found that most studies used valid and reliable measurements and clearly reported and described their results. The main potential problems in the studies reviewed included poor control of confounding variables (e.g., medications), low statistical power, and a failure to report exact p-values.

Table 4

Downs and Black (1998) Checklist for Methodological Quality Adapted to Evaluate Studies Identified in the Systematic Review of the DSM–5 Severity Ratings for ED

Study	a	b	c	d	e	f	g	h	i	j	k	l	m
Dakanalis, Bartoli, et al. (2017)	1	1	1	1	N/A	1	1	0	1	1	N/A	1	1
Dakanalis, Clerici, et al. (2017)	1	1	1	1	N/A	1	1	0	1	1	N/A	1	1
Dakanalis, Colmegna, et al. (2017)	1	1	1	1	1	1	1	0	1	1	N/A	1	1
Dakanalis, Riva, et al. (2017)	1	1	1	1	N/A	1	0	0	1	1	N/A	1	1
Dakanalis, Colmegna, et al. (2018)	1	1	1	1	N/A	1	1	0	1	1	N/A	1	1
Dakanalis, Timko, et al. (2018)	1	1	1	1	N/A	1	1	0	1	1	N/A	1	1
Dakanalis, Zanetti, et al. (2018)	1	1	1	1	N/A	1	1	0	1	1	N/A	1	1
Dalle Grave et al. (2018)	1	1	1	1	N/A	0	0	0	1	1	N/A	0	0
Forrest et al. (2020)	1	1	1	1	N/A	1	1	0	1	1	N/A	0	1
Gianini et al. (2017)	1	1	0	1	N/A	X	1	0	1	1	N/A	0	0
Gorrell et al. (2019)	1	1	1	1	N/A	X	1	0	1	1	N/A	0	0
Grilo et al. (2015a)	0	1	0	1	N/A	X	1	0	1	1	N/A	0	1
Grilo et al. (2015b)	0	0	0	1	N/A	X	0	0	1	1	N/A	0	1
Grilo et al. (2015c)	1	1	1	1	N/A	X	0	0	1	1	N/A	1	1
Jenkins et al. (2016)	1	1	1	1	N/A	X	1	0	1	1	N/A	0	0
Krug et al. (2021)	1	1	1	1	N/A	1	0	0	1	1	N/A	0	1
Krug et al. (2022)	1	1	1	1	N/A	1	0	0	1	1	N/A	0	0
Machado et al. (2017)	1	0	1	1	N/A	X	1	0	1	1	N/A	0	1
Nakai et al. (2017)	1	1	1	1	N/A	X	1	0	1	1	N/A	0	0
Smink et al. (2014)	1	1	1	1	N/A	0	0	0	1	1	N/A	0	0
Smith et al. (2017)	1	1	X	1	N/A	X	1	0	1	1	N/A	0	0
Zayas et al. (2018)	1	1	1	1	N/A	1	0	0	1	1	N/A	0	1

Note. 1 = criteria met; 0 = criteria not met; X = unable to determine; N/A = criteria not applicable. Columns: a = the aims, objectives, hypothesis; b = outlined main outcomes in introduction or method; c = specified participant characteristics that are appropriate for severity indices and ED research; d = described main findings; e = stated characteristics of participants lost to follow-up (if relevant); f = had representative sample of the population (including clinical, but not convenience samples); g = included exact probability values reported (or confidence intervals included); h = noted any "data-dredging" explicitly; i = used appropriate statistical tests; j = employed valid and reliable main outcome measures; k = recruited participants in different groups contemporaneously; l = conducted adequate adjustment for confounding variables; and m = had sufficient power.

Meta-Analysis

As shown in Figure 1, 19 studies were included in the meta-analysis. Twelve authors were emailed and asked for additional data. Four authors provided additional data that were not reported in publications. The number of studies that were included in each analysis varied depending on the ED diagnoses and the variables of interest (e.g., eating concerns).

Initially, this review aimed to conduct a meta-analysis to examine the clinical utility of the number of purging methods as an alternative severity indicator for BN, and overvaluation of weight and shape and drive for thinness as transdiagnostic severity indicators across all EDs. However, due to insufficient data, the clinical utility of these variables as transdiagnostic severity across the different subtypes of EDs could not be examined. Instead, we explored the clinical utility of overvaluation of weight and shape as an alternative severity indicator for BED alone. Because the two shape and weight concern subscales were shown to be strongly correlated in the EDE and EDE-Q (r = .85 to .95; Peterson et al., 2007), we combined these two subscales into one overall weight/shape concerns scale for all subsequent analyses to reduce the number of comparisons between severity groups and the Type I error rate. Furthermore, the four models (i.e., four severity groups) in the current review applied weights on each study depending on the number of cases in each severity categories. As the number of cases in each severity categories was different across studies, each category was weighted differently in each model. Therefore, this would lead to the frequency distribution of the severity groups not sum up to 100%.

Meta-Analyses of the *DSM–5* Severity Categories for EDs

The Frequency Distribution of the DSM–5 Severity Categories

For AN, analysis from 10 studies showed that 38% of patients fell into the mild (95% CI [28, 46]; I^2 = 93%; $p < .01$), 22% into the moderate (95% CI [19, 26]; I^2 = 61%; $p < .01$), 17% into the severe (95% CI [13, 21]; I^2 = 73.50%; $p < .01$), and 21% into the extreme (95% CI [13, 29]; I^2 = 94.51%; $p < .01$) *DSM–5* AN severity groups.

For BN, meta-analysis from 13 studies found that 31% individuals were categorized in the mild (95% CI [23, 39]; I^2 = 95%; $p < .01$), 25% in the moderate (95% CI [28, 30]; I^2 = 49%; $p = .02$), 21% in the severe (95% CI [13, 31]; I^2 = 97%; $p < .01$), and 20% in the extreme (95% CI [15, 25]; I^2 = 88%; $p < .01$) *DSM–5* BN severity groups.

For BED, the meta-analysis from 12 studies found the frequency distribution of BED severity groups were 43% of patients fell into the mild (95% CI [35, 51]; I^2 = 95%; $p < .01$), 39% into the moderate, (95% CI [33, 44]; I^2 = 89%; $p < .01$), 10% into the severe, (95% CI [7, 14]; I^2 = 91%; $p < .01$), and 5% into the extreme (95% CI [3, 9]; I^2 = 92%; $p < .01$) *DSM–5* BED severity groups. The amount of heterogeneity was substantial, with moderate to high amounts of variance for nearly all *DSM–5* severity groups. Meta-analyses and sub-group analyses were conducted and are presented in Supplementary 3 and 4.

The Clinical Usefulness of the DSM–5 Severity Indices

Table 5 shows the pooled mean differences in ED symptomatology between the *DSM–5* severity categories. Eggers' test identified three cases where publication bias may be present. The trim-and-fill analysis was used to provide adjusted SMD for these cases.

Anorexia Nervosa. Data from five studies (Dalle Grave et al., 2018; Gianini et al., 2017; Machado et al., 2017; Smith et al., 2017; Zayas et al., 2018) were pooled to assess the SMD between the *DSM–5* AN severity categories for global eating psychopathology, dietary restraint, eating, and weight/shape concerns. The results showed that individuals in the moderate *DSM–5* AN severity group experienced significantly less global eating psychopathology compared with those in the extreme group (g = –0.16). No other significant differences between the *DSM–5* AN severity groups in the remaining variables were found (Table 5). Significant heterogeneity, with moderate to high amount of variance across studies (I^2 from 23%–.91%), were found. Thus, meta regressions and subgroup analyses were conducted (see Supplementary 5 and 6).

Bulimia Nervosa. The SMD between the *DSM–5* BN severity categories for global eating psychopathology was pooled from four studies (Gianini et al., 2017; Grilo et al., 2015c; Smith et al., 2017; Zayas et al., 2018). The analyses revealed that those in the mild BN group exhibited significantly less global eating pathology compared to those in the severe BN groups (g = 0.35, $p < .01$), but not between the other severity groups.

Seven studies (Dakanalis, Bartoli, et al., 2017; Dakanalis, Clerici, et al., 2017; Dakanalis, Colmegna, et al., 2018; Gianini et al., 2017; Grilo et al., 2015c; Nakai et al., 2017; Zayas et al., 2018) were included in the analyses to assess the pooled SMD between the BN severity groups for the specific ED symptoms of dietary restraint and eating, weight/shape concerns. Individuals in the mild BN group experienced significantly less weight/shape concerns than those in the moderate (g = –0.22, p = .03), severe (g = –0.39, p = .04) and extreme BN groups (g = –0.61, p = .02). Furthermore, individuals in the mild BN group exhibited significantly less dietary restraint than those in the moderate (g = –0.40, p =.01), severe (g = –0.69, p = .04), and extreme groups (g = –1.05, p = .01), and significantly fewer eating concerns than those in the severe (g = –0.81, p = .02) and extreme groups (g = –1.23, $p <$ .01). Compared with those in the severe BN group, individuals in the moderate BN group experienced significantly fewer eating concerns (g = –0.52, p = .01). Those in the moderate BN group also displayed significantly less dietary restraint (g = –0.80, p = .03), weight/shape concerns (g = –0.68, $p < .01$) and eating concerns (g = –1.04, $p < .01$) than the extreme group. Lastly, compared with those in the extreme BN group, individuals in the severe group also had significantly less dietary restraint (g = –0.64, p =.01), weight/shape concerns (g = –0.46, p = .02), and eating concerns (g = –0.69, $p < .01$). Significant heterogeneity with high amount of variance was found between the studies (I^2 from 55%–. 91%). Meta regression and subgroup analyses were conducted and are presented in Supplementary 5 and 6.

Binge Eating Disorder. Pooled data from up to five studies (Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015a, 2015b; Smith et al., 2017) indicated that individuals with BED in the mild group exhibited significantly less global eating psychopathology compared with those in the severe (g = –0.48, p = .04) and extreme BED groups (g = –0.42, p = .03). No other significant differences were found in the remaining group comparisons.

Seven studies (Dakanalis, Colmegna, et al., 2017; Dakanalis, Riva, et al., 2017; Dakanalis, Zanetti, et al., 2018; Gianini et al., 2017; Grilo et al., 2015a, 2015b; Nakai et al., 2017) were included in the meta-analysis to assess the SMD between BED severity groups in dietary restraint, eating, and weight/shape concerns. The pooled data analyses demonstrated some differences between groups. Specifically, individuals in the moderate BED severity category exhibited significantly lower weight/shape concerns compared to those in the severe BED group (g = –0.25, p = .02). Those in the extreme BED group scored significantly higher in dietary restraint compared with those in the moderate (g = –0.98, p = .03) and severe groups (g = –0.49, $p < .01$). Individuals in the mild group had significantly lower eating concerns compared with those in the moderate (g = –0.4, $p < .01$), severe (g = –1.26, $p < .01$), and extreme BED groups (g = –0.54, $p < .01$). Significant heterogeneity with large variance between studies was identified (I^2 from 23%–93%; Table 5). Metaregressions subgroup analyses were conducted (see Supplementary 5 and 6).

Meta-Analysis of Alternative Severity Ratings

The Frequency Distribution of Overvaluation of Weight and Shape for BED

A meta-analysis using four studies (Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015a, 2015b) showed that using the overvaluation of weight and shape severity index, individuals with BED were close to evenly represented in both the low (40%; 95% CI [34, 46]) and high (60%; 95% CI [54, 66]) categories. Significant heterogeneity between studies were identified in low (I^2 = 89.35%, Q = 28.17, $p <$. 0001), and high overvaluation categories (I^2 = 89.35%, Q = 28.17, $p <$. 0001). Metaregression and subgroup analyses were conducted (see Supplementary 3 and 4).

Clinical Usefulness of Overvaluation of Weight and Shape for BED

Analysis of four studies (Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015a, 2015b) found significant differences in ED symptomatology between BED overvaluation of weight and shape severity groups. Particularly, those in the high overvaluation of weight and shape group endorsed significantly more global eating psychopathology (g = –1.23, $p < .01$), weight/shape concerns (g = –0.64, $p < .01$), dietary restraint (g = –0.16, p = .03), and eating concerns (g = –0.81, $p < .01$). No significant heterogeneities were found.

Discussion

This study is the first to systematically review and meta-analytically investigate the clinical utility of the *DSM–5* and alternative severity indices for EDs. The current systematic review included 22 studies that examined the clinical usefulness of the *DSM–5* severity in AN, BN, or BED within either a community or clinical setting. Most of the reviewed studies found modest to no support for the *DSM–5* ED severity indices (e.g., Smith et al., 2017). Similarly, the meta-analytic results, comprising 19 studies, provided limited support for the clinical utility of the *DSM–5* severity

Table 5

Summary of Overall Effect Sizes, Heterogeneity, and Egger's Test P-Value for Comparison Between DSM–5 AN, BN, and BED Severity Groups

			Heterogeneity					
Variables	Group comparisons	Studies(*n*)	Hedge's *g* with [95% CI]	*p*	*Q*	I^2	*p*	Eggers' test p-value
AN global psychopathology	Mild vs. Moderate	5	0.34 [–0.33, 1.02]	.23	26.33	85%	<.001	15
	Mild vs. Severe	5	0.17 [–0.42, 0.77]	.47	12.17	67.1%	.02	.09
	Mild vs. Extreme	5	–0.08 [–0.37, 0.21]	.49	4.81	17%	.31	.14
	Moderate vs. Severe	5	–0.01 [–0.32, 0.30]	.75	4.51	11%	.34	.58
	Moderate vs. Extreme	5	–0.16 [–0.24, –0.08]	.006	0.33	0%	.98	.49
	Severe vs. Extreme	5	–0.14 [–0.45, 0.17]	.29	4.11	3%	.39	.55
AN weight and shape concerns	Mild vs. Moderate	6	0.01 [–0.15, 0.17]	.85	3.03	0%	.70	.68
	Mild vs. Severe	6	–0.01 [–0.37, 0.34]	.92	12.88	61.2%	.02	.52
	Mild vs. Extreme	6	0.02 [–0.38, 0.42]	.92	18.56	73.1%	.002	.43
	Moderate vs. Severe	6	–0.07 [–0.24, 0.10]	.35	2.85	0%	.73	.38
	Moderate vs. Extreme	6	–0.06 [–0.33, 0.20]	.57	7.32	31.7%	.20	.32
	Severe vs. Extreme	6	–0.05 [–0.28, 0.17]	.57	4.85	0%	.43	.64
AN restraint	Mild vs. Moderate	6	–0.01 [–0.31, 0.33]	.95	11.68	52.2%	.04	.75
	Mild vs. Severe	6	–0.06 [–0.54, 0.41]	.75	27.14	81.6%	<.001	.82
	Mild vs. Extreme	6	–0.10 [–0.78, 0.58]	.73	49.22	89.8%	<.001	.62
	Moderate vs. Severe	6	–0.13 [–0.37, 0.10]	.21	5.52	9.4%	.37	.67
	Moderate vs Extreme	6	–0.13 [–0.51, 0.24]	.40	15.06	66.8%	.01	.37
	Severe vs. Extreme	7	–0.06 [–0.38, 0.26]	.64	8.68	42.2%	.12	.52
AN eating concern	Mild vs. Moderate	6	–0.02 [–0.31, 0.28]	.90	9.82	49.1%	.08	.77
	Mild vs. Severe	6	–0.09 [–0.63, 0.45]	.68	29.27	82.9%	<.001	.64
	Mild vs. Extreme	6	–0.09 [–0.86, 0.67]	.76	60.44	91.7%	<.001	.80
	Moderate vs. Severe	6	–0.15 [–0.38, 0.08]	.16	5.15	2.8%	.40	.33
	Moderate vs. Extreme	6	–0.14 [–0.62, 0.35]	.49	23.92	79.1%	.0002	.41
	Severe vs. Extreme	6	–0.06 [–0.38, 0.27]	.68	10.82	53.8%	.06	.47
BN global psychopathology	Mild vs. Moderate	5	–0.54 [–2.04, 0.96]	.37	25.83	84.5%	<.001	.40
	Mild vs. Severe	4	–0.35 [–0.49, –0.20]	.005	0.43	0%	.93	.40
	Mild vs. Extreme	4	–0.92 [–2.17, 0.33]	.10	11.77	74.5%	.01	.43
	Moderate vs. Severe	4	0.23 [–1.49, 1.94]	.70	34.08	91.2%	<.001	.47
	Moderate vs. Extreme	4	–0.30 [–1.04, 0.43]	.28	9.67	69%	.02	.73
	Severe vs. Extreme	4	–0.63 [–1.60, 0.34]	.13	13.99	78.6%	.003	.75
BN weight and shape	Mild vs. Moderate	7	–0.22 [–0.41, –0.03]	.03	6.66	9.9%	.35	.39
	Mild vs. Severe	7	–0.39 [–0.79, –0.01]	.04	25.14	76.1%	<.001	.21
	Mild vs. Extreme	7	–0.61 [–1.10, –0.11]	.02	32.89	81.8%	<.001	.05
	Moderate vs. Severe	7	–0.23 [–0.50, 0.05]	.09	13.84	56.7%	.03	.10
	Moderate vs. Extreme	7	0.68 [–1.14, –0.22] *	.009	30.17	80.1%	<.001	.04
	Severe vs. Extreme	7	–0.46 [–0.80, –0.12]	.02	18.30	67.2%	.005	.10
BN restraint	Mild vs. Moderate	7	–0.40 [–0.67, –0.13]	.01	13.20	54.6%	.04	.27
	Mild vs. Severe		–0.69 [–1.33, –0.04]	.04	65.23	90.8%	<.001	.25
	Mild vs. Extreme	7	–1.05 [–1,80, –0.30]	.01	67.27	91.1%	<.001	.09
	Moderate vs. Severe	7	–0.38 [–0.86, 0.10]	.10	41.76	85.6%	<.001	.09
	Moderate vs. Extreme	7	–0.80 [–1.48, –0.13]	.03	66.13	90.9%	<.001	.08
	Severe vs. Extreme	7	–0.64 [–1.09, –0.19]	.01	30.37	80.2%	<.001	.09
BN eating concerns	Mild vs. Moderate	7	–0.34 [–0.71, 0.03]	.06	22.25	73%	.001	.24
	Mild vs. Severe	7	–0.81 [–1.43, –0.19]	.02	57.02	89.5%	<.001	.35
	Mild vs. Extreme	7	–1.23 [–1.89, –0.57]	.01	58.44	89.7%	<.001	.29
	Moderate vs. Severe	7	–0.52 [–0.86, –0.17]	.01	23.51	74.5%	.001	.24
	Moderate vs. Extreme	7	–1.04 [–1.47, –0.61]	.001	30.38	80.3%	<.001	.29
	Severe vs. Extreme	7	–0.69 [–1.05, –0.33]	.003	20	70%	.003	.76
BED global eating psychopathology	Mild vs. Moderate	5	–0.27 [–0.55, 0.01]	.05	14.17	71.8%	.007	.55
	Mild vs. Severe	3	–0.48 [–0.89, –0.07]	.04	1.95	0%	.38	.68
	Mild vs. Extreme	3	–0.42 [–0.72, –0.11]	.03	0.43	0%	.81	.89
	Moderate vs Severe	3	–0.66 [–2.39, 1.06]	.24	23.5	91.5%	<.001	.40
	Moderate vs. Extreme	3	–0.13 [–0.89, 0.63]	.54	2.60	23.2%	.27	.68
	Severe vs. Extreme	3	0.21 [–0.97, 1.39]	.53	4.35	54.1%	.11	.10
BED weight and shape concerns	Mild vs. Moderate	7	–0.12 [–0.25, 0.01]	.07	6.91	13.1%	.33	.84
	Mild vs. Severe	5	–0.46 [–0.92, 0.01]	.05	8.97	55.4%	.06	.05
	Mild vs. Extreme	4	–0.81 [–1.76, 0.12]	.07	17.36	82.7%	.001	.53
	Moderate vs. Severe	5	–0.25 [–0.43, –0.07]	.02	1.57	0%	.82	.57
	Moderate vs. Extreme	4	–0.50 [–1.08, 0.08]*	.07	6.93	56.7%	.07	.01
	Severe vs. Extreme	4	–0.38 [–0.91, 0.16]	.11	4.45	32.6%	.22	.19
BED dietary restraint	Mild vs. Moderate	7	–0.22 [–0.45, 0.01]	.06	17.48	65.7%	.01	.31

(*table continues*)

Table 5 (*continued*)

Variables	Group comparisons	Studies(*n*)	Heterogeneity: Hedge's *g* with [95% CI]	*p*	*Q*	I^2	*p*	Eggers' test p-value
	Mild vs. Severe	5	–0.57 [–1.60, 0.46]	.20	35.71	88.8%	<.001	.59
	Mild vs. Extreme	4	–1.55 [–3.13, 0.02]	.05	43.24	93.1%	<.001	.42
	Moderate vs. Severe	5	–0.31 [–0.83, 0.22]	.18	10.48	61.8%	.03	.25
	Moderate vs. Extreme	4	–0.98 [–1.83, –0.13]	.03	13.91	78.4%	.003	.50
	Severe vs. Extreme	4	–0.49 [–0.94, –0.03]	.04	3.36	10.6%	.34	.93
BED eating concerns	Mild vs. Moderate	7	–0.40 [–0.67, –0.13]*	.008	22.44	73.3%	.001	.03
	Mild vs. Severe	5	–1.26 [–1.84, –0.67]	.004	16.09	75.1%	.003	.57
	Mild vs. Extreme	4	–0.91 [–3.72, –0.11]	.04	53.35	94.5%	<.001	.04
	Moderate vs. Severe	5	–0.54 [–0.84, –0.24]	.008	4.28	6.5%	.37	.13
	Moderate vs. Extreme	4	–1.06 [–2.25, 0.12]	.06	26.11	88.5%	<.001	.35
	Severe vs. Extreme	4	–0.58 [–1.38, 0.22]	.10	8.92	66.4%	.03	.25

* Indicated biased adjusted summarized mean differences due to Eggers' test indicates the presence of potential publication bias.

ratings for AN and BED. Importantly, the results did, however, provide evidence for the *DSM–5* severity index for BN and the alternative severity rating of overvaluation of weight and shape for BED.

Frequency Distribution of the *DSM–5* Severity Categorizations

For the *DSM–5* AN severity index, the results of the meta-analyses ($n = 10$) found that patients with AN were close to evenly represented across all the *DSM–5* severity groups (i.e., ranging from 18% to 34.27%). This finding is somewhat unsurprising as most studies (e.g., Dakanalis, Timko, et al., 2018; Gianini et al., 2017) that examined the *DSM–5* AN severity groups found a balanced distribution across the *DSM–5* severity groups for AN. Such distribution offers some level of support for the use of BMI to index severity.

In terms of BN, our meta-analysis ($n = 13$) found that those with BN were close to evenly distributed across the four BN severity groups (i.e., ranging from 20% to 31%). As summarized in the Systematic Review section, evidence as to whether the distribution of the *DSM–5* BN severity groups is balanced across severity groups is conflicting (e.g., Dakanalis, Bartoli, et al., 2017; Jenkins et al., 2016). Specifically, six studies (e.g., Dakanalis, Bartoli, et al., 2017) found close to evenly distributed cases across the four severity groups, while other six studies found that the majority of those with BN fell into the mild or moderate categories (e.g., Smith et al., 2017). Such inconsistencies may be due to the small sample sizes and differences in sample characteristics. Using meta-analytic techniques to aggregate the results from studies across multiple settings, the current review was able to provide new insights into the frequency distribution of the *DSM–5* severity groups for BN. Overall, in BN the distribution of the *DSM–5* severity categories appears to be balanced across the four severity categories, thus providing initial support for the validity of the *DSM–5* severity indices in BN.

Finally, for BED, the meta-analysis of 12 studies found that up to 82% of individuals with BED were within the mild to moderate *DSM–5* severity groups. This finding implies that almost all studies were unable to classify individuals with BED into severe or extreme groups (e.g., Gianini et al., 2017). This was surprising as almost all included studies were conducted in clinical settings. This raises questions about the validity of the *DSM–5* severity ratings for BED as they were unable to identify individuals at the severe or extreme end of BED severity.

The Clinical Utility of the *DSM–5* Severity Specifiers

Given the inconsistencies in the literature regarding the clinical utility of the *DSM–5* severity indices for AN, BN, and BED, this study was the first to apply meta-analytic techniques to gain further insights into the degree of ED psychopathology experienced by individuals in each of the severity groups within AN, BN, and BED.

Anorexia Nervosa

Our meta-analysis showed no significant differences in terms of ED psychopathology across the AN severity groups. This finding contrasts with previous studies, not included in the current review, which suggested that those with AN and a low BMI tended to present with greater ED and general psychopathology (Button et al., 2010; Franko et al., 2018). However, the finding that ED psychopathology did not vary according to BMI is consistent with most of the studies (e.g., Machado et al., 2017; Smith et al., 2017) summarized in our systematic review. These studies suggest that BMI may not adequately differentiate the degree of ED symptom severity that individuals with AN experience. Given the centrality of weight restoration in the treatment of AN, it is understandable that the *DSM–5* used current BMI to index AN severity, especially as it is considered convenient and non-invasive (Reas & Rø, 2017). However, our results suggest that BMI might not offer significant clinical utility to index severity, as it does not effectively reflect patients' level of psychopathology. Therefore, it is pertinent to explore alternative severity indices for AN that incorporate important biological (e.g., weight suppression—the differences between the highest weight and current weight), personal, and cognitive factors (e.g., drive for thinness, Krug et al., 2021) to accurately inform clinical practice.

Bulimia Nervosa

Our findings support the *DSM–5* severity ratings for BN. Individuals with BN in the milder severity groups endorsed significantly less dietary restraint, and eating and weight/shape concerns than those in the more severe groups, with a small to large effect

size. This finding is consistent with previous research that was not included in the current review, which found that the frequency of inappropriate compensatory behaviour was positively correlated with ED psychopathology (Peñas-Lledó et al., 2002; Shroff et al., 2006). Our findings also align with several studies included in the review (e.g., Dakanalis, Bartoli, et al., 2017), which revealed significant differences in general and ED psychopathology across the *DSM–5* BN severity groups. Overall, using the frequency of inappropriate compensatory behaviors as an indicator of BN severity seems to offer some degree of clinical utility in reflecting the level of ED symptomatology and treatment needs in those with BN.

Binge Eating Disorder

The current review did not provide robust support for the clinical utility of the frequency of binge eating episodes as a severity index for BED as outlined in the *DSM–5*. Although BED was recently introduced as an independent ED by the *DSM–5* (APA, 2013), there were a handful of studies that investigated the utility of the *DSM–5* severity index for BED (e.g., Forrest et al., 2020). Findings of our meta-analysis are aligned with several studies included in the systematic review (e.g., Nakai et al., 2017; Smith et al., 2017), who found that the severity specifier did differentiate ED symptomatology between a few—but not across all four—BED severity groups. However, several studies included in current the systematic review found that BED *DSM–5* severity groups significantly differed in terms of in ED psychopathology (e.g., Dakanalis, Colmegna, et al., 2017; Forrest et al., 2020). It should be noted that the aforementioned studies categories included individuals with BED with comorbid obesity. Because differences in clinical psychopathology between non-obese and obese patients with BED have been suggested (Goldschmidt et al., 2011), using the frequency of binge eating episodes might effectively capture BED severity in obese patients with BED but not in non-obese patients with BED. Further research is needed to test this hypothesis.

Overall, our results suggest that individuals with BED who engaged in at least one binge-eating episode per week are likely to experienced similar levels of ED psychopathology as those individuals with BED who engaged in more than four episodes of binge-eating per week. This finding suggests that the frequency of binge-eating episodes in isolation, without incorporating clinical judgment, may not offer insight into the severity of ED psychopathology in those with BED. Accordingly, an alternative severity classification using other cognitive, psychological, or biological factors for BED should be sought.

Alternative Severity Indicators

The systematic review component provided support for the use of the alternative transdiagnostic severity ratings of overvaluation of weight and shape (Gianini et al., 2017) and drive for thinness (Krug et al., 2021; Krug et al., 2022) across AN, BN, and BED. These ratings were able to distinguish ED psychopathology across the AN, BN, and BED severity groups more robustly than the *DSM–5* approach. We did not have sufficient data to undertake meta-analyses on the clinical usefulness of these alternative transdiagnostic severity indicators (i.e., drive for thinness, overvaluation of weight and shape) for AN, BN, and BED. However, we were able to apply meta-analytic techniques to examine the clinical utility of overvaluation of weight and shape as a distinctive alternative BED severity rating.

Distribution of Overvaluation of Weight and Shape

Regarding the frequency distribution of overvaluation of weight and shape as an alternative severity specifier for BED, our meta-analysis (n = 4) found that individuals with BED were close to evenly presented in both the low and high overvaluation of weight and shape categories (i.e., 40% versus 60%, respectively). Such close to even distribution is consistent with the frequency distributions reported in previous studies (e.g., Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015a, 2015b). Despite this area of research still being in its infancy, the current results offer provisional evidence for the use of overvaluation of weight and shape as an alternative severity classification for BED.

Clinical Utility of Overvaluation of Weight and Shape as Severity Ratings for BED

Concerning the clinical utility of overvaluation of weight and shape as an alternative severity rating for BED, our meta-analysis found that individuals in the high overvaluation of weight and shape group exhibited significantly greater ED symptomatology than the low overvaluation group, with small to large effect sizes. These findings are consistent with studies that were included in our systematic review (e.g., Forrest et al., 2020; Gianini et al., 2017; Grilo et al., 2015a, 2015b). These studies also found overvaluation of weight and shape to be a better severity indicator for BED than the *DSM–5* approach. This finding is consistent with the notion that weight and shape overvaluation is the cardinal feature for BED. However, given this area of research is still emerging and the current review is based on a limited number of studies, we must temper any conclusions regarding the clinical utility of overvaluation of weight and shape as an alternative severity rating for BED. Further research replicating these findings, while incorporating larger sample sizes and different study designs to address the limitations of existing studies, is needed before clinicians can incorporate overvaluation of weight and shape as the defining severity indicator for BED. Such contributions to the literature are essential to increase diagnostic accuracy and improve our understanding of symptom severity in BED.

Implications

The current review offers important implications for our understanding of the *DSM–5* severity specifiers for EDs. First, both our systematic review and meta-analysis offered limited support for the *DSM–5* severity ratings for BED and AN, respectively. This suggests that the *DSM–5* AN and BED severity ratings do not differentiate between individuals in terms of ED psychopathology. We found that compared with the *DSM–5* severity specifiers, the alternative severity classification based on overvaluation of weight and shape were able to significantly differentiate BED patients in terms of ED psychopathology, thus highlighting the clinical utility of this alternative approach. Consequently, it may be beneficial to consider alternative severity classifications for EDs on a wide range of commonly shared variables (i.e. affect, cognitive, biological, or personality) that can occur across all ED diagnoses.

Second, although BMI and the frequency of binge eating did not meaningfully differentiate the severity of AN and BED, respectively,

this does not entirely rule out the usefulness of having a severity classification for these EDs. Rather, to guide treatment and level of care, clinicians should consider a wider range of symptoms that a patient may present with, rather than relying on singular symptoms to denote disorder severity. The *DSM–5* does allow clinicians to modify a patient's severity classification based on the degree of functional impairment, the intensity of symptoms, and the need for supervision. This flexibility may help to improve the clinical utility of the *DSM–5* severity specifiers as they currently stand. However, research investigating if and how clinicians apply these recommendations, and whether this clinical discretion does, in fact, enhance the clinical utility of the *DSM–5* severity specifiers for EDs is lacking.

Strengths and Limitations of the Current Review

The current review has several limitations which must be acknowledged. First, this review was only able to obtain data to compare the differences in ED psychopathology between the *DSM–5* severity categories for AN, BN, and BED to investigate their clinical utility. However, a good measure of severity should be able to index the intensity of psychopathological features of the disorder and the level of care and treatment outcomes (i.e., predictive validity; Gianini et al., 2017). Therefore, future research should include additional biomedical, clinical, and functional indicators for the *DSM–5* severity ratings in their studies to cast light on the predictive validity of the *DSM–5* severity ratings for EDs.

Second, this review offers evidence supporting the *DSM–5* severity ratings for BN, while offering limited support for the *DSM–5* severity indicators for AN and BED. However, current findings are based on a small number of studies, and almost all studies included in this review were cross-sectional studies, therefore replication of current findings using different study designs or different data collection methods (e.g., multi-method approach) are needed. Moreover, most participants included in the review were female and adult. This is important given the significant differences in ED psychopathology which exist between males and females (Murray et al., 2017) and between adolescents and adults (Rajagopalan et al., 2019). This limitation could be addressed by using subgroup analyses to account for gender and age effects. Due to the limited studies that included males or adolescents, we could not conduct subgroup analyses. Future research investigating the clinical utility of the *DSM–5* and alternative ED severity ratings in males and adolescents is needed.

Third, within the meta-analyses, substantial heterogeneity of the observed effects was identified. However, most of the included studies did not include sufficient information for their samples to allow us to identify the source of such heterogeneity. Therefore, future research should explore other possible moderators such as duration of illness or medication use, which could not be modelled in current analyses due to insufficient data.

The current review also had several strengths. First, the inclusion of unpublished studies reduced any potential impact of publication bias that may have been presented if only peer-reviewed publications were included (Borenstein et al., 2009). Second, the use of meta-analytic methods allowed us to pool effect sizes, thus disentangling conflicts regarding the clinical utility of the *DSM–5* severity indices. Finally, our review also provides insight into the quality of studies that have investigated in this topic, thus providing insight into possible methodological recommendations for future studies.

Future Directions

Given that the *DSM–5* is governed by Western thinking and most studies included in our review were conducted in Western countries, the results cannot be generalized to non-Western countries. A universal BMI cut-off score is not suitable to use across all ethnic groups due to differences in body compositions across ethnicity (Soh et al., 2006). Thus, Western governed BMI cut-off for AN severity may not yield the same sensitivity in other ethnic groups. Future research should investigate how the *DSM–5* and alternative severity ratings for EDs perform for individuals from a range of different ethnic backgrounds.

Furthermore, the OSFED is significantly more prevalent (i.e., 11.2%) compared with full criterial ED (i.e., 6.2%; Galmiche et al., 2019). Individuals with OSFED experience adverse health consequences equivalent to those individuals with a full threshold ED (Krug et al., 2022). Despite this, the *DSM–5* did not establish severity ratings for OSFED. Only Krug et al. (2021) used the frequency of inappropriate compensatory behaviour and drive for thinness as an attempt to index severity for OSFED/purging disorder, and found limited support for such usage. Future research should investigate severity ratings across all OSFED subtypes to assist with clinical management.

Conclusion

This systematic review and meta-analysis assessed the clinical utility of the *DSM–5* severity indices for AN, BN, and BED across all published and unpublished studies since the introduction of the *DSM–5* ED severity indices. Within the retrieved studies, we also assessed several alternative severity indices, such as weight and shape concerns. The main finding from the meta-analyses was that, except for BN, the current literature offers little support of the clinical utility of the *DSM–5* severity specifiers for AN and BED. For BED, using overvaluation of weight and shape as an alternative severity classification offered better discriminative capacity between severity groups in comparison to the current *DSM–5* approach. Further exploration regarding alternative severity ratings for EDs is needed.

References

Abbate-Daga, G., Pierò, A., Gramaglia, C., Gandione, M., & Fassino, S. (2007). An attempt to understand the paradox of anorexia nervosa without drive for thinness. *Psychiatry Research*, *149*(1-3), 215–221. https://doi.org/10.1016/j.psychres.2005.10.017

American Psychiatric Association. (APA). (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). https://doi.org/10.1176/appi.books.9780890420249.dsm-iv-tr

American Psychiatric Association. (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9781585629992

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes bench- marking using the CORE_OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*, 184–196.

Beck, A. T., Brown, G., Steer, R. A., Eidelson, J. I., & Riskind, J. H. (1987). Differentiating anxiety and depression: A test of the cognitive content-specificity hypothesis. *Journal of Abnormal Psychology*, *96*(3), 179.

Bohn, K., & Fairburn, C. G. (2008). The clinical impairment assessment questionnaire (CIA). In C. G. Fairburn (Ed.), *Cognitive behavioral therapy for eating disorders* (2nd ed., pp. 315–317). Guilford Press.

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221–235). Sage.

Button, E. J., Chadalavada, B., & Palmer, R. L. (2010). Mortality and predictors of death in a cohort of patients presenting to an eating disorders service. *International Journal of Eating Disorders*, *43*(5), 387–392. https://doi.org/10.1002/eat.20715

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale.

Dakanalis, A., Bartoli, F., Caslini, M., Crocamo, C., Zanetti, M. A., Riva, G., Clerici, M., & Carrà, G. (2017). Validity and clinical utility of the *DSM–5* severity specifier for bulimia nervosa: Results from a multisite sample of patients who received evidence-based treatment. *European Archives of Psychiatry and Clinical Neuroscience*, *267*(8), 823–829. https://doi.org/10.1007/s00406-016-0712-7

Dakanalis, A., & Clerici, M. (2017). Severity of binge-eating disorder and its effects on treatment outcome. *The Journal of Clinical Psychiatry*, *78*(7), e841. https://doi.org/10.4088/JCP.17lr11589

Dakanalis, A., Clerici, M., Riva, G., & Carrà, G. (2017). Testing the *DSM–5* severity indicator for bulimia nervosa in a treatment-seeking sample. *Eating and Weight Disorders*, *22*(1), 161–167. https://doi.org/10.1007/s40519-016-0324-2

Dakanalis, A., Colmegna, F., Riva, G., & Clerici, M. (2017). Validity and utility of the *DSM–5* severity specifier for binge-eating disorder. *International Journal of Eating Disorders*, *50*(8), 917–923. https://doi.org/10.1002/eat.22696

Dakanalis, A., Colmegna, F., Zanetti, M. A., Di Giacomo, E., Riva, G., & Clerici, M. (2018). Evaluation of the *DSM–5* severity specifier for bulimia nervosa in treatment-seeking youth. *Child Psychiatry and Human Development*, *49*(1), 137–145. https://doi.org/10.1007/s10578-017-0735-y

Dakanalis, A., Riva, G., Serino, S., Colmegna, F., & Clerici, M. (2017). Classifying adults with binge eating disorder based on severity levels. *European Eating Disorders Review*, *25*(4), 268–274. https://doi.org/10.1002/erv.2518

Dakanalis, A., Timko, C. A., Colmegna, F., Riva, G., & Clerici, M. (2018). Evaluation of the *DSM–5* severity ratings for anorexia nervosa in a clinical sample. *Psychiatry Research*, *262*, 124–128. https://doi.org/10.1016/j.psychres.2018.02.009

Dakanalis, A., Zanetti, M. A., Colmegna, F., Riva, G., & Clerici, M. (2018). Classifying binge eating-disordered adolescents based on severity levels. *Journal of Adolescence*, *62*, 47–54. https://doi.org/10.1016/j.adolescence.2017.10.003

Dalle Grave, R., Sartirana, M., El Ghoch, M., & Calugi, S. (2018). *DSM–5* severity specifiers for anorexia nervosa and treatment outcomes in adult females. *Eating Behaviors*, *31*, 18–23. https://doi.org/10.1016/j.eatbeh.2018.07.006

De Leo, D., Frisoni, G. B., Rozzini, R., & Trabucchi, M. (1993). Italian community norms for the Brief Symptom Inventory in the elderly. *British Journal of Clinical Psychology*, *32*(2), 209–213. https://doi.org/10.1111/j.2044-8260.1993.tb01045.x

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*(6), 377–384. https://doi.org/10.1136/jech.52.6.377

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Eddy, K. T., Crosby, R. D., Keel, P. K., Wonderlich, S. A., le Grange, D., Hill, L., Powers, P., & Mitchell, J. E. (2009). Empirical identification and validation of eating disorder phenotypes in a multisite clinical sample. *Journal of Nervous and Mental Disease*, *197*(1), 41–49. https://doi.org/10.1097/NMD.0b013e3181927389

Edler, C., Haedt, A. A., & Keel, P. K. (2007). The use of multiple purging methods as an indicator of eating disorder severity. *International Journal of Eating Disorders*, *40*(6), 515–520. https://doi.org/10.1002/eat.20416

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363–370.

Fairburn, C. G., Cooper, Z., & O'Connor, M. (1993). The eating disorder examination. *International Journal of Eating Disorders*, *6*, 1–8.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, *41*(5), 509–528. https://doi.org/10.1016/S0005-7967(02)00088-8

First, M., Spitzer, R., Gibbon, M., & Williams, J. (1996). *Structured Clinical Interview for DSM–IV Axis I Disorders (SCID-I) research version.* Biometric Research.

Forrest, L. N., Jacobucci, R. C., & Grilo, C. M. (2020). Empirically determined severity levels for binge-eating disorder outperform existing severity classification schemes. *Psychological Medicine*, *30*(1), 1–11. https://doi.org/10.1017/S0033291720002287

Franko, D. L., Tabri, N., Keshaviah, A., Murray, H. B., Herzog, D. B., Thomas, J. J., Coniglio, K., Keel, P. K., & Eddy, K. T. (2018). Predictors of long-term recovery in anorexia nervosa and bulimia nervosa: Data from a 22-year longitudinal study. *Journal of Psychiatric Research*, *96*, 183–188. https://doi.org/10.1016/j.jpsychires.2017.10.008

Furukawa, T. A., Kessler, R. C., Slade, T., & Andrews, G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine*, *33*(2), 357–362. https://doi.org/10.1017/S0033291702006700

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000-2018 period: A systematic literature review. *The American Journal of Clinical Nutrition*, *109*(5), 1402–1413. https://doi.org/10.1093/ajcn/nqy342

Garner, D. M. (2004). *Eating Disorder Inventory-3 (EDI-3). Professional manual*. Psychological Assessment Resources.

Garner, D. M., & Olmsted, M. P. (1991). *Eating Disorder Inventory 2*. Psychological Assessment Resources Odessa.

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*(2), 15–34. https://doi.org/10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6

Garner, D. M., Olmsted, M. P., Polivy, J., & Garfinkel, P. E. (1984). Comparison between weight-preoccupied women and anorexia nervosa. *Psychosomatic Medicine*, *46*(3), 255–266. https://doi.org/10.1097/00006842-198405000-00007

Gianini, L., Roberto, C. A., Attia, E., Walsh, B. T., Thomas, J. J., Eddy, K. T., Grilo, C. M., Weigel, T., & Sysko, R. (2017). Mild, moderate, meaningful? Examining the psychological and functioning correlates of *DSM–5* eating disorder severity specifiers. *International Journal of Eating Disorders*, *50*(8), 906–916. https://doi.org/10.1002/eat.22728

Goldschmidt, A. B., Le Grange, D., Powers, P., Crow, S. J., Hill, L. L., Peterson, C. B., Crosby, R. D., & Mitchell, J. E. (2011). Eating disorder symptomatology in normal-weight vs. obese individuals with binge eating disorder. *Obesity*, *19*(7), 1515–1518. https://doi.org/10.1038/oby.2011.24

Gorrell, S., Hail, L., Kinasz, K., Bruett, L., Forsberg, S., Delucchi, K., Lock, J., & Le Grange, D. (2019). A test of the *DSM–5* severity specifier for bulimia nervosa in adolescents: Can we anticipate clinical treatment

outcomes? *International Journal of Eating Disorders*, *52*(5), 586–590. https://doi.org/10.1002/eat.23034

Grilo, C. M., Ivezaj, V., & White, M. A. (2015a). Evaluation of the *DSM–5* severity indicator for binge eating disorder in a clinical sample. *Behaviour Research and Therapy*, *71*, 110–114. https://doi.org/10.1016/j.brat.2015.05.003

Grilo, C. M., Ivezaj, V., & White, M. A. (2015b). Evaluation of the *DSM–5* severity indicator for binge eating disorder in a community sample. *Behaviour Research and Therapy*, *66*, 72–76. https://doi.org/10.1016/j.brat.2015.01.004

Grilo, C. M., Ivezaj, V., & White, M. A. (2015c). Evaluation of the *DSM–5* severity indicator for bulimia nervosa. *Behaviour Research and Therapy*, *67*, 41–44. https://doi.org/10.1016/j.brat.2015.02.002

Gutiérrez-Zotes, J. A., Bayón, C., Montserrat, C., Valero, J., Labad, A., Cloninger, C. R., & Fernández-Aranda, F. (2004). Temperament and Character Inventory-Revised (TCI-R). Standardization and normative data in a general population sample. *Actas Españolas de Psiquiatría*, *32*(1), 8–15.

Hebebrand, J., Himmelmann, G. W., Heseker, H., Schäfer, H., & Remschmidt, H. (1996). Use of percentiles for the body mass index in anorexia nervosa: Diagnostic, epidemiological, and therapeutic considerations. *International Journal of Eating Disorders*, *19*(4), 359–369. https://doi.org/10.1002/(SICI)1098-108X(199605)19:4<359::AID-EAT4>3.0.CO;2-K

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hirst, R. B., Beard, C. L., Colby, K. A., Quittner, Z., Mills, B. M., & Lavender, J. M. (2017). Anorexia nervosa and bulimia nervosa: A meta-analysis of executive functioning. *Neuroscience & Biobehavioral Reviews*, *83*, 678–690. https://doi.org/10.1016/j.neubiorev.2017.08.011

Hoiles, K. J., Egan, S. J., & Kane, R. T. (2012). The validity of the transdiagnostic cognitive behavioural model of eating disorders in predicting dietary restraint. *Eating Behaviors*, *13*(2), 123–126. https://doi.org/10.1016/j.eatbeh.2011.11.007

Hrabosky, J. I., Masheb, R. M., White, M. A., & Grilo, C. M. (2007). Overvaluation of shape and weight in binge eating disorder. *Journal of Consulting and Clinical Psychology*, *75*(1), 175–180. https://doi.org/10.1037/0022-006X.75.1.175

Jenkins, P. E., Luck, A., Cardy, J., & Staniford, J. (2016). How useful is the *DSM–5* severity indicator in bulimia nervosa? A clinical study including a measure of impairment. *Psychiatry Research*, *246*, 366–369. https://doi.org/10.1016/j.psychres.2016.10.011

Kaufman, J., Birmaher, B., Brent, D., Rao, U. M. A., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(7), 980–988. https://doi.org/10.1097/00004583-199707000-00021

Krug, I., Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the *DSM–5* severity indices for eating disorders. *European Eating Disorders Review*, *29*(3), 482–498. https://doi.org/10.1002/erv.2776

Krug, I., Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2022). *How to assess severity in male with eating disorders? The DSM–5 severity index versus severity based on drive for thinness*. Manuscript in preparation.

Kuehnel, R. H., & Wadden, T. A. (1994). Binge eating disorder, weight cycling, and psychopathology. *International Journal of Eating Disorders*, *15*(4), 321–329. https://doi.org/10.1002/eat.2260150403

Lydecker, J. A., Ivezaj, V., & Grilo, C. M. (2020). Testing the validity and clinical utility of the severity specifiers for binge-eating disorder for predicting treatment outcomes. *Journal of Consulting and Clinical Psychology*, *88*(2), 172–178. https://doi.org/10.1037/ccp0000464

Machado, P. P., Grilo, C. M., & Crosby, R. D. (2017). Evaluation of the DSM–5 severity indicator for anorexia nervosa. *European Eating Disorders Review*, *25*(3), 221–223. https://doi.org/10.1002/erv.2508

Murray, S. B., Nagata, J. M., Griffiths, S., Calzo, J. P., Brown, T. A., Mitchison, D., Blashill, A. J., & Mond, J. M. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, *57*, 1–11. https://doi.org/10.1016/j.cpr.2017.08.001

Nakai, Y., Nin, K., Noma, S., Teramukai, S., Fujikawa, K., & Wonderlich, S. A. (2017). The impact of *DSM–5* on the diagnosis and severity indicator of eating disorders in a treatment-seeking sample. *International Journal of Eating Disorders*, *50*(11), 1247–1254. https://doi.org/10.1002/eat.22777

Peñas-Lledó, E., Bulik, C. M., Lichtenstein, P., Larsson, H., & Baker, J. H. (2015). Risk for self-reported anorexia or bulimia nervosa based on drive for thinness and negative affect clusters/dimensions during adolescence: A three-year prospective study of the TChAD cohort. *International Journal of Eating Disorders*, *48*(6), 692–699. https://doi.org/10.1002/eat.22431

Peñas-Lledó, E., Vaz Leal, F. J., & Waller, G. (2002). Excessive exercise in anorexia nervosa and bulimia nervosa: Relation to eating characteristics and general psychopathology. *International Journal of Eating Disorders*, *31*(4), 370–375. https://doi.org/10.1002/eat.10042

Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., Bardone-Cone, A. M., Klein, M., & le Grange, D. (2007). Psychometric properties of the eating disorder examination-questionnaire: Factor structure and internal consistency. *International Journal of Eating Disorders*, *40*(4), 386–389. https://doi.org/10.1002/eat.20373

Racine, S. E., & Wildes, J. E. (2013). Emotion dysregulation and symptoms of anorexia nervosa: The unique roles of lack of emotional awareness and impulse control difficulties when upset. *International Journal of Eating Disorders*, *46*(7), 713–720. https://doi.org/10.1002/eat.22145

Rajagopalan, A., Bhattacharya, A., Peebles, R., Dakanalis, A., & Timko, C. A. (2019, October). *A literature review: Do DSM–5 severity indicators for eating disorders apply to children and adolescents?* [Paper presentation]. American Academy of Children and Adolescent Psychiatry 66th Annual Meeting, Chicago, Illinois, United States.

Reas, D. L., & Rø, Ø. (2017). Investigating the *DSM–5* severity specifiers based on thinness for adults with anorexia nervosa. *International Journal of Eating Disorders*, *50*(8), 990–994. https://doi.org/10.1002/eat.22729

Rozenblat, V., Ong, D., Fuller-Tyszkiewicz, M., Akkermann, K., Collier, D., Engels, R. C. M. E., Fernandez-Aranda, F., Harro, J., Homberg, J. R., Karwautz, A., Kiive, E., Klump, K. L., Larson, C. L., Racine, S. E., Richardson, J., Steiger, H., Stoltenberg, S. F., van Strien, T., Wagner, G., . . . Krug, I. (2017). A systematic review and secondary data analysis of the interactions between the serotonin transporter 5-HTTLPR polymorphism and environmental and psychological factors in eating disorders. *Journal of Psychiatric Research*, *84*, 62–72. https://doi.org/10.1016/j.jpsychires.2016.09.023

Shroff, H., Reba, L., Thornton, L. M., Tozzi, F., Klump, K. L., Berrettini, W. H., Brandt, H., Crawford, S., Crow, S., Fichter, M. M., Goldman, D., Halmi, K. A., Johnson, C., Kaplan, A. S., Keel, P., LaVia, M., Mitchell, J., Rotondo, A., Strober, M., . . . Bulik, C. M. (2006). Features associated with excessive exercise in women with eating disorders. *International Journal of Eating Disorders*, *39*(6), 454–461. https://doi.org/10.1002/eat.20247

Smink, F. R., van Hoeken, D., Oldehinkel, A. J., & Hoek, H. W. (2014). Prevalence and severity of *DSM–5* eating disorders in a community cohort of adolescents. *International Journal of Eating Disorders*, *47*(6), 610–619. https://doi.org/10.1002/eat.22316

Smith, K. E., Ellison, J. M., Crosby, R. D., Engel, S. G., Mitchell, J. E., Crow, S. J., Peterson, C. B., Le Grange, D., & Wonderlich, S. A. (2017). The validity of *DSM–5* severity specifiers for anorexia nervosa, bulimia

nervosa, and binge-eating disorder. *International Journal of Eating Disorders*, *50*(9), 1109–1113. https://doi.org/10.1002/eat.22739

Soh, N. L., Touyz, S. W., & Surgenor, L. J. (2006). Eating and body image disturbances across cultures: A review. *European Eating Disorders Review*, *14*(1), 54–65. https://doi.org/10.1002/erv.678

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wolfi, G., & Treasure, I. (2011). Medical and psychological consequences of eating disorders. In J. Morris (Ed.), *ABC of eating disorders* (2nd ed., pp. 20–21). John Wiley & Sons.

Zayas, L. V., Wang, S. B., Coniglio, K., Becker, K., Murray, H. B., Klosterman, E., Kay, B., Bean, P., Weltzin, T., Franko, D. L., Eddy, K. T., & Thomas, J. J. (2018). Gender differences in eating disorder psychopathology across *DSM–5* severity categories of anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, *51*(9), 1098–1102. https://doi.org/10.1002/eat.22941

Received July 19, 2021
Revision received January 7, 2022
Accepted January 24, 2022 ■

Chapter 3: Summary and Rationale for the Current Thesis

3.1 Chapter Introduction

EDs significantly affect the biopsychosocial well-being of affected individuals and place a heavy burden on healthcare systems. Chapter 2 (Study 1) critically examined DSM-5 severity ratings and the OWS in BED. The chapter found limited support for DSM-5 severity ratings in AN and BED and only moderate support for BN. Specifically, individuals classified in the mild and moderate groups for AN and BED showed little difference in ED psychopathology compared to those in the severe or extreme categories. While many studies included in the systematic review and meta-analysis (Chapter 2, Study 1) have used valid and reliable measures to assess ED and general psychopathology, issues like poor control of confounding variables (e.g. duration of illness), low statistical power, and failure to report exact *p*-values undermined the strength of the findings.

Several gaps and limitations were identified from Chapter 2 (Study 1), which shaped the direction of the following chapters in this PhD thesis. First, Chapter 2 revealed that, despite the primary goal of severity ratings being to aid clinical practice, our systematic review and meta-analysis (Study 1) did not identify any studies investigating whether mental health professionals, such as psychiatrists and psychologists, incorporate these ratings into their routine practice. As the result, we know little about clinicians' perspectives on the usefulness of these ratings for informing treatment choices, such as the type and frequency of interventions. It is therefore crucial to examine clinicians' views on these ratings and their actual application in clinical settings.

Second, while Chapter 2 found limited support for the DSM-5 severity ratings for AN and BED in terms of ED psychopathology, the utility of these ratings remains uncertain. This is largely because most research has primarily focused on the relationship between severity

groups and psychological factors alone – as indicated in Chapter 2 (Study 1). While these factors are important, they oversimplify the complexity of EDs and neglect the unique challenges individuals face. Moreover, an effective severity rating should encompass not only ED psychopathology or prognosis but also the factors underlying the onset and maintenance of the disorder (APA, 2013; Zimmerman et al., 2018). Such information helps clinicians proactively identify individuals at high risk for ED and supports essential decisions around hospitalisation and treatment intensity (Zimmerman et al., 2018). Therefore, the lack of research examining ED severity ratings alongside a broader set of biopsychological factors – such as quality of life, psychiatric comorbidity, and risk factors – raises important questions about their overall utility.

Third, the primary aim of severity ratings is to guide treatment outcomes, under the assumption that individuals in milder severity categories will have better chances of full remission (Lucas et al., 2022; Zimmerman et al., 2018). However, at the time Study 1 was conducted, our systematic search identified only one study (Dalle Grave et al., 2018) that had explored the relationship between DSM-5 ED severity ratings and treatment outcomes. Dalle Grave et al. examined the relationship between AN DSM-5 severity rating and treatment outcome and found no significant differences in weight recovery or positive outcomes between the DSM–5 AN severity groups across baseline, 6-month, and 12-month follow-ups. Despite the strength of being a longitudinal study, the small sample size ($N = 128$) may have limited its ability to detect meaningful differences among the four DSM-5 severity groups. Therefore, further research with a larger sample is needed.

Fourth, alternative severity ratings such as OWS (Gianini et al., 2017; Grilo et al., 2015a), the number of purging methods (Gianini et al., 2017), and weight suppression (Khalil et al., 2024) have been proposed as potentially superior to DSM-5 severity ratings for EDs. However, research on these alternative severity ratings for EDs remains limited. Chapter 2

(Study 1) identified only one study examining the number of purging behaviours as a severity rating for BN (Gianini et al., 2017), three studies evaluating OWS as an alternative severity rating for BED (Grilo et al., 2015a; 2015b; 2015c) and one study each investigating OWS (Gianini et al., 2017) and DT (Krug et al., 2021) as transdiagnostic severity ratings for EDs. Despite the potential of these alternative measures, our systematic review and meta-analysis did not identify any studies directly comparing the clinical utility of DSM-5 ratings with weight suppression. Furthermore, direct comparisons between DSM-5, ICD-11, and alternative indicators such as weight suppression, OWS, and DT – grounded in established models like the transdiagnostic model of EDs (Fairburn, 2003) – remain sparse. This gap highlights the need for further research to assess the effectiveness and applicability of these severity ratings across diverse populations. To address this, the current thesis aimed to evaluate the clinical utility of both established DSM-5 and ICD-11 severity ratings and alternative measures, including OWS, DT, and weight suppression, to determine their relevance and applicability in capturing ED severity.

Fifth, out of the 22 studies included in our systematic review and meta-analysis (Study 1, Chapter 2), only two (Grilo et al., 2015b; Smink et al., 2014) used community-based samples. However, these studies focused exclusively on either DSM-5 BED severity ratings or the prevalence of DSM-5 severity groups in the community. Consequently, there is limited understanding of how DSM-5, ICD-11, and alternative severity ratings function within community-based samples.

Community-based research is crucial for enhancing the generalisability of findings. Clinical samples typically reflect individuals already in treatment, which may not represent the full spectrum of those affected by EDs. In contrast, community-based studies capture a broader population, including individuals not engaged in clinical care, ensuring findings are more widely applicable (Hart et al., 2011). One challenge in using DSM-5 and ICD-11

severity ratings within community samples, however, is the heavily negatively skewed distribution, with few or no participants in the severe category (Grilo et al., 2015b; Smink et al., 2014). This highlights the importance of employing severity indicators on a continuous scale, which better captures the full range of severity and provides a more nuanced understanding of these ratings in community settings. Therefore, this thesis aimed to address this gap by comparing DSM-5 and alternative severity ratings as continuous variables within a community-based sample, ultimately enhancing the generalisability of the findings.

Ultimately, Chapter 2 highlighted critical limitations and gaps in the literature, informing the design of the thesis's original investigations. These findings stress the need for more rigorous research to enhance understanding of ED severity and improve clinical practice. Critically, the systematic review and meta-analysis identified three key areas where further investigations are warranted to improve our understanding of ED severity ratings.

3.2 Overall Objectives of the Current Thesis

Overall, this thesis had three overarching aims:

1. To explore mental health clinicians' perspectives on the usefulness of ED severity ratings in clinical practice
2. To examine and directly compare the relationship between severity ratings on a wide range of clinical correlates, including risk factors, psychological, biological, cognitive factors and treatment outcomes using a treatment-seeking sample.
3. To evaluate DSM-5 and alternative severity ratings for EDs on a continuous scale within a community-based setting, aiming to enhance understanding of these ratings across the broader spectrum of disordered eating and EDs.

To achieve these aims, in addition to Study 1 (Chapter 2), five additional studies were conducted. **Study 2** used a mixed-method design to evaluate whether clinicians involved in

the care of individuals with EDs apply severity ratings in practice and their views on these ratings.

Study 3a and Study 3b examined the relationship between DSM-5, ICD-11 and alternative OWS severity ratings and various clinical correlates in a treatment-seeking sample of individuals with AN. Specifically, Study 3a compared the ability of DSM-5 and OWS AN severity ratings to index psychological factors (e.g., quality of life, ED symptoms), biological factors (e.g., potassium, magnesium), and cognitive flexibility in individuals with AN. Study 3b, in contrast, explored the relationship between DSM-5, ICD-11, and DT severity ratings and psychiatric comorbidities.

Study 4 applied medical research principles to assess whether DSM-5 and alternative severity rating, based on DT, were associated with established risk factors for AN. This approach stems from the argument that valid severity ratings should not only capture ED psychopathology but also identify factors related to the onset of AN.

Since the goal of severity classifications is clinical utility, treatment outcomes must be a key consideration. **Study 5** was the first to directly compare three severity classification systems, which were DSM-5, ICD-11, and DT in their ability to predict treatment outcomes in treatment-seeking participants with AN and BN.

Lastly, as the earlier studies focused on clinical samples**, Study 6** employed a community-based sample to validate both BMI and alternative severity indicator. Specifically, using ecological-momentary assessment (EMA) during the 7-day period, it examined BMI and weight suppression as continuous variables to determine their ability to predict stated-based body dissatisfaction and disordered eating urges in a large community-based sample ($N = 686$). This approach was a necessary step back as ad not been empirically validated these severity indicators had not been empirically validated prior to their integration into the current severity frameworks.

Chapter 4: Clinical Utility of DSM-5 Severity Ratings for Eating Disorders: Perspectives and Practices of Mental Health Clinicians

4.1 Chapter Introduction

Study 1 (Chapter 2) revealed that most research has concentrated on the relationship between DSM-5 severity ratings and clinical correlates (e.g., ED psychopathology; Dang et al., 2023). While these ratings are designed to guide clinical practice, little is known about whether mental health professionals, such as psychiatrists and psychologists, use them in their daily practice. Moreover, there is limited understanding of clinicians' perspectives on the utility of these ratings in shaping treatment decisions, including determining the type or frequency of care.

To address these gaps, a qualitative methodology is particularly suited for exploring the views of mental health clinicians on ED severity ratings. Qualitative approaches provide an understanding of how clinicians interpret and apply ED severity ratings within different clinical contexts (Yim et al., 2019). Specifically, qualitative approach allows us to investigate how these ratings are used in diagnosis, treatment planning, and communications, offering a deeper insight into their perceived clinical relevance. Another significant advantage of qualitative research is its ability to gather clinicians' recommendations for alternative severity rating approaches (Yim et al., 2019). Through interviews, clinicians can offer specific suggestions for refining the DSM framework or integrating alternative severity indicators, such as weight suppression.

While qualitative methods offer valuable insights into clinicians' experiences and perspectives, quantitative methods are also crucial for measuring the extent to which severity ratings are used in clinical practice (Aouad et al., 2022; Pettersen et al., 2018). Therefore, **Chapter 4 (Study 2)** employed a mixed-methods design, combining both qualitative and quantitative approaches. This allowed the study to explore not only how Australian-based

healthcare professionals (i.e., general practitioners, psychiatrists, psychologist, dietitians) apply severity ratings in practice, their professional views on the clinical relevance of these ratings, and their recommendations for alternative severity indicators, but also how frequently these ratings are utilised in clinical settings.

4.2 Study 2. Exploring Clinician Perspectives on the DSM-5 Eating Disorder Severity Ratings: A Qualitative Study

This chapter constitutes a manuscript currently under review submitted for publication at the *Journal of Treatment and Prevention.*

Response to reviewers' comments had been submitted on the 26th of January 2025

Dang, A., Kirk, H., Kiropoulos, L., & Krug, I. Exploring Clinician Perspectives on the DSM-5 Eating Disorder Severity Ratings: A Qualitative Study

Exploring Clinician Perspectives on the DSM-5 Eating Disorder Severity Ratings: A Qualitative Study

An Dang[a], Hayley Krik[a], Litza Kiropoulos[a] & Isabel Krug[a]

[a] Melbourne School of Psychological Sciences. University of Melbourne, Melbourne, Australia

Abstract

The DSM-5 introduced severity ratings in 2013 for anorexia nervosa (AN), bulimia nervosa (BN), and binge-eating disorder (BED) based on BMI, compensatory behaviour frequency, and binge eating frequency, respectively. While several studies have assessed the validity of these ratings, little is known about their use in clinical practice. This study examined clinicians' use of DSM-5 severity ratings and their views on their clinical value. A sample of 38 clinicians (mean age = 38.5, 97.37% female) completed an online survey on the usage of DSM-5 severity ratings in clinical practice, and six participated in one-on-one interviews for further insights. Results showed that 60% of clinicians do not use these ratings. Qualitative findings obtained from both survey and one-on-one interviews found that clinicians perceive these ratings as invalidating, pathology-reinforcing, and lacking clinical relevance. However, some noted that these ratings can aid in triage and communication among clinicians in tertiary settings. This study highlights substantial concerns about the clinical utility of DSM-5 severity ratings for EDs, suggesting the need for a more comprehensive approach that considers broader psychological, medical, and cognitive indicators.

Keywords: eating disorders, severity ratings, clinician perspectives

Clinical Implications

- Approximately 60% of mental health clinicians do not use the DSM-5 severity ratings for Anorexia Nervosa, Bulimia Nervosa and Binge Eating Disorder in clinical practice.
- Clinicians perceive these ratings as invalidating, pathology-reinforcing, and lacking clinical relevance.
- In some cases, these ratings aid in triage and communication among clinicians in tertiary settings.

In 2013, the American Psychiatric Association (APA) introduced severity ratings for anorexia nervosa (AN), bulimia nervosa (BN), and binge-eating disorder (BED) in the fifth edition of the Diagnostic and Statistical Manual (DSM-5; APA, 2013). These ratings classify individuals with an eating disorder into "mild," "moderate," "severe," or "extreme" categories. For AN, severity is determined by body mass index (BMI), with thresholds defined as mild (≥17.0 kg/m²), moderate (16–16.99 kg/m²), severe (15–15.99 kg/m²), and extreme (<15 kg/m²). For BN, severity is based on the frequency of inappropriate compensatory behaviours, while for BED, it is based on the frequency of binge-eating episodes. The severity categories for both BN and BED are mild (1–3 episodes per week), moderate (4–7 episodes per week), severe (8–13 episodes per week), and extreme (14 or more episodes per week).A universal severity classification system for psychiatric disorders is crucial for guiding research, improving diagnoses, and enhancing communication among healthcare professionals (Zimmerman et al., 2018). It also aids in developing targeted treatments. The specific purpose of the three DSM-5 severity ratings for AN, BN and BED remains unclear, as the DSM-5 does not define what these ratings are meant to reflect (e.g., prognosis, level of symptomatology; Gianini et al., 2017; Reas et al., 2017). Prior to their introduction, there was no research explaining how the frequency thresholds or BMI cut-offs were determined to distinguish between severity levels (Dang et al., 2021). Additionally, the clinical utility of these ratings has been inconsistent, with studies showing that the thresholds do not reliably predict eating disorder or general psychopathology (Gianini et al., 2017; Nakai et al., 2017) or treatment outcomes (Dang et al., 2024; Dalle Grave et al., 2018). This lack of clarity and empirical support calls into question the validity these severity ratings.

For eating disorders (EDs), severity can be linked to a range of physical, cognitive, and psychological symptoms, their intensity, and frequency. It can also reflect psychiatric comorbidities, the disorder's impact on psychosocial functioning, and the patient's quality of

life (Spindler & Milos, 2007). Additionally, the concept of severity should consider expected treatment outcomes, as the goal is clinical utility. Patients with higher severity are expected to show fewer positive outcomes compared to less severe cases and should have a better response to the gold standard treatment (Zimmerman et al., 2018).

Several empirical studies, including a systematic review (*n* =22) and meta-analysis (*n* = 19), attempted to assess the clinical and predictive validity of all three DSM-5 severity ratings (Dang et al 2021). The review found that the DSM-5 severity ratings for BN did adequately reflect ED psychopathology, but this was not the case for AN and BED. Further research by Dang et al. (2022) compared the DSM-5 and alternative overvaluation of weight/shape (OWS) severity ratings for AN, finding that OWS more effectively indexed psychological correlates (e.g., dietary restraint, depressive symptoms), though neither system adequately reflected biological correlates (e.g., potassium, magnesium). Additionally, Mora-Maltas et al. (2023) compared DSM-5 AN and alternative AN severity ratings based on drive for thinness (DT), and duration of illness in indexing cognitive inflexibility, concluding that duration of illness was superior to both DSM-5 and DT severity ratings. In terms of DSM-5 severity rating and treatment response, studies have mostly focused on DSM-5 severity for AN and found no significant differences among adolescents, young adults, and adults with AN in-treatment response (e.g., Dalle Grave et al., 2018; Smink et al., 2014).

Despite an extensive review of the current literature, it appears that most studies have focused solely on examining the correlation between ED severity ratings and various psychological and biological variables (e.g., Dang et al., 2023) . Although the primary purpose of these severity ratings is to assist in clinical practice, there is limited understanding of whether mental health clinicians (e.g., psychiatrists, and psychologists) utilise these ratings in their practice. Additionally, little is known about clinicians' perceptions regarding the usefulness of these ratings in guiding treatment decisions, such as determining the types or

frequency of treatment. Gaining insight into whether clinicians are using these severity ratings in clinical practice and their professional opinions about these severity ratings can provide insight into the DSM-5 severity ratings clinical applicability and utility, and provide recommendations for the further development of the severity ratings in upcoming editions of the DSM.

The current study

The current study had two aims to address the gaps in the literature, which were to: 1) obtain qualitative data (via an online survey) on whether clinicians are using the DSM-5 proposed severity ratings for EDs in their clinical practice, and 2) investigate clinicians' attitudes and opinions about the clinical utility of the DSM-5 severity ratings for EDs via open-ended/free-text survey questions for all participants and one-to-one semi-structured interviews on a subsample. One-to-one semi-structured interviews provide rich data by encouraging open-ended reflection, opinions, and experiences from participants, while also providing enough structure to achieve optimal use of interview time, as well as systematic and comprehensive responses that are focused on relevant topic matter (Jamshed, 2014).

Method

Participants

This study was embedded in a larger project which looked at accessing to eating disorder treatment in Australia. Participants were recruited via several sources, including sending emails with the study information sheet to different Australian psychology, dietetic, and medical clinics. We also recruited through social media posts (Facebook and Twitter), and the websites of Australian ED organisations. To meet eligibility, participants needed be a healthcare professional (e.g., psychologists, psychiatrists, general practitioners, and dieticians) who at the time of assessment were involved in the provision of medical and psychological support for individuals with EDs in the community. Participants also needed to

have sufficient English skills and cognitive ability to complete an online self-report survey and a 20–40-minute one-to-one interview over Zoom.

Thirty-eight clinicians (mean age = 38.5, 97.37% female, 90% Caucasian) completed the online Qualtrics survey. Six clinicians volunteered for additional interviews conducted post-survey, with similar demographic characteristics. Literature on sample size justifications in qualitative health research, particularly one-to-one interviews, is limited and not dependent on the number of interviews (Vasileiou et al., 2018). Given the time constraints of clinicians, the sample size for both quantitative and qualitative data reflects these access limitations. The study adhered to the guidelines of the Declaration of Helsinki and received approval from a university in Melbourne, Australia (blinded for review).

Material

An online survey was delivered via Qualtrics, comprising a custom compilation of questions related to clinician demographics, usage and opinions, via free-text entry, regarding the DSM-5 severity ratings for EDs (e.g., *"Do you apply these severity ratings in your daily clinical practice?"*)

Interviews were semi-structured to allow for open-ended responses on the intended subject matter. Pre-determined open-ended questions aimed to capture participants' reflections on their general experiences working with individuals with EDs. The development of these questions was informed by the literature (DiCicco-Bloom & Crabtree, 2006) and refined in consultation with the research team. These questions also allowed participants the freedom to provide reflections and elaborate on their individual experiences with the DSM-5 severity ratings for EDs in practice.

Procedure

Interested participants accessed the survey via a hyperlink or QR code in the advertisement email. Informed consent and plain language statements about the study were

provided. The survey also asked participants if they would like to be contacted for an optional follow-up interview 1-3 weeks after completing the survey. After providing this information, participants were prompted to complete the online survey. Those who agreed to a follow-up interview received an email 1-3 weeks later to schedule a time slot for the interview with one of the authors. Participants who completed the follow-up interview received a $30 gift voucher as reimbursement.

No participants who completed the survey were excluded from the sample ($N = 38$). Out of all respondents, 17 initially expressed interest in a face-to-face online interview. Three participants became unavailable due to conflicting commitments, and another six were unresponsive after initial contact. Ultimately, eight participants were scheduled for interviews, but two dropped out last minute. This left a total of six participants for the qualitative interview data.

One-on-one interviews were conducted and recorded over Zoom with the author (HK) who identifies as a White female in her mid-twenties, has postgraduate clinical psychology training, and had no pre-existing relationship to participants. To ensure confidentiality, all identifiable information was removed from the interview transcripts, including names, other individuals, and locations. Participants were assigned representative numbers for data reporting. The interviews were conducted between January and June 2023 and ranged in duration from 29 to 54 min ($M = 41$ min). Upon completion, participants were provided with a debriefing statement. Each participant received a $30 eGift Card for participating

Analysis

The qualitative data from one-on-one interviews and free-text responses were analysed using a six-step reflexive thematic analysis, which identifies patterns of meaning in qualitative data (Braun & Clarke, 2006, 2019). Transcripts were imported into NVivo 12 for organisation and analysis, systematically distilling the data into key themes. The process

involved closely reading responses, extracting core meanings, identifying recurring themes, and assigning descriptive labels and definitions. This approach was chosen to let the findings emerge naturally, without the constraints of structured, deductive methods.

Results

Quantitative Findings

Table 1 details the demographics of the 38 clinicians who participated in the online survey. The majority were female (97.37%) with an average of 10 years of experience (mean = 10.49, SD = 7.8), primarily working with ED clients in private practice (76.3%).

Table 1.

Summary of Participant Demographics

	M	*SD*	Range
Age (years)	38.5	9.23	28-64
Experience (years)	10.49	7.84	.5-30

	N	%
Sex		
Female	37	97.37%
Non-Binary (incl. transgender women)	1	2.63%
Current Occupation		
Clinical Psychologist	14	36.8%
Psychologist	8	21.1%
Dietician	8	21%
GP	2	5.3%
Psychiatrist	1	2.6%
Current employment status		
Full-time work	22	57.9%
Part-time work	16	42.1%
Workplace Setting		
Private Practice	29	76.3%
Public	5	13.2%
Academia	2	5.3%
Other	2	5.3%
Exposure to EDPs		
Yes	35	92%
No	3	8%
Ethnic Background		
Caucasian	34	89.5%
Southern Asian	2	5.3%
Hispanic/Latin	1	2.6%
Missing	1	2.6%

Most commonly used ED therapy modality in current workplace setting		
CBT-E	25	65.8%
Psychodynamic	3	7.9%
DBT	2	5.3%
Other	8	21.1%

Note.
EDP = Eating Disorder Plan*
*Eating Disorder Plan is a government-subsidised program in Australia that provides up to 20 sessions with dietitians and 10 sessions with psychologists for assessment and treatment of eating disorders.
CBT-E = Enhanced cognitive behaviour therapy
DBT = Dialectical behaviour therapy

Qualitative findings showed that while only 3% of clinicians were unaware of the DSM-5 severity ratings, a significant majority (60%) did not use these ratings in their clinical practice.

Thematic Analysis

Table 2 presents the demographics of the final six participants in the one-on-one interviews. The group included clinical psychologists (*n*=4), a dietician (*n*=1), and a psychiatrist (*n*=1). Most were female (83%), aged 29-49, and working in private practice. Their experience with EDs ranged from 3.5 to 22 years, with most participants having a client base primarily consisting of ED patients.

Table 2

Demographic Characteristics of Interview Participants

Participant	Sex	Age	Occupation	No. of years experience	Job Status	Current % of ED patients	Primary Work Setting
P1	F	30	Clin. Psyc.	3.5	FT	95	PP
P2	NB	38	Dietician	8	PT	91	PP
P3	F	48	Clin. Psyc.	22	PT	100	PP
P4	F	41	Clin. Psyc.	10	FT	90	PP

P5	F	29	Clin. Psyc.	4	PT	92	PP
P6	F	49	Psychiatrist	20	PT	71	PP

Note.
ED = eating disorders
NB = non-binary
Clin.Psyc = clinical psychologist
FT = fulltime
PT = part-time
PP = private practice

Thematic analysis of both free-text and one-on-one interview responses identified four themes (1) Invalidating, (2) pathology-reinforcing, (3) lacking clinical relevance, and (4) Communication of ED severity to others

Theme 1: Invalidating

Overall, a major theme that emerged from our analysis is that clinicians perceive the DSM-5 severity ratings as invalidating patients' experiences rather than serving as a useful clinical indicator. EDs are serious conditions with significant impacts on both mental andphysical health. Therefore, classifying someone with an ED as having a “mild” disorder fails to convey the illness's seriousness and lacks clinical value.

> *“Once you get diagnosed, you're already quite unwell. So I think it's really invalidating and misleading using terms like mild. Because you know if you're 15% under your body weight you are really medically unstable you are not mildly unwell. You know, you don't just have a head cold, and I also don't think it lends anything clinically.”*

Additionally, reliance on weight-based severity indicators can further invalidate the struggles of individuals with atypical AN, despite the DSM-5 not assigning a severity level to this condition. Atypical AN is characterised by meeting all diagnostic criteria for AN except for maintaining a weight within the normal range (APA, 2013). Under the DSM-5, a BMI > 17 kg/m^2 is classified as mild severity AN, while individuals with atypical AN typically have

a BMI > 18.5 kg/m^2. This distinction reinforces the misconception that atypical AN is less severe than AN. However, many individuals with atypical AN experience significant weight loss through restrictive dietary and/or compensatory behaviours, even though they remain within the normal BMI range due to a previously higher body weight. Focusing solely on weight as a severity indicator overlooks the substantial impact on their quality of life and the severe disordered eating behaviours that those with atypical AN may still be experiencing.

"I think it's unhelpful. I think as well, it really takes away from the impact of atypical anorexia, which I also don't agree with that label."

Theme 2: Pathology-reinforcing

Since individuals with EDs often place excessive importance on weight and shape, another theme that emerged was that classifying them by weight, or the frequency of disordered eating behaviours may unintentionally reinforce this overvaluation of weight and shape.

"Obviously eating disorder clients over-evaluate and put a lot of emphasis on weight and if we kind of do the same it's reinforcing that."

Furthermore, labelling someone as "mild" can lead clients to internalise unhelpful beliefs that they are "not sick enough" to deserve treatment, potentially downplaying the severity of their disorder and hindering their recovery.

"You just have like a 'mild to moderate' eating disorder instead of 'severe'. And because of that, because your eating disorder isn't like bad enough you don't have increased access to treatment. And so, unfortunately I've seen some clients internalize that as "I'm not sick enough" or

"I don't deserve treatment at all" or it feeds into the narrative that maybe my eating disorder isn't a problem here. I'm not sick enough that you know everything's fine. Yeah, so. Okay, I'm not a fan."

Theme 3: Lacking clinical relevance

Overall, most clinicians reported that they do not rely on the DSM-5 severity ratings in clinical practice, believing that a single indicator cannot capture the complexity of EDs.

" I prefer to assess the whole person presentation and board risk factors, not just these categories"

The majority of clinicians also expressed scepticism about the clinical utility of severity ratings for EDs, noting that these ratings are not particularly helpful or reliable, therefore raising concerns about their predictive validity and practical application in clinical settings.

"Hmm. Not very good. Not very useful. So I don't really use them, because when I do the severity doesn't really... often they are sort of frequency based. So, numbers of days and you can see a switch. There is some indication, but I don't think it's a big or it has a poor predictive validity"

They argued that severity labels (e.g., "moderate" or "mild") do not significantly impact clinical practice, as the same treatment approach is used regardless of the label. They also pointed out that the severity of symptoms or medical conditions do not always correspond to BMI.

"I don't think this [severity rating] should be inclusive in the diagnosis, and severity of symptoms or medical state is not always congruent with BMI

Theme 4: Communication of ED severity to others

Clinicians also shared their views on the usage of the ED severity ratings for communicating with clients, families, and other health professionals. Of the 14 clinicians who reported using the severity ratings, six found them helpful for psychoeducation with

clients, triage in referrals to tertiary services, and communication within teams, especially in cases involving severe or refractory patient.

"Sometimes useful to help express severity of ED to clients and provide psychoeducation, rarely use otherwise"

"Familiarity with the severity ratings as per the DSM-5 can help with the triage process when referring to Tertiary ED services and communicating with other team members"

On the other hand, the remaining clinicians shared that they do not think these severity ratings aid communication between practitioners and other related parties, mostly clients:

"There's usually very little reason to communicate that because it won't change the number of sessions or anything like that."

"And if I did, I would never communicate them"

Discussion

Using a mixed design method (qualitative and quantitative), this study aimed to investigate whether mental health clinicians are using the DSM-5 proposed severity ratings for EDs, introduced in 2013, in clinical practice. Additionally, we collected clinicians' opinions on the clinical utility of these severity ratings through open-ended survey questions for all participants and one-on-one semi-structured interviews with a subset of six participants.

The descriptive statistics revealed that 60% of clinicians do not use DSM-5 severity ratings in practice, primarily because they find these ratings inadequate in capturing the complexity of EDs. This aligns with empirical research, including cross-sectional studies (e.g., Dang et al., 2023; Giannini et al., 2017; Krug et al., 2021) and meta-analyses (Dang et al., 2022), which show that the DSM-5 severity indicator fails to reflect ED psychopathology, biological factors, and psychiatric comorbidities that individuals with ED experience.

Many argue that labelling a condition as "mild" can be misleading and invalidating. Studies by Dang et al. (2022; 2023) support this view, showing that individuals in the mild DSM-5 severity group often exhibit higher levels of ED psychopathology and psychiatric comorbidities than those in the severe group. Moreover, basing severity ratings on weight for AN, or the frequency of disordered behaviours for BN and BED, may reinforce patients' preoccupation with weight and shape, perpetuating their psychopathology.

Clinicians worry that such classifications may lead patients to internalise harmful beliefs that they are "not sick enough" for treatment, making them feel unworthy of care or downplaying the seriousness of their condition. This is particularly concerning given that individuals with EDs often lack insight into the impact of disordered eating on their psychological and physical health (Fatt et al., 2022). Overall, many clinicians feel that DSM-5 ED severity ratings lack clinical value and may even worsen ED symptomatology.

While some clinicians use severity ratings to improve communication and aid in triaging or referring to specialised ED services, most find them less useful in direct client discussions. In practice, severity ratings are primarily used for administrative purposes, such as referrals and treatment planning, rather than for clinical decision-making.

Implications

Our findings indicate that most clinicians do not use these ratings, citing a lack of clinical relevance and noting that they may even be invalidating or harmful to individuals with EDs. Given the complex biopsychosocial nature of EDs and their diverse symptom profiles, relying on a single measure– such as BMI, frequency of compensatory behaviours, binge episodes, or alternatives like weight/shape overvaluation– to assess severity is insufficient. Future research might be beneficial to explore dimensional severity models that integrate a broader range of clinical indicators, including weight suppression (difference between highest and current weight; Lowe et al., 2006), psychological distress, functional

impairment, and social and medical impacts, alongside DSM-5 criteria. Such models could provide a more comprehensive and nuanced understanding of ED severity.

Additionally, while no single indicator can fully capture ED severity, a checklist or brief severity screen that encompasses multiple aspects can significantly enhance screening in high-pressure, underfunded settings. This tool would streamline the assessment process, enabling clinicians to swiftly identify key issues and prioritize care despite limited resources. It should include a range of indicators– psychological, medical, and social, including social support– to provide a comprehensive view and facilitate targeted interventions.

Limitations and future directions

This study has several limitations that should be acknowledged. First, the small sample size (*n*=38) restricts the ability to draw definitive conclusions about DSM-5 ED severity ratings. Second, the sample predominantly comprised Caucasians (97.37%) and female clinicians (97.37%), limiting the generalisability of findings to more diverse populations and potentially underrepresenting the perspectives of clinicians from other demographic backgrounds. Third, the sample was skewed, with psychologists in private settings (*n*=22) overrepresented and psychiatrists (*n*=1) underrepresented in the quantitative data. The qualitative sample (*n*=6) mirrored this trend, with a majority of psychologists (*n*=4) and only one psychiatrist included. This imbalance limits the ability to draw conclusions about the perspectives of other health professionals, such as psychiatrists and mental health nurses, on DSM-5 ED severity ratings. The qualitative data also suggest that these severity ratings were primarily used for triaging in tertiary settings. Therefore, future research should aim to include a broader range of health professionals, as well as more diverse demographic groups, in both private and tertiary settings.

Third, and related to the second limitation, the sample was biased because specialised ED clinicians were self-selected into the study, which limited the diversity of clinician

specialties and experience. This is particularly important to consider when assessing whether non-specialised clinicians used these severity ratings to decide if and where patients should be re-referred. Future studies would benefit from including the perspectives of healthcare professionals who manage and treat a broader range of mental health issues, not just EDs.

Conclusion

In conclusion, this study highlights significant concerns regarding the clinical utility of DSM-5 severity ratings for EDs. Most clinicians do not use these ratings, finding them inadequate for capturing the complexity of EDs and potentially harmful to patients. The reliance on measures like weight and frequency of disordered behaviours can perpetuate harmful beliefs and hinder recovery. While the ratings are occasionally useful for administrative purposes, they hold limited value for direct clinical discussions. To improve the assessment and treatment of EDs, a more comprehensive approach that considers a broader range of psychological, medical, and social indicators is needed. Developing a practical checklist or severity screening tool would better support clinicians in high-pressure settings, ensuring that care is appropriately tailored to each patient's unique needs.

References

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589-597.

Dalle Grave, R., Sartirana, M., El Ghoch, M., & Calugi, S. (2018). DSM–5 severity specifiers for anorexia nervosa and treatment outcomes in adult females. *Eating Behaviors*, *31*, 18-23.

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders. *Clinical Psychology: Science and Practice*, *29*(4), 325.

Dang, A., Granero, R., Sanchez, I., Solé, L.G., Toro, J.J.-d., Rosinska, M., Jimenez-Murcia, S., Krug, I. and Fernandez-Aranda, F. (2024), Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 Versus ICD–11 Versus Drive for Thinness: *European Eat Disorders Review*

Dang, A. B., Kiropoulos, L., Anderluh, M., Collier, D., Fernandez-Aranda, F., Karwautz, A., Treasure, J., Wagner, G., & Krug, I. (2024). Do risk factors differentiate DSM–5 and drive for thinness severity groups for anorexia nervosa? *Journal of Eating Disorders*, *12*(1), 5.

Dang, A. B., Kiropoulos, L., Castle, D., Jenkins, Z., Phillipou, A., Rossell, S., & Krug, I. (2024). Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM–5 , the ICD–11, and overvaluation of Weight/Shape severity ratings. *Eating Disorders*, 1-17.

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education*, *40*(4), 314-321.

Fatt, S. J., Mitchison, D., Bussey, K., & Mond, J. (2022). Methods used to assess insight in individuals with eating disorders: a scoping review. *Journal of Mental Health*, 1-12.

Gianini, L., Roberto, C. A., Attia, E., Walsh, B. T., Thomas, J. J., Eddy, K. T., Grilo, C. M., Weigel, T., & Sysko, R. (2017). Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers. *International Journal of Eating Disorders*, *50*(8), 906-916.

Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of Basic and Clinical Pharmacy*, *5*(4), 87.

Mora-Maltas, B., Lucas, I., Granero, R., Vintró-Alcaraz, C., Miranda-Olivos, R., Baenas, I., Sánchez, I., Jiménez-del Toro, J., Sánchez-González, J., & Krug, I. (2023). Cognitive flexibility and DSM–5 severity criteria for eating disorders: assessing drive for thinness and duration of illness as alternative severity variables. *Journal of Eating Disorders*, *11*(1), 155

Nakai, Y., Nin, K., Noma, S. I., Teramukai, S., Fujikawa, K., & Wonderlich, S. A. (2017). The impact of DSM-5 on the diagnosis and severity indicator of eating disorders in a treatment-seeking sample. *International Journal of Eating Disorders*, *50*(11), 1247-1254.

Krug, I., Binh Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the DSM-5 severity indices for eating disorders. *European Eating Disorders Review*, *29*(3), 482-498.

Lowe, M. R., Davis, W., Lucks, D., Annunziato, R., & Butryn, M. (2006). Weight suppression predicts weight gain during inpatient treatment of bulimia nervosa. *Physiology & Behavior*, *87*(3), 487-492

Reas DL, Rø Ø. Investigating the DSM5 severity specifiers based on thinness for adults with anorexia nervosa. *International Journal of Eating Disorders.* 50:990–994.

Smink, F. R., van Hoeken, D., Oldehinkel, A. J., & Hoek, H. W. (2014). Prevalence and severity of DSM-5 eating disorders in a community cohort of adolescents. *International Journal of Eating Disorders*, *47*(6), 610-619.

Spindler, A., & Milos, G. (2007). Links between eating disorder symptom severity and psychiatric comorbidity. *Eating Behaviors*, *8*(3), 364-373.

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, *18*, 1-18.

Zimmerman, M., Morgan, T. A., & Stanton, K. (2018). The severity of psychiatric disorders. *World Psychiatry*, *17*(3), 258-275.

Chapter 5: Comprehensive Comparison of DSM-5, ICD-11, and OWS Severity Ratings for Anorexia Nervosa: Psychological and Biological Correlates

5.1 Chapter Introduction

Chapters 2 and 4 revealed that the current BMI-based severity rating for AN, as outlined in the DSM-5, fails to effectively differentiate levels of ED-related psychopathology. This may explain why mental health clinicians have not widely adopted these severity ratings in practice, as they do not fulfill their intended purpose. Most studies, including those discussed in Chapters 2 and 4, have primarily focused on the DSM-5 severity ratings, with little attention given to ICD-11 or alternative frameworks. Furthermore, those that examine DSM-5 severity ratings have mostly explored its relationship with ED and general psychopathology.

AN is a multifaceted disorder, involving not just ED psychopathology but also significant cognitive impairments (Mora-Maltas et al., 2023), medical complications (Gibson et al., 2019), and psychiatric comorbidities (Himmerich et al., 2019). Despite this, little is known about whether the DSM-5 and ICD-11 BMI-based severity rating can effectively index biological correlates, such as electrolyte disturbances (e.g., potassium, magnesium) that are common in AN (Buh et al., 2024; Udo et al., 2019). Electrolyte disturbances often reflect the severity of disordered eating behaviours. For example, low magnesium levels may require medical admission, even when BMI is within the normal range, suggesting that if BMI were an effective severity rating, it should correlate with these biological markers.

As highlighted in Chapter 1, psychiatric comorbidity is common in individuals with AN, with symptom severity often increasing alongside the number of co-occurring psychiatric conditions (Himmerich et al., 2019). Comorbidities, especially affective and anxiety disorders (Andrés-Pepiñá et al., 2020), can intensify cognitive-affective symptoms

like preoccupation with being underweight, fear of weight gain, and fixation on appearance among those with AN. Additionally, psychiatric comorbidities have been linked to higher mortality rates in AN (Ulfvebrand et al., 2015). Specifically, psychiatric comorbidity was linked to a 1.9-fold increase in 10-year mortality for individuals with AN compared to those without comorbidity, and this risk was notably higher, with a four-fold increase, when comorbidity was diagnosed between the ages of 6 and 25 years (Søeby et al., 2024). Therefore, for a severity rating to effectively replace a comprehensive assessment that captures the multifaceted nature of AN, it should also correlate with the number of psychiatric comorbidities.

Chapters 2 (Study 1) and 4 (Study 2) identify a significant gap in the literature: although OWS is a central diagnostic criterion for AN, with individuals often basing their self-worth on these factors, only one study (Gianini et al., 2017) has explored OWS as an alternative measure of severity for AN. Consequently, there is limited understanding of how DSM-5 severity ratings compare with those of the ICD-11 (another BMI-based severity scale with different cut-off points) or OWS. Such comparisons are crucial, as the BMI cut-off scores in the DSM-5 were arbitrarily determined, whereas the ICD-11 cut-off of a BMI of 14 kg/m^2 is based on its established association with high medical risk, often warranting hospital admission (RANZCP; Hay et al., 2014). Direct comparisons of the DSM-5 and ICD-11 severity rating systems could help determine whether the limited support for BMI as an indicator of AN severity stems from inherent limitations of BMI as a measure or the arbitrary nature of its cut-offs.

In summary, as found in Chapter 2 (Study 1), although a few studies have evaluated the clinical utility of proposed DSM-5 severity rating for AN, the majority have focused primarily on the relationship between these severity groups and ED psychopathology. The next chapter, Chapter 5 (Studies 3a and 3b), aimed to address these gaps, articulated above,

by comparing the DSM-5, ICD-11, and OWS severity ratings for AN in relation to psychological, biological correlates as well as psychiatric comorbidities.

Chapter 5 comprises two studies **(Studies 3a and 3b),** both of which utilised a sample of 312 treatment-seeking patients with AN from St. Vincent's Hospital, Melbourne, Australia. Study 3a directly compared the DSM-5 and OWS severity ratings in indexing psychological and biological variables in individuals with AN; however, it did not include the ICD-11 as the study was completed before its publication. To address this gap, I conducted additional analyses using the same sample to examine how the ICD-11 BMI-based severity rating relates to psychological and biological variables in individuals with AN. These findings are presented in Section 5.3.

Study 3b shifted focus to psychiatric comorbidities, recognising the negative association between clinical outcomes in AN and the number of psychiatric comorbidities, as highlighted in the literature (Herzog et al., 2018). It compared the DSM-5, ICD-11, and OWS severity ratings to explore how these indicators relate to comorbid psychiatric conditions in individuals with AN.

5.2 Study 3a. Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates

Study 3a was published in *European Eating Disorders Review* in January 2023.

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2023). Assessing severity in anorexia nervosa: Do the DSM‐5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?. *European Eating Disorders Review*, 31(4), 447-461. https://doi.org/10.1002/erv.2969

Contribution to the Study

Author	Task contribution
An Dang	Conceptualisation; manuscript writing; editing; data analyses
Litza Kiropoulos	Editing
David Castle	Data collection, editing
Zoey Jenkins	Data collection, editing
Andrea Phillipou	Data collection, editing
Rossell Sussan	Data collection, editing
Isabel Krug	Conceptualisation; manuscript – reviewing and editing; supervision

Received: 15 June 2022 | Revised: 19 November 2022 | Accepted: 10 January 2023

DOI: 10.1002/erv.2969

WILEY

RESEARCH ARTICLE

Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?

An Binh Dang[1] | **Litza Kiropoulos**[1] | **David J. Castle**[2,3] | **Zoe Jenkins**[4,5,6] | **Andrea Phillipou**[4,7,8] | **Susan L. Rossell**[4,7] | **Isabel Krug**[1]

[1]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Victoria, Australia

[2]Centre for Complex Interventions, Centre for Addictions and Mental Health, Toronto, Ontario, Canada

[3]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

[4]Department of Mental Health, St Vincent's Hospital, Melbourne, Victoria, Australia

[5]Iverson Institute, Swinburne University of Technology, Melbourne, Victoria, Australia

[6]Department of Psychiatry, The University of Melbourne, Melbourne, Victoria, Australia

[7]Centre for Mental Health, Swinburne University of Technology, Melbourne, Victoria, Australia

[8]Department of Mental Health, Austin Health, Melbourne, Victoria, Australia

Correspondence
An Binh Dang, Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, VIC, Australia.
Email: dangb1@student.unimelb.edu.au

Funding information
BMI

Abstract

Objective: This study evaluated the severity ratings for anorexia nervosa (AN) in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and an alternative severity rating based on overvaluation of weight/shape, on a range of psychological and biological variables.

Method: A sample of 312 treatment-seeking patients with AN (mean age = 25.3, *SD* = 7.6; mean BMI = 16.8 kg/m^2, *SD* = 2.4) were categorised using both DSM-5 severity levels (mild/moderate/severe/extreme) and weight/shape (low/high) overvaluation. The severity categories were compared on a range of psychological (e.g., eating psychopathology) and biological (e.g., sodium) variables.

Results: Results showed that the overvaluation of weight/shape appeared better at indexing the level of severity in psychological variables among patients with AN compared to the DSM-5 severity rating with moderate to large effect sizes. Moreover, the DSM-5 mild and moderate severity groups experienced significantly higher eating and general psychopathology than the severe and extreme groups. Finally, neither the DSM-5 nor the weight/shape severity groups differed on any of the biological variables.

Conclusions: This study provided no support for the DSM-5 severity rating for AN, while initial support was found for the weight/shape overvaluation approach in indexing psychological but not biological correlates.

KEYWORDS
anorexia nervosa, biological factors, DSM-5, severity rating, weight and shape overvaluation

Abbreviations: AAN, Atypical anorexia nervosa; AN, Anorexia nervosa; AN-BP, AN- Binge Purging; APA, American Psychiatric Association; BDQ, Brief Disability Questionnaire; BED, binge eating disorder; BETRS, Body Image and Eating Disorder Treatment Recovery Service; BMI, body mass index; BN, bulimia nervosa; CFS, Cognitive Flexibility Scale; DASS-21, Depression Anxiety Stress Scale; DSM-5, Diagnostic and Statistical Manual of Mental Disorders-5; ED, eating disorder; EDE-Q, Eating Disorder Examination Questionnaire; OSFED, Otherwise Specified Feeding and Eating Disorders; QLES-Q-SF, Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form.

Highlights

- This study assessed both the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) and the alternative severity indices for anorexia nervosa (AN) using overvaluation of weight and shape on a range of psychological (e.g., eating psychopathology) and biological (e.g., sodium, potassium level) correlates.
- Overvaluation of weight/shape appeared to be better at indexing the level of severity in psychological variables among patients with AN compared to the DSM-5 severity rating with moderate to large effect sizes.
- Neither the DSM-5 nor the weight/shape severity groups were able to index the level of severity in biological variables.

1 | INTRODUCTION

Anorexia nervosa (AN) is a psychological disorder characterised by self-starvation and low body weight (American Psychiatric Association [APA], 2013). Rates of psychiatric comorbidity are high and mortality rates from suicide or physical complications are the highest among any of the psychological disorders (Zipfel et al., 2015). To index disorder severity, the fifth edition of the Diagnostic and Statistical Manual of Mental Health (DSM-5; APA, 2013) introduced the use of Body Mass Index (BMI). The BMI severity indicator for AN is a simple index of weight-for-height that categorises individuals into mild (≥17.0 kg/m^2), moderate (16–16.99 kg/m^2), severe (15–15.99 kg/m^2), and extreme (<15 kg/m^2) severity groups. As stated, the DSM-5 severity specifiers provide a guide for clinicians in rating the intensity, duration, psychopathology, or other disorder-specific markers of severity (APA, 2013). Although the DSM-5 assumes a positive trend between severity index (i.e., BMI) and psychopathology, the severity is subject to clinical judgement based on present symptoms, functional impairment or medical complications (APA, 2013).

Despite its broad usage, recent studies found limited evidence for the clinical utility of the DSM-5 AN severity rating, namely its lack of association with ED-related psychological variables (e.g., body weight/shape concerns; Dang et al., 2022). Consequently, studies have looked for alternative severity ratings such as overvaluation of weight/shape (Gianini et al., 2017) and drive for thinness (Krug et al., 2022), and found that these alternative strategies distinguished severity groups better than the DSM-5 BMI severity rating approach. Overvaluation of weight/shape has been suggested as a diagnostic feature of AN (APA, 2013; Fairburn et al., 2003), but has yet to be investigated as an alternative severity measure for AN.

Additionally, haematological and electrolyte disturbances are highly prevalent in AN due to restricted eating and purging behaviours for the AN- Binge Purging (AN-BP) subtype (Miller et al., 2005), yet to date, no study has assessed the relationship between the DSM-5 severity rating for AN and biological variables (i.e., haematological and electrolyte imbalance). Therefore, this study was the first to explore the clinical utility of the DSM-5 and the alternative severity rating based on overvaluation of weight/shape for AN on a range of psychological and biological factors.

1.1 | The DSM-5 severity ratings for anorexia nervosa for psychological indicators

The use of BMI in DSM-5 to index AN severity reflects the centrality of body weight as a key feature of AN (Criterion A; APA, 2013). The existing literature, however, offers mixed evidence for using BMI to index AN severity. In some studies, having a low BMI has been linked to increased medical (e.g., cardiac problems) and psychological complications (e.g., comorbidity with anxiety) and a worse prognosis (Button & Palmer, 2010; Wolfi & Treasure, 2011). Conversely, other studies have found no relationships between BMI and general psychopathology (e.g., depression or anxiety) that are commonly present in conjunction with AN (Gluschokoff et al., 2019; Hebebrand et al., 1996).

Studies have directly examined the clinical utility of the DSM-5 severity ratings of AN (e.g., Dakanalis et al., 2018; Dalle Grave et al., 2018). Two studies found the DSM-5 AN severity groups to be indistinguishable on parameters such as age of onset and number of hospitalisations (Gianini et al., 2017; Machado et al., 2017). While studies found individuals in the DSM-5 AN severe and extreme groups presented with higher ED symptoms

overall (Dakanalis et al., 2018; Nakai et al., 2017), others found no significant differences across the DSM-5 AN severity groups in AN-related psychopathology (e.g., distress or ED psychopathology; Dalle Grave et al., 2018; Krug et al., 2021).

Given such inconsistency in the literature, Dang et al. (2022) conducted a systematic review and meta-analysis ($N = 22$) to assess the clinical usefulness of the DSM-5 severity ratings for EDs (i.e., AN, binge eating disorder [BED] and bulimia nervosa [BN]). Specifically, for AN, Dang et al.'s meta-analysis ($n = 5$; Dalle Grave et al., 2018; Gianini et al., 2017; Machado et al., 2017; Smith et al., 2017; Zayas et al., 2018) found no significant differences in terms of ED psychopathology across AN severity groups, offering limited support for the usage of BMI as a severity index for AN.

Besides ED psychopathology, the DSM-5 further outlines that level of functional impairment needs to be considered when assessing AN severity (APA, 2013). Such considerations are important in evaluating the level of AN severity since those with AN have been found to experience greater deficits in cognitive flexibility (e.g., Fagundo et al., 2012; Nakazato et al., 2010), higher level of disability in everyday activities (e.g., Flament et al., 2001; Godard et al., 2013), and poorer quality of life (e.g., Mond et al., 2005; Sy et al., 2013) compared to both the general population and those with other psychiatric conditions. However, most studies that examined the clinical utility of the DSM-5 AN severity index often overlooked the impact of AN on cognition, everyday activities and quality of life (Dang et al., 2022). Therefore, an examination of the differences in these variables, along with the level of ED and general psychopathology (i.e., depression, anxiety, stress) across the DSM-5 AN severity groups is needed to inform the clinical utility of the DSM-5 approach.

1.2 | The DSM-5 severity ratings for anorexia nervosa for biological indicators

Besides the psychological impacts, medical complications are also highly prevalent in AN, with more than 50% of all deaths in AN being due to the refusal and inability to consume adequate calories and nutrients or purging behaviours in individuals with AN-BP subtype (Mehler et al., 2018; Sullivan, 1995). These complications can impact almost every organ system and result in electrolyte imbalance and haematologic abnormalities (Gallagher et al., 2021). The severity of electrolyte and haematologic complications is suggested to depend on the extent and duration of weight loss and compensatory behaviours (Forney et al., 2016). For example, repeated vomiting and prolonged abuse of laxatives and diuretics have been found to be associated with a loss of electrolytes, such as potassium and sodium (Forney et al., 2016; Mehler et al., 2018). Additionally, having a low white blood cell count, and low haemoglobin potassium, magnesium, and phosphate levels have been associated with severe consequences such as increased infection rates, cardiac muscle weakness or paralysis and cardiac arrhythmias (Caregaro et al., 2005). Therefore, in clinical practice, besides low weight, individuals with AN may require hospitalisation for electrolyte imbalance or haematologic disturbances (Bargiacchi et al., 2019).

The DSM-5 proposes using BMI to index AN severity to guide clinical decision-making (e.g., need for hospitalisation or supervision); however, the relationships between BMI and electrolyte and haematologic indicators are inconsistent (e.g., Misra et al., 2004). Specifically, while some studies have found low BMI to be positively correlated with a lower white blood cell count (e.g., Devuyst et al., 1993; Walsh et al., 2019) and potassium level (Jacobs & Schneider, 1985), other studies did not find such correlations (e.g., Imbierowicz et al., 2004; Misra et al., 2004; Polli et al., 2008). Studies have also found a very weak link between a patient's BMI and haemoglobin level (Mamou et al., 2016; Misra et al., 2004). This means that those who are at a higher risk for major haematologic complications are not necessarily individuals with a low BMI. Moreover, the relationships between BMI and the level of magnesium and phosphate are still not clear (see Raj et al., 2012).

Low levels of phosphate (i.e., hypophosphatemia) can lead to life-threatening complications, especially during the refeeding period if phosphate is not adequately supplemented (Oudman et al., 2018). Low levels of phosphate have been suggested to be more frequent in those with severe malnutrition (Ornstein et al., 2003; Raj et al., 2012). The National Institute for Health and Clinical Excellence (NICE) guidelines state that individuals with a BMI < 16 kg/m^2 are at risk of developing hypophosphatemia (NICE, 2006), but studies that directly examined the relationship between BMI and phosphate levels are few (Ornstein et al., 2003).

Bradycardia, pulse rate under 60 beats per minute, is the most prevalent cardiac complication in patients with AN, occurring in up to 95% of cases (Mehler & Brown, 2015). While bradycardia has not consistently been suggested to be associated with patients' current BMI, change in BMI (i.e., the amount of weight loss) has been found to be a strong independent predictor of bradycardia (Assalone et al., 2022; Whitelaw et al., 2018). However, we are aware of no study to date that has directly investigated the link between the DSM-5 AN severity groups based on BMI and pulse rate.

1.3 | Overvaluation of weight and shape: Alternative severity indicator for anorexia nervosa

Given the ambiguity surrounding the DSM-5 AN severity rating, researchers have found support for alternative severity classifications for AN, including transdiagnostic indices such as overvaluation of weight/shape (Gianini et al., 2017) and drive for thinness (Krug et al., 2021). These alternative approaches are guided by Fairburn et al.'s (2003) transdiagnostic model of EDs. This model proposes that EDs present with distinctive clinical manifestations (e.g., severe dietary restraint; binge eating episodes; compensatory behaviours), which stem from a common core psychopathology characterised by overvaluation of eating, body shape and weight (Cooper & Fairburn, 1993; Fairburn, 2008). Building on this model, empirical research has found overvaluation of weight/shape to be a promising alternative severity rating specifically for BED (e.g., Dang et al., 2022; Grilo et al., 2015) and transdiagnostically (i.e., across patients with AN, BN and BED; Gianini et al., 2017).

Despite weight/shape overvaluation being suggested as a key diagnostic feature of AN (APA, 2013), it has not yet been examined as an alternative severity rating for AN. Furthermore, no study to date has directly assessed the relationship between overvaluation of weight/shape and biological parameters (i.e., electrolyte and haematological level). It has been suggested that weight/shape overvaluation drives dietary restraint and extreme weight control methods (e.g., purging; Fairburn, 2008; Forrest et al., 2018; Murphy et al., 2010). High levels of dietary restraint, purging or diuretic and laxative abuse, in turn, have been found to lead to electrolytes and haematological imbalance (Forney et al., 2016; Mehler et al., 2018; Pomeroy, 2001). To address these gaps, this study also aimed to investigate the relationship between overvaluation of weight/shape as an alternative AN severity index and its relationship to a range of biological and psychological correlates that are commonly observed in those with AN.

1.4 | The current study

Using a treatment-seeking population of patients with AN, we examined the clinical validity and utility of two AN severity indicators, the DSM-5 severity rating based on BMI and an alternative indicator based on overvaluation of weight/shape. Specifically, the current study explored how these severity indicators differentiated on biological (e.g., electrolyte, haematologic) and psychological indicators, including those that have not been investigated previously (e.g., cognitive flexibility or physical disability) but are common in AN (Forney et al., 2016; Mehler et al., 2018).

Based on the previous literature (e.g., Dang et al., 2022; Gianini et al., 2017), it was hypothesised that overvaluation of weight/shape would provide a clinically more meaningful way to index severity for AN in terms of the biological and psychological variables of interest than the DSM-5 BMI-based severity categories.

2 | METHOD

2.1 | Participants

Data were obtained from the outpatient treatment unit of the Body Image and Eating Disorder Treatment Recovery Service (BETRS) at St Vincent's Hospital in Melbourne, Australia. Participants received an diagnosis (AN-R = 235, 75.3%) after a comprehensive assessment by senior clinicians with consultation from psychiatrists per DSM-5 criteria (APA, 2013). The sample included 312 treatment-seeking patients with AN with a mean age of 25.3 (SD = 7.6) and a mean BMI of 16.8 kg/m2 (SD = 2.4).

2.2 | Measures

Demographic information regarding ethnicity, education, employment, relationship status, age, gender, weight (before and after treatment), height, and ED diagnosis were collected. The age of AN onset was self-reported by the participants at the initial assessment. The duration of the illness was calculated as the duration between the self-reported age of AN onset and age at initial assessment. At the time of assessment, all included participants had blood work measuring sodium, magnesium, potassium, phosphate, haemoglobin, and white blood cell count. Pulse rate was also taken at the time of assessment.

2.2.1 | Eating psychopathology

The Eating Disorder Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994), a 28-item self-report measure, was used to measure ED psychopathology. The EDE-Q records items concerning the past 28 days and includes four subscales: eating restraint, weight concern, shape concern, eating concern, and one global score. Higher scores on the EDE-Q indicate greater levels of disordered eating. The EDE-Q has shown good test-retest reliability (Peterson et al., 2007; Reas et al., 2006), and

convergence with the interview version (i.e., Eating Disorder Interview (EDE); Cooper & Fairburn, 1987; Mond et al., 2004).

2.2.2 | Depression, anxiety and stress

The Depression Anxiety Stress Scale (DASS-21; Lovibond & Lovibond, 1995), a 21 four-point scale questionnaire (0 = *never* to 3 = *always*), was used to measure the three related negative emotional states of depression, anxiety, and stress that participants experienced over the previous week. The DASS-21 has been found to have excellent internal consistency, while its score interpretation has been found to have good construct validity (Henry & Crawford, 2005; Page et al., 2007).

2.2.3 | Brief Disability Questionnaire

To assess the level of disability in everyday activities, the Brief Disability Questionnaire (BDQ; Naismith et al., 2007) an 8-item three-point scale (0 = *never* to 2 = *definitely)*, was used. Items 1–3 measure physical disability while items 4 to 6 assess mental health disability (e.g., motivation or efficiency for work activities). Across items 1–6, a total score range of 8–13 indicates a 'moderate' level of disability, while a score range of 14–22 outlines a 'severe' level of disability. Items 7–8 examine the periods during which respondents were unable to carry out their usual activities due to their disability, providing an estimate of 'functional disability. The BDQ has been found to have a good test-retest reliability and to be associated with varying disabilities across affective, anxiety and somatisation disorders (Ormel et al., 1994; Ware et al., 1996).

2.2.4 | Quality of life

The Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form (Q-LES-Q-SF; Schechter et al., 2007) was used to capture life satisfaction over the previous week. The QLES-Q-SF has 14 self-administered items where participants rate their satisfaction with life in seven domains on a 5-point scale (1 = *very poor* to 5 = *very good*): physical health, feelings, work, household duties, schoolwork, leisure time and social relations. Higher scores indicate greater life satisfaction and enjoyment, with a maximum score of 70. The scale has been found to have good test-retest reliability and internal consistency, and its score interpretations also have shown sound construct validity (Hope et al., 2009; Ritsner et al., 2002).

2.2.5 | Cognitive flexibility

Cognitive flexibility was measured using the Cognitive Flexibility Scale (CFS; Martin & Rubin, 1995). The CFS is a 6-point Likert scale (1 = *strongly disagree* to 6 = *strongly agree*) with a higher score on the CFS indicating greater cognitive flexibility. The CFS has been shown to have a good concurrent and construct validity with a high test-retest reliability (Martin & Rubin, 1995).

2.2.6 | Procedure

At admission to the treatment facilities, all participants were provided with a written informed consent to participate in research, completed self-reported questionnaires and underwent blood work to measure biological data. The study obtained ethics approval from the Human Research Ethics Committee at St Vincent's Hospital, Melbourne. All procedures for data collection were in line with the Declaration of Helsinki.

2.3 | Anorexia nervosa severity categories

2.3.1 | The DSM-5 severity groups based on BMI

Using participants' BMI at assessment, the AN severity subgroups were created using the DSM-5 severity definitions based on BMI, with a total of 136 (43.6%) participants being classified as mild (≥17.0 BMI), 64 (20.5%) as moderate (16–16.99 BMI), 49 (15.7%) as severe (15–15.99 BMI) and 63 (20.2%) as extreme (<15 BMI).

2.3.2 | The alternative severity rating based on overvaluation of shape/weight

Based on previous studies by Gianini et al. (2017) and Grilo et al. (2015), overvaluation of weight/shape was measured using two specific items from the EDE-Q: "Over the past 4 weeks, has your *shape* influenced how you feel about (judge, think, evaluate) yourself as a person?" and "Over the past 4 weeks has your *weight* influenced how you feel about (judge, think, evaluate) yourself as a person?" These two overvaluation items were rated on a 7-point scale (0 = *no importance* to 6 = *supreme importance*) in reference to the past 28 days. Therefore, a cut-off score of 4 (i.e., score ≥4 on either overvaluation item) was used to classify individuals with

AN into low and high levels of overvaluation of weight/shape (Gianini et al., 2017).

Due to missing data for the two overvaluation of weight and shape items, out of 312 participants, we were only able to categorise 240 participants into low ($n = 50$; 20.8%) and high ($n = 190$; 79.2%) overvaluation of weight/shape groups. Additionally, as suggested by previous studies (e.g., Grilo et al., 2015), to reduce bias, we excluded the two overvaluation of weight/shape items when comparing the low and high overvaluation of weight/shape groups on the EDE-Q weight/shape subscales.

2.4 | Statistical analyses

For comparison between the groups (i.e., DSM-5 BMI and overvaluation of weight/shape severity groups), we used chi-square tests (χ^2) for categorical variables. We also conducted analysis of variance (ANOVA) for continuous dependent variables, followed by Bonferroni post-hoc comparisons. All tests were two-tailed with the chosen alpha level at *0.05*. To account for multiple testing, the False Discovery Rate (FDR) procedure was used.

Discriminative capacity between the DSM-5 BMI and overvaluation of weight/shape severity groups was evaluated using both the results of the significant tests (*p*-value), and the effect size coefficients: Cramer's V (V) coefficient was calculated for categorical measure with values of 0.06, 0.15, and 0.30 were interpreted as small, medium, and large effect sizes respectively (Cohen, 1988) and partial eta-squared coefficient (η^2) was used to measure effect for ANOVA with the values of 0.06, 0.10, and 0.25 being interpreted as small, medium, and large effect sizes (Levine & Hullett, 2002). Furthermore, as not all participants completed all assessments, each analysis was based on available data, treating incomplete data as missing. The number of participants included in each analysis was stated in Tables 1 and 2.

3 | RESULTS

3.1 | Demographics and clinical characteristics

Table 1 summarises the demographic characteristics of the whole sample as well as both the DSM-5 and overvaluation of weight/shape severity groups. Most patients with AN were single (76.2%), Caucasian (91.5%), female (94.6%), and had completed secondary or a lower education level (61.3%). The mean age of the AN onset was 17.4 years ($SD = 6.1$), and the duration of the illness was 7.8 years ($SD = 7.4$). The mean BMI was 16.8 kg/m^2 ($SD = 2.4$) at intake/assessment and 17.4 kg/m^2 ($SD = 2.6$) at discharge. The DSM-5 AN severity groups had a significantly different BMI at discharge, with individuals in milder severity groups having higher BMI ($p < 0.001$; $\eta^2 = 0.479$), but the DSM-5 severity groups were indistinguishable in either age of onset or duration of AN. On the other hand, the AN overvaluation of weight/shape severity groups did not differ significantly in age, age of onset, duration of illness and BMI at discharge; however, they differed significantly in BMI at assessment, with individuals in the low overvaluation of weight/shape severity group displaying a significantly lower BMI than those in the high overvaluation of weigh/shape group ($p = 0.001$; $\eta^2 = 0.046$). Furthermore, overvaluation weight/shape AN alternative severity groups significantly differed on the distribution of gender, with more females with high overvaluation of weight/shape than expected by chance.

3.2 | Comparison of the classifications for the severity categories

Table 2 summarises the descriptive statistics and statistical analyses for the individuals in the different DSM-5 and overvaluation of weight/shape AN severity groups on the assessed psychological and biological variables.

3.2.1 | Psychological variables

Concerning the psychological data, the DSM-5 severity rating based on BMI for AN achieved adequate discriminative capacity (i.e., significant with mostly small-moderate effect sizes) for anxiety ($p = 0.02$; $\eta^2 = 0.053$), dietary restraint ($p < 0.001$, $\eta^2 = 0.098$), shape ($p < 0.001$; $\eta^2 = 0.084$), eating ($p = 0.005$, $\eta^2 = 0.064$) and weight concerns ($p < 0.001$; $\eta^2 = 0.108$). Post-hoc results showed that individuals in the mild and the moderate DSM-5 severity AN groups experienced higher dietary restraint and weight/shape concerns than those in the extreme severe group. Individuals in the mild DSM-5 AN severity group also exhibited significantly higher anxiety and eating concerns than those in the extreme group. However, the DSM-5 severity rating for AN was not able to discriminate differences between the DSM-5 AN severity groups on the remaining psychological variables.

In terms of overvaluation of weight/shape as an alternative AN severity classification, we found significant differences between the two overvaluation of weight/shape groups with moderate to large effect size on eating psychopathology [dietary restraint ($p < 0.001$; $\eta^2 = 0.207$), eating concerns ($p < 0.001$; $\eta^2 = 0.275$), shape ($p < 0.001$; $\eta^2 = 0.463$) and weight concerns ($p < 0.001$; $\eta^2 = 0.420$)]

TABLE 1 Demographic information for the DSM-5 and overvaluation of weight/shape severity groups

		DSM-5 classification										OWS classification				
		Total (*n*, %)		Mild (*n*, %)		Moderate (*n*, %)		Severe (*n*, %)		Extreme (*n*, %)	P (V)	Low (*n*, %)		High (*n*, %)		P (V)
Gender	Male	17	5.4%	12	8.8%	2	3.1%	1	2.0%	2	3.2%	0.144 (0.132)	6	12.0%	6	3.2%	**0.011(0.165[†])**
	Female	295	94.6%	124	91.2%	62	96.9%	48	98%	61	96.8.0%		44	88.0%	184	96.8%	
Marital status	Single	237	76.2%	107	78.7%	49	76.6%	36	73.5%	45	72.6%	0.108 (0.154)	35	70%	149	78.4%	0.063 (0.209)
	Widowed	2	0.6%	2	1.5%	0	0%	0	0%	0	0%		0	0%	1	0.5%	
	Divorced/Separated	16	5.2%	2	1.4%	2	3.1%	4	8.2%	8	12.9%		4	8.0%	8	4.3%	
	Married - in couple	35	11.3%	14	10.3%	11	16.9%	5	10.2%	5	8.1%		6	12.0%	22	11.6%	
	Other/Unknown	22	6.8%	11	8.1%	3	4.6%	4	6.5%	4	6.5%		5	10.0%	10	5.3%	
Education Secondary/lower		193	62.6%	88	65.7%	43	68.3%	30	61.1%	32	51.6%	0.121 (0.176)	27	54.0%	120	63.2%	0.052 (0.243)
	Vocational	7	2.3%	4	3.0%	1	1.6%	0	0.0%	2	3.2%		2	4.0%	49	25.8%	
	University	49	25.6%	27	20.1%	14	22.2%	15	30.6%	23	37.1%		12	24.0%	5	2.6%	
	Other/Unknown	29	9.6%	15	11.1%	5	7.9%	4	8.1%	5	8.0%		9	18.0%	16	8.4%	
Ethnicity	Aboriginal	7	2.5%	3	2.3%	1	1.7%	3	6.7%	0	0%	0.053 (0.170)	2	4.0%	4	2.1%	0.891 (0.087)
	Caucasian	261	91.5%	121	94.5%	55	94.8%	40	88.9%	45	83.3%		41	82.0%	162	85.3%	
	Asian	10	3.5%	3	2.4%	1	1.7%	1	2.2%	5	9.3%		2	4.0%	6	3.2%	
	Other/Unknown	7	2.5%	1	0.8%	1	1.7%	1	2.2%	4	7.42%		5	10.0%	19	10.0%	
Employment	Unemployed.	29	9.7%	9	7.6%	3	6.1%	2	6.5%	7	15.9%	0.429 (0.143)	6	12.0%	15	7.9%	0.240 (0.185)
	Student	101	33.8%	44	33.8%	19	30.2%	14	30.4%	24	40.0%		16	32.0%	62	32.6%	
	Employed	80	26.8%	40	30.8%	19	30.2%	11	23.9%	10	16.6%		16	32.0%	47	24.7%	
	Home duties	10	3.3%	5	3.8%	1	1.6%	1	2.2%	2	4.5%		0	0%	8	4.2%	
	Unable to work/Other	79	26.4%	30	23.1%	20	31.7%	14	30.4%	15	25.0%		12	24.0%	58	30.5%	

(Continues)

TABLE 1 (Continued)

DSM-5 classification									OWS classification				
	Mild		Moderate		Severe		Extreme			Low OWS		High OWS	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	P (η^2)	Mean	S.D	Mean	S.D	P (η^2)
Age (yrs-old)	25.3	7.6	27.4	10.5	26.5	8.6	29.2	10.3	0.140 (0.029)	27.5	9.5	26.3	8.9	0.518 (0.000)
Onset ED (yrs-old)	17.4	6.1	16.5	4.5	18.6	5.8	19.1	7.1	0.290 (0.022)	19.4	5.5	17.2	6.0	0.06 (0.004)
Duration ED (yrs)	7.8	7.4	11.3	11.2	7.6	7.0	9.9	9.4	0.170 (0.029)	7.4	7.1	9.3	9.0	0.266 (0.002)
BMI at assessment (kg/m^2)	16.8	2.4	16.4	0.3	15.4	0.3	13.9	0.9	**<0.001 (0.605**[†])	16.0	2.1	17.3	2.6	**0.001** (0.046)
BMI at discharge (kg/m^2)	17.4	2.6	16.4	0.3	16.0	1.2	14.8	2.0	**<0.001 (0.479**[†])	16.8	2.7	17.7	2.6	0.085 (0.047)

Note: Bold: significant comparison effect.

Abbreviations: DSM-5, Diagnostic and Statistical Manual of Mental Disorders; OWS, Overvaluation of Weight/shape; SD, Standard Deviation; V, Cramer's V – effect size for Chi-square statistic; η^2, Effect size for Analysis of Variance (ANOVA).

[†]Effect size in moderate-medium, and large-high range ($\eta^2 > 0.10$).

and general psychopathology variables [depression ($p < 0.001$; $\eta^2 = 0.127$), anxiety ($p < 0.001$; $\eta^2 = 0.128$) and stress symptoms ($p < 0.001$; $\eta^2 = 0.134$)]. Furthermore, as shown in Table 2, individuals in the low weight/shape concern group were found to experience significantly lower eating and general psychopathology compared to those in the high weight/shape concern group.

3.2.2 | Biological variables

Regarding the biological data, there were no meaningful clinically significant findings in the level of potassium, magnesium, phosphate, white blood cells, haemoglobin, and pulse rate amongst both the DSM-5 BMI and the alternative severity rating for AN based on overvaluation of weight/shape concerns.

4 | DISCUSSION

The current study is the first to examine the clinical utility of the DSM-5 AN severity rating system based on BMI and an alternative severity rating for AN using overvaluation of weight/shape, on a wide range of biological and psychological variables. Specifically, we looked at both severity ratings based on their ability to distinguish psychological and biological indicators that are commonly observed in patients with AN. In terms of the psychological variables, we found that the overvaluation of weight/shape appeared to be better at indexing the severity of individuals with AN compared to the DSM-5 severity rating based on BMI. However, both DSM-5 and overvaluation of weight/shape severity groups did not differ significantly in any of the biological indicators of interest.

4.1 | The DSM-5 anorexia nervosa severity rating

Our findings revealed that, while the DSM-5 AN severity groups based on BMI differed significantly in BMI at discharge, they were indistinguishable in age, duration of illness and age of illness onset. These findings replicate other clinical studies (e.g., Machado et al., 2017; Smith et al., 2017). Furthermore, our analysis of the application of the DSM-5 AN severity classification system identified limited significant findings amongst psychological variables with small to moderate effect sizes. Specifically, when examining the distinctiveness between DSM-5 AN severity groups, an unexpected effect was observed where those with a higher BMI (i.e., lower severity) exhibited higher levels of ED psychopathology and

TABLE 2 Clinical Profiles across the anorexia nervosa severity groups based on the DSM-5 and overvaluation of weight/shape

Anorexia nervosa	DSM-5 classification																	OWS score									
	Mild (1)			Moderate (2)			Severe (3)			Extreme (4)								Low OWS			High OWS						
	Mean	SD	*n*	Mean	SD	*n*	Mean	SD	*n*	Mean	SD	*n*	*N*	*F*	*p*	Effect size η^2	Post-hoc	Mean	SD	*n*	Mean	SD	*n*	*N*	*F*	*p*	Effect size η^2
DASS-21: Depression	28.3	11.0	123	23.8	12.1	58	23.9	12.3	44	23.9	13.2	56	281	2.7	0.104	0.035		17.9	13.1	48	28.3	10.6	174	222	32.1	**<0.001**	**0.127**[†]
DASS-21: Anxiety	22.2	10.6	123	17.0	12.6	58	17.7	12.0	44	16.3	11.0	56	281	4.1	**0.02**	0.053	1 > 4	11.7	10.4	48	21.6	10.8	174	222	32.3	**<0.001**	**0.128**[†]
DASS-21: Stress	28.0	9.0	123	24.2	11.5	58	27.3	9.9	44	25.0	10.4	56	281	1.9	0.290	0.026		19.7	10.6	48	28.5	8.9	174	222	34.0	**<0.001**	**0.134**[†]
EDE-Q: Restraint	4.3	1.5	116	4.5	1.6	51	3.7	1.9	31	3.0	2.0	46	244	8.3	**<0.001**	0.098	1 > 4, 2 > 4	2.5	2.0	49	4.4	1.4	185	234	60.4	**<0.001**	**0.207**[†]
EDE-Q: Eating concern	4.1	1.2	115	3.9	1.4	50	3.5	1.6	32	3.1	1.7	46	243	5.3	**0.005**	0.064	1 > 4	2.3	1.6	49	4.2	1.1	186	235	88.5	**<0.001**	**0.275**[†]
EDE-Q: Shape concern[a]	5.0	1.2	115	4.8	1.4	50	4.6	1.4	32	3.9	1.5	46	243	7.1	**<0.001**	0.084	1 > 4,2 > 4	2.9	1.2	49	5.1	1.1	186	235	172.1	**<0.001**	**0.420**[†]
EDE-Q weight concern[a]	4.7	1.3	115	4.5	1.4	50	4.2	1.7	32	3.4	1.6	46	243	9.3	**<0.001**	**0.108**[†]	1 > 4, 2 > 4	2.5	1.4	49	5.1	1.1	186	[a]35	265.9	**<0.001**	**0.463**[†]
BDQ: Total	11.2	5.6	125	11.0	5.6	60	10.3	5.6	43	11.0	5.2	59	287	0.2	0.890	0.003		9.3	6.0	48	11.4	5.3	180	228	5.7	**0.045**	0.025
Q_LES_Q_FF	35.0	9.8	103	37.0	9.7	50	36.6	13.3	36	37.4	10.2	44	233	0.8	0.638	0.012		41.0	12.1	42	34.5	9.3	145	187	14.1	**<0.001**	0.071
CFS	44.4	7.5	74	46.7	11.0	26	47.8	10.7	11	50.8	10.0	17	128	2.5	0.890	0.059		49.6	9.9	23	45.1	8.7	103	126	4.7	0.073	0.036
Pulse rate	68.0	14.0	76	72.2	16.8	39	70.4	21.4	30	65.1	12.9	36	181	1.0	0.527	0.024		67.8	14.6	30	68.9	15.8	102	132	0.1	0.805	0.001
Sodium count	145.3	41.9	62	140.5	2.8	43	130.1	31.1	28	131.2	26.1	36	169	1.6	0.286	0.040		137.3	5.4	30	140.0	36.4	90	120	0.2	0.759	0.001
Magnesium count	2.7	12.4	56	0.9	0.1	39	1.0	0.6	21	1.2	1.4	32	148	0.4	0.815	0.011		0.9	0.4	26	2.0	9.3	80	106	0.3	0.659	0.003
Potassium count	4.2	0.4	61	4.1	0.3	43	4.2	0.4	27	4.1	0.6	36	167	0.3	0.825	0.008		4.1	0.5	30	4.2	0.4	89	119	0.2	0.683	0.001
Phosphate count	1.2	0.2	52	1.2	0.2	38	1.4	0.2	19	1.2	0.2	30	139	2.6	0.116	0.077		1.2	0.2	25	2.9	9.8	72	97	1.0	0.771	0.007
Haemoglobin count	121.6	37.4	52	120.6	37.4	42	92.8	59.4	27	118.6	37.0	36	157	2.0	0.193	0.053		122.5	34.8	28	115.0	43.4	84	112	0.7	0.528	0.006
White blood cell	5.9	1.7	56	5.4	2.2	41	5.3	2.0	26	5.0	2.5	35	158	1.0	0.521	0.028		5.7	2.3	29	5.4	2.0	81	110	0.4	0.659	0.003

Note: Bold: significant comparison effect.

Abbreviations: DSM-5, Diagnostic and Statistical Manual of Mental Disorders; OWS, Overvaluation of Weight/shape; SD, Standard Deviation; η^2, Effect size for Analysis of Variance (ANOVA).

[a]When comparing no versus high overvaluation of weight/shape severity groups, relevant overvaluation items were excluded[a]from these subscales.

[†]Effect size in moderate-medium, and large-high range ($\eta^2 > 0.10$).

anxiety symptoms compared to those with a lower BMI (i.e., higher severity). While unexpected, this is not the first time such patterns have been reported. Specifically, the studies of Nakai et al. (2017) and Smith et al. (2017) reported a similar reserved effect. A possible explanation is that patients with AN who have a higher BMI might perceive themselves as not thin enough, hence being more concerned about their eating, weight, and shape, all of which are core putative factors for EDs (Fairburn et al., 2003; Nakai et al., 2017; Smith et al., 2017).

Despite electrolyte and haematologic abnormalities being common in malnutrition, in the current study the DSM-5 severity groups did not significantly differ across any of the measured biological variables. Our findings are consistent with previous studies, which found BMI to have weak correlations with white blood cell count (Misra et al., 2004; Polli et al., 2008), haemoglobin (Mamou et al., 2016; Misra et al., 2004), potassium (Imbierowicz et al., 2004), and magnesium levels (Raj et al., 2012). Furthermore, no differences in phosphate levels were observed between the DSM-5 AN severity groups despite a BMI of <16 kg/m^2 having been previously suggested to be associated with an increased risk of developing hypophosphatemia (NICE, 2006). Therefore, the current findings imply that BMI levels might not accurately reflect the biological status of individuals with AN, and that those who are at a higher risk for electrolytes and haematologic complications are not necessarily those with a low BMI. The lack of significant associations between the DSM-5 AN severity groups and the assessed biological markers might be due to the reliance on current BMI, omitting prior weight history. Particularly, recent studies have suggested that, in addition to absolute or present weight, weight suppression (i.e., the differences between highest and current weight) is crucial to consider in determining severity of AN (Bachmann et al., 2016; Berner et al., 2017). For example, despite falling into the mild BMI severity level, overweight people who develop AN through rapid weight loss may be medically unstable (Lebow et al., 2015). Therefore, it might be important for future studies to explore the role of weight suppression in indexing AN severity, especially when it comes to differentiating biological markers.

4.2 | Alternative severity rating for anorexia nervosa using overvaluation of weight and shape

This study assessed overvaluation of weight/shape as an alternative severity rating for AN. The results showed that overvaluation of weight/shape seems to be more robust in indexing the severity level of psychological variables in those with AN compared to the DSM-5 severity indicator based on BMI. Specifically, we found that while the high and low overvaluation of weight/shape AN groups did not differ significantly in age, age of onset and duration of illness, they differed significantly in BMI at assessment, with individuals in the low overvaluation of weight/shape groups having a higher current BMI.

Furthermore, individuals with AN in the high overvaluation of weight/shape group experienced significantly higher ED psychopathology, depressive, anxiety, and stress symptoms and had a poorer quality of life compared to individuals with AN in the low overvaluation of weight/shape group. The observed differences between the high and low overvaluation weight/shape AN groups on the psychological measures reflected moderate to large effect sizes.

These findings are consistent with Fairburn et al.'s (2003) transdiagnostic conceptualisation of ED symptoms, which specifies that overvaluation of weight/shape precipitates and maintains efforts of other ED symptoms. Overvaluation of shape and weight is thought to lead to dietary restraint, which subsequently reinforces and maintains other ED behaviours such as repeated checking of body shape and weight or avoidance behaviours, and the engagement in extreme weight control methods (e.g., purging). Engaging in ED behaviours further fosters overvaluation of shape and weight, resulting in a difficult-to-break ED cycle (Fairburn, 2008; Forrest et al., 2018; Murphy et al., 2010). Therefore, the current findings suggest that overvaluation of weight/shape may be more robust in indexing ED and general psychopathology than the DSM-5 severity ratings for AN.

Regarding the biological markers (i.e., electrolyte, haematologic and pulse rate), we found that the high and low overvaluation of weight/shape AN severity groups did not differ significantly in any of the assessed variables. This implies that despite overvaluation of weight/shape severity rating emerges to be better at indexing ED psychopathology than the DSM-5 severity index, both indices failed to reflect differences in biological factors associated with AN. This may suggest that there might be other factors contributing to the link between psychological and physical symptoms experienced by patients with AN. Future studies should seek out behavioural ED symptoms (e.g., purging, fluid restrictions) that could potentially lead to electrolyte and haematologic abnormalities.

4.3 | Clinical and theoretical implication

The current study has important implications for our understanding of the DSM-5 severity indicator based on

BMI and alternative severity ratings for AN. First, our study provided no support for the DSM-5 severity ratings for AN. Second, we also found that the BMI severity groups did not differ significantly in any of the interested biological markers, hence BMI may not capture the physical states of individuals with AN. Altogether, this suggests that the BMI severity classification system proposed by the DSM-5 may have no to limited utility and validity. Therefore, clinicians should consider a wider range of symptoms that an individual may present with (e.g., level of insight, pulse rate), rather than relying solely on BMI to denote disorder severity.

Third, the current findings indirectly undermine the validity of the newly introduced atypical anorexia nervosa (AAN) diagnostic classification in the DSM-5 within the Otherwise Specified Feeding and Eating Disorders (OSFED). Individuals with AAN are those who meet all diagnostic criteria for AN, except that despite significant weight loss, the patient's weight is within or above the normal range (APA, 2013). Since our findings suggest that low weight or low BMI does not accurately reflect the level of illness severity from both psychological and biological aspects, the utility of having AAN as a separate diagnostic classification is unclear. As such, it is concerning that some individuals with AAN are sometimes dismissed by healthcare practitioners as being clinically insignificant, even though research has revealed that the level of psychopathology and impairments of individuals with AAN is as severe as those with full threshold AN (Harrop et al., 2021). Overall, findings from the current study imply that the traditional view of a lower BMI equates with a more severe AN illness is inherently problematic. Conducting comprehensive psychological and medical assessments is therefore needed to inform treatment plans for individuals with AN and AAN.

Fourth, the current study is the first to examine the clinical utility of overvaluation of weight/shape as an alternative severity indicator for AN. We found that this alternative index appears to be clinically more useful at indexing psychological correlates (e.g., ED psychopathology, depression, stress, disability level or quality of life) of AN patients compared to the DSM-5 BMI severity rating. However, this alternative severity rating also failed to correlate with biological aspects that are frequently experienced by individuals with AN. Therefore, it is recommended that rather than solely relying on BMI to index AN severity, as suggested by the DSM-5, important cognitive ED symptoms (e.g., overvaluation of weight/shape, dietary restraint) should also be taken into consideration for diagnosing and assessing the level of severity of AN.

Fifth, since both the DSM-5 and the alternative severity index approach did not capture information regarding the differences in the biological data (electrolytes, haematological levels, and pulse rates) between the severity groups, this does not entirely rule out the usefulness of having a severity classification for these EDs. It suggests the importance of exploring alternative severity markers (e.g., weight suppression) in indexing AN severity, especially in distinguishing biological markers might be a useful avenue for future research.

4.4 | Limitations and future direction

The current study has some limitations. First, while the current study has addressed limitations of previous research by looking at a wider range of psychological variables and being the first to examine the relationships between severity ratings for AN and biological variables this study is limited by the usage of self-report measures and cross-sectional design. Such limitations prevent us from examining the predictive validity of both the DSM-5 and the alternative weight/shape overvaluation severity ratings for AN regarding prognostic outcomes, treatment efficacy, or crossover movements between ED diagnoses. The conclusion of the current study on the utility of the two classifying systems is therefore limited to the variables selected. The results of current findings need to be interpreted with caution and future replications are needed. Future research should consider longitudinal designs with additional clinical validators (e.g., treatment outcomes or indicators of cardiovascular health) to assist both researchers and clinicians to gain insight into the predictive validity of both the DSM-5 and overvaluation of weight and shape alternative severity ratings for AN.

Second, despite studies having suggested distinctive diagnostic characteristics (e.g., elevated emotion dysregulation in the AN-BP subtype; Racine & Wildes, 2013) and biological aspects (e.g., potassium levels are related with the frequency of purging behaviours; Koh et al., 1989), between the two AN subtypes (i.e., AN-BP and AN-R; Racine & Wildes, 2013), the current study did not have sufficient data to examine whether the DSM-5 and overvaluation of weight/shape severity ratings for AN differed between these two AN subtypes. Therefore, future research is warranted to focus on examining the clinical utility of different AN severity ratings for both AN-R and the AN-BP subtypes. Particularly, given that purging methods and frequency are associated with alterations in electrolytes and haematologic levels (Miller et al., 2005), it might be worth exploring the utility of both purging methods and frequency as a severity rating for the AN-BP subtype.

Third, predominantly individuals with enduring AN (i.e., mean duration of illness of 7.8 years) participated in

the current study. It has been shown that the levels of disturbances in both biological and psychological variables are moderated by the duration of the illness (Takakura et al., 2019). Specifically, individuals with AN with a duration of illness of more than 5 years to be more severe disturbances in both biological and psychological variables than those with a duration of illness of less than 5 years (Nakai et al., 2015; Takakura et al., 2019). This implies that the results of the current study might not be generalisable to those with a shorter duration of illness.

Fourth, the current research was conducted in a treatment-seeking facility in Australia with most participants being Caucasian females. Given the differences in ED psychopathology which exist between males and females (Murray et al., 2017), adolescents and adults (Rajagopalan et al., 2019), and individuals across cultures (Soh et al., 2006), current findings are not generalisable to other groups. Future research investigating the clinical utility of the DSM-5 and alternative ED severity ratings in the aforementioned populations is therefore needed.

5 | CONCLUSION

Overall, the current study found that the DSM-5 severity rating based on BMI for AN seems to have no clinical utility in differentiating ED symptomatology, general psychopathology, quality of life, disability level, and biological indicators (e.g., electrolyte levels). While severity ratings based on overvaluation of weight/shape were also not informative for biological variables, the current study provided initial support for this alternative severity rating system to be useful in indexing psychological correlates. Importantly, this study provides evidence for an alternative AN severity rating based on overvaluation of weight/shape that may have more clinical utility than severity based on BMI. Future studies should focus on exploring alternative severity markers for AN especially in distinguishing biological markers.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

An Binh Dang https://orcid.org/0000-0001-7961-6913
Zoe Jenkins https://orcid.org/0000-0002-8551-0792
Isabel Krug https://orcid.org/0000-0002-5275-3595

REFERENCES

American Psychiatric Association (APA). (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9781585629992

Assalone, C., Leonardi, L., Franceschi, R., Fumanelli, J., Maines, E., Marini, M., & Soffiati, M. (2022). Determinants of severe bradycardia in adolescents hospitalized for anorexia nervosa. *Pediatrics International*, *64*(1), e14967. https://doi.org/10.1111/ped.14967

Bachmann, K. N., Schorr, M., Bruno, A. G., Bredella, M. A., Lawson, E. A., Gill, C. M., Miller, K. K., Meenaghan, E., Gerweck, A. V., Slattery, M., Eddy, K. T., Ebrahimi, S., Koman, S. L., Greenblatt, J. M., Keane, R. J., Weigel, T., Misra, M., Bouxsein, M. L., & Klibanski, A. (2016). Vertebral volumetric bone density and strength are impaired in women with low-weight and atypical anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism*. https://doi.org/10.1210/jc.2016-2099

Bargiacchi, A., Clarke, J., Paulsen, A., & Leger, J. (2019). Refeeding in anorexia nervosa. *European Journal of Pediatrics*, *178*(3), 413–422. https://doi.org/10.1007/s00431-018-3295-7

Berner, L. A., Feig, E. H., Witt, A. A., & Lowe, M. R. (2017). Menstrual cycle loss and resumption among patients with anorexia nervosa spectrum eating disorders: Is relative or absolute weight more influential? *International Journal of Eating Disorders*, *50*(4), 442–446. https://doi.org/10.1002/eat.22697

Button, E. J., Chadalavada, B., & Palmer, R. L. (2010). Mortality and pre-dictors of death in a cohort of patients presenting to an eating disorders service. *International Journal of Eating Disorders*, *43*(5), 387–392. https://doi.org/10.1002/eat.20715

Caregaro, L., Di Pascoli, L., Favaro, A., Nardi, M., & Santonastaso, P. (2005). Sodium depletion and hemoconcentration: Overlooked complications in patients with anorexia nervosa? *Nutrition*, *21*(4), 438–445. https://doi.org/10.1016/j.nut.2004.08.022

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cooper, P. J., & Fairburn, C. G. (1993). Confusion over the core psychopathology of bulimia nervosa. *International Journal of Eating Disorders*, *13*(4), 385–389. https://doi.org/10.1002/1098-108X(199305)13:4<385::AID-EAT2260130406>3.0.CO;2-W

Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, *6*(1), 1–8. https://doi.org/10.1002/1098-108X(198701)6:1<1::AID-EAT2260060102>3.0.CO;2–9

Dakanalis, A., Timko, C., Colmegna, F., Riva, G., & Clerici, M. (2018). Evaluation of the DSM-5 severity ratings for anorexia nervosa in a clinical sample. *Psychiatry Research*, *262*, 124–128. https://doi.org/10.1016/j.psychres.2018.02.009

Dalle Grave, R., Sartirana, M., El Ghoch, M., & Calugi, S. (2018). DSM-5 severity specifiers for anorexia nervosa and treatment outcomes in adult females. *Eating Behaviors*, *31*, 18–23. https://doi.org/10.1016/j.eatbeh.2018.07.006

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders. *Clinical Psychology: Science and Practice*, *29*(4), 325–344. https://doi.org/10.1037/cps0000078

Devuyst, O., Lambert, M., Rodhain, J., Lefebvre, C., & Coche, E. (1993). Haematological changes and infectious complications in anorexia nervosa: A case-control study. *QJM: International Journal of Medicine*, *86*(12), 791–799. https://doi.org/10.1093/oxfordjournals.qjmed.a068762

Fagundo, A. B., De la Torre, R., Jiménez-Murcia, S., Agüera, Z., Granero, R., Tárrega, S., & Fernández-Aranda, F. (2012). Executive functions profile in extreme eating/weight conditions: From anorexia nervosa to obesity. https://doi.org/10.1371/journal.pone.0043382

Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363–370. https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, *41*(5), 509–528. https://doi.org/10.1016/S0005-7967(02)00088-8

Flament, M. F., Godart, N. T., Fermanian, J., & Jeammet, P. (2001). Predictive factors of social disability in patients with eating disorders. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *6*(2), 99–107. https://doi.org/10.1007/BF03339758

Forney, K. J., Buchman-Schmitt, J. M., Keel, P. K., & Frank, G. K. (2016). The medical complications associated with purging. *International Journal of Eating Disorders*, *49*(3), 249–259. https://doi.org/10.1002/eat.22504

Forrest, L. N., Jones, P. J., Ortiz, S. N., & Smith, A. R. (2018). Core psychopathology in anorexia nervosa and bulimia nervosa: A network analysis. *International Journal of Eating Disorders*, *51*(7), 668–679. https://doi.org/10.1002/eat.22871

Gallagher, D., Parker, A., Samavat, H., & Zelig, R. (2021). Prophylactic supplementation of phosphate, magnesium, and potassium for the prevention of refeeding syndrome in hospitalized individuals with anorexia nervosa. *Nutrition in Clinical Practice*, *37*(2), 328–343. https://doi.org/10.1002/ncp.10786

Gianini, L., Roberto, C. A., Attia, E., Walsh, B. T., Thomas, J. J., Eddy, K. T., Grilo, C. M., Weigel, T., & Sysko, R. (2017). Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers. *International Journal of Eating Disorders*, *50*(8), 906–916. https://doi.org/10.1002/eat.22728

Gluschkoff, K., Jokela, M., & Rosenström, T. (2019). The general psychopathology factor: Structural Stability and Generalizability to within-individual changes. *Frontiers in Psychiatry*, *10*, 594. https://doi.org/10.3389/fpsyt.2019.00594

Goddard, E., Hibbs, R., Raenker, S., Salerno, L., Arcelus, J., Boughton, N., Connan, F., Goss, K., Laszlo, B., Morgan, J., Moore, K., Robertson, D., Saeidi, S., Schreiber-Kounine, C., Sharma, S., Whitehead, L., Schmidt, U., & Treasure, J. (2013). A multi-centre cohort study of short term outcomes of hospital treatment for anorexia nervosa in the UK. *BMC Psychiatry*, *13*, 287. https://doi.org/10.1186/1471-244X-13-287

Grilo, C. M., Ivezaj, V., & White, M. A. (2015). Evaluation of the DSM-5 severity indicator for binge eating disorder in a community sample. *Behaviour Research and Therapy*, *66*, 72–76. https://doi.org/10.1016/j.brat.2015.05.003

Harrop, E. N., Mensinger, J. L., Moore, M., & Lindhorst, T. (2021). Restrictive eating disorders in higher weight persons: A systematic review of atypical anorexia nervosa prevalence and consecutive admission literature. *International Journal of Eating Disorders*, *54*(8), 1328–1357. https://doi.org/10.1002/eat.23519

Hebebrand, J., Himmelmann, G. W., Heseker, H., Schäfer, H., & Remschmidt, H. (1996). Use of percentiles for the body mass index in anorexia nervosa: Diagnostic, epidemiological, and therapeutic considerations. *International Journal of Eating Disorders*, *19*(4), 359–369. https://doi.org/10.1002/(SICI)1098-108X(199605)19:4<359::AID-EAT4>3.0.CO;2-K

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, *44*(2), 227–239. https://doi.org/10.1348/014466505X29657

Hope, M. L., Page, A. C., & Hooke, G. R. (2009). The value of adding the Quality of Life Enjoyment and Satisfaction Questionnaire to outcome assessments of psychiatric inpatients with mood and affective disorders. *Quality of Life Research*, *18*(5), 647–655. https://doi.org/10.1007/s11136-009-9478-y

Imbierowicz, K., Curkovic, I., Braks, K., Geiser, F., Liedtke, R., & Jacoby, G. E. (2004). Effect of weight-regulating practices on potassium level in patients with anorexia or bulimia nervosa. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, *12*(5), 300–306. https://doi.org/10.1002/erv.591

Jacobs, M., & Schneider, J. (1985). Medical complications of bulimia: A prospective evaluation. *QJM: International Journal of Medicine*, *54*(2), 177–182. https://doi.org/10.1093/oxfordjournals.qjmed.a067839

Koh, E., Onishi, T., Morimoto, S., Imanaka, S., Nakagawa, H., & Ogihara, T. (1989). Clinical evaluation of hypokalemia in anorexia nervosa. *Japanese Journal of Medicine*, *28*(6), 692–696. https://doi.org/10.2169/internalmedicine1962.28.692

Krug, I., Binh Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Menchon, J. M., Fernandez-Aranda, F., & Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the DSM-5 severity indices for eating disorders. *European Eating Disorders Review*, *29*(3), 482–498. https://doi.org/10.1002/erv.2776

Krug, I., Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2022). How to assess severity in male with eating disorders? The DSM–5 se-verity index versus severity based on drive for thinness. Manuscript in preparation.

Lebow, J., Sim, L. A., & Kransdorf, L. N. (2015). Prevalence of a history of overweight and obesity in adolescents with restrictive eating disorders. *Journal of Adolescent Health*, *56*(1), 19–24. https://doi.org/10.1016/j.jadohealth.2014.06.005

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, *28*(4), 612–625. https://doi.org/10.1111/j.1468-2958.2002.tb00828.x

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Machado, P. P., Grilo, C. M., & Crosby, R. D. (2017). Evaluation of the DSM-5 severity indicator for anorexia nervosa. *European*

Eating Disorders Review, *25*(3), 221–223. https://doi.org/10.1002/erv.2508

Mamou, G., Sider, A., Bouscary, D., Moro, M. R., & Blanchet-Collet, C. (2016). Anemia in anorexia nervosa: The best way to deal with it—an overview of literature. *J Hum Nutr Food Sci*, *4*(1), 1081.

Martin, M. M., & Rubin, R. B. (1995). A new measure of cognitive flexibility. *Psychological Reports*, *76*(2), 623–626. https://doi.org/10.2466/pr0.1995.76.2.623

Mehler, P. S., & Brown, C. (2015). Anorexia nervosa – medical complications. *J Eat Disord*, *3*, 11. https://doi.org/10.1186/s40337-015-0040-8

Mehler, P. S., Blalock, D. V., Walden, K., Kaur, S., McBride, J., Walsh, K., & Watts, J. (2018). Medical findings in 1,026 consecutive adult inpatient–residential eating disordered patients. *International Journal of Eating Disorders*, *51*(4), 305–313. https://doi.org/10.1002/eat.22830

Miller, K. K., Grinspoon, S. K., Ciampa, J., Hier, J., Herzog, D., & Klibanski, A. (2005). Medical findings in outpatients with anorexia nervosa. *Archives of Internal Medicine*, *165*(5), 561–566. https://doi.org/10.1001/archinte.165.5.561

Misra, M., Aggarwal, A., Miller, K. K., Almazan, C., Worley, M., Soyka, L. A., Herzog, D. B., & Klibanski, A. (2004). Effects of anorexia nervosa on clinical, hematologic, biochemical, and bone density parameters in community-dwelling adolescent girls. *Pediatrics*, *114*(6), 1574–1583. https://doi.org/10.1542/peds.2004-0540

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*(5), 551–567. https://doi.org/10.1016/S0005-7967(03)00161-X

Mond, J. M., Owen, C., Hay, P. J., Rodgers, B., & Beumont, P. J. V. (2005). Assessing quality of life in eating disorder patients. *Quality of Life Research*, *14*, 171–178. https://doi.org/10.1007/s11136-004-2657-y

Murphy, R., Straebler, S., Cooper, Z., & Fairburn, C. G. (2010). Cognitive behavioral therapy for eating disorders. *Psychiatric Clinics*, *33*(3), 611–627. https://doi.org/10.1016/j.psc.2010.04.004

Murray, S. B., Nagata, J. M., Griffiths, S., Calzo, J. P., Brown, T. A., Mitchison, D., Blashill, A. J., & Mond, J. M. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, *57*, 1–11. https://doi.org/10.1016/j.cpr.2017.08.001

Naismith, S. L., Longley, W. A., Scott, E. M., & Hickie, I. B. (2007). Disability in major depression related to self-rated and objectively-measured cognitive deficits: A preliminary study. *BMC Psychiatry*, *7*(1), 1–7. https://doi.org/10.1186/1471-244X-7-32

Nakai, Y., Fujita, M., Nin, K., Noma, S., & Teramukai, S. (2015). Relationship between duration of illness and cardiac autonomic nervous activity in anorexia nervosa. *BioPsychoSocial Medicine*, *9*(12), 12. https://doi.org/10.1186/s13030-015-0032-6

Nakai, Y., Nin, K., Noma, S., Teramukai, S., Fujikawa, K., Wonderlich, S. A. i., Teramukai, S., Fujikawa, K., & Wonderlich, S. A. (2017). The impact of DSM-5 on the diagnosis and severity indicator of eating disorders in a treatment-seeking sample. *International Journal of Eating Disorders*, *50*(11), 1247–1254. https://doi.org/10.1002/eat.22777

Nakazato, M., Hashimoto, K., Schmidt, U., Tchanturia, K., Campbell, I. C., Collier, D. A., Iyo, M., & Treasure, J. (2010). Serum glutamine, set-shifting ability and anorexia nervosa. *Annals of General Psychiatry*, *9*(1), 29. https://doi.org/10.1186/1744-859X-9-29

National Institute for Health and Clinical Excellence. (2006). Nutrition support in adults Clinical guideline CG32. Retrieved from www.nice.org.uk/page.aspx?o=cg032

Ormel, J., VonKorff, M., Ustun, T. B., Pini, S., Korten, A., & Oldehinkel, T. (1994). Common mental disorders and disability across cultures: Results from the WHO collaborative study on psychological problems in general health care. *JAMA*, *272*(22), 1741–1748. https://doi.org/10.1001/jama.1994.03520220035028

Ornstein, R. M., Golden, N. H., Jacobson, M. S., & Shenker, I. R. (2003). Hypophosphatemia during nutritional rehabilitation in anorexia nervosa: Implications for refeeding and monitoring. *Journal of Adolescent Health*, *32*(1), 83–88. https://doi.org/10.1016/S1054-139X(02)00456-1

Oudman, E., Wijnia, J. W., Oey, M. J., van Dam, M. J., M. J., & Postma, A. (2018). Preventing Wernicke's encephalopathy in anorexia nervosa: A systematic review. *Psychiatry and Clinical Neurosciences*, *72*(10), 774–779. https://doi.org/10.1111/pcn.12735

Page, A. C., Hooke, G. R., & Morrison, D. L. (2007). Psychometric properties of the depression anxiety stress scales (DASS) in depressed clinical samples. *British Journal of Clinical Psychology*, *46*(3), 283–297. https://doi.org/10.1016/S0005-7967(96)00068-X

Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., Bardone-Cone, A. M., Klein, M., & Le Grange, D. (2007). Psychometric properties of the eating disorder examination-questionnaire: Factor structure and internal consistency. *International Journal of Eating Disorders*, *40*(4), 386–389. https://doi.org/10.1002/eat.20373

Polli, N., Scacchi, M., Giraldi, F. P., Sormani, M., Zappulli, D., & Cavagnini, F. (2008). Low insulin-like growth factor I and leukopenia in anorexia nervosa. *International Journal of Eating Disorders*, *41*(4), 355–359. https://doi.org/10.1002/eat.20506

Pomeroy, C. (2001). Anorexia nervosa, bulimia nervosa, and binge eating disorder: Assessment of physical status. https://doi.org/10.1037/10502-008

Racine, S. E., & Wildes, J. E. (2013). Emotion dysregulation and symptoms of anorexia nervosa: The unique roles of lack of emotional awareness and impulse control difficulties when upset. *International Journal of Eating Disorders*, *46*(7), 713–720. https://doi.org/10.1002/eat.22145

Raj, K. S., Keane-Miller, C., & Golden, N. H. (2012). Hypomagnesemia in adolescents with eating disorders hospitalized for medical instability. *Nutrition in Clinical Practice*, *27*(5), 689–694. https://doi.org/10.1177/0884533612446799

Rajagopalan, A., Bhattacharya, A., Peebles, R., Dakanalis, A., & Timko, C. A. (2019). 1.10 a literature review: Do Dsm-5 severity indicators for eating disorders apply to children and adolescents? *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(10 Suppl), S149–S150. https://doi.org/10.1016/j.jaac.2019.08.032

Reas, D. L., Grilo, C. M., & Masheb, R. M. (2006). Reliability of the Eating Disorder Examination-Questionnaire in patients with binge eating disorder. *Behaviour Research and Therapy*, *44*(1), 43–51. https://doi.org/10.1016/j.brat.2005.01.004

Ritsner, M., Kurs, R., Kostizky, H., Ponizovsky, A., & Modai, I. (2002). Subjective quality of life in severely mentally ill patients: A comparison of two instruments. *Quality of Life Research*, *11*(6), 553–561. https://doi.org/10.1023/A:1016323009671

Schechter, D., Endicott, J., & Nee, J. (2007). Quality of life of 'normal'controls: Association with lifetime history of mental illness. *Psychiatry Research*, *152*(1), 45–54. https://doi.org/10.1016/j.psychres.2006.09.008

Smith, K. E., Ellison, J. M., Crosby, R. D., Engel, S. G., Mitchell, J. E., Crow, S. J., Peterson, C. B., Le Grange, D., & Wonderlich, S. A. (2017). The validity of DSM-5 severity specifiers for anorexia nervosa, bulimia nervosa, and binge-eating disorder. *International Journal of Eating Disorders*, *50*(9), 1109–1113. https://doi.org/10.1002/eat.22739

Smith, K. E., Ellison, J. M., Crosby, R. D., Engel, S. G., Mitchell, J. E., Crow, S. J., Peterson, C. B., Grange, D. L., & Wonderlich, S. A. (2017). The validity of DSM-5 severity specifiers for anorexia nervosa, bulimia nervosa, and binge-eating disorder. *International Journal of Eating Disorders*, *50*(9), 1109–1113. http://doi.org/10.1002/eat.22739

Soh, N. L., Touyz, S. W., & Surgenor, L. J. (2006). Eating and body image disturbances across cultures: A review. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, *14*(1), 54–65. https://doi.org/10.1002/erv.678

Sullivan, P. F. (1995). Mortality in anorexia nervosa. *American Journal of Psychiatry*, *152*(7), 1073–1074.

Sy, R., Ponton, K., De Marco, P., Pi, S., & IsHak, W. W. (2013). Quality of life in anorexia nervosa: A review of the literature. *Eating Disorders*, *21*(3), 206–222. https://doi.org/10.1080/10640266.2013.779176

Takakura, S., Aso, C. S., Toda, K., Hata, T., Yamashita, M., & Sudo, N. (2019). Physical and psychological aspects of anorexia nervosa based on duration of illness: A cross-sectional study. *BioPsychoSocial Medicine*, *13*(1), 32. https://doi.org/10.1186/s13030-019-0173-0

Walsh, B. T. (2019). Diagnostic categories for eating disorders: Current status and what lies ahead. *Psychiatric Clinics*, *42*(1), 1–10. https://doi.org/10.1016/j.psc.2018.10.001

Ware, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, *30*(6), 473–483. https://doi.org/10.1097/00005650-199206000-00002

Whitelaw, M., Lee, K. J., Gilbertson, H., & Sawyer, S. M. (2018). Predictors of complications in anorexia nervosa and atypical anorexia nervosa: Degree of underweight or extent and recency of weight loss? *Journal of Adolescent Health*, *63*(6), 717–723. https://doi.org/10.1016/j.jadohealth.2018.08.019

Wolfi, G., & Treasure, I. (2011). Medical and psychological consequences of eating disorders. *ABC of Eating Disorders*, *20*.

Zayas, L. V., Wang, S. B., Coniglio, K., Becker, K., Murray, H. B., Klosterman, E., Bean, P., Weltzin, T., Franko, D. L., Eddy, K. T., & Thomas, J. J. (2018). Gender differences in eating disorder psychopathology across DSM-5 severity categories of anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, *51*(9), 1098–1102. https://doi.org/10.1002/eat.22941

Zipfel, S., Giel, K. E., Bulik, C. M., Hay, P., & Schmidt, U. (2015). Anorexia nervosa: Aetiology, assessment, and treatment. *The Lancet Psychiatry*, *2*(12), 1099–1111. https://doi.org/10.1016/S2215-0366(15)00356-9

How to cite this article: Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2023). Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates? *European Eating Disorders Review*, *31*(4), 447–461. https://doi.org/10.1002/erv.2969

5.3 Additional analyses. Assessing ICD-11 severity in anorexia nervosa: Do the ICD-11 AN severity groups differ in psychological and biological correlates?

This chapter builds on Study 3a to explore the utility of the ICD-11 framework in distinguishing AN severity in adult populations and its ability to predict psychological and biological variables. It is formatted as a short manuscript, following the structure of a typical research paper. However, to avoid redundancy, only new information is included, with overlapping content from Study 3a omitted.

Introduction

In 2019, the ICD-11 introduced severity ratings for AN based-on BMI. These ratings delineate two severity categories: "significantly low BMI" (≥14 kg/m²) and "dangerously low BMI" (<14 kg/m²). The threshold of 14 kg/m² was designed to improve diagnostic precision by recognising the profound impact of extreme underweight on AN prognosis and the corresponding intensity of required treatment (Uher & Rutter, 2013).

Despite its introduction in 2019, research on the clinical utility of the ICD-11 AN severity framework remains limited. To date, aside from Studies 3b and 5 (discussed later in this chapter and Chapter 7), only one study has specifically investigated the clinical utility of the ICD-11 AN severity criteria. Engelhardt et al. (2021) examined a cohort of 469 children and adolescents with AN and found that individuals classified in the “dangerously low BMI” ICD-11 severity category had longer treatment durations and lower BMIs at discharge compared to those in the “significantly low BMI” group. However, according to Engelhardt et al. (2021) these differences were insufficient to guide clinical decisions regarding risk assessment or treatment stratification. Additionally, no studies have directly evaluated the ICD-11 severity rating for AN in adult populations, particularly its ability to predict psychological and biological variables.

The primary aim of Chapter 5 is to evaluate the clinical utility of three severity frameworks for AN: DSM-5, ICD-11, and OWS. As previously mentioned, while Study 3a compared DSM-5 and OWS severity ratings, it did not include ICD-11 severity ratings because the ICD-11 framework was published after the study’s completion. To address this gap, this section presents additional analyses conducted on the same sample to examine how the ICD-11 AN BMI-based severity rating relates to psychological and biological variables. Specifically, it investigates whether individuals with AN in the ICD-11 “significantly low BMI” severity group differ significantly from those in the “dangerously low BMI” severity

group across psychological variables (e.g., eating disorder psychopathology, disability, and symptoms of depression, anxiety, or stress) and biological variables (e.g., potassium and magnesium levels).

Method

This section employed the same methodology outlined in Study 3a, with identical details regarding participants, measures, and procedures. Further information regarding the ICD-11 severity groups and statistical analyses is included below.

ICD-11 Severity Groups Based on BMI

Participants were classified into ICD-11 severity subgroups using their BMI at the time of assessment. Based on the ICD-11 definitions, 284 participants (91%) were classified as having a "significantly low BMI" , while 28 participants (9%) were classified as having a "dangerously low BMI".

Statistical analyses

To compare ICD-11 AN severity groups and their ability to distinguish outcome variables, we used chi-square tests (χ^2) for categorical variables and ANOVA for continuous dependent variables. All tests were two-tailed, with an alpha level set at .05. To adjust for multiple comparisons, the False Discovery Rate procedure was applied.

The discriminative capacity between ICD-11, DSM-5 BMI, and OWS severity groups (Study 3a, Section 5.3) was assessed using both the *p*-values from significant tests and effect size coefficients. Cramer's V (V) was used for categorical measures, with values of 0.06, 0.15, and 0.30 representing small, medium, and large effect sizes, respectively (Cohen, 1988). For ANOVA, partial eta-squared (η^2) was calculated, with values of 0.06, 0.10, and 0.25 indicating small, medium, and large effect sizes (Levine & Hullett, 2002). Additionally, as not all participants completed every assessment, each analysis was based on the available data, treating incomplete data as missing.

Results

Demographics and Clinical Characteristics

Table 14 presents the demographic characteristics of the two ICD-11 AN severity groups, highlighting significant differences in ethnicity, employment and BMI at discharge. A higher proportion of individuals in the "significantly low BMI" AN severity group identified as Caucasian compared to the "dangerously low BMI" AN severity group, beyond what would be expected by chance. Conversely, a greater proportion of individuals in the "dangerously low BMI" group were engaged in home duties compared to those in the "significantly low BMI" group, exceeding chance expectations. Significant differences were also found in BMI at discharge, with individuals in the "significantly low BMI" severity group having a higher BMI ($p < .001$, $\eta^2 = .186$) compared to those in the "dangerously low BMI" AN severity group.

Table 14.

Demographic Information for the ICD-11 Anorexia Nervosa Severity Groups

		ICD-11 AN severity groups				
		Significantly low BMI ≥14.0 kg/m²		**Dangerously Low BMI <14 kg/m2**		**P (V)**
Gender	Male	16	5.6%	1	3.6%	.537(.026)
	Female	268	94.4%	27	96.4%	
Marital status	Single	214	75.6%	23	82.1%	.611(.107)
	Widowed	2	0.7%	0	0%	
	Divorced/Separated	15	5.3%	1	3.6%	
	Married - in couple	34	12.0%	1	3.6%	
	Other/Unknown	18	6.4%	3	7.0%	
Education		82	29.2%	5	18.5%	.452(.148)
	Secondary/lower					
	Vocational	6	2.1%	1	3.7%	
	University	166	59.1%	19	70.3%	
	Other/Unknown	8	2.8%	0	0%	
Ethnicity	Aboriginal	7	2.7%	0	0%	**<.001(.301†)**
	Caucasian	241	92.4%	20	83.3%	
	Asian	9	3.5%	1	4.2%	
	Other/Unknown	4	1.5%	3	12.5%	
Employment	Unemployed	29	10.6%	0	0%	**.040 (.210†)**
	Student	92	33.6%	9	36.9%	
	Employed	76	27.8%	4	16.0%	
	Home duties	7	2.6%	3	12.0%	
	Unable to work/Other	70	25.5%	9	36.0%	

	Significantly low BMI ≥14.0 kg/m²		Dangerously Low BMI <14 kg/m²		
	Mean	S.D	Mean	S.D	P (η^2)
Age (yrs-old)	26.68	9.22	28.46	14.03	.327 (.003)
Onset ED (yrs-old)	19.4	6.09	17.62	6.01	.770 (.000)
Duration ED (yrs)	9.07	8.63	11.32	11.56	.116 (.004)
BMI at assessment (kg/m²)	17.15	2.22	12.95	0.68	**<.001 (.241†)**
BMI at discharge (kg/m²)	17.77	2.39	13.88	1.84	**.<.001 (.186†)**

Note:
ICD-11: International Classification of Diseases 11th edition
V = Cramer's V – effect size for Chi-square statistic.
η^2 = Effect size for ANOVA.
*Bold: significant comparison effect †effect size in moderate-medium, and large-high range (η^2>0.10).

ICD-11 severity categories in indexing outcomes variables

Table 15 summarises the descriptive statistics and results of the statistical analyses for the two ICD-11 AN severity groups across the assessed psychological and biological variables.

Table 15.

Clinical Profiles across the Anorexia Nervosa ICD-11 Severity Groups Based BMI

	ICD-11 AN severity groups									
Anorexia nervosa	**Significantly low BMI ≥14.0 kg/m²**			**Dangerously Low BMI <14 kg/m²**				**Effect size**		
	Mean	SD	*n*	Mean	SD	*n*	*N*	*F*	*p*	η^2
DASS-21: Depression	25.96	12.19	256	18.96	14.39	25	281	7.52	**.006**	.026
DASS-21: Anxiety	19.38	11.11	256	11.40	11.47	25	281	7.09	**.008**	.025
DASS-21: Stress	26.55	22.30	256	20.48	10.26	25	281	8.31	**.004**	.029
EDE-Q: Restraint	4.11	1.79	228	2.75	2.02	16	244	9.66	**.002**	.038
EDE-Q: Eating concern	3.81	1.51	227	2.96	1.61	16	243	5.07	**.025**	.021
EDE-Q: Shape concern	4.79	1.40	227	3.75	1.85	16	243	12.58	**<.001**	.050
EDE-Q Weight concern	4.79	1.54	227	3.57	1.51	16	243	13.08	**<.001**	.051
CFS	46.79	8.65	122	47.31	9.51	6	128	3.21	.053	.008
Pulse Rate	72.00	17.68	164	70.18	19.51	17	181	0.1	.602	.001
Sodium Count	139.44	21.25	155	137.88	6.61	14	169	0.2	.720	.001
Magnesium Count	1.1	1.26	134	1.32	5.51	14	148	0.3	.864	.001
Potassium Count	4.79	5.05	129	4.10	0.59	14	143	0.2	.505	.001
Phosphate Count	2.29	2.29	145	1.19	0.22	22	159	1.0	.537	.001
Haemoglobin Count	110.83	46.86	145	106.89	45.24	14	159	0.7	.692	.001
White blood cell	6.09	4.64	144	4.65	2.35	14	158	0.4	.131	.007

Note.
ICD-11: International Classification of Diseases 11th edition
SD = Standard Deviation.
η^2 = Effect size for ANOVA.
*Bold: significant comparison effect

Psychological variables. Individuals in the ICD-11 "significantly low BMI" group reported significantly higher levels of both general psychopathology (measured via the Depression Anxiety Stress Scale – 21 ([DASS-21]; Lovibond & Lovibond, 1995), and ED psychopathology (measured via the EDE-Q) than those in the “dangerously low BMI” AN severity group. Specifically, those in the "significantly low BMI” group reported higher levels of depressive (p = .006, η^2 = .026), stress (p = .008, η^2 = .025), and anxiety (p = .004, η^2 = .029) symptoms, as well as greater dietary restraint (p = .002, η^2 = .038), and higher concerns related to eating (p = .025, η^2 = .021), weight (p < .001, η^2 = .050), and shape (p < .001, η^2 = .051), compared to the "dangerously low BMI" group, with small effect sizes.

Biological variables. For the biological data, no statistically significant differences were found between the two ICD-11 BMI severity groups for levels of potassium, magnesium, phosphate, white blood cells, haemoglobin, or pulse rate.

Discussion

The current study is the first to examine the clinical utility of the ICD-11 AN severity rating system based on BMI on a wide range of biological and psychological variables. Specifically, we looked at the ICD-11 severity rating on its ability to distinguish psychological and biological indicators that are commonly observed in patients with AN. In terms of the psychological variables, we found those in the milder ICD-11 severity group (i.e., “significantly low BMI”) reported higher level of ED and general psychopathology than those in the more severe (i.e., “dangerously low BMI”) severity group. However, the ICD-11 severity groups did not differ in any of the assessed biological correlates.

ICD-11 Severity Groups and Psychological and Biological Variables

Our findings revealed that, while the ICD-11 severity groups significantly differed in mean age and BMI at discharge, they were indistinguishable in illness duration and age of illness onset. Additionally, our analysis demonstrated limited significant results among

psychological variables, and even these findings showed small effect sizes. Paradoxically, individuals in the “significantly low BMI” category (higher BMI and lower severity) exhibited higher levels of ED and general psychopathology compared to those in the “dangerously low BMI” group. This pattern contradicts expectations of a severity measure, which would ideally associate greater severity with greater psychopathology.

Although no previous studies have specifically explored the relationship between ICD-11 AN severity groups and ED and general psychopathology, similar results have been observed in research on DSM-5 severity classifications for AN, which also rely on BMI. For instance, studies by Nakai et al. (2017) and Smith et al. (2017) reported a similar inverse pattern, with those with a higher BMI reported a higher level of ED psychopathology. A potential explanation for this counterintuitive finding lies in differences in insight between severity groups. Research has shown that individuals with severe AN (e.g., those in the ‘dangerously low BMI’ category) may have less insight into the seriousness of their condition, leading to an underreporting of symptoms (Webb et al., 2022). Conversely, individuals in milder severity groups may have greater awareness of their illness, amplifying distress related to eating behaviours and body image (Konstantakopoulos et al., 2011). This heightened insight might also lead to feelings of inadequacy about their weight or appearance, increasing dietary restraint and anxiety. Furthermore, shorter illness duration, often associated with milder severity, may result in greater familial or societal pressure to seek early treatment, further exacerbating anxiety and ED-related concerns (Hebebrand et al., 2024).

Regarding biological variables, our results showed no significant differences between ICD-11 severity groups across any measured markers, including electrolyte and hematologic parameters such as potassium, magnesium, white blood cell count, and haemoglobin. This aligns with prior research demonstrating weak correlations between BMI and aforementioned

variables (e.g., Mamou et al., 2016; Misra et al., 2004; Polli et al., 2008; Raj et al., 201). Although the ICD-11 threshold of 14 kg/m² was intended to improve diagnostic precision and account for the impact of extreme underweight on prognosis and treatment intensity (Uher & Rutter, 2013), our findings suggest that it does not operate as intended in distinguishing biological risk.

Clinical implication

The findings of this study carry significant clinical implications for assessing severity in AN. Specifically, the results challenge the utility of BMI-based severity classifications in both the DSM-5 and ICD-11. While findings from Study 3a raised concerns about the DSM-5 BMI AN severity rating, this study extends those concerns to the ICD-11 AN severity rating system, which employs different BMI cut-offs. Notably, BMI-based severity groups in both systems did not differ significantly in key biological markers, suggesting that BMI alone may not sufficiently capture the biological variables correlated with physical state of individuals with AN. This is particularly surprising given prior evidence supporting the clinical utility of dangerously low BMI levels (e.g., BMI <14 kg/m² in the ICD-11) for guiding medical interventions, including hospital admission (Hay et al., 2014). However, it is worth noting that Hay and colleagues (2014) also pointed out additional indicators of high medical risk – such as heart rate < 40 bpm, hypokalaemia, hypoglycaemia, and hypophosphatemia – highlighting the importance of a multifaceted approach.

These findings suggest that BMI-based severity classifications in the DSM-5 and ICD-11, when used in isolation, may have limited validity and clinical utility. As this is the first study to examine these classification systems in this context, further research is needed to replicate and validate these results. In the meantime, clinicians are encouraged to take a more comprehensive approach to assessing AN severity, incorporating a broader range of

symptoms – such as insight, vital signs, and other physiological indicators – rather than relying solely on BMI.

Limitations and future directions

Besides the limitations discussed in Study 3a, an additional limitation to acknowledge is the reliance on a self-reported measure of ED psychopathology (i.e., the EDE-Q). While the EDE-Q is a validated tool, it may not provide the most accurate representation of ED psychopathology in a sample of treatment-seeking individuals with AN. Research has shown that lower BMI is associated with poorer insight into the disorder among individuals with AN (Konstantakopoulos et al., 2011), which can undermine the reliability of self-reports, especially concerning behavioural and psychological symptoms. To improve accuracy, future research should consider cross-validating self-reported measures with clinician-rated assessments or structured clinical interviews for a more comprehensive evaluation of eating disorder and general psychopathology.

Integrated Conclusion from Study 3a and Section 5.3

Overall, findings from Studies 3a and Section 5.3 underscore key distinctions in the utility of severity classification systems for AN. While both AN severity ratings per DSM-5 and ICD-11 rely on BMI-based criteria, their focus on physical metric limits their capacity to accurately reflect the psychological severity of the disorder. Paradoxically, both systems classify individuals in milder severity categories as having higher ED psychopathology and comorbidities, highlighting their inadequacy in capturing the psychological burden of AN.

In contrast, the OWS severity classification system demonstrated stronger alignment with the conceptual purpose of a severity rating, showing significant positive associations with ED-specific and general psychopathology. Higher levels of ED and general psychopathology were observed in individuals classified in the more severe OWS group (i.e., high OWS). However, none of the three frameworks – DSM-5, ICD-11, or OWS –

effectively captured variations in biological variables, indicating the need for further refinement to comprehensively assess severity.

These findings must be interpreted in light of the limitations outlined in both Study 3a and this section (Section 5.3). Despite these limitations, the results provide valuable insights into the relative strengths and weaknesses of current severity classification systems, particularly the promise of the OWS framework in indexing psychological severity in AN.

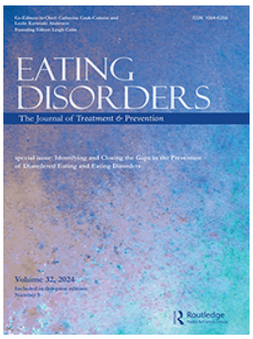
5.4 Study 3b. Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings

Study 3b was published in *Journal of Treatment and Prevention (Eating Disorder)* in May 2024

Dang, A. B., Kiropoulos, L., Castle, D., Jenkins, Z., Phillipou, A., Rossell, S., & Krug, I. (2024). Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings. *Eating Disorders*, 1-17. https://doi.org/10.1080/10640266.2024.2346001

Contribution to the Study

Author	Task contribution
An Dang	Conceptualisation; manuscript writing; editing; data analyses
Litza Kiropoulos	Editing
David Castle	Data collection, editing
Zoey Jenkins	Data collection, editing
Andrea Phillipou	Data collection, editing
Rossell Sussan	Data collection, editing
Isabel Krug	Conceptualisation; manuscript – reviewing and editing; supervision



Routledge
Taylor & Francis Group

Eating Disorders
The Journal of Treatment & Prevention

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/uedi20

Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings

An Binh Dang, Litza Kiropoulos, David Castle, Zoe Jenkins, Andrea Phillipou, Susan Rossell & Isabel Krug

To cite this article: An Binh Dang, Litza Kiropoulos, David Castle, Zoe Jenkins, Andrea Phillipou, Susan Rossell & Isabel Krug (2024) Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings, Eating Disorders, 32:5, 546-562, DOI: 10.1080/10640266.2024.2346001

To link to this article: https://doi.org/10.1080/10640266.2024.2346001

Published online: 06 May 2024.

EATING DISORDERS
2024, VOL. 32, NO. 5, 546–562
https://doi.org/10.1080/10640266.2024.2346001

Routledge
Taylor & Francis Group

Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings

An Binh Dang[a], Litza Kiropoulos[a], David Castle[b,c], Zoe Jenkins[d,e,f], Andrea Phillipou[g,h,i,j], Susan Rossell[d,k], and Isabel Krug[a]

[a]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, VIC, Australia; [b]Department of health, Tasmanian Centre for Mental Health Service Innovation Hobart, Tasmania, Australia; [c]Department of Psychiatry, University of Tasmania, Hobart, Tasmania, Australia; [d]Department of Mental Health, St Vincent's Hospital, Melbourne, Australia; [e]Iverson Institute, Swinburne University of Technology, Melbourne, Australia; [f]Department of Psychiatry, The University of Melbourne, Melbourne, Australia; [g]Centre for Youth Mental Health, Orygen, Melbourne, Australia; [h]Centre for Youth Mental Health, The University of Melbourne, Melbourne, Australia; [i]Department of Psychological Sciences, Swinburne University of Technology, Melbourne, Australia; [j]Department of Mental Health, Austin Health, Melbourne, Australia; [k]Centre for Mental Health, Swinburne University of Technology, Melbourne, Australia

ABSTRACT

This study assessed the rate of a.) the total and b.) specific psychiatric comorbidities among the three severity ratings for Anorexia Nervosa (AN): DSM-5, ICD-11 and overvaluation of weight and shape (OWS). The sample comprised 312 treatment-seeking patients with AN (mean age = 26.9). Weight and height were taken at intake to calculate BMI, the foundation for the DSM-5 and ICD-11 severity indices. The EDE-Q was used to assess OWS, and the Mini International Neuropsychiatric Interview was conducted to assess psychiatric comorbidities. For the DSM-5, the mild severity group showed a higher total number of psychiatric comorbidities, especially for panic, social anxiety, generalised anxiety, and post-traumatic stress disorders compared to the severe and extremely severe groups. ICD-11 and OWS severity groups did not significantly differ in total comorbidities, except for major depressive disorder and obsessive-compulsive disorders being more prevalent in the "significantly low BMI" ICD-11 group. The high OWS group displayed a notably higher rate of major depressive disorder than the low OWS group. The study underscores inconsistent patterns across the three severity systems, emphasising the need to recognise the current limitations of the assessed severity classification systems in AN assessment and guiding treatment.

CONTACT An Binh Dang dangb1@student.unimelb.edu.au Melbourne School of Psychological Sciences, The University of Melbourne, Redmond Barry, Level 7, Melbourne, VIC 3051, Australia

Clinical Implications

- Existing severity classifications for Anorexia Nervosa do not effectively gauge psychiatric comorbidity in treatment-seeking individuals.
- Both BMI-based severity ratings, DSM-5 and ICD-11, show that less severe groups (higher BMI) have a higher prevalence of psychiatric comorbidities.
- The high OWS group exhibits a notably higher rate of major depressive disorder compared to the low OWS group.
- Findings from the study do not advocate for the usage of the DSM-5, ICD-11, and OWS severity ratings in indexing psychiatric comorbidity in clinical settings.

Introduction

Anorexia Nervosa (AN) is currently diagnosed according to the fifth edition of the Diagnostic Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013) and its text-revised version (DSM-5-TR; APA, 2013) or the 11th edition of the International Classification of Diseases (ICD-11; World Health Organisation [WHO], 2019). Both the DSM-5 and ICD-11 introduced Body Mass Index (BMI) specifiers to index AN severity. DSM-5 has four severity groups: mild (≥17.0 kg/m2), moderate (16–16.99 kg/m2), severe (15–15.99 kg/m2), and extreme (<15 kg/m2), while ICD-11 has two groups: significantly low BMI (≥14 kg/m2) and dangerously low BMI (<14 kg/m2). As outlined in the DSM-5 (APA, 2013), severity specifiers are intended to assist clinicians in evaluating the degree of intensity, frequency, duration, symptom count, or other disorder-specific indicators of severity, rather than specifiers related to the course or descriptive features of the disorder.

AN commonly co-occurs with psychiatric comorbidities, including major depressive disorder (MDD; up to 85.6%; Carrot et al., 2017) and post-traumatic stress disorder (PTSD; 17.5%; Ferrell et al., 2022). Compared to those without psychiatric comorbidities, individuals with AN and psychiatric comorbidity tend to exhibit increased eating disorder (ED) symptom severity, maintenance of ED behaviours (e.g., restriction, purging), poorer functioning, and treatment outcomes (Blinder et al., 2006; Bulik et al., 2003; Hambleton et al., 2022; Spindler & Milos, 2007). As such, the presence of psychiatric comorbidities consistently indicates a more complex and severe illness presentation in AN and is indicative of a poorer prognosis. Despite the link between ED severity and psychiatric comorbidity, only a few studies have assessed the clinical utility of DSM-5 AN severity ratings in indexing comorbidity (Dakanalis et al., 2018), and no study to date has examined the clinical utility of the ICD-11 severity index for this purpose. Additionally, while overvaluation of weight/shape (OWS) has been proposed as a potential alternative

severity indicator for AN (e.g., Gianini et al., 2017), its ability to index psychiatric comorbidity is not well understood. Therefore, this study aimed to investigate whether, BMI and OWS, a stand-alone severity indicator can effectively gauge the rate of psychiatric comorbidity, a crucial marker of AN severity, in a sample seeking treatment

DSM-5 severity rating for AN

Although prior research (e.g., Brand-Gothelf et al., 2014) has suggested an increased number of psychiatric comorbidities to higher severity in AN (assessed via ED psychopathology measured through the Eating Disorders Inventory–2 (Garner & Olmsted, 1991)), a recent meta-analysis conducted by Dang et al. (2022) contradicts these findings. Dang et al.'s meta-analysis, involving 19 studies assessing DSM-5 ED severity ratings, with five studies specifically examining the DSM-5 severity index for AN (e.g., Dakanalis et al., 2018; Krug et al., 2021), found no significant differences in ED psychopathology and comorbidity rates across the four DSM-5 AN severity categories. The studies included in Dang et al.'s meta-analysis, however, used self-report measures (e.g., Krug et al., 2021). Only Dakanalis et al. used the Structured Clinical Interview for DSM-5 Disorders ([SCID]; First et al., 2015), but the assessed comorbidities were limited to mood, anxiety, and substance use disorders. Furthermore, no study to date has explored the total number of psychiatric comorbidities across each DSM-5 AN severity group. This leaves the relationship between the DSM-5 AN severity groups and the rate of total psychiatric comorbidity as well as the prevalence of specific psychiatric comorbidities (e.g., PTSD), unexplored.

ICD-11 severity rating for AN

Acknowledging that many individuals meeting the AN prototype narrowly miss the BMI <17.5 kg/m^2 criterion, the ICD-11 omitted the BMI <17.5 kg/m^2 criterion for AN. However, recognising the significance of severe underweight as a predictor of poor prognosis, the ICD-11 proposes severe underweight (e.g., BMI < 14.0 kg/m^2) as a severity qualifier, termed "dangerously low body weight" (Button et al., 2010; Rosling et al., 2011). This adjustment aims to improve diagnostic accuracy and consider the importance of severely underweight in assessing AN prognosis and treatment intensity (Uher & Rutter, 2012). Despite this, the BMI cut-off values for both the DSM-5 and ICD-11 AN severity categories lack a solid research foundation, as noted by Dang et al. (2022). The rationale behind the ICD-11 binary severity classification remains unclear, and to date, no study has evaluated the clinical utility of the two ICD-11 severity ratings for AN.

Overvaluation of weight/shape severity rating

Amidst uncertainty about the clinical utility of the DSM-5 and ICD-11 AN severity ratings, researchers have explored alternative classifications, focusing on cognitive factors such as overvaluation of weight and shape ([OWS], e.g., Dang et al., 2023). Research has found support for OWS as a transdiagnostic severity indicator for AN, bulimia nervosa, and binge eating disorder (e.g., Dang et al., 2022, 2023). For instance, high OWS in AN has been found to be linked to an increase in self-reported depression, anxiety, and stress symptoms (Dang et al., 2023). However, studies that used clinician-administered interviews to assess psychiatric comorbidity rates across the OWS severity groups are lacking.

The current study

This study aimed to address the gaps in the literature articulated above, by determining psychiatric comorbidity rates across the DSM-5, ICD-11, and OWS AN severity groups in a treatment-seeking AN population. Specifically, our aims included assessing: 1) whether there was an association between the DSM-5, ICD-11, and OWS-defined AN severity groups and the total number of psychiatric comorbidity and 2) whether there were differences in the rate of specific psychiatric comorbidities (e.g., PTSD, MDD) across these three severity indices.

Method

Participants

This study included treatment-seeking patients with AN ($N = 312$, mean age: 26.9 years, mean BMI: 16.8 kg/m^2) from an outpatient ED treatment unit in Melbourne, Australia. Our clinicians used the DSM-5 criteria to assess participants' diagnoses of AN at the time of assessment. All patients with AN from the treatment unit were included in the study, with no exclusions, provided they consented. Sixty-six participants (21%) reported having at least one episode of binge eating and inappropriate compensatory behaviours over the past 3 months (see Dang et al., 2023; Phillipou & Beilharz, 2019 for more details).

Measures

Sociodemographic and clinical information

The study gathered key demographic data at intake. Participants also self-reported their age of AN onset, from which illness duration was calculated. Measurements of height and weight were taken with individuals wearing

indoor clothing and without shoes. Height was determined using a stadiometer, while weight was evaluated using the Tanita scale (Jebb et al., 2000), which employs bioelectrical impedance technology to calculate BMI.

Psychiatric comorbidity

To screen for adult psychiatric disorders based on the DSM-5 criteria, trained senior mental health clinicians relied on clinical judgment to select relevant modules from the Mini International Neuropsychiatric Interview [MINI] version 7.0.2 (Sheehan et al., 1998). Their judgment was based on gathered information from the participants' medical and psychiatric history, clinical observations, and intake interviews. Specific psychiatric diagnoses included in the MINI comprise MDD, persistent depressive disorder (PDD), agoraphobia, panic disorder (PD), social anxiety disorder (SAD), obsessive-compulsive disorder (OCD), generalised anxiety disorder (GAD), PTSD, alcohol use disorder (AUD), and substance use disorder (SUD).

AN severity categories

Severity groups based on DSM-5 and ICD-11 severity specifiers

Using the DSM-5 severity definitions based on BMI, with a total of 136 (43.6%) participants being classified as mild (≥17.0 BMI kg/m^2), 64 (20.5%) as moderate (16–16.99 BMI kg/m^2), 49 (15.7%) as severe (15–15.99 BMI kg/m^2) and 63 (20.2%) extreme (<15 BM kg/m^2). Based on the ICD-11 BMI severity cutoff, 284 (91%) and 28 (9%) participants were categorised into the "significantly low" and "dangerously low," respectively.

Subgroups based on overvaluation of shape/weight

Guided by previous studies (e.g., Gianini et al., 2017), OWS was measured using two items from the Eating Disorder Examination-Questionnaire (Fairburn & Beglin, 1994): "Over the past 4 weeks, has your shape influenced how you feel about (judge, think, evaluate) yourself as a person?" and "Over the past 4 weeks has your weight influenced how you feel about (judge, think, evaluate) yourself as a person?" A score of 5 or 6 on either item classified individuals with AN into low and high levels of OWS (Gianini et al., 2017) (see Dang et al., 2023 for more details). Due to missing data, out of 312 participants, we were only able to categorise 240 participants into low (n = 50; 20.8%) and high (n = 190; 79.2%) OWS AN groups.

Statistical planning

Differences in all study variables between AN severity groups were assessed using ANOVA for continuous variables and chi-square tests (Pearson's or

Fisher's exact test, as appropriate) for categorical variables. Post-hoc pairwise comparisons with Bonferroni correction were conducted to identify where the differences lied, and effect sizes (partial η^2 for continuous variables and Cramer's φ for categorical variables) were calculated. The partial η^2 with values of 0.06, 0.10, and 0.25 were interpreted as indicating low-poor, moderate-medium, and large-high effect sizes, respectively (Levine & Hullett, 2002). Additionally, Cramer's φ coefficient with values of 0.06, 0.15, and 0.30 was interpreted as indicating low-poor, moderate-medium, and large-high effect sizes, respectively (Cohen, 1988).

The distribution of the data was based on counts and had an excess of zero responses indicating most participants did not have a psychiatric comorbidity (58%). As such, the DSM-5, ICD-11 and OWS AN group differences in psychiatric comorbidities were examined using zero-inflated Poisson regression. The zero-inflated model expands on the Poisson distribution to allow for the analysis of data with excess zeros. This model produces two sets of results, one related to whether there was a psychiatric comorbidity and one related to the number of psychiatric comorbidities (Cameron & Trivedi, 2001; Coxe et al., 2009; Karazsia & van Dulmen, 2008; Lambert, 1992). For the DSM-5 AN severity groups, the severity groups were entered into the equation as dummy variables to compare the mild, moderate, severe, and extremely severe in the number of psychiatric comorbidities. Furthermore, Pairwise Exclusion of Missing Values was employed to manage missing data.

Results

Demographic information

Table 1 summarises the demographic characteristics of the entire sample and each severity index. Most participants with AN were single (76.2%), Caucasian (84.6%), and female (94.6%). The mean age of AN onset was 17.61 years (SD = 5.83), with an illness duration of 8.79 years (SD = 8.83).

From the MINI, 131 participants (42%) were found to have at least one additional psychiatric disorder. The DSM-5 and ICD-11 AN severity groups exhibited no significant differences in any demographic and clinical-related information. Individuals in the low OWS severity group had a higher BMI at assessment than the high OWS group ($p < .001$). The distribution of each severity group for the three assessed severity indices is presented in Table 1.

Total number of psychiatric comorbidities across an severity indices and groups

Table 2 presents results from the zero-inflated Poisson regressions with AN severity per DSM-5, ICD-11, and OWS as predictors of the total number of psychiatric comorbidities. The zero model showed no significant differences

Table 1. Demographic information for each severity groups within the DSM-5, ICD-11 and overvaluation of weight/shape severity ratings for AN.

			DSM-5 AN severity classification								ICD AN severity classification				Overvaluation of weight/shape (OWS) classification			
	Total		DSM-5 Mild		DSM-5 Moderate		DSM-5 Severe		DSM-5 Extreme		Significantly Low		Dangerously Low		OWS/Low		OWS/High	
	N = 312		*N = 136*		*N = 64*		*N = 49*		*N = 63*		*N = 284*		*N = 28*		*N = 50*		*N = 190*	
Gender																		
Male	17	5.4%	12	8.8%	2	3.1%	1	2.0%	2	3.2%	*16*	*5.6%*	*1*	*3.6%*	*6*	*12.0%*	6	3.2%
Female	295	94.6%	124	91.2%	62	96.9%	48	98.0%	61	96.8%	*268*	*94.4%*	*27*	*96.4%*	*44*	*88.0%*	184	96.8%
Marital status																		
Single	237	76.2%	107	78.7%	49	76.6%	36	73.5%	45	72.6%	214	75.6%	23	82.1%	35	70%	149	78.4%
Widowed	16	1.5%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	2	0.7%	0	0%	0	0%	1	0.5%
Divorced/Separated	16	5.2%	2	1.4%	2	3.1%	4	8.2%	8	12.9%	15	5.3%	1	3.6%	4	8.0%	8	4.3%
Married—in couple	35	11.3%	14	10.4%	11	17.2%	5	10.2%	5	8.1%	34	12.0%	1	3.6	6	12.0%	22	11.6%
Other/Unknown	21	6.8%	11	8.1%	2	3.1%	4	8.2%	4	6.5%	18	6.4%	3	10.7%	5	10.0%	10	5.3%
Education																		
Secondary/lower	87	28.2%	43	32.2%	18	28.6%	17	34.6%	9	14.5%	82	29.2%	5	18.5%	7	14.3%	54	28.7%
Vocational	7	2.3%	4	3.0%	1	1.6%	0	0%	2	3.2%	6	2.1%	1	3.7%	2	4.1%	5	2.7%
University Commenced	106	34.4%	45	33.6%	25	39.7%	13	26.5%	23	37.1%	96	34.2%	10	37.0%	20	40.8%	66	35.1%
University Completed	79	25.6%	27	20.1%	14	22.2%	15	30.6%	23	37.1%	70	24.9%	9	33.3%	12	24.5%	49	26.1%
Other/Unknown	29	9.4%	15	11.1%	5	7.9%	3	6.1%	3	4.8%	10	3.5%	2	7.4%	8	16.4%	14	7.5%
Ethnicity																		
Aboriginal	7	2.5%	3	2.3%	1	1.7%	3	6.7%	0	0%	7	2.7%	0	0%	2	4.0%	4	2.1%
Caucasian	241	84.6%	112	87.5%	51	87.9%	39	86.7%	39	72.2%	227	87.0%	14	58.3%	41	82.0%	162	85.3%
Asian	10	3.6%	3	2.4%	1	1.7%	1	2.2%	11	3.6%	9	3.5%	1	4.2%	2	4.0%	6	3.2%
Other/Unknown	27	9.5%	10	7.8%	5	8.6%	2	4.4%	10	18.5%	16	6.9%	9	37.5%	5	10.0%	19	10.0%
Employment																		
Unemployed.	29	9.7%	11	8.5%	4	6.3%	6	13.0%	8	13.3%	29	10.6%	0	0%	6	12.0%	15	7.9%
Student	101	33.8%	44	33.8%	19	30.2%	15	30.4%	24	40.0%	92	33.6%	9	36.0%	16	32.0%	62	32.6%
Employed	63	26.3%	34	28.6%	14	28.6%	6	19.4%	9	20.5%	76	27.8%	4	16.0%	16	32.0%	47	24.7%
Home Duties	10	3.3%	5	3.8%	1	1.6%	1	2.2%	3	5.0%	7	2.6%	3	12.0%	0	0%	8	4.2%
Unable to work/Other	79	26.5%	30	23.1%	20	31.7%	14	30.4%	15	25%	70	25.5%	9	36.0%	12	24.0%	58	30.5%

Table 2. Zero-inflated Poisson regression for severity groups contrasts predicting psychiatric comorbidity.

	B	*SE*	*z*	*p*	95% CI (*B*)
DV = Number of Psychiatric Comorbidity (Count model)					
Severity groups					
DSM-5 Mild to Moderate	**−0.36**	0.15	−2.95	**.005**	(−0.75; −0.15)
DSM-5 Mild to Severe	**−0.36**	0.15	−2.03	**.009**	(−0.61; −0.01)
DSM-5 Mild to Extreme	**−0.73**	0.14	−4.34	**<.001**	(−1.46; −0.55)
DSM-5 Moderate to Severe	0.14	0.19	0.77	.44	(−0.22; 0.51)
DSM-5 Moderate to Extreme	**−0.55**	0.26	−2.17	**.042**	(−1.05; −0.05)
DSM-5 Severe to Extreme	−0.70	0.26	−2.73	.006	(−1.20; −0.20)
Low BMI to Dangerously Low BMI	−0.40	0.23	1.71	.09	(−1.13; 0.03)
Low OWS to high OWS	0.32	0.16	1.95	.05	(−0.11; 0.60)
DV = No psychiatric comorbidity (Zero-inflated model)					
DSM-5 Mild to Moderate	−0.52	0.31	−1.64	.10	(−1.30; 0.04)
DSM-5 Mild to Severe	−0.72	0.35	−2.03	.34	(−1.36; 0.16)
DSM-5 Mild to Extreme	−0.47	0.43	−1.09	.27	(−1.33; 0.38)
DSM-5 Moderate to Severe	−0.13	0.43	−0.30	.77	(−0.98; 0.72)
DSM-5 Moderate to Extreme	0.15	0.49	0.31	.76	(−0.81; 1.12)
DSM-5 Severe to Extreme	0.28	0.51	0.55	.56	(−0.72; 1.28)
Low BMI to Dangerously Low BMI	−0.18	0.45	−0.41	.68	(0.70; 1.01)
Low OWS to high OWS	0.07	0.35	−0.19	.85	(−1.00;1.05)

among DSM-5 severity groups regarding the likelihood of having no psychiatric comorbidity. The count model showed that there were significant differences among the groups in the total number of psychiatric comorbidities. More specifically, those in the DSM-5 mild group were associated with a higher total number of psychiatric comorbidities in comparison to those in the DSM-5 moderate (OR = 1.43), severe (OR = 1.43) and extremely severe (OR = 2.08) AN groups. In addition, those in the DSM-5 moderate severity group were associated with a higher total number of psychiatric comorbidities compared to those in the extremely severe group (OR = 1.73).

The regression models found no significant differences between the ICD-11 "significantly low BMI" and "dangerously low BMI" groups and between the low OWS and high OWS groups in both the likelihood of not having a psychiatric comorbidity and the total number of psychiatric comorbidities.

Rate of specific psychiatric disorders across the AN severity indices and groups

Table 3 presents the comorbidity rates for the specific psychiatric disorders for individuals with AN, across the different severity groups for DSM-5, ICD-11, and OWS. The chi-square tests of independence revealed that there were significant associations between the DSM-5 AN severity groups and the rates of PD [χ^2 (3, n = 161) = 11.16, p = .011], SAD [χ^2 (3, n = 161) = 12.89, p = .005], GAD [χ^2 (3, n = 153) = 9.99, p = .019] and PTSD [χ^2 (3, n = 160) = 9.64, p = .022]. Post hoc comparisons showed that those in the mild DSM-5 AN severity group had significantly higher rates of PD (p = .004), SAD (p = .003), GAD (p = .005), and PTSD (p = .020) than those in the DSM-5 extremely severe group. In addition, those in the mild DSM-5 AN severity group had

Table 3. Psychiatric comorbidity rate between DSM-5, ICD-11, and OWS severity groups.

	DSM-5 Classification											ICD-11 Classification							Overvaluation of weight and shape						
	Mild		Moderate		Severe		Extreme					Significantly Low		Dangerously Low					Low OWS		High OWS				
Anorexia nervosa	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	*F*/	*P*	h^2/V	n	Mean (SD)	n	Mean (SD)	*F*	*p*	h^2/V	n	Mean (SD)	n	Mean (SD)	*F*	*p*	h^2/V
Age (yrs-old)	135	25.3 (7.6)	63	27.4 (10.5)	49	26.5 (8.6)	63	29.2 (10.3)	2.3	.140	.029†	282	*26.7 (9.1)*	28	28.5 (9.8)	0.96	.327	.003	50	27.5 (9.5)	188	26.3 (8.9)	0.7	.518	.000
Onset ED (yrs-old)	128	17.4 (6.1)	60	16.5 (4.5)	43	18.6 (5.8)	55	19.1 (7.1)	1.6	.290	.022	261	*17.5 (5.7)*	25	19.1 (6.9)	1.8	.176	.006	47	19.4 (5.5)	174	17.2 (6.0)	5.2	0.06	.004
Duration ED (yrs)	129	7.8 (7.4)	59	11.3 (11.2)	44	7.6 (7.0)	56	9.9 (9.4)	2.1	.170	.029	262	*8.8 (8.8)*	26	8.5 (8.9)	0.3	.862	.000	47	7.4 (7.1)	173	9.3 (9.0)	1.9	.266	.002
BMI at assessment (kg/m^2)	136	18.9 (2.2)	64	16.4 (0.3)	49	15.4 (0.3)	63	13.9 (0.9)	120.4	**<.001**	.605†	284	*17.1 (2.2)*	28	12.9 (0.7)	98.3	**<.001**	.241	50	16.0 (2.1)	190	17.3 (2.6)	10.5	**.001**	**.046**
	n	%	n	%	n	%	n	%	χ2	*P*	V	n	%	n	%	χ2	*p*	V	n	%	n	%	χ2	*p*	V
Presence of at least 1 psychiatric disorder	**51**		**31**		**26**		**23**																		
Depressive Disorder	39	39.8	23	23.5	20	20.4	16	16.3	3.4	.332	.146	92	93.9	6	6.1	*5.2*	***.034***	.180	9	14.1	55	85.9	5.5	**.019**	.299
Panic Disorder	26	53.1	9	18.4	10	20.4	4	8.2	*11.2*	***.011***	.263	47	95.9	2	4.1	*3.1*	*.097*	.139	6	15.4	33	84.6	1.5	.229	.117
Social Anxiety	21	60	5	14.3	7	20.0	2	5.7	*12.9*	***.005***	.283	34	97.1	1	2.9	*2.8*	*.124*	.132	3	12.5	21	87.5	1.5	.214	.121
Obsessive Compulsive Disorder	15	48.4	8	25.8	4	12.9	4	12.9	*3.0*	*.399*	.135	31	100	0	0	*4.5*	***.045***	.168	6	24	19	76	0.1	.749	.031
Generalized Anxiety Disorder	26	47.3	13	23.6	12	21.8	4	7.3	*9.9*	***.019***	.256	52	94.5	3	5.5	*2.3*	*.172*	.122	6	15.8	32	84.2	1.4	.241	.117
Post Traumatic Stress Disorder	17	60.7	4	14.3	5	17.9	2	7.1	*9.6*	***.022***	.245	27	96.4	1	3.6	*1.8*	*.311*	.105	3	12.5	21	87.5	1.6	.205	.124
Alcohol Use Disorder	7	36.9	2	15.4	2	15.4	2	15.4	*1.8*	*.610*	.107	11	84.6	2	15.4	*0.3*	*.632*	.046	3	33.3	6	66.7	0.8	.386	.085
Substance Use Disorder	5	55.6	1	11.1	3	33.3	0	0	*4.8*	*.189*	.173	9	100	0	0	*1.1*	*.599*	.084	2	25	6	75	0.1	.826	.021

a significantly higher rate of SAD than those in the moderate DSM-5 group ($p = .026$).

The ICD-11 "significantly low" BMI group had significantly higher rates of MDD ($p = .034$) and obsessive-compulsive disorder (OCD; $p = .045$) compared to the "dangerously low" group. Individuals in the high OWS alternative AN severity group had a higher rate of MDD ($p = .019$) than the low OWS group.

Discussion

The study investigated 1) the relationships between the DSM-5, ICD-11, and OWS-defined severity groups for AN and the total number of psychiatric comorbidities and 2) whether there were significant differences in specific psychiatric comorbidity rates among the three assessed severity ratings. Overall, although there were no differences in overall comorbidity counts within the ICD-11 and OWS AN severity groups, the DSM-5 mild severity group exhibited a significantly higher total number of psychiatric comorbidities compared to the DSM-5 moderate, severe, and extremely severe AN groups. Significant differences in the prevalence of specific psychiatric disorders (e.g., PD, SAD, GAD, and PTSD using the DSM-5 severity index; or OCD and MDD using the ICD-11 severity rating) were also found across the three assessed severity ratings, which will be discussed further below.

Distribution of AN severity groups for DSM-5, ICD-11, and OWS

Despite using a sample of treatment-seeking individuals with AN, both the DSM-5 and ICD-11 mainly categorised most individuals with AN into the milder (higher BMI) groups. Only about 15% and 8% of individuals with AN fell into the DSM-5 severe and ICD-11 "dangerously low body weight" severity groups, respectively, raising doubts about correctly identifying these two groups in AN. Conversely, up to 78% of AN participants were classified as having high OWS—the more severe end of the OWS severity spectrum. Such distribution is consistent with what was found in previous studies (e.g., Dang et al., 2023). Considering that the present sample consists of treatment-seeking participants with an average illness duration of 8.79 years, it seems that the OWS severity rating effectively mirrors the level of AN severity in this specific cohort in comparison to the DSM-5 and ICD-11 severity systems.

Comorbidity: BMI-based severity ratings per DSM-5 and ICD-11

Total number of psychiatric comorbidities

The DSM-5 and ICD-11 systems use BMI to index severity in AN, differing only in categorical cut-off scores, yet they did not yield a consistent pattern of

results in the total number of psychiatric comorbidities across the AN severity groups. Regarding the DSM-5 severity groups, the mild AN group was found to be associated with higher odds of having a greater number of psychiatric comorbidities compared to the moderate, severe, and extremely severe AN groups. The moderate DSM-5 group was also found to have a higher number of psychiatric comorbidities than the DSM-5 extremely severe groups. As for ICD-11, no significant differences were found between the two severity groups in the total number of psychiatric comorbidities.

While no study has examined the relationship between both the DSM-5 and ICD-11 AN severity groups and the total number of psychiatric comorbidities, our findings contradict prior research associating psychiatric comorbidity with increased AN severity, including factors such as suicidal ideations and ED symptomatology (Hambleton et al., 2022; Spindler & Milos, 2007). The discrepancy between our results and prior studies (e.g., Dang et al., 2023), coupled with the unexpected trend of higher comorbidity in the milder DSM-5 severity groups, suggests that the current DSM-5 and ICD-11 severity ratings for AN may need to be reconsidered.

Specific psychiatric comorbidities

The pattern of higher comorbidity in the milder DSM-5 AN severity groups extended to specific psychiatric diagnoses, with the mild DSM-5 AN group exhibiting significantly higher rates of PD, SAD, GAD, and PTSD compared to the extreme DSM-5 AN group. In addition, the mild DSM-5 AN severity group had significantly higher rates of SAD compared to the moderate DSM-5 AN severity group.

When examining individual psychiatric comorbidity rates across the ICD-11 severity groups, we observed that individuals in the 'significantly low BMI' ICD-11 group were more likely to have MDD and OCD than those in the 'dangerously low BMI' group. Despite these differences, both DSM-5 and ICD-11 BMI-based severity classification systems share a common pattern: individuals with a higher BMI tend to have a higher likelihood of experiencing additional psychiatric comorbidities. These disparities in the type of psychiatric comorbidities could stem from the arbitrary BMI cut-off scores in both the DSM-5 and ICD-11 systems.

Relating to the previous literature, no prior study has assessed psychiatric comorbidity rates among the two ICD-11 AN severity groups. However, in line with Dakanalis et al. (2018) findings regarding the DSM-5 severity ratings for AN, we also found no differences in the rates of MDD across the DSM-5 AN severity groups. In contrast to Dakanalis et al., the current study observed significant differences in the rates of PD, SAD, and GAD among the DSM-5 AN severity groups. These differences may stem from our separate assessment of the individual anxiety disorders (SAD versus GAD), unlike Dakanalis et al.

who grouped these disorders under an overall anxiety disorder group. Our findings also deviate somewhat from Krug et al. (2021), who observed no significant differences between the DSM-5 AN severity groups concerning self-reported depression, anxiety, and stress symptoms as measured by the Symptom Checklist-90-Revised (Derogatis, 1983). Discrepancies between our findings and Krug et al.'s could be attributed to their separate evaluation of DSM-5 severity ratings for AN-binge purge and AN-restricted subtypes, while we combined these subtypes for our analysis due to having a smaller sample size. This means that the DSM-5 severity classification for AN may operate differently depending on the specific AN subtypes, highlighting the importance of considering potential separate severity ratings for each AN subtype, such as the number of purging methods as an alternative severity rating for AN-binge purge subtypes (Dang et al., 2022; Gianini et al., 2017).

Comorbidity: overvaluation of weight and shape

No significant differences were found in the total number of psychiatric comorbidities between the low and high OWS AN groups. However, we found that those in the high OWS group had significantly higher rates of MDD than those in the low OWS group. Such findings are consistent with previous studies (Dang et al., 2023; Gianini et al., 2017) that showed increased self-reported symptoms of depression in those with AN in the high OWS group. Building on previous studies that found OWS to effectively capture ED and general psychopathology in AN (Dang et al., 2022, 2023), coupled with the current findings found OWS to be indicative of psychiatric comorbidity, OWS emerges as a more robust indicator of AN severity compared to BMI-based DSM-5 and ICD-11 indices. This is reflected in its ability to index general and ED psychopathology (Dang et al., 2023) and psychiatric comorbidity rates as observed in our study. Overall, there was no consistent pattern of psychiatric comorbidity emerging across the three tested severity systems for AN.

Implications

The current findings have implications for DSM-5, ICD-11, and OWS AN severity ratings. Although the OWS severity systems demonstrated a slight advantage over the BMI-based severity systems, associating higher OWS severity with higher rates of MDD and OCD, the current evidence falls short of substantiating their utility in clinical practice. Consequently, emphasising the importance of comprehensive clinical assessments remains crucial for informed treatment planning of individuals with AN.

Second, both this study and prior research (e.g., Dang et al., 2022; Krug et al., 2021) challenge the traditional assumption that lower BMI indicates more complex and severe AN, such as higher ED psychopathology and psychiatric

comorbidity. It highlights that BMI alone does not reliably correlate with psychiatric comorbidity in AN, as it is just one facet of a multifaceted psychiatric disorder involving fear of weight gain and body distortions (Gadsby, 2017). Using BMI as the primary gauge of AN severity perpetuates the idea that individuals without an extremely low BMI are "not unwell enough," despite experiencing similar disordered eating patterns and facing comparable physical and psychological risks, as shown in previous studies (e.g., Dang et al., 2023; Gianini et al., 2017). Therefore, BMI should not be the sole determinant for clinical decisions in individuals with AN (Phillipou & Beilharz, 2019).

Limitations and future directions

The current study has some limitations which should be acknowledged. First, the cross-sectional design prevents exploring the predictive validity of the DSM-5, ICD-11, and OWS AN severity ratings, including treatment outcomes and ED diagnoses crossover.

Second, senior clinicians used their clinical judgment to select relevant MINI modules to screen for ICD-11 psychiatric disorders. This approach, meant to ease burdens, could introduce clinician selection biases compared to administering the full MINI. Consequently, future research should employ a more standardised method for establishing psychiatric comorbidity rates.

Third, although the ICD-11 severity rating aims to serve as a prognostic indicator for physical complications and mortality risk (WHO, 2019), this study only assessed psychiatric comorbidity. Future studies should assess these severity ratings and further severity indicators (e.g., weight suppression, Berner et al., 2013) on a range of psychological (e.g., perfectionism, Buzzichelli et al., 2018; Drieberg et al., 2019), biological (e.g., potassium, Funayama et al., 2021) or physical indicators (amenorrhoea); using a longitudinal design. For example, there is a consensus that amenorrhoea should not be included as a diagnosis criterion for AN but should be recorded, since it may be an indicator of severity and may help distinguish between constitutional thinness and AN (Attia & Roberto, 2009).

Fourth, our categorisation of individuals with AN into multiple severity groups may have lacked the statistical power to detect significant links between psychiatric comorbidity and severity levels. Future research should consider larger sample sizes to further explore these relationships.

Finally, this was conducted in an Australian treatment facility, with predominantly participants being Caucasian females. Given the recognised variations in ED presentations among males and females (Murray et al., 2017), adolescents versus adults (Rajagopalan, et al., 2019), and individuals from diverse cultural backgrounds (Soh et al., 2008), the results may not be broadly applicable to other demographic groups, warranting further replications.

Conclusion

Our findings suggest that the currently existing severity indicators for AN (e.g. DSM-5 and ICD-11) cannot effectively index psychiatric comorbidity in individuals with AN. We did, however, find that the alternative OWS severity system was slightly more robust in indexing the prevalence of specific psychiatric comorbidities in AN compared to their BMI-based counterparts (i.e., DSM-5 and ICD-11). Specifically, we found a correlation between a high OWS severity and elevated rates of MDD and OCD. However, the existing evidence does not conclusively support the practical application of the OWS severity rating in clinical settings for AN. This means that as we navigate the complexities of AN severity classifications, integrating a holistic approach (e.g., in-depth assessment of both psychological and behavioural symptoms) remains pivotal for advancing the precision and effectiveness of clinical interventions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

An Binh Dang http://orcid.org/0000-0001-7961-6913

References

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Attia, E., & Roberto, C. A. (2009). Should amenorrhea be a diagnostic criterion for anorexia nervosa? *International Journal of Eating Disorders*, *42*(7), 581–589. https://doi.org/10.1002/eat.20720

Berner, L. A., Shaw, J. A., Witt, A. A., & Lowe, M. R. (2013). The relation of weight suppression and body mass index to symptomatology and treatment response in anorexia nervosa. *Journal of Abnormal Psychology*, *122*(3), 694. https://doi.org/10.1037/a0033930

Blinder, B. J., Cumella, E. J., & Sanathara, V. A. (2006). Psychiatric comorbidities of female inpatients with eating disorders. *Psychosomatic Medicine*, *68*(3), 454–462. https://doi.org/10.1097/01.psy.0000221254.77675.f5

Brand-Gothelf, A., Leor, S., Apter, A., & Fennig, S. (2014). The impact of comorbid depressive and anxiety disorders on severity of anorexia nervosa in adolescent girls. *The Journal of*

Nervous and Mental Disease, *202*(10), 759–762. https://doi.org/10.1097/NMD.0000000000000194

Bulik, C. M., Tozzi, F., Anderson, C., Mazzeo, S. E., Aggen, S., & Sullivan, P. F. (2003). The relation between eating disorders and components of perfectionism. *American Journal of Psychiatry*, *160*(2), 366–368. https://doi.org/10.1176/appi.ajp.160.2.366

Button, E. J., Chadalavada, B., & Palmer, R. L. (2010). Mortality and predictors of death in a cohort of patients presenting to an eating disorders service. *International Journal of Eating Disorders*, *43*(5), 387–392. https://doi.org/10.1002/eat.20715

Buzzichelli, S., Marzola, E., Amianto, F., Fassino, S., & Abbate-Daga, G. (2018). Perfectionism and cognitive rigidity in anorexia nervosa: Is there an association? *European Eating Disorders Review*, *26*(4), 360–366. https://doi.org/10.1002/erv.2591

Cameron, A. C., & Trivedi, P. K. (2001). Essentials of count data regression. *A Companion to Theoretical Econometrics*, 331.

Carrot, B., Radon, L., Hubert, T., Vibert, S., Duclos, J., Curt, F., & Godart, N. (2017). Are lifetime affective disorders predictive of long-term outcome in severe adolescent anorexia nervosa? *European Child & Adolescent Psychiatry*, *26*(8), 969–978. https://doi.org/10.1007/s00787-017-0963-5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*.

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121–136.

Dakanalis, A., Timko, C. A., Colmegna, F., Riva, G., & Clerici, M. (2018). Evaluation of the DSM-5 severity ratings for anorexia nervosa in a clinical sample. *Psychiatry Research*, *262*, 124–128. https://doi.org/10.1016/j.psychres.2018.02.009

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders. *Clinical Psychology Science & Practice*, *29*(4), 325–344. https://doi.org/10.1037/cps0000078

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2023). Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates? *European Eating Disorders Review*, *31*(4), 447–461. https://doi.org/10.1002/erv.2969

Derogatis, L. R. (1983). SCL-90-R: Administration, scoring and procedures. Manual II for the R (evised) version and other instruments of the psychopathology rating scale series.

Drieberg, H., McEvoy, P. M., Hoiles, K. J., Shu, C. Y., & Egan, S. J. (2019). An examination of direct, indirect and reciprocal relationships between perfectionism, eating disorder symptoms, anxiety, and depression in children and adolescents with eating disorders. *Eating Behaviors*, *32*, 53–59. https://doi.org/10.1016/j.eatbeh.2018.12.002

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363–370.

Ferrell, E. L., Russin, S. E., & Flint, D. D. (2022). Prevalence estimates of comorbid eating disorders and posttraumatic stress disorder: A quantitative synthesis. *Journal of Aggression, Maltreatment & Trauma*, *31*(2), 264–282. https://doi.org/10.1080/10926771.2020.1832168

First, M., Williams, J., Karg, R., & Association, S. R. A. P. (2015). *Arlington*. VA, American Psychiatric Association.

Funayama, M., Mimura, Y., Takata, T., Koreki, A., Ogino, S., Kurose, S., & Shimizu, Y. (2021). Hypokalemia in patients with anorexia nervosa during refeeding is associated with binge–purge behavior, lower body mass index, and hypoalbuminemia. *Journal of Eating Disorders*, *9*(1), 1–10. https://doi.org/10.1186/s40337-020-00356-7

Gadsby, S. (2017). Distorted body representations in anorexia nervosa. *Consciousness and Cognition*, *51*, 17–33. https://doi.org/10.1016/j.concog.2017.02.015

Garner, D. M., & Olmsted, M. P. (1991). *Eating disorder inventory 2*. Psychological Assessment Resources Odessa.

Gianini, L., Roberto, C. A., Attia, E., Walsh, B. T., Thomas, J. J., Eddy, K. T., Grilo, C. M., Weigel, T., & Sysko, R. (2017). Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers. *International Journal of Eating Disorders*, *50*(8), 906–916. https://doi.org/10.1002/eat.22728

Hambleton, A., Pepin, G., Le, A., Maloney, D., Touyz, S., & Maguire, S. (2022). Psychiatric and medical comorbidities of eating disorders: Findings from a rapid review of the literature. *Journal of Eating Disorders*, *10*(1), 132. https://doi.org/10.1186/s40337-022-00654-2

Jebb, S. A., Cole, T. J., Doman, D., Murgatroyd, P. R., & Prentice, A. M. (2000). Evaluation of the novel Tanita body-fat analyser to measure body composition by comparison with a four-compartment model. *British Journal of Nutrition*, *83*(2), 115–122. https://doi.org/10.1017/S0007114500000155

Karazsia, B. T., & Van Dulmen, M. H. (2008). Regression models for count data: Illustrations using longitudinal predictors of childhood injury. *Journal of Pediatric Psychology*, *33*(10), 1076–1084.

Krug, I., Binh Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the DSM-5 severity indices for eating disorders. *European Eating Disorders Review*, *29*(3), 482–498. https://doi.org/10.1002/erv.2776

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1–14.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, *28*(4), 612–625.

Murray, S. B., Nagata, J. M., Griffiths, S., Calzo, J. P., Brown, T. A., Mitchison, D., Blashill, A. J., & Mond, J. M. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, *57*, 1–11. https://doi.org/10.1016/j.cpr.2017.08.001

Phillipou, A., & Beilharz, F. (2019). Should we shed the weight criterion for anorexia nervosa? *Australian & New Zealand Journal of Psychiatry*, *53*(6), 501–502. https://doi.org/10.1177/0004867418814958

Rajagopalan, A., Bhattacharya, A., Peebles, R., Dakanalis, A., & Timko, C. A. (2019, October). A literature review: Do DSM–5 severity indicators for eating disorders apply to children and adolescents? [Paper presentation]. American Academy of Children and Adolescent Psychiatry 66th Annual Meeting, Chicago, Illinois, United States.

Rosling, A. M., Sparén, P., Norring, C., & von Knorring, A. L. (2011). Mortality of eating disorders: A follow-up study of treatment in a specialist unit 1974–2000. *International Journal of Eating Disorders*, *44*(4), 304–310. https://doi.org/10.1002/eat.20827

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(20), 22–33.

Soh, N. L. W., Touyz, S., Dobbins, T., Surgenor, L., Clarke, S., Kohn, M., Lee, E. L., Leow, V., Rieger, E., Ung, K. E. K., & Walter, G. (2008). Body image disturbance in young North European and East Asian women with and without eating disorders in Australia and in Singapore. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, *16*(4), 287–296.

Spindler, A., & Milos, G. (2007). Links between eating disorder symptom severity and psychiatric comorbidity. *Eating Behaviors*, *8*(3), 364–373. https://doi.org/10.1016/j.eatbeh.2006.11.012

Uher, R., & Rutter, M. (2012). Classification of feeding and eating disorders: Review of evidence and proposals for ICD-11. *World Psychiatry*, *11*(2), 80–92. https://doi.org/10.1016/j.wpsyc.2012.05.005

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

Chapter 6: Do Risk Factors Differentiate DSM-5and Drive for Thinness Severity Groups for Anorexia Nervosa?

6.1 Chapter Introduction

As outlined in Chapter 1, considerable efforts have been made to identify risk factors for developing AN. Factors such as childhood obesity, perfectionism, and teasing about weight and shape have consistently been linked to an increased risk of AN (Barakat et al., 2023; Yamamiya et al., 2024). In medical research, risk factors are typically associated with poor outcomes, including infection, critical care admission, or death (Arboix et al., 2015; Sun et al., 2021).

Despite the findings from Chapters 2 (Study 1), 4 (Study 2), and 5 (Studies 3a and 3b) providing limited support for the utility and validity of BMI-based AN severity ratings per DSM-5 and ICD-11, these studies were cross-sectional. It is argued that if severity ratings are valid, they should not only reflect the intensity of ED psychopathology but also distinguish important clinical information, such as prognosis or factors contributing to the development and maintenance of AN.

In response, **Chapter 6 (Study 4)** presents a published study that employed a retrospective design with 153 sister pairs discordant for AN ($N = 306$). These sister pairs were recruited from four ED hospitals across Europe (Austria, the UK, Spain, and Slovenia). A sister-pair discordant design is particularly effective for studying AN risk factors because it controls for shared genetic and environmental influences (Karwautz et al., 2001). By comparing sisters where one has AN and the other does not, this approach isolates differences that may contribute to the disorder development (Frisell et al., 2021). Specifically, it minimises confounding variables common in non-sibling samples, such as socioeconomic status or family history, allowing for a more precise examination of individual-specific risk

factors, such as personality traits or life events, that uniquely contribute to the onset of AN (Frisell et al., 2021).

Additionally, another potential severity indicator, DT, has only been explored in one study by Krug et al. (2021). Their findings suggested that DT was superior to DSM-5 severity ratings in indexing both ED-specific and general psychopathology. Chapter 6 (Study 4) builds on this work by comparing DSM-5 and DT severity ratings for AN in relation to established ED risk factors. Understanding these relationships could enhance the clinical utility of these severity systems and inform the development of prevention programs and clinical management policies for AN.

6.2 Study 4. Do Risk Factors Differentiate DSM-5and Drive for Thinness Severity Groups for Anorexia Nervosa?

Study 4 was published in the Journal of Eating Disorder in January 2024

Dang, A. B., Kiropoulos, L., Anderluh, M., Collier, D., Fernandez-Aranda, F., Karwautz, A., Treasure, J., Wagner, G., & Krug, I. (2024). Do risk factors differentiate DSM-5 and drive for thinness severity groups for anorexia nervosa?. *Journal of Eating Disorders,* 12(1), 5. https://doi.org/10.1186/s40337-024-00966-5

Contribution to the Study

Author	Task contribution
An Dang	Conceptualisation; manuscript writing; editing; data analyses
Litza Kiropoulos	Editing
Marija Anderluh, David Collier, Fernando Fernandez-Aranda, Andreas Karwautz, and Gudrun Wagner	Data collection, editing
Isabel Krug	Conceptualisation; manuscript – reviewing and editing; supervision

Journal of Eating Disorders

RESEARCH

Open Access

Do risk factors differentiate DSM-5 and drive for thinness severity groups for anorexia nervosa?

An Binh Dang[1*], Litza Kiropoulos[1], Marija Anderluh[2], David Collier[3], Fernando Fernandez-Aranda[4], Andreas Karwautz[5], Janet Treasure[6], Gudrun Wagner[5] and Isabel Krug[1]

Abstract

Background The current study examined whether risk factors for anorexia nervosa (AN) were related to different levels of severity based on (a) the DSM-5/body mass index (BMI) and (b) drive for thinness (DT) severity ratings.

Methods The sample comprised 153 pairs of individuals with a lifetime diagnosis AN per DSM-IV criteria and their non-ED sisters ($N = 306$, mean age = 26.53; mean current BMI = 20.42 kg/m^2). The Oxford risk factor interview was used to establish AN-related risk factors. Individuals were categorised into the DSM-5 severity groups based on their lowest BMI, while the DT subscale from the eating disorder inventory-2 was used to classify individuals with AN into low and high DT groups.

Results Multinominal regression models showed similar risk factors (e.g., perfectionism, having a history of being teased about weight and shape) contributed to the development of AN using the DSM-5 and DT severity ratings. Follow-up analyses across the severity groups for both indices revealed that only childhood perfectionism was found to be more common in the extreme severe DSM-5 BMI severity group compared to the severe DSM-5 group.

Conclusion Overall, this study found little evidence for AN risk factors being related to the DSM-5 and DT severity ratings. However, given the novelty of this study, replication of the current results is warranted.

Keywords Anorexia nervosa, Risk factors, Severity rating, Drive for thinness, And body mass index

Plain English summary

Several risk factors, such as childhood obesity, have been found to contribute to the development of Anorexia Nervosa (AN). Yet, we are unsure if there is a set of risk factors that influence different levels of AN severity. While the DSM-5 suggests using BMI to measure severity, recent support favour the usage of drive for thinness (DT) as an alternative severity measure. Therefore, this study aimed to explore risk factors specifically associated with the development of different AN severity levels using both the DSM-5 BMI and DT severity classification systems. We recruited 153 pairs of individuals with a lifetime diagnosis AN per DSM-IV criteria and their non-ED sisters. The Oxford risk factor interview was used to establish AN-related risk factors. We found childhood perfectionism, weight/shape teasing, childhood obesity, and breast-related embarrassment to be significant risk factors for AN. Additionally, childhood perfectionism was more common in the extreme severe DSM-5 group compared to the severe DSM-5

*Correspondence:
An Binh Dang
dangb1@student.unimelb.edu.au
Full list of author information is available at the end of the article

group. This suggests that adding perfectionism-related aspects to prevention and early intervention programs for AN may be beneficial. Considering the novelty of this study, replication of the current results is needed.

Introduction

Risk factors ranging from genetic to psychological factors have been implicated in the aetiology of anorexia nervosa (AN) to inform prevention and treatment programs [1, 2]. To aid the clinical management of AN, the fifth edition of the Diagnostic and Statistical Manual (DSM-5) of Mental Disorders [3] presented a new severity rating for AN, which categorises individuals with AN into the mild, moderate, severe, and extreme severe severity groups based-on body-mass index (BMI). However, studies have found the clinical usefulness of the DSM-5 severity rating for AN to be limited [4, 5]. As such, researchers have explored alternative severity classifications for AN, including transdiagnostic indices (e.g., drive for thinness [DT]) and weight and shape concerns), and found these alternative severity ratings to be superior in indexing eating disorder (ED) psychopathology in comparisons to the DSM-5 BMI severity rating [4]. Despite many studies having been conducted on both AN risk factors [6] and severity ratings [4], researchers have yet to examine risk factors specifically associated with the development of different AN severity levels.

Risk factors for anorexia nervosa

Several risk factors have been implicated in the aetiology of AN, which include amongst others a family history of an ED [7, 8] or traits such as perfectionism, low self-esteem, or obsessive personality traits [8, 9]. However, most of these studies used self-report assessments, and are thus limited by recall biases [10]. The Oxford Risk Factor Interview (ORFI) [11], a semi-structured interview, has been considered the gold standard ED risk factor assessment tool. This is due to the ORFI's ability to assist participants in recalling information related to the autobiographic anchor point and the use of probes to encourage detailed descriptions of risk factor-related events (e.g., parental separation).

Most studies using the ORFI employed case–control designs, which compared individuals with AN to individuals with no DSM-5 ED diagnosis or no psychiatric disorders [6, 12]. These studies [6, 12] mainly found, similar to the studies using self-report measures [8], that perfectionism, negative self-evaluation, weight and shape concerns, and familial history of AN were risk factors associated with the development of AN. However, these ORFI case–control studies are limited by their inability to control for environmental and family factors that might contribute to the development of AN [13].

To overcome these limitations, studies have started employing samples of sister pairs discordant for AN, i.e., individuals with AN and their non-ED sisters. This design allows for control of cultural, environmental, and family factors so that individual-specific (i.e., non-shared) factors can be more distinctively assessed [14, 15]. Karwautz et al. [14, 15], employing this discordant sister pair design with 90 and 256 participants, respectively, found that individuals with AN, compared to their non-ED sisters, experienced higher rates of weight and shape teasing, exposure to critical comments, and disruptive life events. However, further research is required to elucidate the relationship between varying severity levels of AN and the associated risk factors of AN.

Combining AN risk factors and severity

While the medical literature has extensively investigated the relationships between risk factors and symptom severity in conditions like coronavirus (COVID-19) disease [16] and acute stroke [17], such exploration has been lacking in psychiatric research for conditions such as AN. For example, having a history of obesity and diabetes has been found to increase the risk of having severe COVID-19 (e.g., respiratory failure, intensive care unit admission [16] This means that the medical field has gained a certain understanding of how certain risk factors not only increase the risk of having a disease but also contribute to progressing to a greater severity (e.g., impairments, mortality). Using such information, medical professionals can proactively identify patients at a high risk of developing severe illnesses, which in turn allows them to inform medical decisions such as the need for hospitalisation or treatment frequency. However, in the ED literature, such a severity-risk factor approach has not yet been assessed.

DSM-5 severity rating for AN based on BMI

To date, the most well-known and widely used severity classification system for AN was proposed by the DSM-5 using BMI [4]. This DSM-5 severity rating categorises individuals with AN into "mild" ($\geq$17.0 kg/m^2), "moderate" (16–16.99 kg/m^2), "severe" (15–15.99 kg/m^2) and "extreme" ($<$15 kg/m^2) severity groups. However, the clinical utility of the BMI specifiers for AN has consistently been questioned by recent research [4, 5]. Dang et al. [4] conducted a meta-analysis (N=22) assessing

all the current DSM-5 severity rating studies for AN, bulimia nervosa (BN) and binge eating disorder (BED). In relation to AN, which comprised five studies [18–22], the review did not find significant differences in ED psychopathology across the DSM-5 BMI severity groups. This means that although BMI has been widely utilised in both clinical practice and empirical research to index severity [23], the level of evidence regarding the DSM-5 AN severity rating is still unknown. If the DSM-5 severity rating is valid, besides indexing the intensity of ED psychopathology, it should also be able to distinguish other clinical information such as the prognosis or factors involved with the development (i.e., risk factors) and maintenance of AN.

Drive for thinness as an alternative severity rating for AN

Given the ambiguity surrounding the DSM-5 AN severity rating, researchers have explored alternative severity classifications for AN such as overvaluation of weight and shape [5, 22], and DT [4, 24]. Studies have consistently found that individuals with AN who scored high on DT had more severe attitudinal and behavioural ED symptoms, as well as more psychiatric comorbidity compared to those with low DT [25, 26]. DT has been considered as another promising alternative severity rating for EDs in general and AN specifically. Krug et al. [24] found that DT provided more significant differences than the DSM-5 BMI severity groups in ED and general psychopathology, with the high DT group scoring higher on these variables than the low DT individuals. Given that previous research has only assessed the clinical utility of DT severity for AN on ED and general psychopathology [4], it is important to expand the scope of severity rating research to look at the relationship between AN-related risk factors and DT as an alternative severity indicator for AN.

The current study

In this study, by using a sister-pair discordant design (comprising individuals with a lifetime DSM-IV diagnosis of AN and a sister without an ED diagnosis), we aimed to identify specific risk factors linked to varying levels of AN severity according to either DSM-5 or DT severity rating. This extends the work of Karwautz et al. [15], who employed a similar design (and part of an overlapping sample) but focused on identifying general risk factors for AN development without delving into specific risk factors contributing to different AN severity levels per DSM-5 and DT ratings. To do so, we first assessed which ORFI risk factors were related to the development of AN based on both the DSM-5 BMI and DT severity ratings, while using non-ED sisters as controls to enhance the distinctiveness of individual factors. We then focused only on the significant risk factors from these primary analyses and examined whether there were significant differences in the frequency of the exposure of each of these significant risk factors among the DSM-5 BMI (mild/moderate, severe, and extreme severe) and DT (high vs. low) severity groups. Understanding such relationships would not only provide insights into the clinical utility of these two severity classification systems, but would also provide a guide for developing more proactive prevention programs and informing public clinical management policies for AN.

Method

Design

This study used a discordant sister pair design, which meant that women who met the diagnosis of AN according to DSM-IV criteria [27] were compared to their non-ED sisters. Individuals meeting the criteria for AN in the DSM-IV would generally still qualify for an AN diagnosis under the DSM-5 criteria [28]. This is because the DSM-5 broadens the AN criteria, removing specific weight thresholds and the amenorrhea requirement to encompass a wider spectrum of individuals with the disorder. This approach aligns with previous studies in this field [20]. The data for this study were collected as part of a multicentre European project.

Recruitment procedure

Four different sites specialising in EDs participated in this study: the Institute of Psychiatry, London, UK; the University of Vienna, Austria; the University Hospital of Bellvigte, Barcelona, Spain, and the University Children's Hospital, Ljubljana, Slovenia. Between 1999 to 2002, participants were recruited from both clinical settings and community resources such as websites and volunteer databases from previous research studies (for more information please refer to Giles et al. [27].

Participants

The current study comprised 153 pairs of individuals with AN and their non-ED sisters ($N=306$). Their combined mean age was 26.25 ($SD=8.00$), and their mean current BMI was 20.42 kg/m^2 ($SD=3.60$). Most participants were employed (43.9%). Among the individuals with AN, the mean age of illness onset and duration of illness was 16.98 years ($SD=4.92$) and 6.51 years (SD = 6.59) respectively.

Inclusion criteria for individuals with AN were: (1) female gender; (2) a lifetime diagnosis of AN according to DSM-IV criteria derived from the Longitudinal Interval Follow-up Evaluation [LIFE] interview [29]; and (3) having a sister close in age. The exclusion criteria for both individuals with AN and their non-ED sisters included

not having a current psychotic disorder and learning disability.

Non-ED sisters were included in the study if they did not have a diagnosis of an ED which was screened by the LIFE interview [29]. To control for environmental factors, the sister pairs needed to have an age gap of less than 10 years and must have lived together in the same family for at least 8 years. If the patient had more than one sister, the sister who was closest in age was included.

Measure

Clinical psychologists, with over 5 years of research experience in the field of EDs, conducted all interviews outlined below. These psychologists received training from senior ED clinicians based in different European countries. Although inter-rater reliability was not assessed in this study, prior research [12, 30] has consistently shown adequate reliability and validity for the semi-structured interviews (e.g., LIFE [29], ORFI [11]) employed in the current study.

Social demographic and clinical information

The first part of the EATATE [31], a semi-structured interview to assess ED symptoms, was used to derive information regarding participants' age, occupation, current height, current weight, and highest and lowest weight since reaching adulthood. From this information, the current, lowest, and highest BMI was calculated. In addition, current weight and height were measured at the ED units at intake.

Longitudinal interval follow-up evaluation (LIFE)

This study used an adapted version of the European LIFE [29] to measure lifetime ED history among participants with AN. This included constructing anchor points and timelines for the development of ED symptoms, symptom severity, and psychiatric treatment received. Non-ED sisters were also screened for any ED and ED-related features using the same LIFE interview This was done to ensure that the non-ED sisters have never met any ED diagnoses. This interview has demonstrated adequate reliability and validity in prior research (e.g., 31).

Oxford risk factor interview for eating disorders (ORFI)

The ORFI [11] was used to examine specific risk factors associated with the development of AN. At the beginning of the interview, the clinicians aimed to identify the period before the onset of AN, where the age of onset is defined as the time when the first significant persistent disordered eating pattern began [11, 12, 32]. Drawing from prior research [11, 32], a comprehensive array of potential risk factors was examined and categorised into three overarching domains: individual vulnerability, environmental risk, and dieting vulnerability. Each domain included several subdomains (e.g., childhood characteristics, childhood abuse), representing specific types of exposure within the main domain. For example, to assess negative self-evaluation, participants were asked, "In general, as a child or adolescent, how did you feel about yourself compared to other people? Did you feel that you were the same, better, or worse than the others?" Another example item to assess the risk of deliberate self-harm included "Have you ever tried to hurt yourself without the intention of committing suicide?". Exposure to a risk factor was rated on a 5-point scale ranging from 0 = *no exposure* to 4 = *high severity, long duration,* or *high frequency of exposure*. The data were then recoded into 0 = *no definite exposure* (initially coded 0, 1, or 2) versus 1 = *definite exposure* to reduce the likelihood of false positives, (initially coded 3 or 4). The ORFI has been found to have good inter-rater reliability with a high level of agreement across the risk factor domains (main weighted kappa: 0.66, SD: ± 0.17) [12]. For more information, please refer to Fairburn et al. [11].

Eating disorders inventory-2 (EDI-2)

The EDI-2 [33] is a 91-item self-report measure of cognitive and behavioural characteristics frequently associated with EDs. We used the EDI-2 DT subscale as an alternative severity rating for AN. EDI-2 DT subscale includes seven items (rating from 1 = never to 6 = always), which assesses excessive dieting anxiety, preoccupation with weight, and fear of weight gain (e.g., "I am preoccupied with the desire to lose weight"). The DT subscale has demonstrated excellent internal reliability with $\alpha = 0.85$ [24].

DSM-5 AN severity groups

This study recruited participants with a lifetime diagnosis of AN according to DSM-IV criteria, meaning that some individuals had already recovered from their ED. Therefore, to have an accurate representation of the level of AN severity that the participants had experienced during their illness, we used the lowest lifetime BMI after reaching 18 years old rather than the current BMI to categorise participants into the four BMI AN severity groups.

Using this minimum BMI severity index, we categorised the AN participants according to the four DSM-5 severity groups. However, since there were only a few individuals who were categorised in the mild and moderate groups [*mild* ($n = 10$), *moderate* ($n = 21$)], we had to reclassify the mild and the moderate AN DSM-5 severity groups into one "mild/moderate" group. This reclassification has been reported in previous studies [33] to allow a sufficiently large sample size across the groups to undertake meaningful comparisons. Overall, a total of

31 participants were classified as mild/moderate, 40 as severe and 82 as extreme severe using the DSM-5 severity ratings based on the lowest lifetime BMI.

Drive for thinness severity groups

The alternative DT severity rating system was derived from the EDI-2 [33] DT subscale. The DT subscale score was derived by summing the scores of all seven DT items. For screening purposes, Garner [34] recommended a cut-off score of > 14 to differentiate those with high DT from those with low DT, which subsequently has been applied across various research studies (e.g., 24). Using the same cut-off of 14, we classified individuals with AN into low DT ($n = 60$) and high DT ($n = 69$) individuals.

Statistical analysis

Analyses of variance (ANOVA) and chi-square analysis were used to compare sociodemographic and clinically related information between the non-ED sisters and the two AN severity indices (DSM-5 BMI and DT) for continuous and categorical variables respectively. Eta-squared coefficient (η^2) as the measurement of effect size for ANOVA (values of 0.06, 0.10, and 0.25 were interpreted as low-poor, moderate-medium, and large-high effect sizes respectively; 36). Cramer's V coefficient was used as a measurement of effect size for chi-square analyses (values of 0.06; 0.15, and 0.30 were interpreted as low-poor, moderate-medium, and large-high effect sizes respectively [35]). Each risk factor was coded as 0 for "no – the risk was not presented", and 1 for "yes – the risk was present".

Multinominal logistic regression analyses with "non-ED sisters" as the reference group were employed to investigate which risk factors were significantly associated with higher odds of developing the different AN severity categories based on the DSM-5 BMI and DT severity ratings. To account for multiple testing, we applied the Benjamini and Hochberg correction for multiple comparison method, where *p*-values were ranked, and then adjusted by multiplying each by a factor of m/k, with m representing the number of independent tests and k being the position of a *p*-value in the sorted array [36]. The chosen alpha level was 0.05. The results of the significant tests (*p*-value) and the effect-size estimates were utilised to evaluate the validity of both the DSM-5 BMI and the alternative DT severity ratings for AN.

The odds ratio (OR) was used as the measurement of the effects of being exposed to AN-related risk factors and the development of different levels of AN severity based on the DSM-5 BMI and DT severity ratings. An OR of greater than 1.00 indicates that exposure was associated with higher odds of developing AN. With an OR of less than 1.0 is associated with lower odds of developing AN. Following Chen's [37] rules-of-thumb OR of 1.68, 3.47, and 6.71 are considered as small, medium, and large effect-size respectively. Chi-square tests were conducted to test for the differences in the rate of being exposed to each significant risk factor for the DSM-5 BMI and DT severity groups.

Result

Sociodemographic and clinical-related information

The sociodemographic and clinical-related information is displayed in Table 1. As expected, there were significant differences in lowest ($p < 0.001$), current ($p < 0.001$), and highest ($p < 0.001$) BMI between individuals with AN and their non-ED sisters. There were no significant differences in age between non-ED sisters and individuals with AN. Further information on sociodemographic of the sample can be found in Additional file 1: Table S1.

Using the DSM-5 BMI severity rating for AN, significant differences between the BMI severity groups were found for the age of illness onset ($p = 0.029$). However, post-hoc comparisons did not reveal where the differences between the groups lay (Table 1). No other significant differences emerged in the sociodemographic and clinically related information between the DSM-5 BMI severity groups. For the DT groups, no significant differences in any of the sociodemographic (Additional file 1: Table S1) and clinically related information were observed.

Exposure to AN-related risk factors and odds of being in the DSM-5 severity groups

Using non-ED sisters as a reference group, multinomial logistic regression models (Table 2) showed that the following factors increase the risk of developing an AN diagnosis across the DSM-5 BMI severity groups: childhood perfectionism (mild/moderate, OR = 6.67; severe, OR = 3.13; extreme severe, OR = 10.00); childhood obesity (mild/moderate, OR = 25.00; severe, OR = 20.00; extreme severe, OR = 16.67); having a history of being teased about weight, shape, or appearance (mild/moderate, OR = 4.50; severe, OR = 3.40; extreme severe, OR = 3.86); and having a history of feeling embarrassment about one's breasts (mild/moderate, OR = 7.14; severe, OR = 5.00; extreme severe, OR = 4.35).

Follow-up chi-square analyses were conducted on the risk factors that significantly contributed to the development of AN based on the DSM-5 BMI severity classification (Table 3). Among all the significant risk factors, significant differences between the DSM-5 BMI severity groups were only obtained for childhood perfectionism ($p = 0.039$), with individuals in the extreme severe DSM-5 BMI severity group presenting with more childhood perfectionism compared to individuals in the severe DSM-5

Table 1 Social and clinical related information of non-ED sisters and individuals with AN

		Comparisons between non-ED sisters and individuals with AN			Comparisons between DSM-5 severity groups				Comparisons between Drive for thinness (DT) severity groups		
	Total Mean (SD)	**Non-ED sisters**	**Individuals with AN**	***P* values (η^2)**	**Mild/Moderate**	**Severe**	**Extreme Severe**	***P* values (η^2)**	**Low DT**	**High DT**	***P* values (η^2)**
Clinical information											
Age (yrs-old)	26.25 (8.00)	26.55 (8.30)	25.97 (7.73)	0.516 (0.001)	26.72 (6.67)	25.76 (6.05)	25.78 (8.81)	0.826 (0.003)	25.19 (6.72)	26.30 (7.78)	0.156 (0.008)
Onset ED (yrs-old)	16.98 (4.92)	–	16.98 (4.92)	–	17.71 (5.01)	17.41 (4.18)	15.82 (3.36)	**0.029 (0.046)**	16.52 (4.30)	16.50 (3.82)	0.713 (0.001)
Duration ED (yrs)	6.51 (6.59)	–	6.51 (6.59)	–	6.21 (6.86)	5.36 (3.97)	7.21 (7.11)	0.314 (0.015)	6.13 (6.07)	6.93 (6.11)	0.465 (0.004)
Current BMI (kg/m^2)	20.42 (3.60)	22.08 (3.90)	18.74 (2.41)	**< 0.001 (0.217)**	19.30 (1.52)	19.30 (2.30)	18.31 (2.68)	0.063 (0.036)	18.63 (2.34)	19.07 (2.47)	0.078 (0.013)
Lowest BMI (kg/m^2)	17.11 (3.61)	19.81 (2.82)	14.48 (2.00)	**< 0.001 (0.547)**	17.04 (1.17)	15.45 (0.30)	13.04 (1.36)	**< 0.001 (0.734)**	14.47 (2.18)	14.50 (1.81)	0.933 (0.001)
Highest BMI (kg/m^2)	22.58 (4.19)	23.58 (4.78)	21.56 (3.20)	**< 0.001 (0.059)**	22.22 (2.80)	21.17 (3.65)	21.02 (3.03)	0.082 (0.033)	21.12 (3.18)	21.62 (3.18)	0.075 (0.013)
Social information				*P* values (V)				*P* values (V)			*P* values (V)

Bold significant comparison using adjusted p-value using the Benjamini-Hochberg false discovery rate procedure (.05 level)

Table 2 Distribution of risk factors for across the mild/moderate, the severe and the extreme severe AN DSM-5/BMI severity groups, and high and low drive for thinness groups results using multinominal logistic regression, with non-ED sisters as the reference group

Severity groups	DSM-5 Mild/Moderate		DSM-5 Severe		DSM-5 Extreme		Low DT		High DT	
	OR [95% CI]	*p*	OR [95% CI]	*p*	OR [95% CI]	*p*	OR [95% CI]	*p*	OR [95% CI]	*p*
Individual vulnerability domain										
Subdomain 1 (childhood characteristic)										
Negative self-evaluation	0.58 [0.26; 1.35]	0.455	1.15 [0.55; 2.44]	0.885	1.10 [0.59; 2.04]	0.890	0.62 [0.31; 1.22]	0.415	1.69 [0.84; 3.45]	0.391
Perfectionism	6.67 [2.33; 20.00]	**0.006**	3.13 [1.20; 8.33]	**0.003**	10.00 [4.55; 20.00]	< **0.001**	8.26 [3.57; 19.23]	< **0.001**	6.67 [2.86; 14.29]	< **0.0001**
Shyness	1.96 [0.83; 4.76]	0.372	1.15 [0.53; 2.50]	0.906	0.91 [0.47; 1.79]	0.897	0.90 [0.43; 1.87]	0.894	1.39 [0.69; 2.86]	0.631
Extreme compliance	1.05 [0.42; 2.63]	0.952	1.89 [0.87; 4.00]	0.366	1.52 [0.79; 2.86]	0.451	1.27 [0.62; 2.60]	0.771	1.75 [0.87; 3.57]	0.359
No close friends	0.42 [0.13; 1.37]	0.407	1.01 [0.43; 2.38]	0.978	0.46 [0.21; 1.02]	0.232	1.03 [0.46; 2.32]	0.951	0.38 [0.15; 0.93]	0.169
No opposite gender friends	0.43 [0.15; 1.25]	0.365	0.55 [0.23; 1.30]	0.416	1.10 [0.55; 2.17]	0.895	0.67 [0.30; 1.49]	0.610	0.78 [0.36; 1.69]	0.773
School absence through anxiety	–	–	0.11 [0.01; 0.88]	0.178	0.14 [0.03; 0.66]	0.089	–	–	0.27 [0.07; 1.00]	0.216
Subdomain 2 (premorbid psychiatric disorder)										
Major depression	1.89 [0.75; 4.76]	0.419	2.94 [1.35; 6.67]	0.051	2.00 [1.05; 3.85]	0.180	2.04 [0.97; 4.17]	0.232	2.50 [1.22; 5.00]	0.087
Anxiety disorders	1.23 [0.51; 3.03]	0.839	2.04 [0.96; 4.35]	0.239	1.47 [0.80; 2.63]	0.444	1.19 [0.60; 2.38]	0.827	1.96 [1.01; 3.85]	0.207
Substance abuse	0.65 [0.07; 6.25]	0.889	0.48 [0.55; 4.55]	0.774	1.23 [0.31; 5.00]	0.894	0.74 [0.07; 7.14]	0.883	3.33 [0.75; 14.29]	0.361
Soft drug abuse	3.57 [1.43; 9.09]	0.057	2.13 [0.85; 5.26]	0.368	1.23 [0.55; 2.78]	0.827	0.88 [0.34; 2.27]	0.884	2.13 [0.94; 4.76]	0.249
Subdomain 3 (behavioral problems)										
Marked conduct problem	1.11 [0.80; 100]	0.263	2.94 [0.15; 50.00]	0.734	1.59 [0.09; 25.00]	0.893	–	–	5.56 [0.49; 50.00]	0.411
School absence (truancy)	0.72 [0.10; 5.26]	0.889	1.41 [0.32; 6.25]	0.846	1.23 [0.38; 4.00]	0.897	0.64 [0.13; 3.13]	0.813	1.67 [0.51; 5.56]	0.678
Deliberate self-harm	7.14 [1.75; 25.00]	0.053	5.56 [1.39; 20.00]	0.097	4.17 [1.22; 14.29]	0.127	4.17 [0.97; 16.67]	0.233	8.33 [2.08; 33.33]	**0.030**
Subdomain 4 (physical health)										
Severe personal health problems	2.22 [0.83; 5.88]	0.363	1.37 [0.51; 3.57]	0.781	1.33 [0.62; 2.94]	0.718	1.56 [0.68; 3.70]	0.558	2.00 [0.85; 4.55]	0.362
Childhood obesity	25.00 [7.14; 100.00]	< **0.001**	20.00 [5.88; 100.00]	**< 0.001**	16.67 [4.76; 50.00]	< **0.001**	25.00 [5.26; 100.00]	**0.001**	33.33 [7.69; 100.00]	< **0.001**
Subdomain 5 (parental psychiatric disorder)										
Parental depression	1.14 [0.40; 3.23]	0.877	0.68 [0.23; 2.00]	0.733	1.38 [0.68; 2.78]	0.650	1.82 [0.83; 4.00]	0.377	0.82 [0.34; 2.00]	0.849
Parental anxiety	0.56 [0.06; 4.76]	0.811	2.38 [0.65; 8.33]	0.435	2.08 [0.75; 5.88]	0.419	0.82 [0.20; 3.45]	0.891	2.33 [0.70; 7.69]	0.418
Parental drug abuse	1.69 [0.55; 5.26]	0.635	0.75 [0.20; 2.86]	0.853	1.03 [0.43; 2.50]	0.961	0.71 [0.25; 2.00]	0.775	1.23 [0.47; 3.23]	0.848
Environmental Risk Domain										
Subdomain 1 (parental problem)										
Distress due to low parental contact	0.76 [0.15; 3.85]	0.892	1.18 [0.35; 3.85]	0.886	1.82 [0.72; 4.55]	0.462	0.65 [0.19; 2.27]	0.749	1.55 [0.58; 4.15]	0.651
Distress due to separation from parents	0.84 [0.16; 4.55]	0.890	0.59 [0.11; 3.23]	0.782	0.43 [0.12; 1.56]	0.447	0.35 [0.07; 1.81]	0.442	0.57 [0.14; 2.36]	0.713
Distress due to parental arguments	2.13 [0.92; 5.00]	0.284	0.99 [0.44; 2.22]	0.984	2.63 [1.47; 4.76]	0.142	3.31 [1.70; 6.45]	0.006	1.08 [0.55; 2.12]	0.877
Parental criticisms	1.54 [0.55; 4.35]	0.682	1.03 [0.38; 2.78]	0.968	1.45 [0.69; 3.03]	0.609	1.11 [0.47; 2.60]	0.887	1.56 [0.69; 3.56]	0.558
Parental high expectation	1.39 [0.59; 3.23]	0.724	0.74 [0.33; 1.64]	0.726	1.53 [0.85; 2.78]	0.419	1.89 [0.97 3.70]	0.239	0.92 [0.46; 1.82]	0.885

Table 2 (continued)

Severity groups	DSM-5 Mild/Moderate		DSM-5 Severe		DSM-5 Extreme		Low DT		High DT	
	OR [95% CI]	*p*	OR [95% CI]	*p*	OR [95% CI]	*p*	OR [95% CI]	*p*	OR [95% CI]	*p*
Parental over-involvement	0.88 [0.33; 2.38]	0.880	1.18 [0.46; 3.03]	0.902	0.84 [0.41; 1.72]	0.837	0.74 [0.32; 1.70]	0.731	1.21 [0.56; 2.63]	0.836
Parental neglect	0.31 [0.03; 2.86]	0.564	4.35 [1.37; 14.29]	0.092	0.56 [0.17; 1.92]	0.639	0.69 [0.18; 2.60]	0.813	1.77 [0.63; 5.00]	0.547
Parental affection	0.75 [0.29; 1.96]	0.790	1.39 [0.58; 3.33]	0.719	0.70 [0.35; 1.37]	0.528	0.76 [0.36; 1.62]	0.479	1.32 [0.63; 2.75]	0.730
Subdomain 2 (disruptive events)										
Distress due to parental death	3.33 [0.51; 20.00]	0.451	2.00 [0.30; 12.50]	0.733	1.27 [0.20; 7.69]	0.886	0.73 [0.07; 7.41]	0.889	3.68 [0.81; 16.67]	0.314
Distress due to parental chronic illness	1.20 [0.35; 4.00]	0.894	2.13 [0,81; 5.56]	0.375	0.95 [0.38; 2.44]	0.947	1.17 [0.44; 3.11]	0.892	1.51 [0.60; 3.77]	0.655
Distress due to frequent house moves	2.00 [0.37; 11.11]	0.691	3.13 [0.78; 12.50]	0.364	3.33 [1.04; 10.00]	0.193	2.47 [0.68; 8.93]	0.422	2.29 [0.62; 8.47]	0.444
Subdomain 3 (teasing and bullying)										
Teasing (not concerning weight, shape, eating or appearance)	0.84 [0.23; 3.06]	0.894	1.37 [0.50; 3.75]	0.779	2.29 [1.12; 4.67]	0.123	1.27 [0.53; 3.04]	0.815	1.88 [0.83; 4.27]	0.372
Bullying	0.24 [0.05; 1.04]	0.237	0.27 [0.08; 0.92]	0.181	0.38 [0.17; 0.85]	0.113	0.35 [0.14; 0.89]	0.147	0.22 [0.07; 0.66]	0.055
Subdomain 4 (sexual and physical abuse)										
Sexual abuse	1.31 [0.53; 3.25]	0.795	2.39 [1.13; 5.08]	0.130	1.34 [0.72; 2.51]	0.633	1.41 [0.70; 2.82]	0.618	1.36 [0.68; 2.72]	0.653
Physical abuse	0.57 [0.18; 1.79]	0.614	1.22 [0.53; 2.82]	0.842	0.92 [0.46; 1.84]	0.883	0.72 [0.32; 1.63]	0.707	1.35 [0.65; 2.79]	0.686
Dieting Vulnerability Domain										
Subdomain 1 (dieting risk)										
Influenced by parent eating disorders	1.63 [0.27; 9.71]	0.810	1.35 [0.24; 7.69]	0.890	0.30 [0.03; 2.81]	0.562	1.13 [0.23; 5.46]	0.924	0.84 [0.15; 4.69]	0.889
Family member dieting for shape and weight	1.96 [0.76; 5.03]	0.419	1.05 [0.43; 2.58]	0.953	0.89 [0.42; 1.87]	0.895	0.91 [0.40; 2.07]	0.873	1.15 [0.53; 2.50]	0.893
Critical comments by family about weight, shape or eating	1.05 [0.45; 2.46]	0.947	1.77 [0.84; 3.76]	0.373	1.19 [0.65; 2.15]	0.811	1.16 [0.59; 2.27]	0.857	1.35 [0.70; 2.62]	0.644
Repeated comments by others about shape and weight	0.75 [0.24; 2.33]	0.832	1.74 [0.73; 4.17]	0.446	1.56 [0.76; 3.19]	0.462	1.18 [0.51; 2.74]	0.886	1.61 [0.73; 3.53]	0.472
Teasing about shape, weight, or appearance	4.50 [1.86; 10.99]	0.**010**	3.40 [1.48; 7.81]	**0.038**	3.86 [1.95; 7.63]	**0.002**	4.93 [2.40; 10.10]	<**0.001**	3.61 [1.75; 7.46]	**0.007**
Dieting at school common	2.48 [0.76; 8.06]	0.379	2.20 [0.73; 6.58]	0.416	2.04 [0.83; 5.03]	0.361	2.23 [0.83; 5.95]	0.367	1.82 [0.66; 5.05]	0.495
Subdomain 2 (pudicity domain)										
Preparation for menarche	0.45 [0.14; 1.43]	0.421	0.86 [0.36; 2.04]	0.895	0.64 [0.31; 1.33]	0.476	0.96 [0.43; 2.15]	0.950	0.57 [0.24; 1.37]	0.453
Menarche mainly unpleasant	2.86 [0.96; 8.33]	0.236	2.04 [0.74; 5.88]	0.416	3.03 [1.32; 6.67]	0.066	3.32 [1.22; 9.01]	0.110	4.90 [1.96; 12.35]	**0.008**
Scared by early breast development	0.29 [0.09; 0.93]	0.182	0.56 [0.23; 1.35]	0.449	0.41 [0.19; 0.85]	0.113	0.30 [0.13; 0.73]	0.061	0.39 [0.17; 0.89]	0.458
Teasing about breast as early age	0.32 [0.10; 1.09]	0.254	0.51 [0.19; 1.32]	0.417	0.22 [0.09; 0.57]	**0.021**	–	–	0.40 [0.15; 1.01]	0.233
Breast as source of embarrassment	7.14 [2.78; 16.67]	**0.001**	5.00 [2.08; 11.11]	**0.004**	4.35 [2.13; 9.09]	**0.001**	3.39 [1.48; 7.75]	**0.037**	5.41 [2.48; 11.76]	**0.001**

Adjusted p-value using the Benjamini-Hochberg false discovery rate procedure

Bold significant comparison using adjusted p-value using the Benjamini-Hochberg false discovery rate procedure (.05 level)

Table 3 Follow-up chi-square analyses on the risk factors that significantly contributed to the development of AN based on the DSM/BT and DT classification

	DSM-5								DT				
	Mild/moderate (1)		Severe (2)		Extreme (3)		χ^2 (*p* value; Cramer's V)	Post-hoc	Low		High		χ^2 (*p* value; Cramer's V)
	N	%	*n*	%	*n*	%			*n*	%	*n*	%	
Perfectionism	11	17.2	11	17.2	42	65.6	6.86 (0.032; 0.212)	3 > 2 (.039)	31	47	35	53	0.43 (0.515; 0.042)
Childhood obesity	11	25.6	12	27.9	20	46.5	1.45 (0.480; 0.098)	–	15	40.5	22	59.5	1.70 (0.193; 0.083)
Teasing about shape, weight, or appearance	13	23.2	14	25	29	51.8	0.48 (0.787; 0.056)	–	34	50	34	50	0.01 (0.937; 0.005)
Breast as source of embarrassment	4	15.4	9	34.6	13	50	1.30 (0.521; 0.092)	–	22	39.3	34	60.7	3.59 (0.058; 0.121)

BMI severity group. No other significant differences between the DSM-5 BMI severity groups in the rate of being exposed to the remaining significant AN risk factors were found.

Exposure to AN-related risk factors and odds of being in the high/low DT severity groups

For DT, multinomial logistic regression models (using non-ED sisters as the comparison group to either low or high DT) found childhood perfectionism (low DT, OR = 8.26; high DT, OR = 6.67); childhood obesity (low DT, OR = 25.00; high DT, OR = 33.33); having a history of being teased about weight, shape, or appearance (low DT, OR = 4.93; high DT, OR = 3.61); and feeling embarrassed about breasts (low DT, OR = 3.39; high DT, OR = 5.41) to be significant risk factors for both the low and high DT groups. Conversely, in comparison to non-ED sisters, having a history of self-harm was found to only increase the risk of being assigned to the high DT AN severity group (OR = 8.33), but not the low DT group. Additionally, in comparison to non-ED sisters, those in the low DT group were more likely to feel distress due to parental arguments (OR = 3.31; Table 2).

Follow-up chi-square analyses were conducted on the risk factors that significantly contributed to the development of AN based on the DT classification (Table 3). However, these analyses showed no significant differences between the low and high DT groups in the frequency of exposure to each of these risk factors.

Discussion

Using a discordant AN sister pair design, the current study identified risk factors that may contribute to the development of different AN severity levels based on the DSM-5 BMI and DT severity classification systems. Several factors (e.g., perfectionism, teasing about weight and shape – see details below) were found to contribute to the development of AN based on both severity indices. However, when looking at the differences in the rate of being exposed to these significant risk factors across the DSM-5 BMI and DT severity groups, besides childhood perfectionism, no other significant risk factor was found.

Risk factors for AN based on the DSM-5 and DT severity indices

Comparisons between non-ED sisters and AN DSM-5 and DT severity groups

Compared to non-ED sisters, we found that almost the same AN risk factors were associated with the severity spectrum of the DSM-5 BMI and DT classification systems. In accordance with the literature, we found that childhood obesity [38, 39], perfectionism [39, 40]; being teased about weight and shape [41] and feeling embarrassed about breasts [42] were associated with having a subsequent diagnosis of AN using these two severity indices.

This overlap in AN risk factors across the DSM-5 BMI and DT severity indices may imply that these two severity classification systems could be correlated with each other. However, in the current study, opposite patterns in the strength of effects across the two severity ratings were found for several risk factors (e.g., childhood obesity). For example, in comparison to non-ED sisters, childhood obesity was associated with a higher risk of being classified into the mild/moderate DSM-5 BMI severity groups. Contrastingly, those reported having childhood obesity were found to have a higher likelihood of being classified into the high DT (as opposed to low DT) severity group.

These patterns of results are somewhat consistent with the literature [43, 44]. Specifically, studies have shown that a higher childhood BMI was directly associated with an elevated adulthood BMI (i.e., mild/moderate DSM-5 BMI severity) and indirectly through both DT in childhood and adulthood (high DT group; 44). This may suggest that different underlying mechanisms might mediate the relationships between AN risk factors and AN severity levels based on these two severity classifications. Further research using longitudinal designs is needed to establish a temporal relationship of the potential different mechanisms that might lead to the different DSM-5 BMI and DT severity groups when compared to a non-ED control sample.

Differences across DSM-5 and DT severity groups in the significant risk factors

Across all DSM-5 and DT severity groups, those with an AN diagnosis were more likely to exhibit childhood perfectionism. However, when assessing differences across the severity groups, only those in the extreme DSM-5 BMI severity group were significantly more likely to present with childhood perfectionism compared to those in the DSM-5 severe group. This finding aligns with previous studies [45, 46], which found that comorbid perfectionistic tendencies in adulthood were higher in AN individuals with a lower BMI (i.e., severe AN group) than those with a higher BMI (i.e., mild severity AN group). Contrastingly, a recent study by Krug et al. [24] did not find significant differences in adult clinical perfectionism across both the DSM-5 BMI and DT severity groups among individuals with AN. These discrepancies in findings may be either due to the heterogeneous samples across these studies or may have been confounded by the level of care obtained by individuals across these studies. For instance, Dakanalis et al. [45] assessed an inpatient clinical AN sample, while both Krug et al. [24] and the current study comprised a mixture of inpatient,

outpatient, and community participants. Another thing to consider is that the current study assessed childhood and not adulthood perfectionism, despite perfectionism being a stable trait [47], and childhood perfectionism is strongly associated with adulthood perfectionism [48].

No significant differences between either the DSM-5 BMI or DT AN severity groups in the rates of being exposed to the remaining significant risk factors were revealed in our follow-up analyses. Of interest, no significant differences in the frequency of childhood obesity across both the DSM-5 BMI and DT AN severity groups were found. Such findings are unexpected because in the current study childhood obesity was the strongest significant predictor for a subsequent AN diagnosis (OR ranges from 17 to 25). Furthermore, studies have consistently found that individuals with AN with premorbid obesity presented with significantly more severe physical (e.g., cardiovascular failure) and psychological sequelae, and poorer prognosis compared to AN individuals with premorbid normal weight [49, 50]. The current non-significant findings might be attributable to a small sample size across the three DSM-5 BMI and the two DT groups, lacking sufficient power to detect differences across severity groups.

Only Dang et al. [4] and Krug et al. [24] have compared the DSM-5 with the DT severity ratings for all EDs. Findings from both studies support the clinical utility of the DT severity rating over the DSM-5 in indexing ED and general psychopathology across most ED subtypes. In the current study, our regression analyses, however, did not reveal any significant association between the evaluated risk factors and the severity of AN, as determined by either the DSM-5 BMI or DT ratings. Previous studies [5, 22, 24] examined the clinical validity (i.e., significant differences in the level of symptomatology across severity groups), whereas the current study focused on the retrospective validity (i.e., ability to distinguish risk factors that are related to the development of different AN severity levels) of these severity ratings. Therefore, the DT severity rating for AN might only be superior to the DSM-5 BMI severity rating in distinguishing the level of AN-related concurrent symptomatology [4, 24], but may not be as effective in capturing retrospectively assessed risk factors. Nevertheless, this study was the first to examine the relationship between the DSM-5 BMI and DT severity indices and AN risk factors, hence replication of these results is needed.

Implications

The staging model by Treasure et al. [51] can be applied to develop prevention and early treatment programs for AN. Based on our findings such an approach should use two steps. First, it would be important to identify individuals at risk of developing AN where significant AN risk factors (i.e., as per the current results, childhood obesity, being teased about weight, shape, and appearance, and feeling embarrassed about breasts) occur. It might be beneficial for these risk factors to be included in prevention programs for AN across all severity spectrums. As shown to be beneficial in previous research [52], cognitive dissonance-based interventions could aim to increase individuals' appreciation of different body types, reducing pressure to be thin, and body dissatisfaction.

Second, we found that individuals with AN who exhibited perfectionistic traits during childhood were more likely to develop extreme severe AN per the DSM-5 BMI severity rating in adulthood. Such findings, therefore, underscore the importance of incorporating perfectionism in early prevention programs to prevent this trait from intensifying throughout adolescence and adulthood. While it might be advantageous to focus on perfectionism across all AN severity levels (e.g., identifying perfectionism triggers), implementing more comprehensive and targeted early perfectionism interventions (e.g., cognitive remediation therapy [CRT]; [53] might be highly beneficial for individuals who are at risk of developing extreme severe AN. Accordingly, CRT has demonstrated promising results in improving ED symptomatology among individuals with severe and enduring AN [54, 55]. Given the novelty of our current results, future research is warranted to explore the efficacy of additional perfectionism interventions, including CRT modules, in mitigating the progression of AN severity.

Finally, except for perfectionism, we found no other risk factors linked to varying DSM-5 AN severity levels, once again casting doubt on the validity of the proposed BMI-based severity rating by the DSM-5. While recognising the importance of severity ratings in prevention strategies and treatment planning, relying solely on BMI may oversimplify the complex nature of AN, encompassing elements such as fear of weight gain and body distortions [56]. This critique does not dismiss the use of a severity index for AN but underscores the necessity for an alternative perspective based on empirical data. Subsequent research should explore different indicators, such as duration of illness or perfectionism, to refine severity assessments, hence enhancing the efficacy of prevention and treatment strategies. Notably, shorter durations of illness have been proposed as a valuable severity indicator for AN [57, 58], correlating with higher BMI [57], more favourable treatment outcomes [58, 59], lower levels of perfectionism [60], and increased cognitive flexibility [61].

Strengths and limitations

A strength of this research was our discordant case–control interview-based design to identify a wide range of psychosocial risk factors for AN. This means the current study was able to control for environmental factors which allowed to establish individual risk factors.

The current study also has a few limitations that need to be acknowledged. First, along with the exploratory nature of our research, we needed to categorise our difficult-to-recruit 153 AN sample into three DSM-5 BMI and two DT groups. This may not have provided sufficient power to detect a significant relationship between the assessed risk factors and the different DSM-5 severity levels. Future studies should employ a larger sample to ensure a more balanced distribution across the different DMS-5 BMI severity groups to gain a more accurate insight into the relationship between AN severity groups and risk factors.

Second, despite using the ORFI [11] to maximise reporting accuracy, potential recall biases are inherent to retrospective reporting. We did not include other informants in assisting with reporting family history but rather solely relying on the family history stated by the participants. Future research should include family members in the study process to bypass the inherent recall bias of retrospective assessment. Future studies should also consider a longitudinal design to assess the temporal relationship of the ORFI risk factors using different informants.

Third, our study departed from DSM-5 guidelines, categorising AN severity based on lowest BMI instead of current BMI. The lowest BMI was utilised because we recruited participants with a lifetime AN diagnosis, potentially including recovered and non-recovered cases. The inclusion of both groups poses a confounding factor, as non-recovered individuals might reach lower BMIs, leading to shifts between DSM-5 severity groups. Furthermore, the divergence between the DSM-5 severity groups based on lowest BMI and our DT severity groups, which were based on DT levels at the time of assessment, alongside with having a mixture of inpatient, outpatient, and community-based individuals with AN recruited across Europe might have confounded the results of the current study. Therefore, future studies should use the current BMI as per the DSM-5, instead of the lowest BMI, to examine the relationship between AN severity groups and risk factors in a sample of individuals with a current DSM-5 diagnosis of AN.

Fourth, our study exclusively included a female sample, while research has shown that different risk factors contribute to the development of AN in males and females [62]. As a result, the findings of the current study cannot be generalised beyond female-specific sample.

Lastly, in highlighting childhood perfectionism as a significant risk factor for extreme severe AN based on the DSM-5 severity rating, it is crucial to acknowledge a potential limitation in the scope of this finding. The DSM-5 AN severity level can shift from DSM-5 mild to DSM-5 extreme with just a one-point change in BMI. The DSM-5 extreme category, on the other hand, has a broad range (e.g., our participants' BMI in the extreme group had BMI ranged from 8.6 to 15 kg/m^2), highlighting potential heterogeneity within the DSM-5 extreme severe group that requires a more detailed exploration for certainty about our study's implications. Future research should specify the DSM-5 extreme severe group into subgroups (e.g., very extreme severe DSM-5 group) for a detailed understanding of diverse characteristics and implications within the extreme severity spectrum.

Conclusion

Our findings provided limited support for the DSM-5 BMI and DT AN severity ratings in identifying related risk factors for varying AN severity levels. However, BMI is an objective physical severity measure whereas DT is a cognitive measure of severity. In treatment, weight changes rapidly (especially in youth), but cognitive symptoms linger, highlighting the need to consider their distinct characteristics within a treatment context. Regardless, the current findings do highlight the significance of childhood perfectionism as a risk factor in developing extreme severe AN according to the DSM-5 BMI severity rating. It may be beneficial for prevention and early intervention programs to include a focus on managing perfectionism to reduce the severity of AN. As our study is the first of its kind, further replication, using several informants and longitudinal designs is needed to provide further insight into the relationship between risk factors for AN and DSM-5 BMI and other alternative severity ratings systems, including weight and shape concerns and, the number of purging methods.

Abbreviations

AN	Anorexia nervosa
ANOVA	Analyses of variance
BED	Binge eating disorder
BMI	Body mass index
BN	Bulimia nervosa
CRT	Cognitive remediation therapy
DSM	Diagnostic statistical manual
DT	Drive for thinness
ED	Eating disorder
EDI-2	Eating disorder inventory-2
LIFE	Longitudinal interval follow-up evaluation
OR	Odd ratio
ORFI	Oxford risk factor interview

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40337-024-00966-5.

Additional file 1. Social and clinical related information of healthy sisters and individuals with AN.

Acknowledgements
Not applicable.

Author contributions
AD: Conceptualization, Methodology, formal analyses, investigation, writing—original draft. LK: Writing—review and editing, supervision. MA, DC, FFA, AK, JT and GW: Data collection, writing—review and editing. IK: Methodology, writing—review and editing, supervision, project administration.

Funding
Financial support was received from the European Union (Framework-V Multicenter Research Grant, QCK1-1999-916), Fondo de Investigacion Sanitaria (CIBER; CB06/03/0034) and Generalitat de Catalunya (2005SGR00322). The CIBER Fisiopatologia de la Obesidad y Nutricion is an initiative of ISCIII. All these funding sources had no further role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Availability of data and materials
All data analysed during this study are included in this article (tables/figures). Raw data can be requested on reasonable requests.

Declarations

Ethical approval and consent to participate
Ethical approvals were obtained from all participated hospitals and research organisations. Consents were obtained from all participants in accordance with the Declaration of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Melbourne School of Psychological Sciences, The University of Melbourne, Redmond Barry Building, Level 7, Room 707, Melbourne, VIC, Australia. [2]University Children's Hospital, University Medical Center Ljubljana SI, Ljubljana, Slovenia. [3]Eating Disorders Unit and SGDP Research Centre, Institute of Psychiatry, King's College, London, UK. [4]Department of Psychiatry, University Hospital of Bellvitge- IDIBELL, Barcelona, Spain. [5]Department of Child and Adolescent Psychiatry, Medical University of Vienna, Vienna, Austria. [6]Section of Eating Disorders, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

Received: 19 August 2023 Accepted: 5 January 2024
Published online: 11 January 2024

References

1. Jacobi C. Psychosocial risk factors for eating disorders. Eat Disord Rev. 2005;1:59–84.
2. Stice E, Whitenton K. Risk factors for body dissatisfaction in adolescent girls: a longitudinal investigation. Dev Psychol. 2002;38(5):669.
3. American Psychiatric Association D, Association AP. Diagnostic and statistical manual of mental disorders: DSM-5. Washington, DC: American Psychiatric Association; 2013.
4. Dang AB, Giles S, Fuller-Tyszkiewicz M, Kiropoulos L, Krug I. A systematic review and meta-analysis on the DSM-5 severity ratings for eating disorders. Clin Psychol Sci Pract. 2022;29(4):325–44.
5. Dang AB, Kiropoulos L, Castle DJ, Jenkins Z, Phillipou A, Rossell SL, et al. Assessing severity in anorexia nervosa: do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates? Eur Eat Disord Rev. 2023;31(4):447–61.
6. Machado BC, Gonçalves SF, Martins C, Hoek HW, Machado PP. Risk factors and antecedent life events in the development of anorexia nervosa: a Portuguese case-control study. Eur Eat Disord Rev. 2014;22(4):243–51.
7. Stice E, Van Ryzin MJ. A prospective test of the temporal sequencing of risk factor emergence in the dual pathway model of eating disorders. J Abnorm Psychol. 2019;128(2):119.
8. Bulik CM, Slof-Op't Landt MC, van Furth EF, Sullivan PF. The genetics of anorexia nervosa. Annu Rev Nutr. 2007;27:263–75.
9. Solmi M, Radua J, Stubbs B, Ricca V, Moretti D, Busatta D, et al. Risk factors for eating disorders: an umbrella review of published meta-analyses. Braz J Psychiatry. 2020;43:314–23.
10. Schellings G, Van Hout-Wolters B. Measuring strategy use with self-report instruments: theoretical and empirical considerations. Metacogn Learn. 2011;6:83–90.
11. Fairbun CG. Risk factors for bulimia nervosa. A community based case-control study. Arch Gen Psychiatry. 1997;54:509–17.
12. Fairburn CG, Cooper Z, Doll HA, Welch SL. Risk factors for anorexia nervosa: three integrated case-control comparisons. Arch Gen Psychiatry. 1999;56(5):468–76.
13. Pike K, Hilbert A, Wilfley DE, Fairburn C, Dohm F-A, Walsh B, et al. Toward an understanding of risk factors for anorexia nervosa: a case-control study. Psychol Med. 2008;38(10):1443–53.
14. Karwautz A, Rabe-Hesketh S, Hu X, Zhao J, Sham P, Collier D, et al. Individual-specific risk factors for anorexia nervosa: a pilot study using a discordant sister-pair design. Psychol Med. 2001;31(2):317–29.
15. Karwautz A, Wagner G, Waldherr K, Nader I, Fernandez-Aranda F, Estivill X, et al. Gene–environment interaction in anorexia nervosa: relevance of non-shared environment and the serotonin transporter gene. Mol Psychiatry. 2011;16(6):590–2.
16. Sun L, Sun Z, Wu L, Zhu Z, Zhang F, Shang Z, et al. Prevalence and risk factors for acute posttraumatic stress disorder during the COVID-19 outbreak. J Affect Disord. 2021;283:123–9.
17. Arboix A. Cardiovascular risk factors for acute stroke: risk profiles in the different subtypes of ischemic stroke. World J Clin Cases. 2015;3(5):418.
18. Machado PP, Grilo CM, Crosby RD. Evaluation of the DSM-5 severity indicator for anorexia nervosa. Eur Eat Disord Rev. 2017;25(3):221–3.
19. Smith KE, Ellison JM, Crosby RD, Engel SG, Mitchell JE, Crow SJ, et al. The validity of DSM-5 severity specifiers for anorexia nervosa, bulimia nervosa, and binge-eating disorder. Int J Eat Disord. 2017;50(9):1109–13.
20. Zayas LV, Wang SB, Coniglio K, Becker K, Murray HB, Klosterman E, et al. Gender differences in eating disorder psychopathology across DSM-5 severity categories of anorexia nervosa and bulimia nervosa. Int J Eat Disord. 2018;51(9):1098–102.
21. Dalle Grave R, Sartirana M, El Ghoch M, Calugi S. DSM-5 severity specifiers for anorexia nervosa and treatment outcomes in adult females. Eat Behav. 2018;31:18–23.
22. Gianini L, Roberto CA, Attia E, Walsh BT, Thomas JJ, Eddy KT, et al. Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers. Int J Eat Disord. 2017;50(8):906–16.
23. Hay PJ, Touyz S, Sud R. Treatment for severe and enduring anorexia nervosa: a review. Aust N Z J Psychiatry. 2012;46(12):1136–44.
24. Krug I, Binh Dang A, Granero R, Agüera Z, Sánchez I, Riesco N, et al. Drive for thinness provides an alternative, more meaningful, severity indicator than the DSM-5 severity indices for eating disorders. Eur Eat Disord Rev. 2021;29(3):482–98.
25. Abbate-Daga G, Pierò A, Gramaglia C, Gandione M, Fassino S. An attempt to understand the paradox of anorexia nervosa without drive for thinness. Psychiatry Res. 2007;149(1–3):215–21.
26. Penas-Lledo E, Fernandez-Aranda F, Jimenez-Murcia S, Granero R, Penelo E, Soto A, et al. Subtyping eating disordered patients along drive for thinness and depression. Behav Res Ther. 2009;47(6):513–9.

27. Giles S, Hughes EK, Fuller-Tyszkiewicz M, Treasure J, Fernandez-Aranda F, Karwautz AF, et al. Bridging of childhood obsessive-compulsive personality disorder traits and adult eating disorder symptoms: a network analysis approach. Eur Eat Disord Rev. 2022;30(2):110–23.
28. Nakai Y, Nin K, Noma S, Teramukai S, Fujikawa K, Wonderlich SA. The impact of DSM-5 on the diagnosis and severity indicator of eating disorders in a treatment-seeking sample. Int J Eat Disord. 2017;50(11):1247–54.
29. Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P, et al. The longitudinal interval follow-up evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. Arch Gen Psychiatry. 1987;44(6):540–8.
30. Warshaw MG, Dyck I, Allsworth J, Stout RL, Keller MB. Maintaining reliability in a long-term psychiatric study: an ongoing inter-rater reliability monitoring program using the longitudinal interval follow-up evaluation. J Psychiatr Res. 2001;35(5):297–305.
31. Anderluh MB, Tchanturia K, Rabe-Hesketh S, Treasure J. Childhood obsessive-compulsive personality traits in adult women with eating disorders: defining a broader eating disorder phenotype. Am J Psychiatry. 2003;160(2):242–7.
32. Fairburn CG, Doll HA, Welch SL, Hay PJ, Davies BA, O'Connor ME. Risk factors for binge eating disorder: a community-based, case-control study. Arch Gen Psychiatry. 1998;55(5):425–32.
33. Garner DM, Olmsted MP, Bohr Y, Garfinkel PE. The eating attitudes test: psychometric features and clinical correlates. Psychol Med. 1982;12(4):871–8.
34. Garner DM, Olmsted MP, Polivy J, Garfinkel PE. Comparison between weight-preoccupied women and anorexia nervosa. Psychosom Med. 1984;46(3):255–66.
35. Cohen J. Statistical power analysis for the behavioral sciences. Academic Press; 2013.
36. Jafari M, Ansari-Pour N. Why, when and how to adjust your P values? Cell Journal (Yakhteh). 2019;20(4):604.
37. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Commun Stat Simul Comput. 2010;39(4):860–4.
38. Pike KM, So M, Hilbert A, Maekawa H, Shimanouchi T, Wilfley D, et al. Risk factors for anorexia nervosa and bulimia nervosa in Japan and compared to a US sample. Int J Eat Disord. 2021;54(2):155–67.
39. Barakat S, McLean SA, Bryant E, Le A, Marks P, Touyz S, et al. Risk factors for eating disorders: findings from a rapid review. J Eat Disord. 2023;11(1):8.
40. Dahlenburg SC, Gleaves DH, Hutchinson AD. Anorexia nervosa and perfectionism: a meta-analysis. Int J Eat Disord. 2019;52(3):219–29.
41. Lie SØ, Rø Ø, Bang L. Is bullying and teasing associated with eating disorders? A systematic review and meta-analysis. Int J Eat Disord. 2019;52(5):497–514.
42. Ålgars M. Shapes and sizes: body image, body dissatisfaction and disordered eating in relation to gender and gender identity. 2012.
43. Laraia BA, Leung CW, Tomiyama AJ, Ritchie LD, Crawford PB, Epel ES. Drive for thinness in adolescents predicts greater adult BMI in the growth and health study cohort over 20 years. Obesity. 2021;29(12):2126–33.
44. Keski-Rahkonen A, Bulik CM, Neale BM, Rose RJ, Rissanen A, Kaprio J. Body dissatisfaction and drive for thinness in young adult twins. Int J Eat Disord. 2005;37(3):188–99.
45. Dakanalis A, Timko CA, Colmegna F, Riva G, Clerici M. Evaluation of the DSM-5 severity ratings for anorexia nervosa in a clinical sample. Psychitry Res. 2018;262:124–8.
46. Halmi KA, Sunday SR, Strober M, Kaplan A, Woodside DB, Fichter M, et al. Perfectionism in anorexia nervosa: variation by clinical subtype, obsessionality, and pathological eating behavior. Am J Psychiatry. 2000;157(11):1799–805.
47. Melero S, Morales A, Espada JP, Fernández-Martínez I, Orgilés M. How does perfectionism influence the development of psychological strengths and difficulties in children? Int J Environ Res Public Health. 2020;17(11):4081.
48. McGrath DS, Sherry SB, Stewart SH, Mushquash AR, Allen SL, Nealis LJ, et al. Reciprocal relations between self-critical perfectionism and depressive symptoms: evidence from a short-term, four-wave longitudinal study. Can J Behav Sci. 2012;44(3):169.
49. Matthews A, Kramer RA, Mitan L. Eating disorder severity and psychological morbidity in adolescents with anorexia nervosa or atypical anorexia nervosa and premorbid overweight/obesity. Eat Weight Disord Stud Anorexia Bulimia Obes. 2022;8:1–10.
50. Meierer K, Hudon A, Sznajder M, Leduc M-F, Taddeo D, Jamoulle O, et al. Anorexia nervosa in adolescents: evolution of weight history and impact of excess premorbid weight. Eur J Pediatr. 2019;178:213–9.
51. Treasure J, Cardi V, Leppanen J, Turton R. New treatment approaches for severe and enduring eating disorders. Physiol Behav. 2015;152:456–65.
52. Dakanalis A, Clerici M, Stice E. Prevention of eating disorders: Current evidence-base for dissonance-based programmes and future directions. Eat Weight Disord Stud Anorexia Bulimia Obes. 2019;24:597–603.
53. Tchanturia K, Hambrook D. Cognitive remediation therapy for anorexia nervosa. Treat Eat Disord Clin Handb. 2010;7:130.
54. Dingemans AE, Danner UN, Donker JM, Aardoom JJ, Van Meer F, Tobias K, et al. The effectiveness of cognitive remediation therapy in patients with a severe or enduring eating disorder: a randomized controlled trial. Psychother Psychosom. 2013;83(1):29–36.
55. Hay P, Chinn D, Forbes D, Madden S, Newton R, Sugenor L, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment of eating disorders. Aust N Z J Psychiatry. 2014;48(11):977–1008.
56. Gadsby S. Explaining body size beliefs in anorexia. Cogn Neuropsychiatry. 2017;22(6):495–507.
57. Marzola E, Martini M, Brustolin A, Abbate-Daga G. Inpatients with severe-enduring anorexia nervosa: understanding the "enduringness" specifier. Eur Psychiatry. 2021;64(1):e44.
58. Fernández-Aranda F, Treasure J, Paslakis G, Agüera Z, Gimenez M, Granero R, et al. The impact of duration of illness on treatment nonresponse and drop-out: exploring the relevance of enduring eating disorder concept. Eur Eat Disord Rev. 2021;29(3):499–513.
59. Glasofer DR, Muratore AF, Attia E, Wu P, Wang Y, Minkoff H, et al. Predictors of illness course and health maintenance following inpatient treatment among patients with anorexia nervosa. J Eat Disord. 2020;8:1–10.
60. Takakura S, Aso CS, Toda K, Hata T, Yamashita M, Sudo N. Physical and psychological aspects of anorexia nervosa based on duration of illness: a cross-sectional study. BioPsychoSocial Med. 2019;13:1–7.
61. Mora-Maltas B, Lucas I, Granero R, Vintró-Alcaraz C, Miranda-Olivos R, Baenas I, et al. Cognitive flexibility and DSM-5 severity criteria for eating disorders: assessing drive for thinness and duration of illness as alternative severity variables. J Eat Disord. 2023;11(1):155.
62. Murray SB, Nagata JM, Griffiths S, Calzo JP, Brown TA, Mitchison D, et al. The enigma of male eating disorders: a critical review and synthesis. Clin Psychol Rev. 2017;57:1–11.

Publisher's Note

Chapter 7: Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 versus ICD–11 versus Drive For Thinness

7.1 Chapter Introduction

The overarching goal of this thesis is to evaluate the proposed severity ratings for ED. A key focus of these severity classifications is their clinical utility, particularly in relation to expected treatment outcomes, which are among the most critical aspects of severity assessment (Zimmerman et al., 2018}. Generally, patients with greater illness severity are expected to exhibit fewer positive treatment outcomes compared to those with less severe cases (Lucas et al., 2022).

Despite the importance of understanding this relationship, our systematic review and meta-analysis (Chapter 2, Study 1) identified only one study (Dalle Grave et al., 2018) that examined the relationship between DSM-5 ED severity ratings and treatment outcomes. Dalle Grave et al. found no significant differences in weight recovery or positive outcomes between the DSM-5 AN severity groups across baseline, 6-month, and 12-month follow-ups. While the longitudinal design of Dalle Grave et al.'s study was a strength, its small sample size (N = 128) likely limited its power to detect meaningful differences among the four DSM-5 severity groups. Furthermore, the study did not compare DSM-5 severity ratings to ICD-11 or other alternative severity classifications for AN. This underscores a critical gap in understanding how different severity systems– DSM-5, ICD-11, and alternatives– relate to treatment outcomes.

To address this gap, **Chapter 7 (Study 5)** presents a published study involving 628 treatment-seeking participants with AN (n = 266) and BN (n = 362) recruited from an ED unit in Spain. This study directly compares the DSM-5, ICD-11, and DT severity ratings in

their ability to index treatment outcomes associated CBT-E, providing a broader and more comprehensive evaluation of these severity classifications.

7.2 Study 5. Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 versus ICD–11 versus Drive For Thinness

This study was published in *European Eating Disorders Review* in December 2024.

Dang, A., Granero, R., Sanchez, I., Solé, L.G., Toro, J.J.-d., Rosinska, M., Jimenez-Murcia, S., Krug, I. and Fernandez-Aranda, F. (2024), Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 Versus ICD–11 Versus Drive for Thinness. *European Eating Disorders Review*. https://doi.org/10.1002/erv.3156

Contribution to the Study

Author	Task contribution
An Dang	Conceptualisation; manuscript writing; editing; data analyses
Roser Granero	Data collection, data analyses
Isabel Sanchez, Laura Galvez Sole, Jessica Jimenez-del Toro, Magda Rosinska, Susana	Data collection
Fernando Fernandez-Aranda	Data collection, editing
Isabel Krug	Conceptualisation; manuscript – reviewing and editing; supervision

European Eating Disorders Review

WILEY

RESEARCH ARTICLE

Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 Versus ICD–11 Versus Drive for Thinness

An Dang[1] | Roser Granero[2,3] | Isabel Sanchez[2,4] | Laura Gálvez Solé[5] | Jessica Jimenez-del Toro[5] | Magda Rosinska[6] | Susana Jimenez-Murcia[4,6,7,8,9] | Isabel Krug[1] | Fernando Fernandez-Aranda[4,6,7,8]

[1]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia | [2]Ciber Fisiopatología Obesidad y Nutrición (CIBERobn), Instituto Salud Carlos III, Barcelona, Spain | [3]Department of Psychobiology and Methodology, Universitat Autònoma de Barcelona—UAB, Barcelona, Spain | [4]Department of Clinical Psychology, University Hospital of Bellvitge-IDIBELL and CIBERobn, Barcelona, Spain | [5]Clinical Psychology Department, University Hospital of Bellvitge, Barcelona, Spain | [6]Psychoneurobiology of Eating and Addictive Behaviors Group, Neurosciences Program, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain | [7]CIBER Fisiopatología Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III, Barcelona, Spain | [8]Department of Clinical Sciences, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain | [9]Psychological Services, University of Barcelona, Barcelona, Spain

Correspondence: An Dang (dangb1@student.unimelb.edu.au) | Fernando Fernandez-Aranda (f.fernandez@ub.edu)

Received: 14 April 2024 | **Revised:** 2 August 2024 | **Accepted:** 31 October 2024

Handling Editor: Nadia Micali

Funding: This research paper is funded by Centres de Recerca de Catalunya, Ministerio de Asuntos Económicos y Transformación Digital, Gobierno de España, Delegación del Gobierno para el Plan Nacional sobre Drogas, Instituto de Salud Carlos III, Department of Health of the Generalitat de Catalunya, Biomedical Research Centre in Physiopathology of Obesity and Nutrition, FEDER funds/European Regional Development Fund, Centre for Mental Health Network (Centro de Investigación Biomédica en Red de Salud Mental: CIBERSAM).

Keywords: drive for thinness | DSM–5 | ICD–11 | severity ratings | treatment outcomes

ABSTRACT

Objective: This study aimed to compare treatment outcomes associated with three severity indicators—DSM–5 for anorexia nervosa (AN) and bulimia nervosa (BN), ICD–11 for AN, and drive for thinness (DT) for AN and BN—within a treatment-seeking population.

Method: A total of 628 female participants diagnosed with either AN (n = 266; mean age = 26.71) or BN (n = 362; mean age = 29.49) were recruited from an ED unit. Upon admission, participants were classified based on DSM–5 (AN and BN) and ICD–11 severity categories. They underwent CBT-E comprehensive manualised programs, according to ED subtype, and were categorised into 'full remission', 'partial remission', 'non-remission' or 'drop-out' based on DSM–5 diagnostic criteria at discharge.

Results: Significant associations were found only for ICD–11 AN severity groups (p = 0.03, Cramer's V = 0.18), with 'dangerously low BMI' associated with poorer outcomes than 'significantly low BMI'. No other significant relationships were found with DSM–5 or DT severity groups.

Conclusion: This study reveals the limitation of using a single severity indicator and emphasises the necessity of a comprehensive assessment to capture the complexities of AN and BN. Future research should also evaluate the validity of these severity measures across various factors, including biological correlates and psychopathology.

[Corrections added on 3 Dec 2024, after first online publication: The last author's first and last names and affiliation 4 have been corrected and a co-Corresponding Author has been added, in this version.]

Summary

- Treatment outcomes were significantly associated with severity based on the ICD–11 index for AN, but not with severity based on the DSM–5 and DT severity groups for both AN and BN.
- This study underscores the limitation of relying on a single severity indicator to comprehensively capture the multifaceted nature of AN and BN.
- It is important to explore alternative tailored severity indicators to capture the complexities of EDs in assessment and treatment.

1 | Introduction

Both anorexia nervosa (AN) and bulimia nervosa (BN) are diagnosed using the DSM–5 (American Psychiatric Association [APA] 2013) and its text revised version (DSM–5-TR; APA 2022) or the ICD–11(World Health Organization [WHO] 2019). While both systems use a severity indicator for AN based on body mass index (BMI), only the DSM–5 proposes a severity indicator for BN based on the weekly frequency of purging behaviours. These severity ratings aim to index the level of general and eating disorder (ED) psychopathology that is experienced by individuals with AN and BN, and thus, if valid, should serve as a crucial tool for tailoring treatment decisions, including duration, frequency and type of treatments (Regier, Kuhl, and Kupfer 2013). Although lower BMI and more frequent engagement in inappropriate compensatory behaviours predict poor long-term outcomes and higher mortality rates in AN and BN, few studies have directly examined the relationship between severity ratings and treatment outcomes (e.g., Dalle Grave et al. 2018). Some studies propose evaluating severity based on cognitive factors like drive for thinness (DT)—a common feature in both AN and BN, characterised by fear of weight gain and striving for an unhealthily low body weight—rather than solely relying on the DSM-5 severity index (Krug et al. 2021). However, no study to date has assessed DT as an alternative severity in indexing treatment outcomes for AN and BN. Therefore, this study aims to examine the relationship between treatment outcomes and three different severity indicators based on DSM–5 (for AN and BN), ICD–11 (for AN), DT (for AN and BN) using a treatment-seeking population.

2 | Severity Ratings

2.1 | DSM-5 Anorexia Nervosa Severity Rating

The DSM–5 AN severity index categorises AN into four levels based on BMI: mild ($\geq$ 17.0 kg/m^2), moderate (16–16.99 kg/m^2), severe (15–15.99 kg/m^2) and extreme ($<$ 15 kg/m^2). Several studies have examined the relationship between these categories and general and ED psychopathology (e.g., Krug et al. 2021), physiological correlates (e.g., potassium; Dang et al. 2023), psychiatric comorbidity (e.g., generalised anxiety disorders Dang, Kiropoulos, Anderluh, et al. 2024; Dang, Kiropoulos, Castle, et al. 2024), and as well as cognitive impairment (Mora-Maltas et al. 2023). However, little to no support was found for the DSM–5 AN severity rating in indexing the abovementioned variables. Dang et al. (2022) conducted a meta-analysis of 19 studies, including five assessing the DSM–5 AN severity indicator in indexing baseline ED psychopathology (e.g., Dakanalis et al. 2018), and found no significant differences across the severity groups.

Regarding BMI and treatment outcomes, various studies (e.g., Keel and Brown 2010; Vall and Wade 2015; Wild et al. 2016) have consistently shown that a lower baseline BMI is associated with poorer outcomes, including lower BMI at discharge, lower quality of life, higher psychiatric comorbidity, higher risk of relapse, higher ED symptoms, but also more dysfunctional personality traits and higher emotional dysregulation (Fernandez-Aranda et al. 2021; Miranda-Olivos et al. 2023; Mallorqui-Bague et al. 2023). Only two studies (Dalle Grave et al. 2018; Smink et al. 2014) have assessed the relationship between the DSM–5 AN severity groups and treatment outcomes with limited support.

Smink et al. (2014) found no association between the four DSM–5 AN severity groups and the proportion of clinically recovered cases in a Dutch cohort study spanning pre-adolescence to young adulthood (N = 2230). Similarly, Dalle Grave et al. (2018), using a sample of 128 participants with AN undergoing Enhanced Cognitive Behavioural Therapy for ED ([CBT-E]; Fairburn 2008), found no significant differences between the DSM–5 AN severity groups in weight recovery and good outcomes across baseline, 6- and 12-month follow-ups. While Dalle Grave et al.'s study has several strengths, including longitudinal follow-up, the sample size of 128 individuals may have limited the power to detect significant differences. Therefore, replication with a larger sample with AN is needed.

2.2 | ICD–11 Severity Rating for AN Based on BMI

Recognising the importance of severe underweight as a predictor of poor prognosis and mortality, the ICD–11 introduced a severity qualifier, 'dangerously low body weight', with a cut-off of BMI $<$ 14.0 kg/m^2 (Button, Chadalavada, and Palmer 2010; Rosling et al. 2011). This adjustment aims to enhance diagnostic accuracy and considers the role of severely underweight status in predicting AN prognosis and treatment intensity (Uher and Rutter 2012). Despite this, only two studies (Engelhardt et al. 2021; Dang, Kiropoulos, Anderluh, et al. 2024; Dang, Kiropoulos, Castle, et al. 2024) have explored the clinical utility of the ICD–11 severity rating for AN. Engelhardt et al. (2021) studied 469 children and adolescents with AN and found that those meeting the 'dangerously low BMI' criteria had longer treatment durations and lower BMIs at discharge compared to those in the 'significantly low BMI' group. However, these differences were not substantial enough to assist in daily clinical practice for risk assessment or treatment stratification. Another study by Dang et al. (N = 312) did not support the clinical usefulness of the AN ICD–11 severity ratings in differentiating severity groups based on psychiatric comorbidity. Furthermore, the rationale behind the ICD–11's binary severity classification remains unclear, and no study has explored the relationship between the AN ICD–11 severity groups and treatment outcomes in treatment-seeking adults with AN.

2.3 | DSM-5 Bulimia Nervosa Severity Rating

For BN, the DSM–5 employed the weekly frequency of inappropriate compensatory behaviours to index severity, which classifies individuals into mild (1–3 episodes/week), moderate (4–7 episodes/week), severe (8–13 episodes/week) and extreme (14 or more episodes/week). Regarding the DSM–5 BN severity rating, Dang et al.'s (2022) meta-analysis, included seven BN studies and found that individuals with milder BN severity exhibited significantly lower levels of baseline dietary restraint, eating and concerns related to weight/shape compared to those in the more severe BN groups, with small to large effect sizes. This supports the notion that the DSM–5 BN severity index serves as an indicator of ED psychopathology at baseline among those with BN.

However, the ability of the DSM–5 BN severity index to predict treatment outcomes remains less explored, with only two studies examining its relationship in adults (Dakanalis et al. 2017; $N = 281$) and adolescents (Gorrell et al. 2019; $N = 110$). Dakanalis et al. found a significant difference in post-treatment abstinence rates from inappropriate compensatory behaviours across the four DSM–5 BN severity groups, with notably higher abstinence rates in the milder groups. In contrast, Gorrell et al. focused on both behavioural (e.g., abstinence from compensatory behaviours) and attitudinal outcomes (e.g., ED global score measured using the Eating Disorder Examination-Questionnaire [EDE-Q; Fairburn and Beglin 1994]) at post-treatment, 6- and 12-month follow-up. They found no associations between the DSM–5 BN severity groups and the tested outcomes. However, additional analyses revealed a significant linear trend in EDE-Q global scores across the DSM–5 BN severity groups at the end of treatment and follow-up, following the expected linear pattern. Based on these findings, the authors concluded that the DSM–5 BN severity indicators may be more informative in predicting attitudinal rather than behavioural treatment outcomes.

2.4 | Drive for Thinness Severity Rating for AN and BN

Researchers have explored alternative classifications amidst uncertainty about the clinical utility of current ICD–11 and DSM–5 severity ratings for AN and BN, focusing on cognitive factors such as DT (e.g., Dang, Kiropoulos, Anderluh, et al. 2024; Dang, Kiropoulos, Castle, et al. 2024; Krug et al. 2021). DT—an intense fear of weight gain leading to disordered eating patterns—has been proposed as a central feature of EDs (Chernyak and Lowe 2010; Peñas-Lledó et al. 2015). Studies indicate that individuals with AN and BN who have lower DT tend to exhibit less severe attitudes and behaviours related to EDs and have lower comorbidity compared to those with high DT scores (Abbate-Daga et al. 2007; Peñas-Lledó et al. 2009; Ramacciotti et al. 2002; Vervaet, Van Heeringen, and Audenaert 2004).

Krug et al. (2021) compared DT with DSM–5 AN and BN severity ratings and found that the DT alternative severity rating was superior for indexing general and ED psychopathology in treatment-seeking populations. Mora-Maltas et al. (2023) later compared cognitive flexibility among 161 individuals with AN across severity groups based on DSM–5, DT and illness duration. They found more deficits in cognitive flexibility in those in the long illness duration (> 12 years) compared to those in the short duration illness (< 12 years) group, but no statistical differences between DT and DSM–5 AN severity groups.

No study has yet examined the relationship between DT severity groups and treatment outcomes in AN and BN. Nevertheless, elevated DT has been linked to poorer recovery rates in AN (Prost-Lehmann et al. 2018) and BN (Keski-Rahkonenal et al. 2013), as well as a higher likelihood of treatment dropout in both disorders (Schnicker, Hiller, and Legenbauer 2013). These findings suggest that elevated DT may lead to the belief that further weight gain is intolerable, potentially causing premature treatment termination and poor outcomes. Despite efforts to explore DT as a severity indicator, no study has directly examined the relationship between DT severity categories and treatment outcomes.

2.5 | The Current Study

Despite DSM–5 AN and BN severity ratings being introduced over a decade ago, few studies have directly assessed their ability to predict treatment outcomes (e.g., Dalle Grave et al. 2018). These studies offer conflicting findings, target diverse demographic groups and use various proxies to assess the predictive validity of these severity categories on treatment outcomes (e.g., behavioural vs. attitudinal changes). No study has yet investigated whether both the ICD–11 AN severity rating and the DT alternative severity rating for AN and BN can reliably predict treatment outcomes in adults with these disorders.

CBT-E, a current evidence-based treatment for AN, aims to reduce restrictive dietary patterns and facilitate weight gain, and for BN, it aims to disrupt dieting cycles, binge eating episodes, and compensatory behaviours, ultimately aiming for symptom abstinence in both disorders. In treating AN and BN, behavioural symptoms like binge eating and purging typically diminish before attitudinal symptoms such as overvaluation of weight and shape. Addressing both behavioural and attitudinal features is crucial for signalling recovery in both disorders, constituting a key treatment goal (Dalle Grave et al. 2013; Fernández-Aranda et al. 2021; Wonderlich et al. 2014). This study, therefore, used attitudinal and behavioural changes as treatment outcome proxies.

Improving the ability to predict treatment responses is crucial for guiding interventions for individuals with EDs. This study aimed to investigate the relationship between severity classifications (DSM–5, ICD–11 and DT) at baseline and treatment outcomes for AN and BN. We assessed (1) the distribution of severity rating groups and (2) the relationship between treatment outcomes across these severity indicators in a sample seeking treatment for AN and BN. Based on previous literature, we hypothesised that the DT alternative severity rating would be a more effective measure for assessing treatment outcomes compared to DSM–5 and ICD–11 severity ratings.

3 | Method

3.1 | Participants

Our study included 628 female participants diagnosed with either AN ($n = 266$) or BN ($n = 362$) who received treatment at a specialised Spanish adult ED unit from 1998 to 2018. Ethics approval for this study was obtained from the Research Ethics Committee of the institution where the study was conducted. Written and signed informed consent was obtained from all participants. AN and BN diagnoses were established during intake assessments through one-to-one semi-structured interviews conducted by clinical psychologists and psychiatrists using DSM-IV-TR criteria. The average assessment time was 45 min. Although data were collected from 1998 to 2018, significant efforts ensured accuracy and consistency, such as having the same trained clinicians throughout the study for diagnostic outcomes. Diagnoses were re-analysed and recoded based on DSM-5 criteria, which broadened the criteria by eliminating specific weight thresholds for both disorders and removing the amenorrhoea requirement for AN, resulting in a more inclusive classification (Dang, Kiropoulos, Anderluh, et al. 2024; Dang, Kiropoulos, Castle, et al. 2024).

Among participants diagnosed with AN (AN-restrict = 162; AN-binge/purge = 104), the average age was 26.71 years (SD = 9.41). The mean age of illness onset was 19.82 years (SD = 7.75) with a mean illness duration (calculated from self-reported AN symptom onset to the diagnostic assessment in the current study) of 6.84 years (SD = 7.03), and a mean BMI of 15.96 kg/m^2 (SD = 1.89).

For those diagnosed with BN, the average age was 29.49 years (SD = 9.35), the mean age of illness onset was 19.11 years (SD = 7.36), with a mean BN duration of 10.54 years (SD = 8.46), and a mean BMI of 25.01 kg/m^2 (SD = 6.68).

4 | Measure

4.1 | Clinical and Demographic Information

Height was assessed barefoot using a stadiometer, and BMI was computed with a Tanita BC-420MA scale (Tanita Corporation, Tokyo, Japan). Demographics, education, occupation and ED-related clinical data were gathered via a standardised face-to-face interview.

4.2 | Measure

4.2.1 | Severity Groups

Upon admission, individuals were classified based on DSM–5 (AN and BN) and ICD–11 (AN) severity ratings, using BMI for AN and clinical interviews for BN.

The DT severity groups for both AN and BN were derived from the DT subscale of the Eating Disorder Inventory (EDI-2; Garner 1991; Spanish version; Garner 1998). The EDI-2 comprises 91 items designed to evaluate cognitive and behavioural traits associated with ED, including DT. The DT sub-scale assesses excessive concern with dieting, preoccupation with weight, and fear of weight gain. Respondents rate each question (e.g., 'I'm preoccupied with the desire to be thin' or 'I think about dieting') on a six-point scale, ranging from 'always' to 'never', with scores from 0 to 3. The total score for the DT subscale was determined by adding the scores of all seven DT items. A cut-off score of 14, recommended by Garner et al. (1984), was used to categorise individuals into low and high DT severity groups, as applied in prior studies (Krug et al. 2021; Penas-Lledo et al. 2009). The internal consistency of the EDI DT subscale in our sample was $\alpha = 0.85$. Please refer to Garner (1991) for specific seven items of the DT subscale.

4.2.2 | Treatment

For AN, participants received a day hospital treatment that addressed both psychological and nutritional-dietary aspects based on the CBT-E programme (Fairburn 2008) that has been shown to be effective (Agüera et al. 2015; for more details, see Fernández-Aranda et al. 1998). Participants attended the hospital 5 days a week (Monday to Friday) from 9 AM to 3 PM, where breakfast and lunch were consumed and monitored.

For BN, participants attended 16-weekly manualised 90-min group therapy sessions that were conducted by experienced psychologists. This programme is published in Spanish (Fernández-Aranda et al. 1998) and has been shown to be effective in addressing ED psychopathology (Agüera et al. 2015).

Both treatments included problem-solving, cognitive restructuring, emotion regulation, self-esteem, body image enhancement and relapse prevention training, as well as psychoeducation, monitoring and normalisation of eating patterns. All treatments were conducted by experienced psychologists.

4.2.3 | Treatment Outcome

Upon discharge, patients were reevaluated and categorised into four groups: 'full remission', 'partial remission', 'non-remission' and 'drop-out'. Full remission was defined according to DSM–5 criteria as the complete absence of symptoms meeting DSM–5 criteria for either AN or BN for at least four consecutive weeks. Partial remission is defined by the DSM–5 as a significant improvement in symptoms with residual symptoms such as fear of gaining weight or disturbances in self-perception of weight and shape. Non-remission indicated poor outcomes. Voluntary treatment discontinuation (i.e., not attending treatment for three consecutive sessions) was classified as 'dropout'. These categories were determined by senior clinical staff through consensus judgement, considering various treatment outcome factors such as nutritional patterns, frequency of binge episodes, compensatory behaviours (e.g., vomiting, laxative use), weight restoration, attitude changes towards weight and shape and ED-related cognitions. These

definitions of treatment outcomes were based on established treatment outcomes in the DSM–5 and have been previously employed in other studies (e.g., Fernández-Aranda et al. 2021; Mairhofer et al. 2021; Steward et al. 2016). Objective assessments, such as weight status and frequency of binge/purge episodes per week, were obtained from food diaries and weight monitoring records.

4.3 | Procedure

4.3.1 | Statistical Analyses

Analyses were conducted using analysis of variance (ANOVA) and the linear-by-linear chi-square tests for continuous and ordinal variables respectively (Agresti 2012). As recommended by Sharpe (2015), post-hoc residual analyses were conducted to explore the relative contribution of cells to the chi square analyses by inspecting the standardised residuals to identify those that exceeded ± 1.96.

To evaluate a significant association in treatment outcomes (dropout, non-remission, partial remission, full remission) among the four DSM–5 severity groups for AN and BN, a chi-square analysis was conducted on the two four-level factors (DSM–5 severity groups and treatment outcomes). Subsequently, post-hoc chi-square analyses were conducted to investigate pairwise comparisons of severity groups (e.g., DSM–5 mild vs. DSM–5 moderate) on treatment outcomes. If significant differences were identified, standardised residuals would be calculated. To account for multiple comparisons, Holmes corrections were applied.

As our goal in comparing multiple severity groups was to find these severity groups associated with different treatment outcomes, a significant chi-square test gives a false alarm, if none of the post-hoc tests are significant (Tian et al. 2018). Hence, in line with Chen et al.'s recommendations, we based our findings on the results of the post-hoc pairwise comparisons.

5 | Results

5.1 | Demographic and Baseline ED Psychopathology

The results from Table 1 indicate that the DSM–5 AN severity groups were only significantly different in terms of education status ($p = 0.014$) and age ($p = 0.012$). Specifically, we found that individuals in the DSM–5 moderate group were more likely to have a primary school level of education compared to those in the DSM–5 extremely severe groups. Additionally, those in the DSM–5 extremely severe group were significantly older than the individuals in the remaining three severity groups. Further information can be found in Supporting Information S1.

The ICD–11 severity groups differed significantly in age ($p = 0.047$) and duration of illness ($p = 0.007$). Participants in the 'dangerously low BMI' group were older and had a longer illness duration compared to those in the 'significantly low BMI' group.

No significant differences in any demographic variables were found across the four DSM–5 BN and the two DT AN severity groups. The high DT BN severity group was found to be more likely to be married compared to the low DT group ($p = 0.006$).

Although not the focus of the paper, significant differences in baseline general and ED psychopathology across the DSM–5 AN and BN, and the DT AN and BN severity groups were found (see Krug et al. 2021). In contrast, the ICD–11 AN severity groups were not significantly different in general and ED psychopathology—these analyses are presented in Supporting Information S2.

5.2 | Frequency of Each Severity Group

The frequency distribution of each severity group per the DSM–5, ICD–11 and DT severity classification is displayed in Figure 1. There was a somewhat equal distribution in the DSM–5 severity groups for AN (mild = 33.1%, moderate = 27.4%, severe = 16.5% and extreme = 22.9%) and BN (mild = 28.6%, moderate = 21.2%,

TABLE 1 | Discriminative capacity for the severity groups at baseline: Total sample.

	Anorexia nervosa ($n = 266$)						BN ($n = 362$)			
	DSM–5		EDI-2 DT		ICD–11		DSM–5		EDI-DT	
	p	ES	*p*	ES	*p*	ES	*p*	ES	*p*	ES
Marital status	0.424	0.106	0.639	0.058	0.729	0.049	0.052	**0.133**[a]	**0.006**[b]	**0.168**[a]
Education	**0.014**[b]	**0.174**[a]	0.420	0.081	0.202	0.110	0.311	**0.101**[a]	0.797	0.035
Employment	0.234	0.127	0.101	0.101	0.422	0.049	0.952	0.031	0.143	0.077
Age (years-old)	**0.012**[b]	0.041	0.827	0.000	**0.047**[b]	0.015	0.629	0.005	0.506	0.001
Onset (years-old)	0.105	0.023	0.101	0.010	0.997	0.000	0.910	0.002	0.846	0.000
Duration (years)	0.299	0.014	0.078	0.012	**0.007**[b]	0.027	0.874	0.002	0.466	0.001

Abbreviations: BN, bulimia nervosa; EDI-2 DT, EDI-2 drive for thinness.
[a]Bold: Effect size within a high/large range ($\eta_2 > 0.10$ or $V > 0.15$).
[b]Bold: Significant comparison.

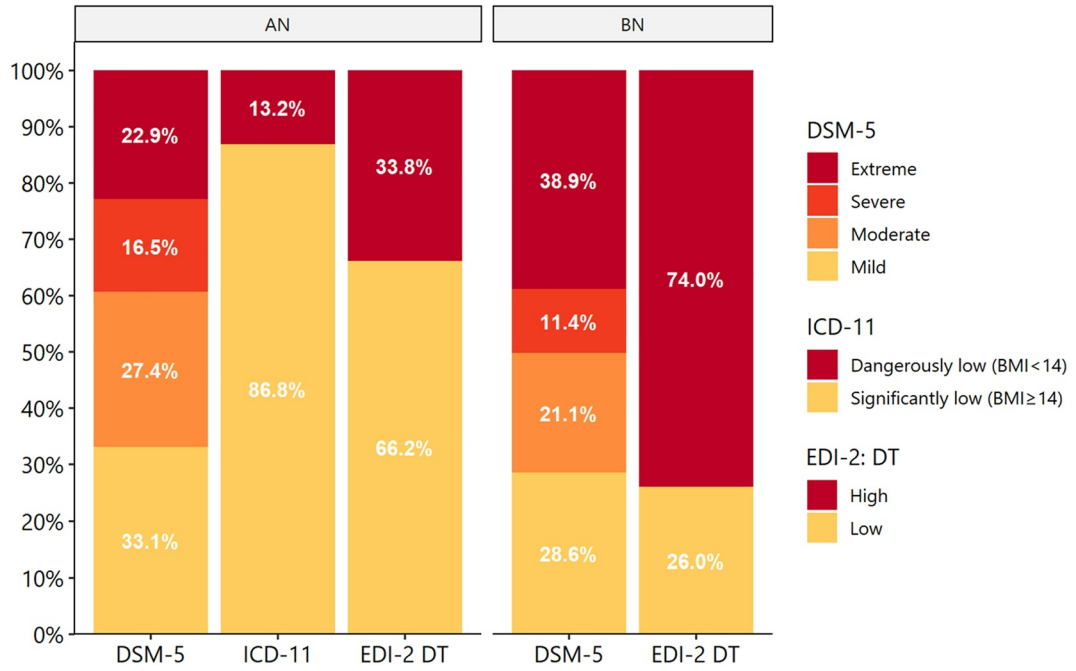


FIGURE 1 | Prevalence of the severity groups in the study. AN, anorexia nervosa; BN, bulimia nervosa; EDI-2 DT, EDI-2 drive for thinness.

severe = 11.4% and extreme = 38.9%). However, 86.8% of participants with AN were categorised into the 'significantly low' ICD–11 severity group. Regarding the DT severity groups, while most participants with AN (66.2%) were categorised into the low DT group, only 26% of those with BN were categorised into the low DT group.

5.3 | Treatment

Results from the chi-square test of independence (Table 2) revealed a significant association between DSM–5 AN severity groups and treatment outcome (χ^2 (9, $n = 266$) = 19.35, $p = 0.02$, Cramer's $V = 0.16$). However, subsequent post-hoc pairwise comparisons of severity groups (e.g., DSM–5 mild vs. DSM–5 moderate) using Holmes correction did not uncover significant differences among the DSM–5 severity groups in treatment outcomes (Table 2). Hence, in line with Chen et al.'s recommendations, we based our findings on the results of the post-hoc pairwise comparisons.

Regarding the ICD–11 AN severity groups, it was found that there was a significant relationship between the two ICD–11 AN severity groups and treatment outcome χ^2 (3, $n = 270$) = 8.89, $p = 0.03$, Cramer's $V = 0.18$, with those in the 'significantly low weight' group having a lower likelihood to drop out from treatment than those in the 'dangerously low weight' group (SR = −1.98). Additionally, those in the 'significantly low weight' group were more likely to achieve full remission than those in the 'dangerously low weight' group (SR = 2.94).

Finally, we did not find any association between treatment outcomes and the remaining tested severity indicators: DSM–5 BN χ^2 (9, $n = 350$) = 16.28, $p = 0.06$; DT AN χ^2 (3, $n = 266$) = 0.22, $p = 0.97$ and DT BN χ^2 (3, $n = 362$) = 1.30, $p = 0.73$. Per Tian et al.'s (2018) recommendation, subsequent post-hoc pairwise comparisons of DSM–5 BN severity groups (e.g., DSM–5 mild vs. DSM–5 moderate) were conducted. With Holmes correction, no significant associations in treatment outcomes between the DSM–5 BN, DT AN and DT BN severity groups were identified.

6 | Discussion

This study examined how pre-treatment severity classifications (DSM–5, ICD–11 and DT) predicted treatment outcomes in adults with AN and BN undergoing CBT-E. Most individuals with AN fell into milder severity groups, while those with BN were mostly in severe categories. Significant associations were found only for ICD–11 AN severity groups, with the 'dangerously low BMI' group showing poorer outcomes than the 'significantly low BMI' group. No other significant relationships were found between treatment outcomes for AN and BN and the DSM–5 and DT severity groups.

6.1 | Frequency Distribution of DSM–5, ICD–11 and DT Severity Groups

6.1.1 | Anorexia Nervosa

Our findings revealed that most individuals diagnosed with AN fell into less severe categories across the three classification systems: DSM–5 AN (mild and moderate severity groups, 60.5%), ICD–11 AN ('significantly low weight', 86.8%) and low DT AN group (66.2%). Similar distributions have been noted in prior studies for both DSM–5 and ICD–11 severity ratings (Dang et al. 2023; Dakanalis et al. 2018; Smith et al. 2017).

Similarly, 66.2% of our participants were classified into the low DT AN severity group, aligning with previous research showing

TABLE 2 | Discriminative capacity for the severity groups at baseline on the therapy outcome.

	Anorexia nervosa: DSM–5 severity															
	Mild/ G1 (n = 88)		Moderate/ G2 (n = 73)		Severe/ G3 (n = 44)		Extreme/ G4 (n = 61)									
	n	%	n	%	n	%	n	%	p	C–V	G1–G2 C–V	G1–G3 C–V	G1–G4 C–V	G2–G3 C–V	G2–G4 C–V	G3–G4 C–V
Dropout	30	34.1	28	38.4	19	43.2	27	44.3	**0.022**[a]	**0.156**[b]	0.200[b]	0.229[b]	0.268[b]	0.250[b]	0.257[b]	0.037
Non-rem.	7	8.0	13	17.8	7	15.9	11	18.0								
Partial-rem.	16	18.2	6	8.2	10	22.7	13	21.3								
Full-rem.	35	39.8	26	35.6	8	18.2	10	16.4								
	Bulimia nervosa: DSM–5 severity															
	Mild/G1 (n = 100)		Moderate/ G2 (n = 74)		Severe/ G3 (n = 40)		Extreme/ G4 (n = 136)									
	N		n		n		n		p	C–V	G1–G2 C–V	G1–G3 C–V	G1–G4 C–V	G2–G3 C–V	G2–G4 C–V	G3–G4 C–V
Dropout	41	41.0	42	56.8	21	52.5	70	51.5	0.061	0.125[b]	0.160[b]	0.164[b]	0.200[b]	0.175[b]	0.160[b]	0.175[b]
Non-rem.	9	9.0	4	5.4	5	12.5	20	14.7								
Partial-rem.	23	23.0	14	18.9	4	10.0	29	21.3								
Full-rem.	27	27.0	14	18.9	10	25.0	17	12.5								

	Anorexia nervosa: EDI-2 DT severity						Bulimia nervosa: EDI-2 DT severity						Anorexia nervosa: ICD–11 severity					
	Low (n = 176)		High (n = 90)				Low (n = 94)		High (n = 268)				Signif.low (n = 231)		Danger. low (n = 35)			
	n	%	n	%	p	C–V	n	%	n	%	p	C–V	n	%	n	%	p	C–V
Dropout	70	39.8	34	37.8	0.974	0.029	50	53.2	129	48.1	0.729	0.060	85	36.8	19	54.3	**0.031**[a]	**0.183**[b]
Non-rem.	24	13.6	14	15.6			10	10.6	31	11.6			32	13.9	6	17.1		
Partial-rem.	30	17.0	15	16.7			19	20.2	52	19.4			38	16.5	7	20.0		
Full-rem.	52	29.5	27	30.0			15	16.0	56	20.9			76	32.9	3	8.6		

Note: G1: DSM–5 Mild group; G2: DSM–5 Moderate group; G3: DSM–5 Severe group; G4: DSM–5 Extreme Severe group.
[a]Bold: Significant comparison.
[b]Bold: Effect size into the mild/moderate to the high/large range ($V > 0.15$).

lower DT scores in AN samples (Abbate-Daga et al. 2007; Penas-Lledo et al. 2009; Ramacciotti et al. 2002). The prevalence of individuals with low DT may be influenced by ego-syntonic functioning, resulting in the underreporting of cognitive factors and fat phobia symptoms (Rieger et al. 2001). Overall, our study found little evidence to support the effectiveness of the tested severity ratings for AN in identifying individuals at the severe or extreme end of AN severity within a treatment-seeking sample.

6.1.2 | Bulimia Nervosa

As regards to BN, using both the DSM–5 criteria and the alternative DT severity ratings, we observed that a significant proportion of our participants fell into the extreme severe DSM–5 BN (38.9%) and high DT BN (74%) severity groups. This distribution of individuals at the severe end of the BN spectrum appears notably elevated in our study compared to findings synthesised by Dang et al.'s (2022) meta-analysis. Dang et al. noted a more balanced distribution across DSM–5 severity groups. The higher prevalence of individuals in the more severe BN spectrum in our study could potentially be indicative of the intense nature, including a longer duration of illness, of BN experienced by individuals seeking treatment at a tertiary ED treatment unit.

6.2 | Treatment Outcomes and Severity Groups

6.2.1 | DSM–5 Severity Rating for Anorexia Nervosa

Besides age and education status, no significant differences in any tested sociodemographic variables were found. It is worth noting that individuals in the extreme severe DSM-5 AN severity group were significantly older than those in the milder DSM-5 AN severity groups. The reasons for these differences are unclear, but older age may reflect the chronicity of the illness (Fernández-Aranda et al. 2021), which might indicate a higher severity of the disorder as defined by BMI.

While we found a significant association between the DSM–5 AN severity groups and treatment outcomes overall using the chi-square test, subsequent pairwise post-hoc analyses did not reveal significant differences across the groups. This aligns with existing literature on the predictive validity of the DSM–5 AN severity rating (Dalle Grave et al. 2018; Smink et al. 2014). Our study suggests that the DSM–5 severity groups based on BMI may not be clinically useful in predicting treatment outcomes for individuals undergoing CBT-E. This raises questions about the validity of the current arbitrary BMI cut-off for determining AN DSM–5 severity groups, especially given the established link between low BMI and poor long-term outcomes (Button, Chadalavada, and Palmer 2010; Lowe, Miller-Kovach, and Phelan 2001). Therefore, future research should explore alternative BMI cut-off scores that may offer better predictive accuracy in AN treatment outcomes.

6.2.2 | ICD–11 Severity Rating for Anorexia Nervosa

In terms of the ICD–11 AN severity rating, we found that the individuals in the 'significantly low group' were less likely to drop out from treatment and achieve full remission than those with a 'dangerously low weight'. These findings are consistent with previous studies that have found that having a BMI < 14.0 kg/m^2 significantly correlated with having a poorer prognosis (Button, Chadalavada, and Palmer 2010; Engelhardt et al. 2021; Rosling et al. 2011). As such, the current study provides supporting evidence for BMI as an indicator of CBT-E treatment outcomes in AN, which is in line with previous research (Engelhardt et al. 2021).

It is worth to note that the debate centres once more on selecting a cut-off score that can effectively distinguish all pertinent AN-clinical variables. While the ICD–11 proposed a severity rating that is indicative of treatment outcomes, it fell short in capturing differences in the baseline level of ED psychopathology (Supporting Information S2). Yet, for practical implementation in daily clinical practice for risk assessment or treatment stratification, an effective severity indicator must encompass a diverse range of clinical indicators, including baseline general and ED psychopathology.

While both the DSM–5 and ICD–11 severity ratings for AN are based on BMI, our findings revealed a significant relationship between the ICD–11 and treatment outcomes, but not for the DSM–5 AN severity groups. This discrepancy may be attributed to the different number of severity groups within each rating system. Specifically, the presence of only two severity groups in the ICD–11 system may enhance the ability to detect significant findings compared to the four severity groups in the DSM–5 AN classification. Furthermore, both systems recognise a BMI below 14 kg/m^2 as indicating extremely severe AN, but they differ in how changes in BMI affect severity categorisation. A one-point change in BMI can shift individuals from DSM–5 moderate to DSM–5 severe categories, whereas the 'significantly low BMI' category in ICD–11 (the less severe of the two ICD–11 severity groups) encompasses all individuals with BMI greater than 14 kg/m^2, potentially leading to more detectable significant differences. Regardless of the different pattern in results between the DSM–5 and ICD–11 severity ratings for AN, this study refrains from supporting the integration of these two severity ratings into clinical practice.

6.2.3 | Bulimia Nervosa

While previous research supports the effectiveness of DSM–5 BN severity ratings in assessing baseline ED psychopathology (e.g., Dang et al. 2022), our study found no significant differences among these groups regarding treatment outcomes. This contrasts with prior findings that linked DSM–5 BN severity ratings to abstinence from compensatory behaviours in adults with BN and a decrease in ED-related attitudinal outcomes in adolescents with BN (Dakanalis et al. 2017; Gorrell et al. 2019).

One potential explanation for this discrepancy could be our study's stricter criteria for defining good treatment outcomes, which incorporate both attitudinal and behavioural aspects. This differs from previous studies that focused solely on behavioural outcomes or considered attitudinal and behavioural aspects separately. Overall, our study did not support the use of DSM–5 BN severity ratings to predict treatment outcomes in adults with BN. This aligns with the assumption in the current DSM–5 BN severity rating that different purging behaviours are interchangeable, although studies suggest that the type and number of purging behaviours are associated with different psychopathologies and treatment responses (Bruce and Steiger 2013; Gianini et al. 2017).

6.3 | Drive for Thinness as an Alternative Severity Rating for Anorexia and Bulimia Nervosa

Prior research has supported using DT as a severity indicator for AN and BN to assess baseline ED and general psychopathology (Krug et al. 2021). However, our study found no significant associations between DT severity groups and treatment outcomes in individuals with AN and BN. This contrasts with previous findings suggesting that higher DT is linked to poorer treatment outcomes in both AN and BN (Prost-Lehmann et al. 2018; Keski-Rahkonen et al. 2013; Schnicker, Hiller, and Legenbauer 2013).

The differences in our findings regarding DT severity groups and treatment outcomes compared to previous studies may be due to different study designs. Previous studies based their AN treatment outcome criteria on participants' weight and scores on the Eating Attitudes Scale–26 (Garner, Olmstead, and Polivy 1983; Prost-Lehmann et al. 2018) or participants' treatment dropout status (Schnicker, Hiller, and Legenbauer 2013), while our focus was on the absence of DSM–5 AN criteria symptoms (both attitudinal and behavioural). For BN, Keski-Rahkonen et al. (2013) conducted a cohort study involving participants with varying treatment histories, types, and durations, while Schnicker, Hiller, and Legenbauer (2013) exclusively examined BN participants who either completed or dropped out of CBT-E treatment, without directly considering BN behavioural and attitudinal symptoms as in our study.

Another possible explanation for the discrepancies between our results and earlier studies may be that previous research examined the relationship between DT and treatment outcome on a continuous scale, whereas we used the cut-off score proposed by Garner et al. (1984) to delineate the two distinct groups. This method might reduce statistical power compared to using continuous variables. Nevertheless, our study is the first to investigate the relationship between an alternative severity indicator, DT and treatment outcome.

6.4 | Implications

This study is the first to directly compare the efficacy of current severity ratings for AN and BN in predicting CBT-E treatment outcomes among adults seeking treatment. The potential to use baseline clinical information to predict outcomes can inform clinicians, patients and families about the expected treatment duration and types. While the results support the ICD–11 AN severity rating to some extent, their inability to capture other clinical information highlights the ongoing need for comprehensive clinical assessments. Considering additional indicators such as overvaluation of weight/shape or perfectionism (key features in EDs per Fairburn, Cooper, and Shafran 2003) could offer deeper insights into these disorders' complex psychopathological aspects. This approach, encompassing the various dimensions of AN and BN, can yield a more focused and comprehensive understanding, leading to more effective clinical decisions and treatment strategies.

6.5 | Strength and Limitations

This study has strengths worth noting. Firstly, it includes participants from both inpatient and outpatient clinics at a reputable ED unit in Spain, enhancing the generalisability of our findings across these settings. Secondly, participants received interventions from trained clinical psychologists following consistent CBT-E treatment procedures.

Conversely, there are also several limitations to the current study that need to be discussed. First, the demographic composition of our sample comprises females who are predominately white, raising concerns about the generalisability of our findings to diverse ethnicities, age groups and genders. Therefore, future research investigating the clinical utility of the DSM–5, ICD–11 and alternative ED severity ratings in a more diverse sample with different social and demographic backgrounds is needed.

Second, this study examines the relationship between DSM–5, ICD–11 and DT severity groups and treatment outcomes. While treatment outcome is a crucial aspect that these severity ratings are expected to index, other relevant domains—such as disordered eating, cognitive impairments, level of disability and quality of life—are also important for assessing their validity but were not examined in this study. Therefore, future research should include more comprehensive studies that explore a broader range of domains.

Third, recognising the complexity of eating disorders, a rigid, one-size-fits-all structured interview approach may miss important aspects of the disorder, hence semi-structured interviews were conducted by senior clinicians to establish ED diagnosis. It is recognised that semi-structured interviews can be susceptible to bias, and incorporating a second rater in future studies could help minimise this potential bias.

Fourth, our exploration of the relationship between the three assessed severity ratings based on DSM–5, ICD–11 and DT and CBT-E treatment outcomes was conducted cross-sectionally, restricting our ability to track changes in symptomatology over time. Furthermore, the evaluation of treatment outcomes, in the current study, relied on combining changes in attitudinal, behavioural and physical symptoms into a categorical variable, impeding our ability to observe changes in each of these specific

aspects over time. As such, future research should be conducted longitudinally and measure the changes of each aspect of ED symptomatology separately rather than looking at as altogether.

Lastly, the study only assessed purging behaviours specified in the DSM–5 for individuals with BN, excluding other compensatory behaviours like insulin misuse or diet pill use, which could introduce confounding factors. Additionally, the study lacked statistical power to explore how distinct compensatory patterns impact treatment outcomes within different DSM–5 BN severity categories. For example, individuals with laxative abuse have been found to have higher baseline general and eating disorder psychopathology compared to those who engage in other purging behaviours (Tozzi et al. 2006), suggesting that individuals with laxative abuse may respond differently to treatment compared to those engaging in excessive exercise, even if they fall within the same DSM–5 BN severity group. Future research should include a wider range of compensatory behaviours and investigate whether specific patterns moderate the relationship between DSM–5 BN severity groups and treatment outcomes.

6.6 | Conclusion

In conclusion, our study aimed to assess the predictive validity of DSM–5, ICD–11 and DT severity ratings that are currently proposed for AN and BN within an all-female treatment-seeking patient group. Interestingly, only the ICD–11 AN severity groups exhibited a significant association with treatment outcomes. Specifically, the higher BMI group classified as 'significantly low' demonstrated more favourable outcomes compared to the lower BMI group labelled as 'dangerously low'. Our current findings underscore the limitation of relying on a single severity indicator to comprehensively capture the multifaceted nature of AN and BN. Consequently, there is a compelling need to redirect attention towards exploring and developing severity indicators tailored for specific purposes, reflecting the nuanced complexities inherent in the assessment and treatment of ED.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

Abbate-Daga, G., A. Pierò, C. Gramaglia, M. Gandione, and S. Fassino. 2007. "An Attempt to Understand the Paradox of Anorexia Nervosa Without Drive for Thinness." *Psychiatry Research* 149, no. 1–3: 215–221. https://doi.org/10.1016/j.psychres.2005.10.017.

Agresti, A. 2012. *Categorical Data Analysis*. Vol. 792. John Wiley & Sons. https://doi.org/10.1002/0471249688.

Agüera, Z., X. Romero, J. Arcelus, et al. 2015. "Changes in Body Composition in Anorexia Nervosa: Predictors of Recovery and Treatment Outcome." *PLoS One* 10, no. 11: e0143012. https://doi.org/10.1371/journal.pone.0143012.

American Psychiatric Association (APA). 2013. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Washington, DC. https://doi.org/10.1176/appi.books.9780890425596.

American Psychiatric Association (APA). 2022. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed., text rev. https://doi.org/10.1176/appi.books.9780890425787.

Bruce, K. R., and H. Steiger. 2013. "Prognostic Implications of Personality Disorders in Eating Disorders." In *Personality Disorders and Eating Disorders*, 320. New York: Routledge. https://doi.org/10.4324/9780203957097.

Button, E. J., B. Chadalavada, and R. L. Palmer. 2010. "Mortality and Predictors of Death in a Cohort of Patients Presenting to an Eating Disorders Service." *International Journal of Eating Disorders* 43, no. 5: 387–392. https://doi.org/10.1002/eat.20715.

Chernyak, Y., and M. R. Lowe. 2010. "Motivations for Dieting: Drive for Thinness Is Different From Drive for Objective Thinness." *Journal of Abnormal Psychology* 119, no. 2: 276–281. https://doi.org/10.1037/a0018398.

Dakanalis, A., M. Clerici, G. Riva, and G. Carra. 2017. "Testing the DSM–5 Severity Indicator for Bulimia Nervosa in a Treatment-Seeking Sample." *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity* 22, no. 1: 161–167. https://doi.org/10.1007/s40519-016-0324-2.

Dakanalis, A., C. A. Timko, F. Colmegna, G. Riva, and M. Clerici. 2018. "Evaluation of the DSM–5 Severity Ratings for Anorexia Nervosa in a Clinical Sample." *Psychiatry Research* 262: 124–128. https://doi.org/10.1016/j.psychres.2018.02.009.

Dalle Grave, R., S. Calugi, M. Conti, H. Doll, and C. G. Fairburn. 2013. "Inpatient Cognitive Behaviour Therapy for Anorexia Nervosa: A Randomized Controlled Trial." *Psychotherapy and Psychosomatics* 82, no. 6: 390–398. https://doi.org/10.1159/000350058.

Dalle Grave, R., M. Sartirana, M. El Ghoch, and S. Calugi. 2018. "DSM–5 Severity Specifiers for Anorexia Nervosa and Treatment Outcomes in Adult Females." *Eating Behaviors* 31: 18–23. https://doi.org/10.1016/j.eatbeh.2018.07.006.

Dang, A. B., S. Giles, M. Fuller-Tyszkiewicz, L. Kiropoulos, and I. Krug. 2022. "A Systematic Review and Meta-Analysis on the DSM–5 Severity Ratings for Eating Disorders." *Clinical Psychology: Science and Practice* 29, no. 4: 325–344. https://doi.org/10.1037/cps0000078.

Dang, A. B., L. Kiropoulos, M. Anderluh, et al. 2024. "Do Risk Factors Differentiate DSM–5 and Drive for Thinness Severity Groups for Anorexia Nervosa?" *Journal of Eating Disorders* 12, no. 1: 5. https://doi.org/10.1186/s40337-024-00966-5.

Dang, A. B., L. Kiropoulos, D. Castle, et al. 2024. "Psychiatric Comorbidity and Severity in Anorexia Nervosa: A Comparative Study of the DSM–5, the ICD–11, and Overvaluation of Weight/Shape Severity Ratings." *Eating Disorders* 32, no. 5: 1–17. https://doi.org/10.1080/10640266.2024.2346001.

Dang, A. B., L. Kiropoulos, D. J. Castle, et al. 2023. "Assessing Severity in Anorexia Nervosa: Do the DSM-5 and an Alternative Severity Rating Based on Overvaluation of Weight and Shape Severity Differ in Psychological and Biological Correlates?" *European Eating Disorders Review* 31, no. 4: 447–461. https://doi.org/10.1002/erv.2969.

Engelhardt, C., M. Föcker, K. Bühren, et al. 2021. "Age Dependency of Body Mass Index Distribution in Childhood and Adolescent Inpatients With Anorexia Nervosa With a Focus on DSM–5 and ICD–11 Weight Criteria and Severity Specifiers." *European Child & Adolescent Psychiatry* 30, no. 7: 1081–1094. https://doi.org/10.1007/s00787-020-01595-4.

Fairburn, C. G. 2008. "Eating Disorders: The Transdiagnostic View and the Cognitive Behavioral Theory." In *Cognitive Behavior Therapy and Eating Disorders*, edited by C. G. Fairburn, 7–22. New York: Guilford Press.

Fairburn, C. G., and S. J. Beglin. 1994. "Assessment of Eating Disorders: Interview or Self-Report Questionnaire?" *International Journal of Eating Disorders* 16, no. 4: 363–370.

Fairburn, C. G., Z. Cooper, and R. Shafran. 2003. "Cognitive Behaviour Therapy for Eating Disorders: A 'Transdiagnostic' Theory and

Treatment." *Behaviour Research and Therapy* 41, no. 5: 509–528. https://doi.org/10.1016/s0005-7967(02)00088-8.

Fernández-Aranda, F., M. Bel, S. Jimenez, M. Vinuales, J. Turon, and J. Vallejo. 1998. "Outpatient Group Therapy for Anorexia Nervosa: A Preliminary Study." *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity* 3: 1–6. https://doi.org/10.1007/bf03339981.

Fernández-Aranda, F., J. Treasure, G. Paslakis, et al. 2021. "The Impact of Duration of Illness on Treatment Nonresponse and Drop-Out: Exploring the Relevance of Enduring Eating Disorder Concept." *European Eating Disorders Review* 29, no. 3: 499–513. https://doi.org/10.1002/erv.2822.

Garner, D. 1998. *Inventario de Trastornos de la Conducta Alimentaria* [Eating Disorder Inventory (EDI-2)]. Madrid: Masson.

Garner, D. M. 1991. *Eating Disorder Inventory-2: EDI2; Professional Kit*. Odessa, FL: Psychological Assessment Resources.

Garner, D. M., M. P. Olmstead, and J. Polivy. 1983. "Development and Validation of a Multidimensional Eating Disorder Inventory for Anorexia Nervosa and Bulimia." *International Journal of Eating Disorders* 2, no. 2: 15–34. https://doi.org/10.1002/1098-108x(198321)2:2<15::aid-eat2260020203>3.0.co;2-6.

Garner, D. M., M. P. Olmsted, J. Polivy, and P. E. Garfinkel. 1984. "Comparison Between Weight-Preoccupied Women and Anorexia Nervosa." *Psychosomatic Medicine* 46, no. 3: 255–266. https://doi.org/10.1097/00006842-198405000-00007.

Gianini, L., C. A. Roberto, E. Attia, et al. 2017. "Mild, Moderate, Meaningful? Examining the Psychological and Functioning Correlates of DSM-5 Eating Disorder Severity Specifiers." *International Journal of Eating Disorders* 50, no. 8: 906–916. https://doi.org/10.1002/eat.22728.

Gorrell, S., L. Hail, K. Kinasz, et al. 2019. "A Test of the DSM-5 Severity Specifier for Bulimia Nervosa in Adolescents: Can We Anticipate Clinical Treatment Outcomes?" *International Journal of Eating Disorders* 52, no. 5: 586–590. https://doi.org/10.1002/eat.23034.

Keel, P. K., and T. A. Brown. 2010. "Update on Course and Outcome in Eating Disorders." *International Journal of Eating Disorders* 43, no. 3: 195–204. https://doi.org/10.1002/eat.20810.

Keski-Rahkonen, A., A. Raevuori, C. M. Bulik, et al. 2013. "Depression and Drive for Thinness Are Associated With Persistent Bulimia Nervosa in the Community." *European Eating Disorders Review* 21, no. 2: 121–129. https://doi.org/10.1002/erv.2182.

Krug, I., A. Binh Dang, R. Granero, et al. 2021. "Drive for Thinness Provides an Alternative, More Meaningful, Severity Indicator Than the DSM-5 Severity Indices for Eating Disorders." *European Eating Disorders Review* 29, no. 3: 482–498. https://doi.org/10.1002/erv.2776.

Lowe, M., K. Miller-Kovach, and S. Phelan. 2001. "Weight-Loss Maintenance in Overweight Individuals One to Five Years Following Successful Completion of a Commercial Weight Loss Program." *International Journal of Obesity* 25, no. 3: 325–331. https://doi.org/10.1038/sj.ijo.0801521.

Mairhofer, D., M. Zeiler, J. Philipp, et al. 2021. "Short-Term Outcome of Inpatient Treatment for Adolescents With Anorexia Nervosa Using DSM–5 Remission Criteria." *Journal of Clinical Medicine* 10, no. 14: 3190. https://doi.org/10.3390/jcm10143190.

Mallorquí-Bagué, N., M. Lozano-Madrid, R. Granero, et al. 2023. "Cognitive and Clinical Gender-Related Differences Among Binge-Spectrum Eating Disorders: Analysis of Therapy Response Predictors." *European Eating Disorders Review* 31, no. 3: 377–389. https://doi.org/10.1002/erv.2961.

Miranda-Olivos, R., Z. Agüera, R. Granero, et al. 2023. "The Role of Food Addiction and Lifetime Substance Use on Eating Disorder Treatment Outcomes." *Nutrients* 15, no. 13: 2919. https://doi.org/10.3390/nu15132919.

Mora-Maltas, B., I. Lucas, R. Granero, et al. 2023. "Cognitive Flexibility and DSM–5 Severity Criteria for Eating Disorders: Assessing Drive for Thinness and Duration of Illness as Alternative Severity Variables." *Journal of Eating Disorders* 11, no. 1: 155. https://doi.org/10.1186/s40337-023-00875-z.

Peñas-Lledó, E., C. M. Bulik, P. Lichtenstein, H. Larsson, and J. H. Baker. 2015. "Risk for Self-Reported Anorexia or Bulimia Nervosa Based on Drive for Thinness and Negative Affect Clusters/Dimensions During Adolescence: A Three-Year Prospective Study of the TC h AD Cohort." *International Journal of Eating Disorders* 48, no. 6: 692–699. https://doi.org/10.1002/eat.22431.

Penas-Lledo, E., F. Fernández-Aranda, S. Jiménez-Murcia, et al. 2009. "Subtyping Eating Disordered Patients Along Drive for Thinness and Depression." *Behaviour Research and Therapy* 47, no. 6: 513–519. https://doi.org/10.1016/j.brat.2009.03.003.

Prost-Lehmann, C., R. Shankland, L. R. França, A. Laurent, and V. Flaudias. 2018. "Symptomatology Long-Term Evolution After Hospitalization for Anorexia Nervosa: Drive for Thinness to Explain Effects of Body Dissatisfaction on Type of Outcome." *Psychiatry Research* 266: 212–217. https://doi.org/10.1016/j.psychres.2018.05.024.

Ramacciotti, C. E., L. Dell'Osso, R. A. Paoli, et al. 2002. "Characteristics of Eating Disorder Patients Without a Drive for Thinness." *International Journal of Eating Disorders* 32, no. 2: 206–212. https://doi.org/10.1002/eat.10067.

Regier, D. A., E. A. Kuhl, and D. J. Kupfer. 2013. "The DSM-5: Classification and Criteria Changes." *World Psychiatry* 12, no. 2: 92–98. https://doi.org/10.1002/wps.20050.

Rieger, E., S. W. Touyz, T. Swain, and P. J. Beumont. 2001. "Cross-Cultural Research on Anorexia Nervosa: Assumptions Regarding the Role of Body Weight." *International Journal of Eating Disorders* 29, no. 2: 205–215. https://doi.org/10.1002/1098-108x(200103)29:2<205::aid-eat1010>3.0.co;2-1.

Rosling, A. M., P. Sparén, C. Norring, and A. L. von Knorring. 2011. "Mortality of Eating Disorders: A Follow-Up Study of Treatment in a Specialist Unit 1974–2000." *International Journal of Eating Disorders* 44, no. 4: 304–310. https://doi.org/10.1002/eat.20827.

Schnicker, K., W. Hiller, and T. Legenbauer. 2013. "Drop-Out and Treatment Outcome of Outpatient Cognitive–Behavioral Therapy for Anorexia Nervosa and Bulimia Nervosa." *Comprehensive Psychiatry* 54, no. 7: 812–823. https://doi.org/10.1016/j.comppsych.2013.02.007.

Sharpe, D. 2015. "Your Chi-Square Test Is Statistically Significant: Now What?" *Practical Assessment, Research and Evaluation* 20, no. 8: n8.

Smink, F. R., D. van Hoeken, A. J. Oldehinkel, and H. W. Hoek. 2014. "Prevalence and Severity of DSM-5 Eating Disorders in a Community Cohort of Adolescents." *International Journal of Eating Disorders* 47, no. 6: 610–619. https://doi.org/10.1002/eat.22316.

Smith, K. E., J. M. Ellison, R. D. Crosby, et al. 2017. "The Validity of DSM-5 Severity Specifiers for Anorexia Nervosa, Bulimia Nervosa, and Binge-Eating Disorder." *International Journal of Eating Disorders* 50, no. 9: 1109–1113. https://doi.org/10.1002/eat.22739.

Steward, T., G. Mestre-Bach, Z. Agüera, et al. 2016. "Enduring Changes in Decision Making in Patients With Full Remission From Anorexia Nervosa." *European Eating Disorders Review* 24, no. 6: 523–527. https://doi.org/10.1002/erv.2472.

Tian, C., X. Manfei, T. Justin, W. Hongyue, and N. Xiaohui. 2018. "Relationship Between Omnibus and Post-Hoc Tests: An Investigation of Performance of the F Test in ANOVA." *Shanghai Archives of Psychiatry* 30, no. 1: 60.

Tozzi, F., L. M. Thornton, J. Mitchell, et al. 2006. "Features Associated With Laxative Abuse in Individuals With Eating Disorders." *Psychosomatic Medicine* 68, no. 3: 470–477. https://doi.org/10.1097/01.psy.0000221359.35034.e7.

Uher, R., and M. Rutter. 2012. "Classification of Feeding and Eating Disorders: Review of Evidence and Proposals for ICD-11." *World Psychiatry* 11, no. 2: 80–92. https://doi.org/10.1016/j.wpsyc.2012.05.005.

Vall, E., and T. D. Wade. 2015. "Predictors of Treatment Outcome in Individuals With Eating Disorders: A Systematic Review and Meta-Analysis." *International Journal of Eating Disorders* 48, no. 7: 946–971. https://doi.org/10.1002/eat.22411.

Vervaet, M., C. Van Heeringen, and K. Audenaert. 2004. "Is Drive for Thinness in Anorectic Patients Associated With Personality Characteristics?" *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association* 12, no. 6: 375–379. https://doi.org/10.1002/erv.586.

Wild, B., H.-C. Friederich, S. Zipfel, et al. 2016. "Predictors of Outcomes in Outpatients With Anorexia Nervosa–Results From the ANTOP Study." *Psychiatry Research* 244: 45–50. https://doi.org/10.1016/j.psychres.2016.07.002.

Wonderlich, S. A., C. B. Peterson, R. D. Crosby, et al. 2014. "A Randomized Controlled Comparison of Integrative Cognitive-Affective Therapy (ICAT) and Enhanced Cognitive-Behavioral Therapy (CBT-E) for Bulimia Nervosa." *Psychological Medicine* 44, no. 3: 543–553. https://doi.org/10.1017/s0033291713001098.

World Health Organization. 2019. International Statistical Classification of Diseases and Related Health Problems. 11th ed. https://icd.who.int/.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Chapter 8: Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample?

8.1 Chapter Introduction

The DSM-5 and ICD-11 use BMI-based cut-offs to classify the severity of AN, but their approaches differ significantly. The ICD-11 BMI severity cut-off at 14 kg/m² is based on clinical guidelines and research evidence, while the DSM-5 four severity groups are set using arbitrary BMI thresholds. This arbitrary basis may partly explain why the DSM-5 often falls short in capturing the full complexity of AN psychopathology.

In Chapter 7 (Study 5), our findings indicate that while the ICD-11 severity rating may help predict treatment outcomes, it does not adequately distinguish baseline levels of ED psychopathology. By contrast, the DSM-5 fails to effectively capture both ED psychopathology (Krug et al., 2021) and treatment outcomes (Dang et al., 2024). These inconsistencies between the two frameworks raise questions about the validity of their respective BMI cut-offs. To address these concerns, we re-examined the BMI-based approach by analysing BMI as a continuous variable within a community-based sample.

An alternative indicator, weight suppression, defined as the difference between an individual's highest lifetime weight and current weight, has gained attention as a potential measure of AN severity (Khalil et al., 2024). Both clinicians and researchers, including psychologists and dietitians interviewed in Study 2 (Chapter 4), regard weight suppression as a potential marker of AN. Weight suppression reflects the active weight-control behaviours characteristic of AN, offering insight into an individual's dietary restraint and OWS, both central aspects of AN (Lowe et al., 2018).

The final study of this thesis **(Study 8),** conducted with a sample of 686 adults from the community, utilised EMA to examine BMI and weight suppression as continuous

variables in predicting both average and daily fluctuations in disordered eating urges and body dissatisfaction. EMA was chosen for its ability to capture real-time data in natural environments, reducing recall bias and enhancing ecological validity (Raugh et al., 2019). The EMA approach also allowed for the identification of dynamic patterns and triggers that static measures often overlook, providing insight into the relationship between BMI, weight suppression, and disordered eating behaviours. Ultimately, it offers a clearer understanding of how these two indicators might contribute differently to the daily clinical presentation of AN in a community-based sample.

8.2 Study 6. Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample?

This study is currently under the second round of review with the *European Eating Disorders Review*

Dang, A., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample?

Contribution to the Study

Author	Task contribution
An Dang	Conceptualisation; manuscript writing; editing; data collection; data analyses
Matthew Fuller-Tyszkiewicz	Data analyses, editing
Litza Kiropoulos	Editing
Isabel Krug	Conceptualisation; manuscript – reviewing and editing; supervision

Do Weight Suppression and BMI Predict Average Levels of Body Dissatisfaction and Disordered Eating Urges in a Community Sample?

Declaration

Ethical approval from the ethics committee of an Australian university; and consents were obtained from all participants from

Funding Declaration: No funding declaration

Abstract

The current study used an ecological momentary assessment (EMA) design to assess whether BMI and weight suppression [the difference between highest and current weight] predicted experiences of state-based body dissatisfaction and disordered eating urges, including dietary restriction, excessive exercise, binge eating, and unhealthy eating. The study collected data from 686 adults (75% female) through six daily EMA surveys throughout seven days (42 possible assessments). Multilevel modelling analyses showed that individuals with a low BMI (p = .005) and weight suppression (p = .004) tend to report having a higher body dissatisfaction average. Furthermore, those with higher BMI (p <. 001) and a higher weight suppression (p = .039) were more likely to report having urges to engage in unhealthy eating. All two weight-based severity ratings (i.e., BMI and weight suppression) had a null effect in predicting state-based averages in body dissatisfaction and disordered eating urges in a predominantly female community-based sample. Our results emphasise the limited usage of the weight-based indicators in indexing ED-related state-based variables.

Keywords: eating disorders; weight suppression; BMI; body dissatisfaction

1. Introduction

Eating disorders (ED) are common among university students, with up to 9% experiencing clinical disorders (Hay et al., 2023) and up to 40% showing subthreshold symptoms (Sparti et al., 2019). This highlights the need to understand symptom precursors for prevention and management. Both BMI and weight suppression (the difference between highest and lowest adult weight; Lowe, 1993) are linked to body dissatisfaction and disordered eating urges—key ED risk factors. However, most studies on BMI and weight suppression's predictive validity are cross-sectional and focus on clinical samples (Dang et al., 2022; 2023). Given ED symptoms' variability and prevalence in community samples (Nagata et al., 2020), this study addresses gaps by examining how BMI and weight suppression predict changes in body dissatisfaction and disordered eating urges in a community sample using a micro-longitudinal design.

1.1. Body mass index

The role of BMI in predicting body dissatisfaction and disordered eating is inconsistent across both clinical and community settings. Several studies (e.g., Duncan et al., 2017; Ro et al., 2012) reported a positive association between higher BMI and increased rates of disordered eating and ED risk. Conversely, others (e.g., Argyrides et al., 2020; Stice et al., 2002, 2017) found no link between BMI and body dissatisfaction or disordered eating.

Furthermore, both the DSM-5 and ICD-11 use BMI as a severity rating for anorexia nervosa (AN), characterised by extreme food restriction, intense fear of weight gain, and a distorted body image. The DSM-5 categorises AN severity into four groups based on BMI: mild (≥17.0 kg/m²), *moderate (16–16.99 kg/m²), severe (15–15.99 kg/m²)*, and extreme (<15 kg/m²). The ICD-11 categorises individuals with AN into "significantly low body weight" and "dangerously low body weight" using a cut-off of BMI < 14.0 kg/m². Although no studies have directly assessed the predictive validity of the ICD-11, a recent meta-analysis

(Dang et al., 2022) of five studies (Dalle Grave et al., 2018; Gianini et al., 2017; Machado et al., 2017; Smith et al., 2017; Zayas et al., 2018) evaluating the severity index for AN, including the DSM-5 severity classification, found limited support for using BMI to index ED psychopathology. Additionally, over 80% of the studies in this field used clinical-based samples (Dang et al., 2022).

Overall, despite BMI being proposed as a severity rating for AN, studies have found limited and inconsistent evidence to support its predictive validity for changes in body dissatisfaction and disordered eating urges in both clinical and community-based settings. Since ED symptomatology exists along a severity spectrum and clinically significant disordered eating is often observed within community samples (Nagata et al., 2020), this study aimed to re-examine the predictive validity of BMI for average changes in these factors in a naturalistic environment, rather than in the cross-sectional designs used in previous studies.

1.2. Weight suppression

A key critique of using BMI to predict disordered eating and body dissatisfaction is its neglect of an individual's weight history (Dang et al., 2021). Weight suppression, linked to future weight gain and the development or worsening of ED symptoms, may be crucial in predicting the level of body dissatisfaction and disordered eating urges (Jennings et al., 2017; Gorrell et al., 2019). Being at a reduced body weight (i.e., weight suppressed) is thought to slow metabolism and create biological pressure to regain lost weight, increasing the risk of subsequent weight gain (Carter et al., 2008; Ochner et al., 2013), potentially leading to efforts to maintain weight suppression through restrictive eating and compensatory behaviours, which may contribute to symptoms (Juarascio et al., 2019).

Several studies support the biological and psychological significance of weight suppression in both clinical (e.g., Berner et al., 2013; Stice et al., 2020) and

community-based samples (e.g., Burnette & Mazzeo, 2020) studies. Gorrell et al. (2019) conducted a narrative review ($N = 31$) looking at the relationship between weight suppression and disordered eating and weight outcomes across both community and clinical-based samples (e.g., AN, Bulimia Nervosa). They concluded that weight suppression is a useful index of ED symptom severity in clinical and community samples (Gorrell et al., 2019). Subsequent studies have further supported its role in indexing ED severity in both adults (Garber et al., 2019) and adolescents (Mathews et al., 2021) with clinical diagnosed ED. However, no study has directly examined the predictive validity of weight suppression in predicting average daily changes in body dissatisfaction and disordered eating urges, nor compared its predictive validity directly to BMI.

1.4. Ecological Momentary Assessment

Studies assessing the predictive validity of BMI, including those using DSM-5 and ICD-11 cutoffs for severity rating, and weight suppression, have primarily relied on retrospective self-report methods. These methods provide insight into person-level characteristics associated with ED symptomatology. However, these methods are susceptible to retrospective bias, specifically since a few studies have found evidence of low correspondence between trait and state averages for the same variables (e.g., Portingale et al., 2023). To address this, studies on ED have begun using Ecological Momentary Assessment (EMA) to monitor real-time changes in disordered eating and body dissatisfaction, which is defined as the negative subjective evaluation of one's body (Lantz et al., 2018; Portingale et al., 2022), while also minimising recall bias (Shiffman et al., 2008).

To date, only one EMA study (Fuller-Tyszkiewicz et al., 2020; N = 260) has examined the relationship between state-based variables (e.g., mood, body dissatisfaction, appearance comparisons) and trait-based ED severity, measured by the Eating Attitudes Test-26 (EAT-26; Garner et al., 1982). Fuller-Tyszkiewicz et al. (2020) found that state-based variables predicted 34% of changes in trait-based ED severity measured by the EAT-26 total

score. However, their study focused on an all-female community sample, and ED severity was not based on BMI and weight suppression – which is the most known severity indicator for restricted types of eating disorders (e.g., AN). BMI and weight suppression are considered snapshots of an individual's current physiological state, yet no study has explored whether they predict state-based ED-related variables (e.g., body dissatisfaction) in a community-based sample (Dang et al., 2021; 2022).

1.5. Limitations and Gaps in the Literature

The current literature lacks exploration of the potential of BMI and weight suppression in assessing disordered eating in real-life settings. Moreover, studies examining the predictive validity of BMI in indexing body dissatisfaction and disordered eating have often used the DSM-5 proposed arbitrary BMI cut-off, lacking a research-based foundation. Therefore, it is important to take a step back and investigate whether BMI and weight suppression can effectively predict state-based variable averages, such as body dissatisfaction and disordered eating urges, using an EMA approach within a community sample.

1.6. The current study

Using a micro-longitudinal EMA design, this study employed a community-based sample to assess and compare the predictive validity of BMI and weight suppression as continuous variables in predicting state-based body dissatisfaction and disordered eating urges. This study used disordered eating urges as proxies for actual disordered eating behaviours, leveraging the connection between urges and behaviour engagement as shown in previous studies (Tasca et al., 2009). Daily disordered eating urges were analysed separately for dietary restraint, excessive exercise, binge eating, and unhealthy eating reflecting their distinct purposes. Dietary restraint directly addresses body dissatisfaction (Fitzsimmons-Craft et al., 2016), while excessive exercise serves as an alternative disordered eating practice (Gonçalves & Gomes,

2012). Binge-eating is often linked to body dissatisfaction and is seen as a distraction accompanied by purging behaviours (Meyer & Waller, 1999). In the current study, urges for overeating, akin to loss of control eating central to binge eating (Goldschmidt et al., 2014), were also assessed as a proxy for binge eating behaviours.

The study also aimed to assess whether baseline ED psychopathology (measured by the EAT-26) moderated the relationships between BMI, weight suppression, and state-based body dissatisfaction and disordered eating urges. This is important as research (e.g., Coffino et al., 2016) highlights how trait ED symptoms influence the link between BMI, weight suppression, and urges like binge eating or dietary restraint. It is intuitive that individuals with heightened ED symptoms, low BMI, or high weight suppression may experience stronger daily disordered eating urges. Understanding these moderation effects clarifies how BMI and weight suppression differ based on ED symptom severity. Overall, based on the current literature, it was hypothesised that:

H1: A lower continuous BMI, but higher weight suppression assessed at baseline would predict a higher state-based body dissatisfaction, and more frequent disordered eating urges (dietary restraint, excessive exercise, binge eating and unhealthy eating).

H2: The EAT-26 total score assessed at baseline would moderate the relationships between BMI and weight suppression, and state-based body dissatisfaction and disordered eating urges (dietary restraint, excessive exercise, binge eating and unhealthy eating).

2. Method

2.1. Participants

After obtaining approval from the Human Research Ethics Committee, participants were recruited through the University's Research Experience Program and various social media platforms. Out of 1932 baseline survey respondents, 969 completed the EMA phase. To avoid bias from systematic completion differences, 283 participants who completed less than 50% of EMA surveys were excluded (Shiffman et al., 2008). The remaining participants

provided a comparable number of surveys to those in similar EMA studies (e.g., Fuller-Tyszkiewicz et al., 2019). The final sample comprised 686 participants (mean age = 19.95, SD = 4.54; mean BMI = 22.14 kg/m², SD = 4.33) aged 18 to 76, with 37 reporting a lifetime ED diagnosis. Most were female (75%), single (69%), had completed year 12 or less (83%), and identified as Asian (55%). Detailed demographic characteristics are in Table 1.

Table 1.
Demographic characteristics of the total sample.

Demographic variables	**Statistics**
Age (M ± SD)	19.95 ± 4.54
Gender (%)	
Female	516 (75%)
Male	161 (23%)
Others	9 (1%)
Sexual orientation (%)	
Heterosexual	523 (78%)
Homosexual	28 (4%)
Bisexual	96 (14%)
Asexual	10 (1%)
Others	16 (2%)
Ethnicity (%)	
Caucasian	242 (35%)
Southern Asian/ Southeast Asian	200 (29%)
Easter Asian	178 (26%)
Others	66 (10%)
Marital Status (%)	
Single	474 (69%)
In a relationship	198 (29%)
Married	12 (1.7%)
Others	2 (0.3%)
Education (%)	
Year 12 or below	605 (83%)
Certificate	10 (1.3%)
Diploma	23 (3.3%)
Bachelor's Degree	71 (9.7)
Postgraduate	20 (2.7%)
Eating disorder diagnosed (%)	
No	649 (95%)
Yes	37 (5%)
ED Risk Status	
Not At risk	556 (82%)
At risk	130 (18%)

Note. ED: Eating Disorder

2.2. Procedure

2.2.1. Phase 1: Trait-Based Assessment

After signing up for the study, participants were first provided with a web link to an online survey platform with information about the study, including a consent form. Once consent had been obtained, participants were asked to provide their email and a self-generated unique identifier (ID) to be able to link the trait (Phase 1) and state-based (Phase 2) data. Participants were then asked to fill in a baseline survey, which gathered demographic information and the variables (BMI and weight suppression) of interest. Upon completion, participants were emailed instructions on how to download and use the custom-built EMA smartphone app SEMA3.

2.2.2. Phase 2: Ecological Momentary Assessment

The morning after completing Phase 1, SEMA3 sent push notifications six times per day at semi-random intervals (every 1-2 hours) between 9:00 AM and 10:00 PM for seven days (up to 42 assessments). Participants completed a brief 1-2 minute survey on body dissatisfaction and disordered eating urges. Random intervals aimed to reduce response bias, while the brief survey minimized participant burden (Fuller-Tyszkiewicz et al., 2020). Participants had a 30-minute window to complete each survey before it expired and was marked as missing. Researchers provided support and sent up to two reminders over seven days to those with less than 50% compliance (Chia et al., 2018). Participants were debriefed upon study completion. Participants recruited from the community were entered into a draw to win one of five AUD100 e-gift cards, (see flowchart, Supplementary 1).

2.3. Measures

2.3.1. Baseline Measures

2.3.1.1. Demographics. Participants provided self-reported information on various demographic-related variables, including their age, gender, sexual orientation, current height, highest weight, current weight, and lowest weight (used to calculate BMI, and weight

suppression), primary language, cultural background, marital status, the highest level of education completed, and whether they had a lifetime diagnosis of an ED (current or prior).

2.3.1.2. Eating pathology. Participants' attitudes and behaviours related to eating were evaluated using the EAT-26 (Garner et al., 1982). The EAT-26 includes 26 items scored on a 4-point Likert scale (0–3), with a possible total of 0–78 points. A total score of ≥ 20 was used to differentiate individuals with clinically significant disordered eating from those with subclinical levels (Garner, et al.,1982). In the present study, the internal consistency of the EAT-26 was strong (Cronbach's alpha = .93).

2.3.2. Ecological Momentary Assessment Measures

2.3.2.1. Body dissatisfaction. Participants were asked about their current satisfaction with their appearance (i.e., *"How satisfied are you with your appearance right now?*") using an 11-point scale ranging from 0 (completely dissatisfied) to 10 (completely satisfied). The responses were reverse scored so that a higher score indicated a higher level of body dissatisfaction. This single-item measurement was adapted from previous EMA studies in ED literature that examined body dissatisfaction (Fuller-Tyszkiewicz et al., 2020; Portingale et al., 2023, Liu et al., 2022, Martin et al., 2023).

2.3.2.2. Urges to engage in disordered eating***.*** Participants were asked to report whether they had experienced urges to engage in specific disordered eating behaviours since the previous survey. These behaviours included: (a) "*Urge to consciously restrict food intake to control weight/shape*" (dietary restraint), (b) *"Urge to engage in at least 15 minutes of exercise to control weight/shape*" (exercise), (c) "*Urge to eat a large amount of food relative to what others would eat in the same situation/time*" (binge eating), and (d) *Urge to eat unhealthy food*". Participants

responded with a code of 1 for "yes" and 0 for "no". Similar items have been used in previous studies to assess DE behaviours and have demonstrated validity about disordered eating constructs (Fitzsimmons-Craft et al., 2016, Martin et al., 2023; Portingale et al., 2022).

2.4. Statistical analysis

2.4.1. Preliminary analyses

Prior to the main analyses, the quality of the baseline and EMA data was evaluated. There was no missing data among BMI, weight suppression, eating psychopathology (measured by the EAT-26) as participants were required to answer all questions during Phase 1 via the Qualtrics platform.

There were no missing data in Phase 2 of the EMA, but participants varied in the number of completed surveys out of 42. Preliminary tests checked for biases by analyzing compliance rates using Pearson's correlation or ANOVA with baseline demographic and predictor variables to find differences between those who completed over 50% and those with less than 50% compliance.

Second, all EMA-assessed variables used as outcomes in the hypothesis testing (e.g., state body dissatisfaction) were examined for time-related and reactivity effects. This involved inspecting whether reports of these outcomes differed based on the time of day (in hourly blocks), day of the week (weekday vs. weekend), and/or the order of assessment throughout the EMA phase. Any significant time-related or reactivity effects were included as covariates in the final models to account for their effects on the modelled outcome variables.

2.4.2. Main analyses

Multilevel modelling was employed for hypothesis testing to account for the non-independence of data resulting from repeated assessments during the EMA phase. We employed a bottom-up approach to data analysis, gradually increasing the complexity of our models. First, null models - only including Level 1 outcome variables (state body

dissatisfaction and disordered eating urges) - were tested to explore between-person differences in average levels for these constructs relative to within-person fluctuations. Multilevel modelling was only utilised when significant variance was observed across individuals, based on intraclass correlation coefficients (ICC) exceeding 5%.

Next, the Level 2 predictors (BMI and weight suppression) were individually included in the models to evaluate their effects on predicting the average of Level 1 outcome variable. Finally, we assessed the moderating effect of the Level 2 trait EAT-26 total score on the relationships between both BMI and weight suppression and the tested Level 1 outcome variables (body dissatisfaction and disordered eating urges). All data pre-processing and analyses were conducted using RStudio version 4.0.2. P value of <.05 was considered statistically significant.

To compare BMI and weight suppression in their ability to predict changes in interested state-based variables using standardised beta coefficients, participants' BMI, weight suppression and weight span were converted into z scores ($M = 0$ and $SD = 1$).

3. Results

3.1. Preliminary analyses

3.1.1. Compliance

On average, participants completed 33.21 ($SD = 6.15$) out of 42 possible EMA surveys, indicating a compliance rate of 79.07%. There were no significant associations between compliance rates for EMA surveys and BMI ($p = .546$), weight suppression ($p = .091$), EAT-26 total score ($p = .277$), or various demographic-related variables, including age ($p = .254$), ethnicity ($p = .414$), education ($p = .517$), gender ($p = .865$), sexual orientation ($p = .764$), marital status ($p = .051$) and ED lifetime diagnosis ($p = .734$).

3.1.2. Time-related and reactivity effect

Significant reactivity effects were found for all Level 1 outcome variables (see Table 2). Specifically, throughout phase 2, scores for body dissatisfaction (p <.001) increased, while urges for dietary restraint *(*p <.001), excessive exercise (p <.001), and unhealthy eating (p < .001) decreased. It was also observed that lower urges for dietary restraint (p <.001), excessive exercise (p <.001), unhealthy eating (p <.001), and higher urges for binge eating (p = .026) were reported later in the day. A significantly higher body dissatisfaction was reported on weekdays compared to weekends (p = .007). Therefore, to account for these effects, the order of assessment, time of day, and day of week were included as covariates in the main analyses.

Table 2.

Effects of Order of Assessment, Time of Day, and Day of Week on Outcomes

Predictors	Body dissatisfaction			Negative mood			Urges for exercise			Urge for dietary restraint			Urges for binge eating			Urges for healthy eating		
	b	*t*	*p*	b (95% CIs)	*t*	*p*	b	*t*	*p*	b	*t*	*p*	b	*t*	*p*	b	*t*	*p*
Order of assessment	0.009	7.97	<.001	0.009	8.08	<.001	-0.001	-7.83	<.001	-0.001	-6.17	<.001	-0.00008	-.0723	.0469	-0.001	-6.38	<.001
Time of day	-0.006	-1.96	.050	-0.020	-6.44	<.001	-0.003	-6.81	<.001	-0.002	-5.279	<.001	0.001	2.39	.016	-0.002	3.95	<.001
Day of week	0.066	2.676	.007	-0.044	-1.73	.083	0.004	1.16	.245	0.001	0.264	.792	-0.001	-0.21	.834	-0.007	-1.53	.127

Note. *b* = unstandardised coefficients, t = t-values for continuous variable, significant p values are bolded.

3.2. Main analyses

3.2.1. Descriptive Statistics

Descriptive statistics of both Level 1 and Level 2 variables (Table 3), including means, standard deviations, and the possible range of continuous variables, as well as frequencies for categorical variables, are presented in Table 3. There was no concern with multicollinearity among the predictors as the correlations between these variables were small (see Supplementary Figure 2). On average, moderate levels of body dissatisfaction were observed in the EMA assessments. Urges for binge eating were reported relatively infrequently (7%), while urges for dietary restraint (35%) and exercise (28%) were more frequently reported. Additionally, 19% of the participants were found to have exhibited clinically significant levels of ED symptoms based on their scores on the EAT-26 (≥ 20).

Table 3.
Descriptive statistics for Level 1 and Level 2 variables

Variables	M ± SD	ICC_{BP}	Range
Level 1			
Body Dissatisfaction	5.09 ± 2.34	0.51	0-10
Urges for dietary restraint (n, %)	177 (35%)	0.39	0-1
Urges for exercise (n, %)	190 (28%)	0.29	0-1
Urges for healthy eating (n, %)	182 (27%)	0.24	0-1
Urges for binge eating (n, %)	49 (7%)	0.21	0-1
Level 2			**Min-Max**
BMI	22.14 ± 4.33	-	15.1-58.6
Weight Suppression	4.96 ± 5.08	-	0-66
EAT-26	11.56 ± 11.63	-	0-65

Note: EAT-26: Eating Attitudes Test-26

3.2.2. Multilevel modelling

3.2.2.1. (H1) BMI and Weight Suppression Predicting Daily Changes in body dissatisfaction, and disordered eating urges

A lower BMI predicted greater average levels of body dissatisfaction (p = .005), and lower urges of having unhealthy eating (p <.001; Table 4).

A lower level of weight suppression was found to be associated with higher average levels of body dissatisfaction (p = .004) and lower urges to engage in unhealthy eating (p = .039) throughout the one-week EMA period. No other significant associations were identified between BMI and weight suppression with the other state-based EMA variables.

3.2.2.2. (H2) Moderating Effects of EAT-26 Total Score

Moderation analyses revealed a significant moderation effect of the overall EAT-26 score on the relationship between BMI and urges for dietary restraint (p = .021). Specifically, BMI has a positive relationship with dietary restraint at low levels of disordered eating symptomatology and a negative relationship with dietary restraint at high levels of disordered eating symptomatology.

Table 4
Fixed Effects for BMI and Weight Suppression, Predicting Outcomes and Moderation Effects of EAT-26 on the tested relationships.

Predictors	Body dissatisfaction			Urge for dietary restraint			Urges for excessive exercise			Urge for dietary restraint			Urges for unhealthy eating		
	b (95% CIs)	*se*	*p*	b (95% CIs)	*se*	*p*	b (95% CIs)	*se*	*p*	b (95% CIs)	*se*	*p*	b (95% CIs)	*Se*	*p*
BMI	-0.13	0.049	**.005**	0.17	0.12	.134	0.15	0.09	.102	0.17	0.24	.129	0.26	0.08	**<.001**
Weight Suppression	-0.14	-0.005	**.004**	0.08	0.12	.498	0.13	0.09	.158	0.21	0.11	.067	0.16	0.08	**.039**
BMI x EAT-26 total	-0.001	-0.004	.797	-0.02	0.01	**.021**	-0.02	0.01	.078	-0.002	0.01	.866	0.01	0.01	.516
Weigh Suppression x EAT 26 total	-0.004	-0.004	.342	-0.01	0.01	.182	0.001	0.01	.890	-0.004	0.01	.670	0.01	0.01	.282

Note. EAT-26: Eating Attitudes Test-26

4. Discussion

The current EMA study examined whether BMI and weight suppression would predict state-based body dissatisfaction and disordered eating urges averages in a sample of predominately female university students throughout a period of seven days. Our findings showed limited support for our two weight-based ratings in predicting state-based body dissatisfaction and disordered eating urge averages. Notably, lower BMI and weight suppression were associated with higher body dissatisfaction, while lower BMI was linked to higher urges for unhealthy eating, and higher weight suppression was associated with increased urges for unhealthy eating. Further moderation analyses for trait-based ED symptomatology (assessed through the total EAT-26 score) on the tested relationships were mainly found to be non-significant. Each of these findings will be discussed in more detail in the subsequent paragraphs.

4.1. Predictive validity of BMI on state body dissatisfaction and disordered eating urges

Consistent with previous studies (e.g., Porras-Garcia et al., 2020), we also found that those with a lower BMI tended to report greater daily body dissatisfaction body dissatisfaction relative to those with a higher BMI. Although the exact mechanisms behind this relationship remain unclear, previous studies have suggested that individuals with a lower BMI tend to overestimate their body size and shape, while those with a higher BMI might provide an underestimation (Cornelissen et al., 2015; Porras-Garcia et al., 2020). However, there are also other studies (e.g., Jones et al., 2019; Kogure et al., 2019) that found contradicting results to ours where individuals with a higher BMI were more likely to experience higher levels of body dissatisfaction. Such disparities can be attributed to differences in the methodology or the sample characteristics (e.g., university students vs. general or clinical ED populations). Specifically, our study focused on university students with a mean BMI within the normal range (22.14 kg/m^2), while other research was conducted with individuals whose mean BMI fell within the overweight range (e.g., 29.1 kg/m^2; Kogure

et al., 2019). Therefore, the current finding regarding BMI as an important predictor of body dissatisfaction can only be generalised to populations that have similar characteristics to ours.

Another important finding was that a lower BMI was related to greater urges to engage in unhealthy eating. This finding contradicts the existing literature, which has shown a significant positive association between BMI and unhealthy eating behaviours (Guertin et al., 2021; Heerman et al., 2017). It is important to note, however, that previous studies have primarily focused on individuals' actual engagement in unhealthy eating behaviours, whereas our study assessed only the presence of urges to engage in such disordered eating behaviours. The relationship between urges and actual behaviours can be influenced by factors such as preoccupation with food and eating as well as dietary restraint. Nevertheless, the complexity of such relationships should be further investigated in future research.

4.2. Weight suppression predicting body dissatisfaction and disordered eating urges.

Inconsistent with our predictions and the literature (e.g., Berner et al., 2013; Burnette et al., 2018), lower weight suppression was found to be associated with a higher average daily level of body dissatisfaction. Such discrepancy from previous studies can be once again attributed to the lower mean weight suppression observed in our university student sample ($M = 4.96$) compared to similar community-based studies, where mean weight suppression ranges were from 8.19 to 12.85 (e.g., Chen et al., 2022; Burnette et al., 2018). Differences in mean weight suppression between our study and prior research may be due to our predominantly female (75%) and Asian (55%) participant makeup, compared to earlier studies with fewer Asian participants (e.g., 7.8% in Burnette et al.) and more males (e.g., 48% in Chen et al., 2022). Therefore, our study's findings mainly apply to individuals with similar characteristics. This suggests a need for future studies to explore if there is a specific

cut-off threshold of weight suppression where it might have a reverse effect (i.e., lower weight suppression leading to higher body dissatisfaction) from what is commonly expected in the literature.

Weight suppression was only found to significantly predict higher daily urges to eat unhealthy food, with a small effect size ($\beta = 0.16$). This finding aligns with the concept that those who suppress their weight, often through dieting and inappropriate compensatory actions, are more inclined to fixate on food, especially unhealthy choices (Butryn et al., 2011). However, weight suppression did not associate with any of the remaining tested disordered eating urges, thus offering limited evidence for its validity in predicting disordered eating urges within natural environments.

Overall, while previous literature has supported the usage of weight suppression to be an alternative way to predict the level of body dissatisfaction and disordered eating urges, weight suppression did not outperform BMI in indexing the above-mentioned factors. Several factors could explain this. First, in this study, BMI was treated as a continuous variable, enhancing statistical power compared to the use of categorical cut-offs, as proposed by the DSM-5, in previous studies (e.g., Dang et al., 2022). These prior studies demonstrated that arbitrary BMI cut-off values were insufficient in capturing variations in ED symptomatology. Second, the current study did not control for the duration of weight loss, which in previous studies has been found to moderate the impact of weight suppression on indexing body dissatisfaction and disordered eating urges. For instance, individuals who lost 4 kg over 4 weeks are likely to face higher disordered eating urges, body dissatisfaction, and psychological and medical risks compared to those who lost the same amount over 6 months. Therefore, future research could replicate the current study using categorical severity indicators after having identified optimal cut-offs for indexing symptom levels across groups.

4.4. The moderating effect of EAT-26 on the relationship between BMI and disordered eating urges

Our second hypothesis was partially supported, in that ED symptomatology – measured by the EAT-26 – moderated the BMI-dietary restraint relationship. While in the current study, BMI alone was not a strong predictor of urges to engage in dietary restraint, there was a significant BMI-ED symptomatology interaction. BMI positively influenced dietary restraint urges at low ED symptomatology levels (i.e., low EAT-26 trait score) but negatively affected it at high ED symptomatology levels (i.e., high EAT-26 trait score). However, we did not identify any other moderating effects of EAT-26 on the tested relationships between our predictors (BMI and weight suppression) and our state-based outcome variables (e.g., body dissatisfaction and other disordered eating urges). Hence, future studies should explore other potential moderating factors (e.g., overvaluation of weight and shape, drive for thinness) that have been found to influence the relationships between BMI and weight suppression with state-based body dissatisfaction and disordered eating urges (e.g., Chernyak & Lowe, 2010), as limited knowledge currently exists in this area.

4.5. Implications

While the findings from the current study need to be replicated, several research and clinical implications can be drawn. Our results provide limited support for using BMI and weight suppression as reliable indicators of state-based averages of body dissatisfaction and disordered eating severity in a community sample. No single physiological indicator currently exists that can accurately inform clinicians about the level of body dissatisfaction and disordered eating urges in the community. Therefore, comprehensive screening in the community remains essential. Practical steps for assessing weight-related aspects may include measuring important biological (e.g., potassium, magnesium, e.g., Quesnel et al., 2023) and cognitive factors (e.g., drive for thinness, overvaluation of weight and shape, e.g., Krug et al., 2021) that predict the

need for early biological and psychological interventions. This is especially important since delayed in interventions has been found to contribute to a protracted illness Course (Austin et al., 2020). Ideally, a combination of self-report questionnaires, clinical interviews, and physiological measures should be used simultaneously to develop tailored and effective prevention and early intervention plans for body dissatisfaction and disordered eating in non-clinical populations.

4.6. Strengths, Limitations and Future Direction

The current study has several notable strengths. First, it employs EMA, which offers naturalistic, micro-longitudinal data in a field often dominated by experimental and cross-sectional research (e.g., Dang et al., 2023; Krug et al., 2021). This enhances ecological validity, generalisability, and allows for the analysis of temporal patterns. Second, it is the first EMA study to examine the predictive validity of both BMI and weight suppression for body dissatisfaction and disordered eating urges within a university sample, providing new insights into these variables in daily life. Lastly, the study benefits from a larger sample size compared to similar EMA research, which strengthens its findings.

The current findings must be interpreted while keeping several limitations in mind. First, to balance between assessment scope and participant burden, the EMA-assessed constructs were captured using single-item questions, which may have missed nuances associated with the assessed constructs, including the effect and cognitions underlying disordered eating urges. This measurement approach, while having been consistently used in other EMA research (Portingale et al., 2022), could have resulted in a less precise estimation of the true effect of the predictor variables. More nuanced questions (e.g., EMA-adapted Body-Image Ideals Questionnaire; Cash, 2000) should be included in upcoming EMA studies assessing the severity of body dissatisfaction and disordered eating.

Second, to facilitate data collection in a non-clinical population, we assessed disordered eating urges rather than actual engagement in such behaviours. It is crucial to

recognize that the urge to engage in disordered eating does not always directly translate into behaviour, as various cognitive factors, such as drive for thinness or preoccupation with food, can moderate this connection. This distinction is important because previous studies have shown that drive for thinness (Krug et al., 2021) and weight and shape concerns (Dang et al., 2022; Gianini et al., 2017) are superior indicators of ED psychopathology in clinical populations. Future studies should account for these potential moderators while also assessing the frequency of actual disordered eating behaviours, not just the urges. Additionally, body dissatisfaction has been linked to factors like self-esteem (e.g., Cruz-Sáez et al., 2020), social comparisons (Ryding et al., 2020), and negative affect (e.g., Mclean et al., 2010), which play significant roles in the development and maintenance of disordered eating behaviours. Future research should integrate aforementioned to better understand and address disordered eating behaviours in both clinical and non-clinical populations, using both EMA and longitudinal assessments to inform targeted interventions.

This study used self-reported weight and height to calculate BMI and weight suppression, which may affect the reliability of these predictors for body dissatisfaction and disordered eating urges in a community setting. Our sample showed lower weight suppression compared to other studies (Chen et al., 2022). Future research should use more accurate weight and height measurements, such as medical records or direct measurements by healthcare professionals. Additionally, given the variations in ED presentations between genders (Murray et al., 2017) and the high proportion of university students in this sample, studies should include more diverse community samples to better understand the relationship between BMI, weight suppression, body dissatisfaction, and disordered eating.

With a mean participant age of 19.95 years (SD = 4.54), the concept of developmental weight suppression (Singh et al., 2021)—the difference between highest premorbid and current z-BMIs—could be promising. This metric has shown greater sensitivity to ED

psychopathology compared to traditional calculations used in this study. Future research should explore the clinical utility of developmental weight suppression, as our study did not include data on the ages at which participants reached their highest past weight.

Conclusion

This study expands the ED severity literature using a micro-longitudinal design in a community setting. Our findings provide limited support for BMI and weight suppression in predicting state-based body dissatisfaction and disordered eating urges. However, the study highlights the need for more longitudinal and diverse research to better understand ED severity. Clinicians should use comprehensive assessments of body dissatisfaction and disordered eating rather than relying solely on weight-based severity indicator.

References

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

Austin, A., Flynn, M., Richards, K., Hodsoll, J., Duarte, T. A., Robinson, P., ... & Schmidt, U. (2021). Duration of untreated eating disorder and relationship to outcomes: A systematic review of the literature. *European Eating Disorders Review*, 29(3), 329-345.https://doi.org/10.1002/erv.2745.

Berner, L. A., Shaw, J. A., Witt, A. A., & Lowe, M. R. (2013). The relation of weight suppression and body mass index to symptomatology and treatment response in anorexia nervosa. *Journal of Abnormal Psychology*, *122*(3), 694. https://doi.org/10.1037/a0033930

Bodell, L. P., Racine, S. E., & Wildes, J. E. (2016). Examining weight suppression as a predictor of eating disorder symptom trajectories in anorexia nervosa. *International Journal of Eating Disorders*, *49*(8), 753–763. https://doi.org/10.1002/eat.22545

Burnette, C. B., Simpson, C. C., & Mazzeo, S. E. (2018). Relation of BMI and weight suppression to eating pathology in undergraduates. *Eating Behaviors*, 30, 16-21. https://doi.org/10.1016/j.eatbeh.2018.05.003

Butryn, M. L., Juarascio, A., & Lowe, M. R. (2011). The relation of weight suppression and BMI to bulimic symptoms. *International Journal of Eating Disorders, 44*(7), 612-617. https://doi.org/10.1002/eat.20881

Cash, T. F. (2000). Users' manuals for the Multidimensional Body-Self Relations Questionnaire, the Body-Image Ideals Questionnaire, the Appearance Schemas Inventory, the Situaional Inventory of Body-Image Dysphoria, and the Appearance Schemas Inventory.

Carter FA, McIntosh VV, Joyce PR, Bulik CM. Weight suppression predicts weight gain over treatment but not treatment completion or outcome in bulimia nervosa (2008). *Journal of Abnormal Psychology*.117(4). doi: 10.1037/a0013942

Chia, J. L. P., Fuller-Tyszkiewicz, M., Buck, K., Chamari, K., Richardson, B., & Krug, I. (2018). An ecological momentary assessment of the effect of fasting during Ramadan on disordered eating behaviors. *Appetite,* 127, 44–51. https://doi.org/10.1016/j.appet.2018.04.017

Chen, J. Y., Piers, A. D., Lesser, E. L., & Lowe, M. R. (2022). The effect of weight suppression on eating behavior: Does the intentionality of weight loss matter?. *Appetite, 174,* 106017. https://doi.org/10.1016/j.appet.2022.106017

Chernyak, Y., & Lowe, M. R. (2010). Motivations for dieting: Drive for Thinness is different from Drive for Objective Thinness. *Journal of Abnormal Psychology*, *119* (2), 276. https://doi.org/10.1037/a0018398

Coffino, J. A., Orloff, N. C., & Hormes, J. M. (2016). Dietary restraint partially mediates the relationship between impulsivity and binge eating only in lean individuals: the importance of accounting for body mass in studies of restraint. *Frontiers in Psychology,* 7, 1499. https://doi.org/10.3389/fpsyg.2016.01499

Cornelissen, K. K., Bester, A., Cairns, P., Tovée, M. J., & Cornelissen, P. L. (2015). The influence of personal BMI on body size estimations and sensitivity to body size change in anorexia spectrum disorders. *Body Image, 13*, 75-85. https://doi.org/10.1016/j.bodyim.2015.01.001

Cruz-Sáez, S., Pascual, A., Wlodarczyk, A., & Echeburúa, E. (2020). The effect of body dissatisfaction on disordered eating: The mediating role of self-esteem and negative affect in male and female adolescents. Journal of Health Psychology, 25(8), 1098-1108. https://doi.org/10.1177/1359105317748734

Dalle Grave, R., Sartirana, M., El Ghoch, M., & Calugi, S. (2018). Dsm-5 severity specifiers for anorexia nervosa and treatment outcomes in adult females. *Eating Behaviors*, *31*, 18–23. https://doi.org/10.1016/j.eatbeh.2018.07.006

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the *DSM–5* severity ratings for eating disorders. *Clinical Psychology: Science and Practice,* *29*(4), 325–344. https://doi.org/10.1037/cps0000078

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2023). Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates?. *European Eating Disorders Review*, 31(4), 447-461. https://doi.org/10.1002/erv.2969

Drutschinin, K., Fuller-Tyszkiewicz, M., Paoli, T. D., Lewis, V., & Krug, I. (2018). The daily frequency, type, and effects of appearance comparisons on disordered eating. *Psychology of Women Quarterly*, *42*(2), 151-161. https://doi.org/10.1177/0361684317732001

Fitzsimmons-Craft, E. E., Ciao, A. C., & Accurso, E. C. (2016). A naturalistic examination of social comparisons and disordered eating thoughts, urges, and behaviours in college women. *International Journal of Eating Disorders*, *49*(2), 141-150. https://doi.org/10.1002/eat.22486

Fuller-Tyszkiewicz, M., Krug, I., Smyth, J.M., Fernandez-Aranda, F., Treasure, J., Linardon, J., Vasa, R. and Shatte, A., (2020). State-based markers of disordered eating symptom severity. *Journal of Clinical Medicine*, *9*(6), 1948. https://doi.org/10.3390/jcm9061948

Fuller-Tyszkiewicz, M., Richardson, B., Lewis, V., Smyth, J., & Krug, I. (2018). Do women with greater trait body dissatisfaction experience body dissatisfaction states differently?

An experience sampling study. *Body image*, *25*, 1–8. https://doi.org/10.1016/j.bodyim.2018.01.004

Garber, A. K., Cheng, J., Accurso, E. C., Adams, S. H., Buckelew, S. M., Kapphahn, C. J., Kreiter, A., Le Grange, D., Machen, V.I., Moscicki, A.B. and Saffran, K. (2019). Weight loss and illness severity in adolescents with atypical anorexia nervosa. *Pediatrics,* 144(6). https://doi.org/10.1542/peds.2019-2339

Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitudes test: psychometric features and clinical correlates. *Psychological Medicine*, 12(4), 871-878. https://doi:10.1017/S0033291700049163

Gianini, L., Roberto, C. A., Attia, E., Walsh, B. T., Thomas, J. J., Eddy, K. T., ... Sysko, R. (2017). Mild, moderate, meaningful examining the psychological and functioning correlates of dsm-5 eating disorder severity specifiers. *International Journal of Eating Disorders*, *50*(8), 906–916. https://doi.org/10.1002/eat.22728

Goddard, E., Hibbs, R., Raenker, S., Salerno, L., Arcelus, J., Boughton, N., Connan, F., Goss, K., Laszlo, B., Morgan, J. and Moore, K. (2013). A multi-centre cohort study of short term outcomes of hospital treatment for anorexia nervosa in the UK. *BMC Psychiatry*, *13*,1-14. https://doi.org/10.1186/1471-244X-13-287

Goldschmidt, A. B., Wall, M. M., Loth, K. A., Bucchianeri, M. M., & Neumark-Sztainer, D. (2014). The course of binge eating from adolescence to young adulthood. *Health Psychology, 33*(5), 457. https://doi.org/10.1037/a0033508

Gonçalves, S. F., & Gomes, A. R. (2012). Exercising for weight and shape reasons vs. health control reasons: The impact on eating disturbance and psychological functioning. *Eating Behaviors,* 13(2), 127-130. https://doi.org/10.1016/j.eatbeh.2011.11.011

Gorrell, S., Reilly, E. E., Schaumberg, K., Anderson, L. M., & Donahue, J. M. (2019). Weight suppression and its relation to eating disorder and weight outcomes: a narrative review. *Eating Disorders*, *27*(1), 52–81. https://doi.org/10.1080/10640266.2018.1499297

Guertin, C., & Pelletier, L. (2021). Planning and self-monitoring the quality and quantity of eating: How different styles of self-regulation strategies relate to healthy and unhealthy eating behaviors, bulimic symptoms, and BMI. *Appetite*, *156*, 104839. https://doi.org/10.1016/j.appet.2020.104839

Hart, L. M., Granillo, M. T., Jorm, A. F., & Paxton, S. J. (2011). Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. *Clinical Psychology Review*, *31*(5), 727–735. https://doi.org/10.1016/j.cpr.2011.03.004

Hay, P., Aouad, P., Le, A., Marks, P., Maloney, D., Touyz, S., & Maguire, S. (2023). Epidemiology of eating disorders: population, prevalence, disease burden and quality of life informing public policy in Australia—a rapid review. *Journal of Eating Disorders,* 11(1), 23. https://doi.org/10.1186/s40337-023-00738-7

Heerman, W. J., Jackson, N., Hargreaves, M., Mulvaney, S. A., Schlundt, D., Wallston, K. A., & Rothman, R. L. (2017). Clusters of healthy and unhealthy eating behaviors are associated with body mass index among adults. *Journal of Nutrition Education and Behavior*, *49*(5), 415-421. https://doi.org/10.1016/j.jneb.2017.02.001

Jennings, K. M., Gregas, M., & Wolfe, B. (2018). Trajectories of change in body weight during inpatient treatment for anorexia nervosa. *Journal of the American Psychiatric Nurses Association,* 24(4), 306-313. https://doi.org/10.1177/1078390317726142

Juarascio, A., Lantz, E. L., Muratore, A. F., & Lowe, M. R. (2018). Addressing Weight Suppression to Improve Treatment Outcome for Bulimia Nervosa. *Cognitive and Behavioral Practice*, 25(3), 391–401. https://doi.org/10.1016/j.cbpra.2017.09.004

Jones, C. L., Fowle, J. L., Ilyumzhinova, R., Berona, J., Mbayiwa, K., Goldschmidt, A. B., ... & Keenan, K. E. (2019). The relationship between body mass index, body dissatisfaction, and eating pathology in sexual minority women. *International Journal of Eating Disorders,* 52(6), 730-734. https://doi.org/10.1002/eat.23072

Kogure, G. S., Ribeiro, V. B., Lopes, I. P., Furtado, C. L. M., Kodato, S., de Sá, M. F. S., ... & Dos Reis, R. M. (2019). Body image and its relationships with sexual functioning, anxiety, and depression in women with polycystic ovary syndrome. *Journal of Affective Disorders,* 253, 385-393. https://doi.org/10.1016/j.jad.2019.05.006

Krug, I., Binh Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J.M., Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the dsm-5 severity indices for eating disorders *European Eating Disorders Review*, *29*(3), 482–498. https://doi.org/10.1002/erv.2776

Liu, D., Huang, Y., Huang, C., Yang, S., Wei, X., Zhang, P., Guo, D., Lin, J., Xu, B., Li, C. Zhang, H. (2022). Calorie restriction with or without time-restricted eating in weight loss. *New England Journal of Medicine*, 386(16), 1495-1504. https://10.1056/NEJMoa2114833

Lowe, M. R. (1993). The effects of dieting on eating behavior: A three-factor model. Psychological Bulletin, 114(1), 100–121. https://doi.org/10.1037/0033-2909.114.1.100

Machado, P. P., Grilo, C. M., & Crosby, R. D. (2017). Evaluation of the DSM-5 severity indicator for anorexia nervosa. *European Eating Disorders Review*, *25*(3), 221–223. https://doi.org/10.1002/erv.2508

McLean, S. A., Paxton, S. J., & Wertheim, E. H. (2010). Factors associated with body dissatisfaction and disordered eating in women in midlife. International Journal of Eating Disorders, 43(6), 527-536.

Martin, G., Portingale, J., Fuller-Tyszkiewicz, M., & Krug, I. (2023). Do appearance comparisons mediate the effects of Porras-Garcia, B., Ferrer-Garcia, M., Yilmaz, L., Sen, Y. O., Olszewska, A., Ghita, A., Serrano-Troncoso, E., Treasure, J., Gutiérrez-Maldonado, J. (2020). Body-related thinspiration and fitspiration on body dissatisfaction, happiness, and disordered eating urges in women's daily lives?. *Body Image, 46*, 108-116. https://doi.org/10.1016/j.bodyim.2023.05.006

Mason, T. B., Smith, K. E., Crosby, R. D., Engel, S. G., & Wonderlich, S. A. (2021). Examination of momentary maintenance factors and eating disorder behaviors and cognitions using ecological momentary assessment. *Eating disorders*, *29*(1), 42–55. https://doi.org/10.1080/10640266.2019.1613847

Matthews, A., Kramer, R. A., & Mitan, L. (2022). Eating disorder severity and psychological morbidity in adolescents with anorexia nervosa or atypical anorexia nervosa and premorbid overweight/obesity. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 1-10. https://doi.org/10.1007/s40519-021-01168-7

Meyer, C., & Waller, G. (1999). The impact of emotion upon eating behavior: The role of subliminal visual processing of threat cues. *International Journal of Eating Disorders, 25*(3), 319-326. https://doi.org/10.1002/(SICI)1098-108X(199904)25:3<319::AID-EAT10>3.0.CO;2-9

Munsch, S., Meyer, A. H., Milenkovic, N., Schlup, B., Margraf, J., & Wilhelm, F. H. (2009). Ecological momentary assessment to evaluate cognitive-behavioral treatment for binge eating disorder. *International Journal of Eating Disorders*, *42*(7), 648–657. https://doi.org/10.1002/eat.20657

Nagata, J. M., Murray, S. B., Compte, E. J., Pak, E. H., Schauer, R., Flentje, A., Capriotti, M.R., Lubensky, M.E., Lunn, M.R., Obedin-Maliver, J. (2020). Community norms for

the Eating Disorder Examination Questionnaire (EDE-Q) among transgender men and women. *Eating Behaviors*, *37*, 101381. https://doi.org/10.1016/j.eatbeh.2020.101381

Obert, J., Pearlman, M., Obert, L., & Chapin, S. (2017). Popular weight loss strategies: a review of four weight loss techniques. *Current gastroenterology reports, 19*, 1-4. https://doi.org/10.1007/s11894-017-0603-8

Ochner, C. N., Barrios, D. M., Lee, C. D., & Pi-Sunyer, F. X. (2013). Biological mechanisms that promote weight regain following weight loss in obese humans. Physiology & Behavior, 120, 106-113. https://doi.org/10.1016/j.physbeh.2013.07.009

Porras-Garcia, B., Ferrer-Garcia, M., Yilmaz, L., Sen, Y. O., Olszewska, A., Ghita, A., ... & Gutiérrez-Maldonado, J. (2020). Body-related attentional bias as mediator of the relationship between body mass index and body dissatisfaction. European Eating Disorders Review, 28(4), 454-464. https://doi.org/10.1002/erv.2730

Portingale, J., Fuller-Tyszkiewicz, M., Liu, S., Eddy, S., Liu, X., Giles, S., & Krug, I. (2022). Love me tinder: The effects of women's lifetime dating app use on daily body dissatisfaction, disordered eating urges, and negative mood. *Body Image*, *40*, 310–321. https://doi.org/10.1016/j.bodyim.2022.01.005

Quesnel, D.A., Cooper, M., Fernandez-del-Valle, M. et al. Medical and physiological complications of exercise for individuals with an eating disorder: A narrative review. J Eat Disord 11, 3 (2023). https://doi.org/10.1186/s40337-022-00685-9

Ryding, F. C., & Kuss, D. J. (2020). The use of social networking sites, body image dissatisfaction, and body dysmorphic disorder: A systematic review of psychological research. *Psychology of Popular Media*, 9(4), 412. https://doi.org/10.1037/ppm0000264

Sala, M., Brosof, L. C., & Levinson, C. A. (2019). Repetitive negative thinking predicts eating disorder behaviors: A pilot ecological momentary assessment study in a treatment-

seeking eating disorder sample. *Behaviour Research and Therapy*, *112*, 12–17. https://doi.org/10.1016/j.brat.2018.11.005

Schaefer, L. M., Engel, S. G., & Wonderlich, S. A. (2020). Ecological momentary assessment in eating disorders research: recent findings and promising new directions. *Current Opinion in Psychiatry*, *33*(6), 528–533. https://10.1097/YCO.0000000000000639

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annu. Rev. Clin. Psychol.*, *4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Singh, S., Apple, D. E., Zhang, F., Niu, X., & Lowe, M. R. (2021). A new, developmentally-sensitive measure of weight suppression. Appetite, 163, 105231. https://doi.org/10.1016/j.appet.2021.105231

Smith, K. E., Ellison, J. M., Crosby, R. D., Engel, S. G., Mitchell, J. E., Crow, S. J., Peterson, C.B., Le Grange, D., Wonderlich, S. A. (2017). The validity of DSM-5 severity specifiers for anorexia nervosa, bulimia nervosa, and binge-eating disorder. *International Journal of Eating Disorders*, *50*(9), 1109–1113. https://doi.org/10.1002/eat.22739

Sparti, C., Santomauro, D., Cruwys, T., Burgess, P., & Harris, M. (2019). Disordered eating among Australian adolescents: Prevalence, functioning, and help received. *International Journal of Eating Disorders*, 52(3), 246-254. https://doi.org/10.1002/eat.23032

Tasca, G. A., Demidenko, N., Krysanski, V., Bissada, H., Illing, V., Gick, M., M., Weekes, K., Balfour, L. (2009). Personality dimensions among women with an eating disorder: Towards reconceptualizing DSM. *European Eating Disorders Review*, *17*(4), 281-289. https://doi.org/10.1002/erv.938

Voss, C., Liu, J., Chang, A., Kosmas, J. A., Biehl, A., Flynn, R. L., ... & Graham, A. K. (2023). Weight Loss Expectations of Adults with Binge Eating: Cross-sectional Study With a

Human-Centered Design Approach. *JMIR Formative Research*, 7(1), e40506. https://10.2196/40506

Wildes, J. E., & Marcus, M. D. (2012). Weight suppression as a predictor of weight gain and response to intensive behavioral treatment in patients with anorexia nervosa. *Behaviour Research and Therapy*, *50*(4), 266–274. https://doi.org/10.1016/j.brat.2012.02.006

Zayas, L. V., Wang, S. B., Coniglio, K., Becker, K., Murray, H. B., Klosterman, E., ... others (2018). Gender differences in eating disorder psychopathology across DSM-5 severity categories of anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, *51*(9), 1098–1102. https://doi.org/10.1002/eat.22941

Chapter 9: General Discussion

This chapter begins with a summary of findings from the six studies conducted as part of this thesis (Section 9.1). In Section 9.2, I will discuss how the identified research gaps were addressed through these studies, focusing on aims, methodologies, and findings that assess the clinical utility of severity ratings across key aspects of EDs, including psychological, biological, risk factors, and treatment outcomes. Following this, I will discuss the clinical implications (Section 9.3), highlight the strengths of the current research (Section 9.4), address its limitations and propose directions for future research (Section 9.5), and conclude the thesis (Section 9.6).

9.1. Summary of Findings

This PhD aimed to investigate the clinical utility of DSM-5 and ICD-11 severity ratings for EDs alongside alternative ED severity indicators (e.g., OWS, DT), focusing on their relevance to clinical practice and intended function. In the following paragraph, I will outline the gaps identified in existing research and how each study within this thesis aimed to address them. A more detailed discussion of the findings is presented in the next section (Section 9.2).

The first aim of the thesis was to synthesise existing research on ED severity ratings, providing a comprehensive understanding of the topic. To achieve this, the project commenced with a systematic review (n = 22) and meta-analysis (n = 19) (Chapter 2, Study 1) that evaluated the current evidence for ED severity ratings. Study 1 (Chapter 2) revealed moderate support for the DSM-5 BN severity rating, which is based on the weekly frequency of inappropriate compensatory behaviours, but limited evidence for the DSM-5 AN (based on BMI) and BED (based on the weekly frequency of binge eating episodes) severity ratings in reflecting ED psychopathology (measured via EDE or EDE-Q). Additionally, the study found

evidence supporting the OWS as an alternative BED severity rating, as individuals with high OWS exhibited greater ED psychopathology compared to those with low OWS.

Study 1 (Chapter 2) identified three additional research gaps. First, although severity ratings were designed to support clinical practice, little is known about whether mental health professionals use them in clinical practice or their perceptions of the utility of these severity ratings in their daily practices. To address this, Study 2 (Chapter 4) recruited Australian-based healthcare professionals involved in ED care, including general practitioners, psychiatrists, psychologists, and dietitians. Quantitative data revealed that 60% of participants did not use these ratings in clinical practice. Qualitative findings, gathered through one-on-one interviews and open-text responses, suggested clinicians perceive these ratings as: (1) invalidating, (2) reinforcing pathology, (3) lacking clinical relevance, and (4) communication of ED severity to others.

The second gap identified in Study 1 (Chapter 2) was the limited focus on alternative severity ratings (e.g., OWS, DT) and their associations with broader clinical correlates, such as risk factors, psychological and biological variables, and treatment outcomes. To address this, a series of three studies were conducted and described in Chapter 5 (Studies 3a and 3b), Chapter 6 (Study 4), and Chapter 7 (Study 5).

To assess the clinical utility of DSM-5, ICD-11, and OWS severity ratings for AN, with a focus on psychological factors (both ED-specific and general psychopathology), psychiatric comorbidities, and biological correlates, the first set of studies (Studies 3a and 3b) utilised a sample of 312 individuals with AN from St Vincent's Hospital, Melbourne. Specifically, Study 3a and the additional analysis (Section 5.3) explored the relationship between DSM-5, ICD-11, and OWS severity ratings in indexing biological variables, ED, and general psychopathology. Study 3b then examined the relationship between the DSM-5, ICD-11, and OWS severity ratings for AN and the rate of psychiatric comorbidity.

Findings from Studies 3a and 3b revealed that both DSM-5 and ICD-11 BMI severity ratings for AN followed a paradoxical pattern that contradicted what an effective severity measure should indicate. Specifically, individuals classified as having milder AN by DSM-5 and ICD-11 criteria (higher BMI) was found to exhibit significantly greater level of ED and general psychopathology (Study 3a), and higher rate of psychiatric comorbidities (e.g., particularly panic disorder, social anxiety disorder, generalised anxiety disorder, and post-traumatic stress disorder, Study 3b) compared to those in more severe (lower BMI) DSM-5 and ICD-11 AN severity groups. In contrast, OWS ratings followed expected patterns of an effective severity rating for AN: individuals in the higher OWS group (more severe) exhibited significantly greater levels of ED and general psychopathology (Study 3a), as well as a higher prevalence of major depressive disorder (Study 3b), compared to those in the lower OWS group. Additionally, regarding biological correlates, Study 3a revealed no significant differences in biological variables (e.g., potassium, magnesium, phosphate) across DSM-5, ICD-11, and OWS severity groups. Overall, Studies 3a and 3b indicated that OWS ratings better reflected psychological severity compared to DSM-5 and ICD-11, while none effectively captured biological correlates of AN, highlighting the need for more comprehensive severity measures.

To further assess the predictive value of DSM-5, ICD-11, and DT severity ratings for treatment outcomes, particularly given their decade-long introduction in the field, Study 5 (Chapter 7) focussed on comparing treatment outcomes for AN and BN based on these severity ratings. A total of 628 female participants (AN: n = 266; BN: n = 362) were recruited from a Spanish ED treatment centre. Participants were categorised into DSM-5, ICD-11, and DT severity groups at admission and underwent AN- or BN-specific CBT-E. Significant associations were found only for the ICD-11 AN severity groups, where the "dangerously low BMI" group predicted poorer outcomes (e.g., a higher likelihood of dropout) compared

to the "significantly low BMI" group. No significant associations were observed for the DSM-5 or DT AN and BN severity ratings, raising concerns about the validity of DSM-5 severity criteria, which may rely on arbitrary BMI cut-offs rather than evidence-based thresholds.

The third and final gap identified by Study 1 (Chapter 2), was the predominant focus on clinical samples in the existing research on ED severity ratings, which limits their generalisability. Clinical samples typically consist of individuals already receiving treatment, whereas community-based samples capture a broader spectrum of individuals affected by EDs, including those who may not seek treatment (Hart et al., 2011). Additionally, findings from Study 5 (Chapter 7) highlighted inconsistencies in how DSM-5 and ICD-11 AN severity ratings, both based on BMI, predicted treatment outcomes. Specifically, DSM-5 AN severity groups showed no significant association with treatment outcomes, whereas the ICD-11 "dangerously low BMI" (more severe) group was associated with poorer outcomes compared to the "significantly low BMI" (less severe) group. This discrepancy raises questions about the adequacy of BMI cut-offs as indicators of AN severity.

To address these issues, Study 6 (Chapter 8) recruited 686 community-based adults, 37 of whom reported having an ED diagnosis. EMA was employed to examine BMI and weight suppression, a promising alternative severity indicator for EDs identified in prior research (Khalil et al., 2024), as continuous variables. This approach avoided skewed distributions where severe cases are often underrepresented in community samples. Study 6 (Chapter 8) aimed to explore how BMI and weight suppression predicted both average levels and daily fluctuations in disordered eating urges and body dissatisfaction. Multilevel modelling analyses revealed that lower BMI and lower weight suppression were associated with greater average body dissatisfaction. Conversely, higher BMI and higher weight suppression were linked to stronger urges to engage in unhealthy eating behaviours.

However, neither BMI nor weight suppression significantly predicted state-based body dissatisfaction or other disordered eating urges (i.e., urges to binge eat, dietary restraint, and excessive exercise). These findings underscored the limited effectiveness of weight-based severity indicators in capturing ED-related state-based variables within a community context.

9.2. Discussion of Key Findings

This section outlines how the identified gaps were addressed through a series of studies. It begins by summarising the aims and methodologies of each study, followed by an integrated discussion of the findings. Rather than reviewing each severity index individually, the focus will be on synthesising the results to assess the clinical utility of severity ratings across various aspects of EDs, including psychological, biological, and concurrent factors, as well as risk factors and treatment outcomes.

Objective 1: Provide a comprehensive and systematic synthesis of the current knowledge on ED severity ratings

The thesis began with a systematic review and meta-analysis (Chapter 2, Study 1) examining the current evidence of the DSM-5 and alternative severity ratings for EDs. The ICD-11 severity rating for AN was not included, as it had not been introduced at the time of the study. The systematic review included 22 studies with 9,196 participants (6,990 females) from both community and clinical settings. The meta-analysis analysed 19 studies, some of which examined severity ratings across all EDs (e.g., Gianini et al., 2017), while others focused on specific disorders (e.g., Grilo et al., 2017 focused solely on BED).

Findings from both the systematic review and meta-analysis were consistent. Specifically, limited support was found for the clinical utility of the DSM-5 severity ratings for AN and BED in indexing ED psychopathology, typically assessed using the EDE and the EDE-Q. However, the findings supported the DSM-5 severity ratings using weekly frequency

of inappropriate compensatory behaviours for BN, that is individuals with BN in the milder DSM-5 severity categories reported significantly less dietary restraint and fewer eating and weight/shape concerns compared to those in more severe categories, with effect sizes ranging from small to large.

The work from Study 1, a systematic review and meta-analysis, revealed that although a substantial number of studies have examined the topic of DSM-5, ICD-11 and alternative severity indices in EDs, the understanding of ED severity ratings has been largely confined to their relationship with ED psychopathology. Although this relationship is a crucial first step, further research is needed to explore the utility of DSM-5, ICD-11 and alternative severity ratings across other aspects (e.g., biological correlates, treatment outcomes), given the multifaceted nature of EDs.

The systematic review from Study 1 (Chapter 2) also provided insights into the strengths and weaknesses of the methodologies used in research on ED severity. Specifically, the quality appraisal component of the review highlighted that most studies in this field employed valid and reliable measures. However, it was also revealed that the methodological rigor of many studies was often compromised by issues such as low statistical power and the failure to report exact p-values. To advance the field, future research must address these limitations by adopting methodologically robust approaches. Moreover, the review noted that all included studies relied exclusively on self-reported ED symptoms, leaving a critical gap in understanding how clinicians apply these severity ratings in real-world practice.

Overall, this comprehensive synthesis was instrumental in identifying key limitations and knowledge gaps in the current research on ED severity indices, which directly informed the empirical investigations in this thesis. Notably, the review highlighted three critical areas for further exploration in ED severity ratings. The objectives of the five subsequent studies (Chapters 4, 5, 6, 7, and 8) were designed to address these gaps, as outlined by the systematic

review and meta-analysis (Chapter 2, Study 1). The following sections provide a recap of the overarching objectives of the thesis, which were developed in response to research gaps identified in Study 1 (Chapter 2). Additionally, the following sections will summarise and discuss in detail how each of the five studies aligns with and addresses these objectives, contributing to a more comprehensive understanding of ED severity ratings.

Objective 2: Mental health clinicians' usage and perceptions about the usefulness of the proposed ED severity ratings in daily clinical practice.

To address the second objective, Chapter 4 (Study 2) presented an original mixed method design, using both qualitative and quantitative analyses. This study explored whether healthcare professionals within multidisciplinary teams treating individuals with EDs in Australia use ED severity ratings in clinical practice, as well as their attitudes toward these ratings. Quantitative data (n = 38) collected from general practitioners, psychiatrists, psychologists, and dietitians revealed that 60% of clinicians did not use these ratings in practice, 36% did, and 3% were unaware of them.

The qualitative component, comprising one-on-one interviews and free-text responses, indicated that 60% of those not using the ratings viewed the ratings as (1) invalidating, (2) pathology-reinforcing, (3) lacking clinical relevance, and (4) communication of ED severity to others. Specifically, for theme 4, while clinicians reported not communicating ED severity to treating clients, some noted that these ratings can aid in triage and communication among clinicians in tertiary settings. The findings from Study 3 (Chapter 4) are largely consistent with previous research, including cross-sectional studies (e.g., Dang et al., 2023; Giannini et al., 2017; Krug et al., 2021) and our meta-analysis (Chapter 2, Study 1; Dang et al., 2022), which demonstrated that the DSM-5 severity indicators do not adequately reflect ED psychopathology, biological factors, or the psychiatric comorbidities experienced by individuals with EDs. This underscores their lack of clinical relevance.

Clinicians also reported that ED severity labels (e.g., "mild" or "moderate") did not significantly influence their clinical approach (theme 3: lack of clinical relevance), as treatment plans were generally the same regardless of the label. Additionally, they noted little correlation between BMI-based severity ratings for AN and the actual severity of symptoms or medical conditions associated with AN. The qualitative findings from Study 2 also highlighted concerns that the DSM-5 severity ratings might do more harm than good (theme 1: invalidating; theme 2: reinforcing pathology). This is particularly concerning given that individuals with EDs often have limited insight into their condition (Konstantakopoulos et al., 2011) and are more likely to present with unrelated physical complaints, such as stomach pain or headaches– a pattern identified in the current literature (Ivancic et al., 2021; Mills et al., 2023; Waller et al., 2014). In primary care settings (e.g., appointments with general practitioners), where consultations are typically limited to 15 minutes, relying on BMI or the frequency of compensatory behaviours as primary indicators of ED severity increases the likelihood of these cases being overlooked (Robinson et al., 2024; Waller et al., 2014). Even when an ED is identified, the label of "mild" may further reduce motivation to seek psychological or dietary support, creating an additional barrier to treatment, especially given the high prevalence of ambivalence toward change among individuals with EDs (Robinson et al., 2024).

Among the 36% of clinicians who used severity ratings in Study 2, most highlighted their administrative utility (theme 4: communication of ED severity to others), including improving communication, triaging, and facilitating referrals to specialised ED services. Clinicians noted that DSM-5 severity ratings streamline communication across services and support decisions about hospital admissions or referrals to community-based care. However, they emphasised that these ratings are only one factor considered in the decision-making process. In addition to BMI and the frequency of binge eating or inappropriate compensatory

behaviours, other key factors discussed by clinicians included the rate of weight loss, weight suppression, electrolyte and haematological levels, psychiatric comorbidities, and the level of social support. This perspective aligns with existing recommendation by RANZCP guideline (Hay et al., 2014), which emphasises the importance of considering a broad range of factors in ED assessment.

Overall, the findings from Study 2 show that DSM-5 severity ratings for EDs are not widely adopted in clinical practice due to concerns about their limited relevance, reinforcement of pathology, and potential to invalidate experiences. While some clinicians use these ratings to aid communication, they view them as just one part of a broader clinical assessment, indicating that the ratings fall short of their intended purpose.

Objective 3: Relationship between ED severity ratings (DSM-5, ICD-11 and alternative severity ratings) across a wide range of other ED-related features (e.g., risk factors, biological and psychological correlates) and treatment outcomes

To address the third objective, Chapters 5, 6, and 7 examined the relationship between DSM-5, ICD-11, and alternative severity ratings across a wide range of ED-related variables. Specifically, we investigated psychological and biological variables (Chapter 5, Study 3a and Section 5.3), psychiatric comorbidity (Chapter 5, Study 3b), risk factors (Chapter 6, Study 4), and treatment outcomes (Chapter 7, Study 5). These chapters aimed to explore the relationships between ED severity ratings and psychological factors (including psychiatric comorbidity), biological correlates, psychological, risk factors, and treatment outcomes. In the following paragraphs, I will discuss in detail the utility of the DSM-5, ICD-11, OWS, and DT severity indices in indexing psychological and biological correlates, risk factors, and treatment outcomes in EDs.

Psychological variables.

In Chapter 5 (Studies 3a and 3b) and Section 5.3 (additional analysis), the DSM-5, ICD-11, and OWS AN severity ratings were evaluated for their ability to capture psychological variables commonly associated with AN. Using a sample of 312 treatment-seeking participants with AN (mean age = 26.9), Studies 3a and 3b assessed the DSM-5 AN, ICD-11 AN, and OWS severity ratings across various dimensions, including ED psychopathology, general psychopathology, psychiatric comorbidity, and biological correlates (to be discussed in a later section). Specifically, Study 3a and Section 5.3 (additional analysis) explored the relationship between these severity ratings and ED psychopathology (measured via the EDE-Q), such as eating concerns and dietary restraint, as well as general psychopathology, including variables such as level of disability (measured with the Brief Disability Questionnaire; Naismith et al., 2007) and quality of life (measured with the Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form; Schechter et al., 2007). Study 3b examined the relationship between the DSM-5, ICD-11, and OWS severity ratings for AN and the rate of psychiatric comorbidity (assessed via the Mini-International Neuropsychiatric Interview [MINI], version 7.0.2; Sheehan et al., 1998).

In the next subsection, I will discuss findings regarding the relationship between the examined psychological variables and the BMI-based DSM-5 and ICD-11 AN severity ratings, as well as the OWS alternative severity ratings for AN, as assessed in Studies 3a and 3b, alongside the additional analyses outlined in Section 5.3.

BMI-based DSM-5 and ICD-11 AN severity ratings. Results from Study 3a revealed that within the DSM-5 BMI-based severity ratings for AN, individuals in the mild and moderate severity groups reported significantly higher dietary restraint and weight/shape concerns compared to those in the extreme severe severity group. Additionally, participants in the mild severity group exhibited greater eating concerns and anxiety symptoms than those in the extreme severe severity group. These findings align with prior research (Nakai et al.,

2017; Smith et al., 2017). However, Study 3a found no significant differences between the severity groups in depressive symptoms, stress, disability, quality of life, or cognitive flexibility.

A similar pattern was observed when examining psychiatric comorbidity across DSM-5 AN severity group in Study 3b. We found that the mild DSM-5 AN severity group had a higher total number of psychiatric comorbidities, particularly panic, social anxiety, generalised anxiety, and post-traumatic stress disorder, compared to the severe and extreme severe AN groups. A meaningful severity rating for AN is expected to capture its impact on psychological well-being and functioning (Zimmerman et al., 2018). However, our findings from both Studies 3a and 3b indicated that the DSM-5 BMI-based severity ratings did not follow this expected pattern. Paradoxically, individuals in the milder DSM-5 AN severity group (with higher BMI) reported greater ED psychopathology (Study 3a) and psychiatric comorbidity (Study 3b) than those in the more severe groups

Regarding the ICD-11 AN severity groups, findings from Section 5.3 revealed similar patterns to those observed in the DSM-5 in Study 3a. Participants in the "significantly low BMI" AN category (higher BMI) reported greater general (depressive, anxiety and stress symptoms) and ED psychopathology (i.e., eating, shape, and weight concerns) compared to those in the "dangerously low BMI" (lower BMI) AN group. Similarly, as observed in Study 3b, individuals in the 'significantly low BMI' ICD-11 AN category (higher BMI) were found to be more likely to have psychiatric comorbidities, such as major depressive disorder and obsessive-compulsive disorder. These findings mirrored the paradox observed in DSM-5, as described above, where higher BMI AN category exhibit higher level of ED and general psychopathology as well as psychiatric comorbidities than lower BMI AN severity categories.

Overall, results from Studies 3a and 3b revealed a consistent pattern across both the DSM-5 and ICD-11 severity classification systems, despite their differences in BMI cut-offs. Individuals in milder AN severity groups (higher BMI) exhibited higher levels of ED psychopathology and psychiatric comorbidity than those in the more severe AN groups (with lower BMI). Several factors may contribute to this observation.

First, individuals in extreme AN severity groups (e.g., BMI <15 kg/m² in DSM-5 or BMI <14 kg/m² in ICD-11) may experience significant cognitive and emotional impairments due to the severe physiological effects of starvation (Calugi et al., 2018). These effects, including reduced emotional processing capacity, may mask psychological distress and ED symptoms such as dietary restraint (Brockmeyer et al., 2012; Calugi et al., 2018; Hebebrand et al., 2022). This masking effect could be further exacerbated by reliance on self-reported tools like the EDE-Q, which, although validated, depend on the individual's subjective ability to reflect and report accurately (Berg et al., 2012).

Second, research has shown that individuals with severe AN may have less insight into the severity of their disorder, making them more likely to underreport symptoms (Webb et al., 2022). In contrast, those in milder severity groups might have greater awareness of their condition, amplifying distress about eating behaviours and body image. This heightened insight may lead to feelings of "not being thin enough," intensifying dietary restraint and anxiety. Additionally, research indicates that milder severity is often associated with a shorter illness duration, which can increase societal or familial pressure to seek early intervention, further exacerbating anxiety and ED-related concerns (Hebebrand et al., 2024).

Lastly, the broad BMI thresholds used to define mild AN severity groups (BMI ≥15 kg/m² in DSM-5 and BMI ≥14 kg/m² in ICD-11) may introduce heterogeneity by encompassing a wide range of clinical presentations. In contrast, the narrower BMI thresholds for the moderate and severe categories in DSM-5 (e.g., a one-point BMI change

can shift a classification from moderate to severe, with moderate defined as 16–16.99 kg/m² and severe as 15–15.99 kg/m²) likely contribute to the observed paradoxical patterns.

Overall, while acknowledging the limitations of Studies 3a and 3b (to be discussed in detail in Section 9.4), including the use of self-reported ED psychopathology measures, our findings suggest that BMI-based severity indicators in both DSM-5 and ICD-11 have limited utility in indexing the ED psychopathology and psychiatric comorbidity experienced by individuals with AN. These findings highlight the need for more comprehensive approaches to assess and classify severity in AN, moving beyond BMI as the sole severity determinant.

OWS alternative severity rating for AN. OWS, an alternative severity indicator for AN, was also evaluated in its ability in indexing psychological variables in Chapter 5 (Studies 3a and 3b). Using the same sample of 312 participants described earlier, in Study 3a, we found that individuals in the high OWS group (indicating greater ED severity) exhibited significantly higher levels of ED psychopathology (e.g., dietary restraint and concerns about eating, weight, and shape) and general psychopathology, including symptoms of depression, anxiety, and stress as measured by the DASS-21, compared to those in the low OWS group (indicating less ED severity).

Study 3b further explored the relationship between OWS and psychiatric comorbidities. A similar pattern emerged, with participants in the high OWS group showing a higher prevalence of major depressive disorder compared to those in the low OWS group. However, among the nine psychiatric disorders screened in Study 3b, significant associations between OWS severity groups and psychiatric comorbidities were observed only for major depressive disorder.

These findings underscore the central role of OWS in the development and maintenance of AN, as theorised by the transdiagnostic model of EDs proposed by Fairburn et al. (2003). OWS influences both psychological and behavioural aspects of the disorder.

Previous studies have similarly established that higher OWS is strongly linked to greater ED-specific psychopathology and broader psychological difficulties, positioning it as a key factor in the onset and persistence of ED symptoms and comorbid conditions (e.g., Forrest et al., 2018). For instance, extreme weight-control behaviours (e.g., strict dieting, excessive exercise, self-induced vomiting, and misuse of laxatives or diuretics) and related psychological concerns (e.g., heightened eating concerns) are directly tied to the belief that shape, and weight are central to self-worth (Fairburn et al., 2003).

Furthermore, the elevated levels of anxiety, depression, and stress commonly associated with high OWS, as observed in Study 3a, suggest that OWS exacerbates vulnerability to general psychological distress. This highlights its value as a marker of severity for both ED-specific and broader psychological correlates.

Overall, findings from Chapter 5 (Studies 3a and 3b) and Section 5.3 revealed that while the DSM-5 and ICD-11 provide BMI-based severity classifications for AN, their reliance on physical metrics limits their ability to accurately index psychological severity. Both systems paradoxically associated milder severity categories with greater ED psychopathology (Study 3a) and higher rates of psychiatric comorbidities (Study 3b), highlighting their inadequacy in reflecting the psychological burden of AN. In contrast, OWS effectively captured the assessed psychological features of AN and demonstrated significant associations with both ED-specific and general psychopathology.

Biological factors

Chapter 5 (Study 3a) and Section 5.3 (additional analyses) investigated the relationship between DSM-5, ICD-11, and OWS severity classifications and key biological markers, including potassium, magnesium, calcium, phosphate, and white and red blood cell counts. These biological correlates, derived from blood samples collected at intake, were

analysed in a sample of 312 participants with AN (same sample as described above) to assess their associations with BMI-based severity groups.

The focus on these biological markers stems from their strong association with malnutrition and purging behaviours commonly observed in individuals with AN (Beaneas et al., 2024). Electrolyte abnormalities in AN are well-documented, with frequent findings including hypokalaemia (low potassium), hypomagnesemia (low magnesium), hypocalcaemia (low calcium), hypophosphatemia (low phosphate), and reduced white and red blood cell counts (Cass, 2020; Gibson, 2019). Consequently, a robust severity indicator for AN must integrate both the psychological dimensions of the disorder and the biological correlates tied to its prevalent medical complications.

BMI Based DSM-5 and ICD-11 AN Severity Ratings. Regarding BMI-based DSM-5 and ICD-11 AN severity indicators, analyses shown in Chapter 5 (Study 3a) and Section 5.3 (additional analyses) revealed no clinically significant differences in potassium, magnesium, phosphate, white blood cell counts, haemoglobin levels, or pulse rate across DSM-5 and ICD-11 severity groups. This highlights the limitations of BMI-based severity ratings in indexing biological risk in AN.

Despite the high prevalence of electrolyte and hematologic abnormalities commonly observed in individuals with AN (e.g., Beaneas et al., 2024; Cass et al., 2020), largely due to malnutrition and purging behaviours, these findings are in line with prior research indicating weak correlations between BMI and biological variables such as white blood cell count (Misra et al., 2004; Polli et al., 2008), haemoglobin (Mamou et al., 2016), potassium (Imbierowicz et al., 2004), and magnesium (Raj et al., 2012). The current findings support studies (Dalle Grave et al., 2024; Kambanis et al., 2024), suggesting that electrolyte and hematologic abnormalities in AN are more closely linked to the frequency of disordered

eating behaviours, such as restraint and purging, rather than to low BMI itself, which is a downstream consequence of these behaviours.

Altogether, the findings from Chapter 5 (Study 3a) and Section 5.3 (additional analyses) suggest that individuals at higher risk for electrolyte and hematologic complications may not necessarily have the lowest BMI, and BMI alone may not reliably reflect the biological status of individuals with AN.

This limitation becomes even more apparent when considering another significant criticism of the DSM-5 and ICD-11 BMI-based severity criteria: their failure to account for the rate of weight loss. Rapid or substantial weight loss – irrespective of an individual's starting BMI – is independently associated with severe physical complications requiring hospitalization (Whitelaw et al., 2018). For example, individuals with higher premorbid weight may experience substantial weight loss through disordered eating behaviours, such as restraint and purging, compared to those with lower premorbid weight. This places them at greater medical risk, even when they are categorised as "mild" under DSM-5 and ICD-11 BMI-based severity criteria (Skowron et al., 2020; Lebow et al., 2015)

Overall, findings from Chapter 5 (Study 3a) and Section 5.3 (additional analyses) highlight the limitations of BMI-based severity classifications for AN per DSM-5 and ICD-11. While BMI is a readily measurable physical indicator, in AN it may fail to capture the complex interplay between weight status and the medical risks associated with the disorder.

OWS Alternative Severity Rating for AN. Regarding the alternative OWS severity rating, in Study 3a, no significant differences were found between the two OWS-based AN severity groups in any of the assessed biological markers. It is not surprising that OWS did not correlate with the biological markers commonly associated with AN. As suggested in previous research (e.g., Buh et al., 2024; Hundemer et al., 2022), electrolyte imbalances are primarily driven by malnutrition, purging behaviours (such as vomiting or laxative use), and

refeeding syndrome, rather than by psychological factors like OWS. While OWS captures the cognitive and emotional dimensions of AN, its relationship with electrolyte imbalances is likely indirect and could be mediated by behaviours such as restrictive eating and purging rather than any direct physiological effect (Dalle Grave et al., 2024; Kambanis et al., 2024). Similarly, as found in Study 3a, BMI, a physical indicator, does not correlate with biological markers like electrolyte imbalances in this context for those with AN. Therefore, it is unsurprising that OWS, a psychological measure, also lacks an association with these biological correlates. Overall, the findings from Study 3a suggest that while OWS appears to be slightly more effective in indexing psychological variables (e.g., ED and general psychopathology) in AN, it may not be helpful in indexing biological correlates that are highly prevalent in those with AN.

Risk Factors

Chapter 6 (Study 4) investigated whether various risk factors were associated with differing levels of AN severity, as defined by DSM-5 BMI and DT severity ratings, in a sample of 153 sister pairs discordant for AN. Risk factors were assessed using the ORFI. An effective severity rating should not only assess ED psychopathology and prognosis but also incorporate key factors contributing to the onset and maintenance of AN (APA, 2013; Zimmerman et al., 2018). This broader perspective is crucial for identifying high-risk individuals and guiding decisions regarding hospitalisation and treatment intensity (Zimmerman et al., 2018).

Study 4 identified childhood obesity, childhood perfectionism, teasing about weight and shape, and embarrassment about breast development as significant retrospective correlates of an AN diagnosis across both DSM-5 and DT severity indices. The links between these risk factors and the development of an AN diagnosis align with existing literature (e.g., Stice, 2017; Yamamiya, 2024) and underscore their established role in AN development.

The primary aim of Study 4, however, was to identify risk factors specifically associated with varying severity levels of AN under the DSM-5 BMI and DT AN severity classification systems. Among all the factors examined across the DSM-5 and DT severity ratings for AN, only childhood perfectionism significantly increased the likelihood of classification into the DSM-5 extreme severity group (BMI < 15 kg/m²). This finding suggested that childhood perfectionism not only predisposes individuals to AN but also may increases the risk of developing its most severe form under the DSM-5 criteria.

Perfectionism, characterised by a relentless pursuit of flawlessness and self-criticism, is a well-documented maintaining mechanism in Fairburn et al.'s (2003) transdiagnostic model of EDs and has been validated in prior research (e.g., Dahlenburg et al., 2019; Stackpole et al., 2023). Studies have found that perfectionism amplifies individuals' overvaluation of weight and shape, leading to cycles of dietary restriction, feelings of failure, and compensatory behaviours that exacerbate the psychological and physical severity of AN (e.g., Howard et al., 2023; Longo et al., 2024).

Additionally, perfectionism has been found to interact with other maintaining mechanisms, such as low self-esteem and mood intolerance, intensifying distress and reinforcing rigid behavioural patterns (Longo et al., 2024; Dahlenburg et al., 2019). While this study focused on childhood perfectionism, previous research has shown it to be a stable trait strongly correlated with adulthood perfectionism (Sand et al., 2021). Consequently, addressing childhood perfectionism in early intervention and prevention programs could be pivotal in reducing the risk of extreme AN severity according to DSM-5 criteria.

Overall, in Study 4, childhood perfectionism was the only risk factor associated with the DSM-5 extreme AN severity level. No other established risk factors differentiated severity levels under the DT alternative severity classification system for AN. Although previous research (e.g., Krug et al., 2021; Mora-Maltas et al., 2023) suggests that DT may be a superior

severity rating for EDs due to its capacity to index ED and general psychopathology more effectively than DSM-5, its ability to distinguish risk factors for differing AN severity levels was not supported in this study. However, as only one ORFI risk factor (e.g., childhood perfectionism) was linked to DSM-5 extreme severity, conclusions about the relationship between DSM-5 severity ratings and established AN risk factors remain limited.

Treatment response

One of the most critical aspects of a severity rating is its ability to guide clinical decision-making and predict prognosis (Lucas et al., 2022). Using a sample of 628 treatment-seeking participants with AN (n = 266) and BN (n = 362) recruited from an ED unit in Spain, Study 5 (Chapter 7) directly compared DSM-5, ICD-11, and DT severity classifications in their ability to predict treatment outcomes associated with CBT-E for AN and BN. In the following section, I will discuss each of the severity ratings in its ability to predict treatment prognosis for those with AN and BN.

Anorexia Nervosa Severity Ratings and Treatment Outcomes

Three distinct severity ratings– two based on BMI (DSM-5 and ICD-11) and one based on cognitive factor DT– were examined in Study 5 to predict the prognosis of individuals with AN undergoing CBT-E treatment.

BMI-Based AN Severity Ratings and Treatment Outcomes. Results from Study 5 (Chapter 7) revealed differing patterns in how the DSM-5 and ICD-11 AN severity ratings, which are grounded in BMI, predicted treatment outcomes for AN. Specifically, no significant associations were found between the DSM-5 AN severity groups and CBT-E treatment outcomes for individuals with AN. These findings align with previous research highlighting the limited clinical utility of the DSM-5 severity rating for predicting AN treatment outcomes (Dalle Grave et al., 2018; Smink et al., 2014).

In contrast, the ICD-11 AN severity rating demonstrated a significant association with treatment outcomes. Individuals in the "significantly low BMI" AN ICD-11 severity group were less likely to drop out from treatment and more likely to achieve full remission compared to those in the "dangerously low BMI" ICD-11 AN severity category. These findings are consistent with prior studies demonstrating that a BMI of <14.0 kg/m² is significantly associated with poorer prognosis (Button et al., 2010; Engelhardt et al., 2021; Rosling et al., 2011).

One reason for the differing patterns observed between the DSM-5 and ICD-11 severity ratings may lie in the structural differences between the two classification systems. The DSM-5 utilises four severity levels for AN (mild, moderate, severe, extreme), each based on relatively narrow BMI thresholds (e.g., moderate: BMI 16.0–16.99 kg/m²; severe: BMI 15.0–15.99 kg/m²). This approach may result in subtle distinctions between groups that lack clinical significance for predicting treatment outcomes. Moreover, the extreme DSM-5 AN severity category, which encompasses individuals with BMIs <15.0 kg/m², includes a broad range of individuals with varying degrees of severity. This heterogeneity within the group may reduce its utility in distinguishing clinically meaningful differences in treatment outcomes.

In contrast, the ICD-11 adopts a simpler, two-tiered severity system for AN, focusing on more critical BMI thresholds: "significantly low BMI" (BMI ≥14.0 kg/m²) and "dangerously low BMI" (BMI <14.0 kg/m²). By emphasising these critical BMI thresholds, as suggested in the literature (e.g., Errichiello et al., 2016) the ICD-11 may better capture the severity of medical risk, particularly for individuals experiencing severe malnutrition (e.g., BMI <14.0 kg/m²) and complications like refeeding syndrome, thus providing a more robust predictor of AN treatment outcomes. This more simplified approach may explain why in Study 5 the ICD-11 AN severity rating was able to index CBT-E treatment outcomes for AN.

Specifically, individuals in the "significantly low BMI" ICD-11 AN severity group were less likely to drop out from treatment or achieve full remission compared to those in the "dangerously low BMI" ICD-11 AN severity group, findings that are consistent with prior studies linking BMI <14.0 kg/m² to poorer prognosis (e.g., Errichiello et al., 2016). Despite being the first study to examine the relationship between ICD-11 AN severity and treatment outcomes, findings from Study 5 (Chapter 7) highlighted the potential utility of the ICD-11 severity system in assessing outcomes in individuals with AN.

Drive for Thinness AN Severity Rating and Treatment Outcomes. Study 5 (Chapter 7) also evaluated DT as an alternative cognitive-based severity rating for AN, specifically assessing its ability to predict treatment outcomes. Results from Study 5 revealed no significant differences across the two DT severity groups in predicting treatment outcomes for AN. These findings contrast with previous literature suggesting that higher DT is associated with poorer treatment outcomes in AN (e.g., Prost-Lehmann et al., 2018).

The absence of significant differences in treatment outcomes across the two DT severity groups in Study 5 suggests that cognitive factors, such as DT, may not independently predict how individuals with AN respond to CBT-E treatment. While DT is a crucial psychological characteristic in the development and maintenance of AN (Ramacciotti et al., 2002; Vansteelandt et al., 2007), its relationship with treatment outcomes may be more complex. This relationship is likely mediated by other factors, including the severity of malnutrition, comorbid psychiatric conditions, and treatment engagement (Jérolon et al., 2022), all of which warrant further investigation in relation to AN severity indices and treatment outcomes. Furthermore, DT may not adequately capture the physiological consequences of AN, which have been shown to influence treatment success (Puckett et al., 2021).

In summary, the findings from Study 5 suggest that while BMI-based severity ratings in both the DSM-5 and ICD-11 have limited effectiveness in predicting AN treatment outcomes, the ICD-11 provides a slightly better prediction for CBT-E treatment outcomes. Additionally, while cognitive factors such as DT play a significant role in the development and maintenance of AN (Stice et al., 1998; 2001), they may not independently predict treatment outcomes. These findings highlight the importance of adopting a more comprehensive approach that integrates both cognitive and physiological factors when assessing prognosis in AN (Puckett et al., 2021).

Bulimia Nervosa Treatment Outcomes and DSM-5 and DT Severity Ratings

In the same study (Study 5), we examined both the DSM-5 BN severity ratings, based on the weekly frequency of compensatory behaviours, and the alternative DT severity groups in predicting treatment outcomes for individuals with BN undergoing group-based CBT-E. Despite findings from our systematic review and meta-analysis (Chapter 2, Study 1), which found the DSM-5 BN severity rating as predictive of ED psychopathology, Study 5 revealed that neither the DSM-5 BN severity rating nor DT severity groups were associated with treatment outcomes. These findings contrast with previous studies, which reported associations between the DSM-5 BN severity rating and abstinence from compensatory behaviours in adults with BN, as well as improvements in ED-related attitudes in adolescents (Dakanalis et al., 2017; Gorell et al., 2019). Notably, no prior studies have evaluated DT severity groups in predicting BN treatment outcomes.

As discussed in Study 5, one potential explanation for the discrepancy between our findings and previous studies (e.g., Dakanalis et al., 2017; Gorell et al., 2019) lies in the stricter criteria for defining good treatment outcomes used in our study. These criteria incorporated both attitudinal (e.g., attitudes towards weight and shape) and behavioural (e.g., compensatory behaviours) components. In contrast, prior studies focused primarily on

behavioural outcomes or assessed attitudinal and behavioural aspects separately, which may account for the divergence in findings.

Overall, Study 5 demonstrated that the DSM-5 BN severity rating, based on the frequency of compensatory behaviours such as vomiting, laxative use, and excessive exercise, along with the DT severity rating, fails to fully capture the clinical complexity and psychological factors influencing BN treatment outcomes. While the DSM-5 BN severity rating is grounded in the frequency of compensatory behaviours, and DT focuses on the cognitive and emotional dimension of BN– the desire for thinness– both measures neglect critical elements such as comorbid psychological conditions (e.g., depression, anxiety), emotional dysregulation, social support, treatment history, and motivation for change (Gorrell et al., 2022). These factors are central to successful BN treatment, particularly in CBT-E. This suggests that a single severity rating based solely on compensatory behaviours (DSM-5) or a cognitive factor like DT may not adequately predict treatment response.

Objective 4: To evaluate DSM-5 and alternative severity ratings for EDs on a continuous scale within a community-based setting, aiming to enhance understanding of these ratings across the broader spectrum of disordered eating and EDs.

To address objective 4, Study 6 recruited 686 community-based adults (mean age = 19.95) and used EMA to evaluate BMI and weight suppression as continuous variables in their ability to capture daily changes in body dissatisfaction level and disordered eating urges (e.g., exercise, binge eating, unhealthy eating, and dietary restraint).

Study 6 sought to address two key limitations in previous ED severity research. First, most prior studies (e.g., Dalle Grave et al., 2018; Krug et al., 2021) focused on clinical samples, which primarily included individuals already in treatment. This focus limited the generalisability of findings to broader populations. By using a community-based sample, Study 6 captured a wider spectrum of ED severity, including individuals not engaged in care,

thereby enhancing the generalisability of its results. While some earlier studies (e.g., Grilo, 2015; Smink, 2014) also utilised community samples, these were typically limited to DSM-5 BED severity ratings or prevalence estimates. This left gaps in understanding DSM-5, ICD-11, and alternative severity ratings for AN and BN within community settings.

Second, as highlighted in the systematic literature synthesis (Chapter 2, Study 1) and feedback from Australian ED healthcare professionals (Chapter 4, Study 2), weight suppression has been identified as a promising alternative severity indicator for restrictive EDs such as AN. However, weight suppression has remained underexplored as a severity measure. Study 6 addressed this gap by examining weight suppression alongside BMI, coding both as continuous variables to overcome the challenges of negatively skewed distributions, where severe cases are often underrepresented in community samples.

Findings from Study 6, obtained through multilevel modelling, indicated that lower BMI and weight suppression were associated with higher daily averages of body dissatisfaction. Additionally, higher BMI and greater weight suppression were more likely to be associated with stronger urges to engage in unhealthy eating. However, BMI and weight suppression did not predict daily changes in other disordered eating urges.

The findings from Study 6, regarding a lower BMI and lower weight suppression predicting greater body dissatisfaction, contradict existing literature (e.g., Burnette et al., 2018; Jones et al., 2019). Specifically, previous studies have found the opposite pattern, showing that individuals with a higher BMI (e.g., Jones et al., 2019; Kogure et al., 2019) and greater weight suppression (e.g., Berner et al., 2013; Burnette et al., 2018) were more likely to experience higher levels of body dissatisfaction. The discrepancies between Study 6 and the findings from prior studies may stem from the characteristics of the Study 6 sample, which consisted mainly of university students with a mean BMI (22.14 kg/m^2) in the normal

range, compared to other studies that included populations with higher mean BMIs in the overweight range (e.g., 29.1 kg/m²; Kogure et al., 2019).

Similarly, the mean weight suppression in this sample (4.96 kg) was half that reported in previous research (e.g., 8.19–12.85 kg; Chen et al., 2022; Burnette et al., 2018). These differences suggest that the findings from Study 6 may only be applicable to populations with similar characteristics. These results also highlight the need for future research to explore whether there is a specific threshold for BMI or weight suppression that might reverse the expected relationship, such as lower weight suppression leading to higher body dissatisfaction.

Another notable finding from Study 6 was that individuals with a higher BMI and higher weight suppression were more likely to report experiencing urges to engage in unhealthy eating. This aligns with previous research suggesting that weight suppression, often achieved through restrictive dieting or compensatory behaviours, can lead to an increased preoccupation with food, particularly unhealthy options (Butryn et al., 2009). This preoccupation may stem from both biological and psychological factors related to weight suppression behaviours (e.g., Rosenbaum & Leibel, 2010). Biologically, weight suppression can trigger hunger signals and metabolic adaptations, making it harder to resist cravings (Rosenbaum & Leibel, 2010). Psychologically, the cognitive strain of dieting can amplify the allure of restricted foods, leading to disordered eating patterns like binge eating (Lowe et al., 2006). Over time, these factors may not only challenge weight maintenance but also increase the likelihood of weight regain and a higher BMI.

One additional factor to note regarding the findings of Study 6 is that the study assessed only the urges to engage in such behaviours. The relationship between urges and actual behaviours can be influenced by factors such as preoccupation with food and dietary restraint (Yu et al., 2016). Future research should further investigate these complexities.

In addition, it is worth noting that BMI and weight suppression did not show associations with any of the other disordered eating urges tested, offering limited evidence for their ability to index disordered eating urges in natural environments. ED behaviours in community settings are often more sporadic and context-dependent, influenced by factors such as stress or mood (Mason et al., 2021; 2022). Static measures such as BMI and weight suppression may not accurately capture these real-time fluctuations (e.g., urges to exercise), even when using EMA. However, the lack of significant findings between BMI and weight suppression for assessed disordered eating urges in Study 6 underscores the limitations of BMI and weight suppression as standalone indicators disordered eating urges in non-clinical populations

9.3. Implications

This thesis evaluated the utility of various severity rating systems for EDs– DSM-5, ICD-11, OWS, and DT– across six studies encompassing clinical (Studies 1 to 5) and community (Study 6) samples. The findings highlighted the limitations of single-dimension severity ratings in capturing the complexity of EDs, which involve an intricate interplay of biological, psychological, emotional, physical, and risk-related factors. Findings from this thesis are summarised in Table 28. As shown in Table 28, no single severity rating adequately captured all relevant aspects of these multifaceted disorders. In this section, I will discuss the clinical implications of these findings, particularly regarding the limitations of current severity systems and opportunities for improvement.

Table 28.

Summary of Findings on the relationship between Severity Ratings for AN, BN, and BED and Assessed Variables,

Eating Disorder	Severity Ratings	Assessed Variables			
		Psychological Factors	Biological Factors	Risk Factors	Treatment Outcomes
Anorexia Nervosa	DSM-5	✗	✗	?	✗
	ICD-11	✗	✗	?	✓
	OWS	✓	✗	?	?
	DT	✓	?	✗	✗
Bulimia Nervosa	DSM-5	✓	?	?	✗
	DT	?	?	?	✗
Binge Eating Disorder	DSM-5	✗	?	?	?
	OWS	✗	?	?	?

Note. DT: Drive for Thinness
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, fifth edition
ICD-11: International Classification of Diseases, 11th edition
OWS: Overvaluation of weight and shape
✓ = Support
✗ = No support
? = Has not been assessed

Refrain from Solely Relying on Severity Ratings in Clinical Practice

First, none of the currently tested severity ratings should be used in isolation to assess severity or guide clinical decision-making. Findings from the current thesis showed that, single-indicator measures, such as the DSM-5 BMI-based AN severity ratings, ICD-11 AN severity ratings, and DT AN and BN severity ratings, failed to capture the multidimensional nature of ED severity. For instance, the DSM-5 BMI-based AN severity rating performed the poorest, as it failed to capture psychological and biological correlates (Studies 3a and 3b), risk factors (Study 4), or treatment outcomes (Study 5). Although the ICD-11 AN, DSM-5 BN, and OWS-based alternative AN severity ratings demonstrated strengths in certain areas, they also exhibited notable gaps. For instance, DSM-5 BN severity ratings (Study 1) and OWS-based AN severity ratings (Studies 3a and 3b) were able to reflect psychological correlates, such as ED and general psychopathology. However, only the ICD-11 AN severity rating demonstrated the ability to predict treatment outcomes for AN (Study 5), setting it apart from the DSM-5 AN, DSM-5 BN, and DT AN and BN alternative severity ratings.

Despite these findings, a limitation shared by all tested severity ratings (e.g., DSM-5 AN, ICD-11 AN, and OWS) was their inability to effectively account for biological correlates such as potassium or calcium levels – markers critical for evaluating medical complications and addressing risks associated with medical instability. This is a significant limitation, as biological factors are essential for informing clinical interventions, particularly in cases requiring urgent medical management.

Overall, none of the tested single-indicator severity ratings could simultaneously capture the full range of variables relevant to EDs – psychological, biological, and treatment outcomes. Based on these findings, this thesis does not recommend the use of current severity ratings as definitive tools in clinical practice. Instead, assessment and treatment decisions should continue to follow broader, evidence-based guidelines (e.g., RANZCP or Australia &

New Zealand Academy for Eating Disorders treatment principles), which incorporate multiple dimensions of severity for a more holistic approach to ED care.

Barriers to Accessing Care Created by Severity Ratings

Beyond their limited utility for guiding clinical practice, current severity ratings (e.g., DSM-5, ICD-11) may inadvertently create barriers to care for individuals with EDs. Across six studies, this thesis found limited support for BMI-based severity ratings, particularly for AN, as they fail to capture the full complexity of illness severity. Notably, the findings from Studies 3a and 3b revealed that individuals with higher BMIs reported greater levels of ED-specific and general psychopathology as well as psychiatric comorbidities than those with lower BMIs, challenging the traditional assumption that lower BMI signifies more severe illness. This disconnects risks excluding individuals with higher BMIs, including those with atypical AN, from receiving high-intensity care, despite their significant psychological impairments (Dang et al., 2024; Harrop et al., 2023).

Concerns about the detrimental implications of BMI-centric severity ratings were also echoed by clinicians interviewed in Study 2. Interviewed healthcare professionals highlighted that the "mild" severity label, often assigned to individuals with higher BMIs, could be particularly damaging for those with AN. This is especially impactful as this group is already less likely to seek treatment (Hamilton et al., 2022; Kästner et al., 2021) and being labelled as "mild" may reinforce the belief that they are "not sick enough" to warrant care. Therefore, severity labelling not only invalidates their experiences but may also discourage engagement with potentially life-saving treatment.

Moreover, BMI-based severity ratings perpetuate the misconception that atypical AN is inherently less severe than AN – another point that was also raised by clinicians interviewed in Study 2. This perspective marginalises individuals with atypical AN, and overlooks evidence (e.g., Dang et al., 2024) demonstrating that they can experience

symptoms as severe, or even more severe, than those with AN, despite having a higher BMI. By framing severity solely around BMI, these measures fail to account for the broader psychological and behavioural complexities of EDs, further entrenching barriers to appropriate care.

Integrating Alternative Ratings and Multidimensional Screening Tools

The findings of this thesis highlight the significant limitations of current severity ratings for EDs, such as DSM-5, OWS, and DT, in effectively guiding prevention efforts, treatment planning, and progress monitoring. As mentioned previously, although OWS (Studies 3a and 3b) alternative severity rating provide valuable insights into psychological factors among those with AN they fail to adequately predict other critical clinical indicators, including biological correlates and treatment outcomes (Study 3a). These shortcomings emphasise the need to integrate alternative severity ratings into a more comprehensive framework for assessing the complexity of EDs.

Despite these limitations, severity ratings can still serve a useful purpose in clinical practice, particularly in resource-constrained environments. As clinicians interviewed in Study 2 noted, severity ratings can support triaging and referring individuals with EDs to tertiary services. Given the widespread shortage of ED-trained clinician, as well as limitations in time and funding in tertiary ED services, there is a clear need for tools that streamline assessment and intervention processes to maximise efficiency in such settings.

One promising solution is the development of a multidimensional severity screening tool guided by current assessment guidelines (e.g., The American Psychiatric Association Practice Guideline for the Treatment of Patients with Eating Disorders, 2023 or the RANZCP, 2014). This tool should integrate psychological, medical, and social dimensions of severity to provide a more holistic understanding of individual needs. Key indicators could include levels of OWS (given its correlations with ED-specific and general psychopathology

as found in the current thesis), social support, treatment history, illness duration, nutritional and fluid intake, and the rate of weight change (both loss and gain) (Crone et al., 2023). Additionally, where possible, initial triage should also initial triage could incorporate brief physical examinations, such as measurements of seated and standing pulse rates to identify bradycardia or tachycardia caused by cardiac deconditioning, blood pressure readings (seated and standing), and body temperature. These physical markers are essential for identifying individuals who may require urgent medical attention or hospitalisation (RANZCP; Hay et al., 2014).

In summary, implementing a multidimensional severity screening tool could significantly enhance the management of EDs, particularly in high-pressure, underfunded settings. Such potential tool would address the gaps left by current single-indicator severity ratings while improving the efficiency and precision of clinical assessment and intervention planning.

9.4. Strengths

This thesis has several methodological and conceptual strengths, informed by the gaps identified in the initial systematic review and meta-analysis (Chapter 2, Study 1). A key strength lies in the large, diverse participant groups comprising the studies. For example, Study 1 synthesised findings from research on ED severity involving a total of 9,196 participants. Studies 2, 3, 4, and 5 extended this by incorporating data from individuals with EDs, clinicians, and community members, capturing a broad range of experiences and presentations. Furthermore, the ED groups were recruited from various tertiary inpatient and outpatient ED services in Australia and internationally, increasing the generalisability of the findings.

Another notable strength is the application of rigorous, contemporary analytic methods. The meta-analytic approach in Study 1 synthesised available studies, offering a

comprehensive overview of the findings while addressing limitations related to small sample sizes. Study 4 (Chapter 6) employed multinomial regression to assess risk factors and severity ratings, calculating probabilities for each outcome and examining whether previously identified risk factors differentiated DSM-5 AN and DT severity groups. Additionally, Study 5 used EMA to analyse whether BMI and weight suppression could index daily variables related to body dissatisfaction and daily disordered eating urges (Chapter 8). Complementing the quantitative method (i.e., survey), the qualitative approach in Study 3 (Chapter 4) provided valuable insights into clinicians' experiences and perspectives regarding the proposed DSM-5 severity ratings for EDs. Through interviews or free-text response, this approach offered a deeper understanding of how the ratings are applied in clinical practice, their perceived usefulness for treatment planning, and potential barriers to implementation–areas that quantitative methods may not fully capture.

9.5. Limitations and Future Directions

9.5.1 Methodology Limitations and Future Directions

The findings of this thesis should be considered in light of several methodological limitations, including aspects of design, measurement tools, and sample composition. A significant limitation is the reliance on cross-sectional designs in all studies except Study 6. While this approach facilitates research feasibility in the field of EDs, it restricts the ability to examine how severity ratings relate to symptom changes, treatment progress, or long-term outcomes. For example, in study 5, the cross-sectional design precluded an evaluation of whether individuals classified as "moderate" based on a severity rating exhibit greater symptom improvement over the course of treatment compared to those classified as "severe." Additionally, this design does not capture the dynamic nature of clinical decision-making, where severity ratings might guide adjustments to treatment or progress monitoring. For instance, a patient with AN initially classified as "extreme severe" based on DSM-5 BMI

severity rating may require hospitalisation. However, as their condition improves through weight restoration or reduced psychological distress, they might be reclassified as "moderate," prompting a shift to outpatient care or less frequent monitoring. Additionally, this cross-sectional design only allows for identifying correlations, not causal relationships, making it difficult to determine whether severity ratings can reliably predict treatment outcomes, such as symptom improvement or relapse or medical complications. Therefore, future research should be conducted longitudinally, measure the changes of each aspect of ED symptomatology (e.g., psychological and biological factors), and clinical decision-making processes as mentioned above.

Second, the generalisability of the findings in this thesis is limited by the predominance of adult White female participants making the results mostly applicable to individuals with similar demographic characteristics. This limitation is particularly notable given the large focus on BMI of the current thesis, as research demonstrates that BMI cut-offs for health outcomes vary across ethnic groups due to differences in body composition (Paeratakul et al., 2002). For example, Asians typically have a BMI 2–3 kg/m² lower than Whites for the same age, sex, and body fat percentage, owing to differences in muscle mass (Wulan et al., 2010). Consequently, it may be possible that Asians may experience less ED-related psychopathology at lower BMI levels than Whites, underscoring the need for caution when applying these findings to other ethnic groups. Similarly, differences between adults and adolescents in the development and manifestation of EDs may affect the applicability of these findings across age groups. Moreover, the exclusion of male participants limits the generalisation of findings to men. This is particularly important as BMI can be misleading in male populations, given their typically higher muscle mass compared to women (Nieves et al., 2005). This can result in a higher BMI even without excess body fat, leading to an underestimation of ED severity in men. Males may appear to have a "normal" BMI despite

experiencing significant psychological distress related to body image, weight, and shape (Zayas et al., 2018). The lack of analysis of the relationship between different severity ratings and ED psychopathology in a male sample further restricts the understanding of how BMI reflects true ED severity in men, where muscle mass and body composition differences can distort results. Future research, therefore, would benefit from investigating how the severity ratings for EDs perform on individuals from a range of different age, gender and ethnic backgrounds.

Third, while the overarching aim of this thesis was to examine ED severity ratings broadly, much of the research focused specifically on AN severity ratings. This emphasis was primarily due to the availability of data, that is the high prevalence of AN cases in the treatment settings from which the sample was drawn. As shown in Table 28, limited information exists regarding the DSM-5 and alternative severity ratings (e.g., DT, OWS, and weight suppression) in indexing biological or risk factors in BN and BED. Furthermore, the predictive validity of the DSM-5 BED and OWS severity ratings in indexing treatment outcomes for BED remains unexamined. Future research is needed to address these gaps by expanding the examination of a wide range of severity ratings across different ED subtypes, including BED on a wide range of outcomes variables (e.g., treatment outcomes).

Fourth, apart from Studies 2 and 3b, the reliance on self-report measures to assess key variables represents another limitation. Although validated measures were employed, self-report may not always be the most appropriate method for assessing certain constructs. All studies in this thesis utilised self-report questionnaires for ED symptoms (e.g., EDE-Q and EDI-2), which require participants to recall their ED attitudes and behaviours over the preceding 28 days. Previous research has shown that while participants reliably recall attitudinal features of their ED (e.g., “Has your weight influenced how you think about (judge) yourself as a person?”), recall of behavioural features (e.g., “Have you gone for long

periods of time (8 waking hours or more) without eating anything at all to influence your shape or weight?") is less reliable (Rose et al., 2013). Additionally, the focus on AN in the current thesis further amplifies the limitations associated with the use of self-reported measures, particularly regarding reliability. Research has shown that lower BMI is linked to poorer insight into the disorder among individuals with AN (Webb et al., 2022). This diminished insight can reduce the accuracy of self-reports, especially for behavioural and psychological symptoms. This factor may also help explain the unexpected findings in Chapter 5 (Studies 3a and 3b), where individuals with lower BMI reported lower ED-related and general psychopathology than those with higher BMI. Therefore, future research should focus on using a combination of both interview or questionnaire and biological assessment to increase the reliability of measured constructs.

Fifth, the use of a community-based sample in Chapter 8 (Study 6) offers both strengths and limitations. This approach enabled the examination of BMI and weight suppression, measured on a continuous scale, in relation to daily fluctuations across a wider range of disordered eating urges. However, its generalisability to clinical populations remains limited. Community participants may often display milder symptoms and lower psychological distress compared to treatment-seeking individuals, which may not fully capture the severity and complexity observed in clinical settings. As a result, the findings from Study 6 provide limited insight into the utility of BMI and weight suppression in predicting disordered eating behaviours in a clinical sample with EDs. Future research should prioritise clinical samples to better assess how BMI and weight suppression on a continuous scale align with ED and general psychopathology.

9.5.2 Theoretical Limitations and Future Directions

Beyond the methodological improvements outlined earlier, several conceptual considerations should guide future research on ED severity ratings. These considerations

include expanding the scope of research to address underexplored areas and adopting more comprehensive approaches to ED severity assessment.

Single-indicator severity

This thesis focused on evaluating single-indicator severity ratings for EDs. However, several promising alternative severity indicators (e.g., number of purging method), as reviewed in the Introduction, remain underexplored. For instance, while weight suppression was examined in Study 6, its evaluation was limited to a non-clinical sample. Future research should prioritise investigating weight suppression in clinical ED populations, particularly among those diagnosed with AN and BN, where it has been shown to correlate strongly with both ED-specific and general psychopathology (Khalil et al., 2024). Another promising alternative single-indicator severity rating is the number of purging methods, which has been proposed as an alternative severity rating for BN (e.g., Gianini et al., 2017).

Existing research (Dieffenbacher et al., 2024; Gianini et al., 2017) suggests that the number of purging methods (e.g., vomiting, diuretic use), rather than the DSM-5 BN severity rating based on purging frequency, provided a more robust measure of psychopathology. Specifically, both studies found that individuals who engage in multiple purging methods exhibit higher levels of both eating disorder-specific and general psychopathology compared to those who use only a single method. Despite its promise, an important gap remains regarding whether the number of purging methods can reliably index treatment outcomes for BN. Addressing this issue is particularly significant given the findings from Study 5, which demonstrated that the DSM-5 BN severity groups failed to predict treatment outcomes in CBT-E for BN.

The concept of transdiagnostic severity indicators – measures that can index symptom severity across and within multiple ED diagnostic categories – has gained increasing attention (e.g., Gianini et al., 2017). However, this thesis did not evaluate

alternative severity measures, such as weight suppression, OWS, or DT from a transdiagnostic perspective. Specifically, Studies 3a and 3b assessed OWS as an alternative severity indicator exclusively for AN, focusing on its ability to differentiate biological correlates, ED-specific and general psychopathology, and psychiatric comorbidities. Study 5 evaluated DT as a severity rating, but only in relation to CBT-E treatment outcomes for individuals with AN and BN. As such, future research should explore how potential transdiagnostic severity indicators (e.g., weight suppression) perform across a broader range of ED subtypes. These investigations should extend beyond the "core" ED diagnoses of AN, BN, and BED to include other diagnostic categories without severity ratings, such as OSFED and ARFID. Expanding the scope in this way could provide valuable insights into the potential usage of these ratings in ED assessment and treatment decision-making.

Multidimensional Severity Ratings

As highlighted throughout this thesis, single-indicator severity measures fail to capture the multifaceted nature of ED presentations. This limitation underscores the need for multidimensional severity ratings that consider a combination of symptoms and clinical indicators to better predict critical outcomes in EDs, such as recovery and relapse rates. Factors like BMI, core illness features, psychosocial influences, functional impairment (e.g., number of days off from work or school) and levels of current support, collected through structured interviews or questionnaires, could be integrated into a comprehensive framework to better assess ED severity and prognosis. The development of such a model should incorporate input from individuals with lived experience, clinicians, and research experts. This collaborative approach would ensure the model relevance, practicality, and alignment with the complexities of ED presentations. By moving beyond single measures such as BMI (for AN) or inappropriate compensatory behaviours frequency (for BN) and bingeing

frequency (for BED), a multidimensional framework could significantly enhance clinical assessment and inform tailored treatment planning.

9.6. Conclusions

This thesis presents the first comprehensive evaluation of DSM-5 and ICD-11 severity ratings, along with alternative severity ratings based on OWS, DT, and weight suppression (tested only in a community setting) for EDs. Findings from the six studies in this PhD thesis underscore the inadequacies of current severity ratings in capturing the biological and psychological complexity of EDs. These limitations can lead to inaccurate assessments, hinder effective treatment planning, and restrict access to appropriate care. The findings emphasise the urgent need to move beyond single-indicator measures, such as BMI for AN or binge-purge frequency for BN and BED, toward a multivariate approach. This approach should integrate combinations of symptoms (e.g., purging methods) and clinical indicators (e.g., rate of weight loss, treatment history) to better predict critical outcomes such as recovery and relapse. By adopting this more nuanced framework, future severity ratings could enhance their clinical utility and ultimately improve care for individuals with eating disorders.

References

Abbate-Daga, G., Pierò, A., Gramaglia, C., Gandione, M., & Fassino, S. (2007). An attempt to understand the paradox of anorexia nervosa without drive for thinness. *Psychiatry Research*, 149(1-3), 215-221.

Agh, T., Kovács, G., Pawaskar, M., Supina, D., Inotai, A., & Vokó, Z. (2015). Epidemiology, health-related quality of life and economic burden of binge eating disorder: a systematic literature review. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 20, 1-12.

Agüera, Z., Romero, X., Arcelus, J., Sánchez, I., Riesco, N., Jiménez-Murcia, S., González-Gómez, J., Granero, R., Custal, N., & Montserrat-Gil de Bernabé, M. (2015). Changes in body composition in anorexia nervosa: Predictors of recovery and treatment outcome. *PloS one,* 10(11), e0143012.

Alfalahi, M., Mahadevan, S., Balushi, R. a., Chan, M. F., Saadon, M. A., Al-Adawi, S., & Qoronfleh, M. W. (2022). Prevalence of eating disorders and disordered eating in Western Asia: a systematic review and meta-Analysis. *Eating Disorders*, 30(5), 556-585.

Allen, K. L., Byrne, S. M., & McLean, N. J. (2012). The dual-pathway and cognitive-behavioural models of binge eating: prospective evaluation and comparison. *European Child & Adolescent Psychiatry*, 21, 51-62.

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

Amiri, S., & Khan, M. A. (2023). Prevalence of non-suicidal self-injury, suicidal ideation, suicide attempts, suicide mortality in eating disorders: a systematic review and meta-analysis. *Eating Disorders,* 31(5), 487-525.

Anbari, F., Mikaeili, N., & Hajloo, N. (2020). Modification of the dual pathway model of binge eating among women: evaluation of fear of negative evaluation and self-esteem as two new extensions. *Journal of Advanced Pharmacy Education and Research*, 10(1-2020), 89-95.

Anderson, L. M., Smith, K. M., Schaefer, L. M., Crosby, R. D., Cao, L., Engel, S. G., Crow, S. J., Wonderlich, S. A., & Peterson, C. B. (2020). Predictors and moderators of treatment outcome

in a randomized clinical trial for binge-eating disorder. *Journal Of Consulting and Clinical Psychology,* 88(7), 631.

Aouad, P., Bryant, E., & Martenstyn, J. (2022). Are DSM–5 eating disorder severity indicators clinically meaningful? A commentary on A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders. *Clinical Psychology: Science and Practice*, 29(4), 345–348

Arboix, A. (2015). Cardiovascular risk factors for acute stroke: Risk profiles in the different subtypes of ischemic stroke. *World Journal of Clinical Cases*: WJCC, 3(5), 418.

Atwood, M. E., & Friedman, A. (2020). A systematic review of enhanced cognitive behavioral therapy (CBT-E) for eating disorders. *International Journal of Eating Disorders*, 53(3), 311-330.

Baenas, I., Etxandi, M., & Fernandez-Aranda, F. (2024). Medical complications in anorexia and bulimia nervosa. *Medicina Clínica* (English Edition), 162(2), 67-72.

Bagaric, M., Touyz, S., Heriseanu, A., Conti, J., & Hay, P. (2020). Are bulimia nervosa and binge eating disorder increasing? Results of a population-based study of lifetime prevalence and lifetime prevalence by age in South Australia. *European Eating Disorders Review*, 28(3), 260-268.

Bakalar, J. L., Shank, L. M., Vannucci, A., Radin, R. M., & Tanofsky-Kraff, M. (2015). Recent advances in developmental and risk factor research on eating disorders. *Current Psychiatry Reports,* 17, 1-10.

Barakat, S., Burton, A. L., Cunich, M., Hay, P., Hazelton, J. L., Kim, M., Lymer, S., Madden, S., Maloney, D., & Miskovic-Wheatley, J. (2023). A randomised controlled trial of clinician supported vs self-help delivery of online cognitive behaviour therapy for Bulimia Nervosa. *Psychiatry Research*, 329, 115534.

Barakat, S., McLean, S. A., Bryant, E., Le, A., Marks, P., Touyz, S., & Maguire, S. (2023). Risk factors for eating disorders: findings from a rapid review. *Journal of Eating Disorders*, 11(1), 8.

Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders*, 45(3), 428-438.

Berner, L. A., Shaw, J. A., Witt, A. A., & Lowe, M. R. (2013). The relation of weight suppression and body mass index to symptomatology and treatment response in anorexia nervosa. *Journal of Abnormal Psychology*, 122(3), 694.

Bhattacharya, A., DeFilipp, L., & Timko, C. A. (2020). Feeding and eating disorders. *Handbook of clinical neurology,* 175, 387-403.

Blythin, S. P., Nicholson, H. L., Macintyre, V. G., Dickson, J. M., Fox, J. R., & Taylor, P. J. (2020). Experiences of shame and guilt in anorexia and bulimia nervosa: A systematic review. *Psychology and Psychotherapy: Theory, Research and Practice*, 93(1), 134-159.

Boltri, M., & Sapuppo, W. (2021). Anorexia nervosa and autism spectrum disorder: a systematic review. *Psychiatry Research,* 306, 114271.

Borg, S. L., Schaefer, L. M., Hazzard, V. M., Herting, N., Peterson, C. B., Crosby, R. D., Crow, S. J., Engel, S. G., & Wonderlich, S. A. (2022). Relationships between childhood abuse and eating pathology among individuals with binge-eating disorder: examining the moderating roles of self-discrepancy and self-directed style. *Eating Disorders*, 30(4), 355-369.

Braun, J., El-Gabalawy, R., Sommer, J. L., Pietrzak, R. H., Mitchell, K., & Mota, N. (2019). Trauma exposure, DSM-5 posttraumatic stress, and binge eating symptoms: results from a nationally representative sample. *The Journal of Clinical Psychiatry*, 80(6), 14848.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology,* 3(2), 77-101.

Brausch, A. M., & Gutierrez, P. M. (2009). The role of body image and disordered eating as risk factors for depression and suicidal ideation in adolescents. Suicide and life-threatening behavior, 39(1), 58-71.

Brewerton, T. D., Dansky, B. S., Kilpatrick, D. G., & O'Neil, P. M. (2000). Which comes first in the pathogenesis of bulimia nervosa: dieting or bingeing? *International Journal of Eating Disorders*, 28(3), 259-264.

Brewerton, T. D., Perlman, M. M., Gavidia, I., Suro, G., Genet, J., & Bunnell, D. W. (2020). The association of traumatic events and posttraumatic stress disorder with greater eating disorder and comorbid symptom severity in residential eating disorder treatment centers. *International* Journal of Eating Disorders, 53(12), 2061-2066.

Brockmeyer, T., Holtforth, M. G., Bents, H., Kämmerer, A., Herzog, W., & Friederich, H.-C. (2012). Starvation and emotion regulation in anorexia nervosa. *Comprehensive Psychiatry*, 53(5), 496-501.

Brown, T. A., & Keel, P. K. (2023). Eating disorders in boys and men. *Annual Review of Clinical Psychology*, 19(1), 177-205.

Bruce, K. R., & Steiger, H. (2013). Prognostic implications of personality disorders in eating disorders. In *Personality Disorders and Eating Disorders* (pp. 247-262). Routledge.

Buh, A., Scott, M., Kiska, R., Fung, S. G., Solmi, M., Kang, R., Salman, M., Lee, K., Milone, B., & Wafy, G. (2024). Impact of electrolyte abnormalities and adverse outcomes in persons with eating disorders: A systematic review protocol. *PloS one,* 19(8), e0308000.

Bulik, C. M. (2002). Anxiety, depression and eating disorders. *Eating Disorders and Obesity: A Comprehensive Handbook*, 2(1), 193-198.

Burnette, C. B., Simpson, C. C., & Mazzeo, S. E. (2017). Exploring gender differences in the link between weight suppression and eating pathology. *Eating Behaviors,* 27, 17-22.

Butryn, M. L., Thomas, J. G., & Lowe, M. R. (2009). Reductions in internal disinhibition during weight loss predict better weight loss maintenance. *Obesity*, 17(5), 1101-1103.

Button, E. J., Chadalavada, B., & Palmer, R. L. (2010). Mortality and predictors of death in a cohort of patients presenting to an eating disorders service. *International Journal of Eating Disorders*, 43(5), 387-392.

Calati, R., De Ronchi, D., Bellini, M., & Serretti, A. (2011). The 5-HTTLPR polymorphism and eating disorders: A meta-analysis. *International Journal of Eating Disorders*, 44(3), 191-199.

Calugi, S., Chignola, E., El Ghoch, M., & Dalle Grave, R. (2018). Starvation symptoms in patients with anorexia nervosa: a longitudinal study. *Eating Disorders*, 26(6), 523-537.

Calvo-Rivera, M. P., Navarrete-Páez, M. I., Bodoano, I., & Gutiérrez-Rojas, L. (2022). Comorbidity between anorexia nervosa and depressive disorder: A narrative review. *Psychiatry Investigation*, 19(3), 155.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *Bmj*, 321(7262), 694-696.

Carbone, E. A., D'Amato, P., Vicchio, G., De Fazio, P., & Segura-Garcia, C. (2021). A systematic review on the role of microbiota in the pathogenesis and treatment of eating disorders. European Psychiatry, 64(1), e2.

Carlson, J. L., & Lemly, D. C. (2024). Medical Considerations and Consequences of Eating Disorders. *Focus*, 22(3), 301-306.

Carpita, B., Muti, D., Cremone, I. M., Fagiolini, A., & Dell'Osso, L. (2022). Eating disorders and autism spectrum: Links and risks. *CNS spectrums,* 27(3), 272-280.

Cass, K., McGuire, C., Bjork, I., Sobotka, N., Walsh, K., & Mehler, P. S. (2020). Medical complications of anorexia nervosa. *Psychosomatics*, 61(6), 625-631.

Castellini, G., Sauro, C. L., Mannucci, E., Ravaldi, C., Rotella, C. M., Faravelli, C., & Ricca, V. (2011). Diagnostic crossover and outcome predictors in eating disorders according to DSM-IV and DSM-V proposed criteria: a 6-year follow-up study. *Psychosomatic Medicine*, 73(3), 270-279.

Catone, G. (2023). Anorexia Nervosa and Concurrent Psychiatric Comorbidity. *In Eating Disorders* (pp. 1-26). Springer.

Chapa, D. A., Bohrer, B. K., & Forbush, K. T. (2018). Is the diagnostic threshold for bulimia nervosa clinically meaningful? *Eating Behaviors*, 28, 16-19.

Chernyak, Y., & Lowe, M. R. (2010). Motivations for dieting: Drive for Thinness is different from Drive for Objective Thinness. *Journal of Abnormal Psychology*, 119(2), 276.

Chidiac, C. W. (2019). An update on the medical consequences of anorexia *Nervosa. Current Opinion in Pediatrics*, 31(4), 448-453.

Claudat, K., Brown, T. A., Anderson, L., Bongiorno, G., Berner, L. A., Reilly, E., Luo, T., Orloff, N., & Kaye, W. H. (2020). Correlates of co-occurring eating disorders and substance use disorders: a case for dialectical behavior therapy. *Eating Disorders*, 28(2), 142-156.

Coffino, J. A., Udo, T., & Grilo, C. M. (2019). The significance of overvaluation of shape or weight in binge-eating disorder: Results from a National Sample of US adults. *Obesity*, 27(8), 1367-1371.

Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, 6(1), 1-8.

Dahill, L. M., Touyz, S., Morrison, N. M., & Hay, P. (2021). Parental appearance teasing in adolescence and associations with eating problems: a systematic review. *BMC Public Health*, 21, 1-13.

Dahlenburg, S. C., Gleaves, D. H., & Hutchinson, A. D. (2019). Anorexia nervosa and perfectionism: A meta-analysis. *International Journal of Eating Disorders*, 52(3), 219-229.

Dakanalis, A., Carrà, G., Calogero, R., Zanetti, M. A., Gaudio, S., Caccialanza, R., Riva, G., & Clerici, M. (2015). Testing the cognitive-behavioural maintenance models across DSM-5 bulimic-type eating disorder diagnostic groups: a multi-centre study. *European Archives of Psychiatry and Clinical Neuroscience*, 265, 663-676.

Dakanalis, A., Clerici, M., Riva, G., & Carra, G. (2017). Testing the DSM-5 severity indicator for bulimia nervosa in a treatment-seeking sample. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity,* 22, 161-167.

Dakanalis, A., Mentzelou, M., Papadopoulou, S. K., Papandreou, D., Spanoudaki, M., Vasios, G. K., Pavlidou, E., Mantzorou, M., & Giaginis, C. (2023). The association of emotional eating with overweight/obesity, depression, anxiety/stress, and dietary patterns: a review of the current clinical evidence. *Nutrients,* 15(5), 1173.

Dakanalis, A., Timko, C. A., Colmegna, F., Riva, G., & Clerici, M. (2018). Evaluation of the DSM-5 severity ratings for anorexia nervosa in a clinical sample. *Psychiatry Research*, 262, 124-128.

Dalle Grave, R., Sartirana, M., & Calugi, S. (2024). Overvaluation of Shape and Weight Module. In *Cognitive Behavior Therapy for Binge Eating Disorder* (pp. 175-187). Springer.

Dalle Grave, R., Sartirana, M., El Ghoch, M., & Calugi, S. (2018). DSM-5 severity specifiers for anorexia nervosa and treatment outcomes in adult females. *Eating Behaviors*, 31, 18-23.

Dang, A., Granero, R., Sanchez, I., Solé, L. G., Toro, J. J. d., Rosinska, M., Jimenez-Murcia, S., Krug, I., & Fernando, F. A. (2024). Unveiling Severity Indicators for Anorexia and Bulimia Nervosa Treatment Success: DSM–5 Versus ICD–11 Versus Drive for Thinness. European Eating Disorders Review.

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the DSM–5 severity ratings for eating disorders. *Clinical Psychology: Science and Practice*, 29(4), 325.

Dang, A. B., Kiropoulos, L., Anderluh, M., Collier, D., Fernandez-Aranda, F., Karwautz, A., Treasure, J., Wagner, G., & Krug, I. (2024). Do risk factors differentiate DSM-5 and drive for thinness severity groups for anorexia nervosa? *Journal of Eating Disorders,* 12(1), 5.

Dang, A. B., Kiropoulos, L., Castle, D., Jenkins, Z., Phillipou, A., Rossell, S., & Krug, I. (2024). Psychiatric comorbidity and severity in anorexia nervosa: a comparative study of the DSM-5, the ICD-11, and overvaluation of Weight/Shape severity ratings. *Eating Disorders,* 1-17.

Dang, A. B., Kiropoulos, L., Castle, D. J., Jenkins, Z., Phillipou, A., Rossell, S. L., & Krug, I. (2023). Assessing severity in anorexia nervosa: Do the DSM-5 and an alternative severity rating based on overvaluation of weight and shape severity differ in psychological and biological correlates? *European Eating Disorders Review*, 31(4), 447-461.

Dang, T. B., Hughes, E. K., Dang, A. B., Lai, H. Y., Lee, J., Liu, S., Portingale, J., Fuller-Tyszkiewicz, M., & Krug, I. (2024). Taking a Deeper Dive Into OSFED Subtypes: A Meta-Analysis and Systematic Review. *International Journal of Eating Disorders.*

Dawes, A. J., Maggard-Gibbons, M., Maher, A. R., Booth, M. J., Miake-Lye, I., Beroes, J. M., & Shekelle, P. G. (2016). Mental health conditions among patients seeking and undergoing bariatric surgery: a meta-analysis. *Jama,* 315(2), 150-163.

Day, S., Bussey, K., Trompeter, N., & Mitchison, D. (2022). The impact of teasing and bullying victimization on disordered eating and body image disturbance among adolescents: a systematic review. *Trauma, Violence, & Abuse*, 23(3), 985-1006.

De Filippo, E., Marra, M., Alfinito, F., Di Guglielmo, M. L., Majorano, P., Cerciello, G., De Caprio, C., Contaldo, F., & Pasanisi, F. (2016). Hematological complications in anorexia nervosa. *Eur J Clin Nutr,* 70(11), 1305-1308. https://doi.org/10.1038/ejcn.2016.115

de Rijk, E. S., Almirabi, D., Robinson, L., Schmidt, U., van Furth, E. F., & Slof-Op't Landt, M. C. (2024). An overview and investigation of relapse predictors in anorexia nervosa: A systematic review and meta-analysis. *International Journal of Eating Disorders*, 57(1), 3-26.

Del Casale, A., Adriani, B., Modesti, M. N., Virzì, S., Parmigiani, G., Vento, A. E., & Speranza, A. M. (2023). Anorexia nervosa and familial risk factors: a systematic review of the literature. *Current Psychology*, 42(29), 25476-25484.

Deloitte Access Economics. (2024). *Paying the Price, The Economic and Social Impact Of Eating Disorders In Australia*. https://butterfly.org.au/wp-content/uploads/2024/10/deloitte-au-eco-paying-the-price-second-edition-180724-new-Oct-24.pdf

Derogatis, L. (2002). SCL-90-R symptom checklist 90 revised (manual). Madrid: TEA Ediciones.

Devoe, D. J., Dimitropoulos, G., Anderson, A., Bahji, A., Flanagan, J., Soumbasis, A., Patten, S. B., Lange, T., & Paslakis, G. (2021). The prevalence of substance use disorders and substance use in anorexia nervosa: a systematic review and meta-analysis. *Journal of Eating Disorders*, 9, 1-15.

Dhopatkar, N., Keeler, J. L., Mutwalli, H., Whelan, K., Treasure, J., & Himmerich, H. (2023). Gastrointestinal symptoms, gut microbiome, probiotics and prebiotics in anorexia nervosa: A review of mechanistic rationale and clinical evidence. *Psychoneuroendocrinology,* 147, 105959.

Di Angelantonio, E., Bhupathiraju, S. N., Wormser, D., Gao, P., Kaptoge, S., De Gonzalez, A. B., Cairns, B. J., Huxley, R., Jackson, C. L., & Joshy, G. (2016). Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet*, 388(10046), 776-786.

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, 40(4), 314-321.

Diedrichs, P. C. (2017). Sociocultural environment and internalization of the thin ideal as eating disorder risk factors. In (pp. 782-786): Springer: Singapore.

Dieffenbacher, A. L., Meule, A., & Voderholzer, U. (2024). Purging frequency and number of purging methods as indicators of severity in bulimia nervosa: Interactive effects with binge eating episodes. *European Eating Disorders Review.*

Dobrescu, S. R., Dinkler, L., Gillberg, C., Råstam, M., Gillberg, C., & Wentz, E. (2020). Anorexia nervosa: 30-year outcome. *The British Journal of Psychiatry,* 216(2), 97-104.

Donato, K., Ceccarini, M. R., Dhuli, K., Bonetti, G., Medori, M. C., Marceddu, G., Precone, V., Xhufi, S., Bushati, M., & Bozo, D. (2022). Gene variants in eating disorders. Focus on anorexia nervosa, bulimia nervosa, and binge-eating disorder. *Journal Of Preventive Medicine and Hygiene*, 63(2 Suppl 3), E297.

Edakubo, S., & Fushimi, K. (2020). Mortality and risk assessment for anorexia nervosa in acute-care hospitals: a nationwide administrative database analysis. *BMC psychiatry*, *20*, 1-8.

Eddy, K. T., Crosby, R. D., Keel, P. K., Wonderlich, S. A., le Grange, D., Hill, L., Powers, P., & Mitchell, J. E. (2009). Empirical identification and validation of eating disorder phenotypes in a multisite clinical sample. *The Journal of Nervous And Mental Disease*, *197*(1), 41-49.

Eddy, K. T., Tabri, N., Thomas, J. J., Murray, H. B., Keshaviah, A., Hastings, E., Edkins, K., Krishna, M., Herzog, D. B., & Keel, P. K. (2017). Recovery from anorexia nervosa and bulimia nervosa at 22-year follow-up. *The Journal of Clinical Psychiatry*, *78*(2), 17085.

Edler, C., Haedt, A. A., & Keel, P. K. (2007). The use of multiple purging methods as an indicator of eating disorder severity. *International Journal of Eating Disorders*, *40*(6), 515-520.

Engelhardt, C., Föcker, M., Bühren, K., Dahmen, B., Becker, K., Weber, L., Correll, C. U., Egberts, K. M., Ehrlich, S., & Roessner, V. (2021). Age dependency of body mass index distribution in childhood and adolescent inpatients with anorexia nervosa with a focus on DSM-5 and ICD-11 weight criteria and severity specifiers. *European Child & Adolescent Psychiatry*, *30*, 1081-1094.

Errichiello, L., Iodice, D., Bruzzese, D., Gherghi, M., & Senatore, I. (2016). Prognostic factors and outcome in anorexia nervosa: a follow-up study. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *21*, 73-82.

Eshetu, B., Worede, A., Fentie, A., Chane, E., Fetene, G., Wondifraw, H., Shimelis, M., Girma, M., Hadgu, R., & Demeke, K. (2023). Assessment of Electrolyte Imbalance and Associated Factors Among Adult Diabetic Patients Attending the University of Gondar Comprehensive Specialized Hospital, Ethiopia: A Comparative Cross-Sectional Study. *Diabetes, Metabolic Syndrome and Obesity*, 1207-1220.

Fairburn, C. (1993). The eating disorder examination. *Binge eating: Nature, assessment, and treatment/Guilford*.

Fairburn, C. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press.

Fairburn, C. G. (2008). Eating disorders: The transdiagnostic view and the cognitive behavioral theory.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363-370.

Fairburn, C. G., & Beglin, S. J. (2008). Eating disorder examination questionnaire. *Cognitive behavior therapy and eating disorders*, *309*, 313.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, *41*(5), 509-528.

Fairburn, C. G., Jones, R., Peveler, R. C., Carr, S. J., Solomon, R. A., O'Connor, M. E., Burton, J., & Hope, R. (1991). Three psychological treatments for bulimia nervosa: A comparative trial. *Archives of General Psychiatry*, *48*(5), 463-469.

Fan, S., Guo, W., Xiao, D., Guan, M., Liao, T., Peng, S., Feng, A., Wang, Z., Yin, H., & Li, M. (2023). Microbiota-gut-brain axis drives overeating disorders. *Cell Metabolism*, *35*(11), 2011-2027. e2017.

Fatt, S. J., George, E., Hay, P., Jeacocke, N., Gotkiewicz, E., & Mitchison, D. (2024). An umbrella review of body image concerns, disordered eating, and eating disorders in elite athletes. *Journal of Clinical Medicine*, *13*(14), 4171.

Fatt, S. J., Mitchison, D., Bussey, K., & Mond, J. (2022). Methods used to assess insight in individuals with eating disorders: a scoping review. *Journal of Mental Health*, 1-12.

Fernández-Aranda, F., Bel, M., Jimenez, S., Vinuales, M., Turon, J., & Vallejo, J. (1998). Outpatient group therapy for anorexia nervosa: a preliminary study. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *3*, 1-6.

Fernández-Aranda, F., Treasure, J., Paslakis, G., Agüera, Z., Giménez, M., Granero, R., Sánchez, I., Serrano-Troncoso, E., Gorwood, P., & Herpertz-Dahlmann, B. (2021). The impact of duration of illness on treatment nonresponse and drop-out: Exploring the relevance of enduring eating disorder concept. *European Eating Disorders Review*, *29*(3), 499-513.

First, M. B. (2024). Psychiatric classification. In *Tasman's Psychiatry* (pp. 1465-1491). Springer.

Fletcher, L., Trip, H., Lawson, R., Wilson, N., & Jordan, J. (2021). Life is different now–impacts of eating disorders on Carers in New Zealand: a qualitative study. *Journal of Eating Disorders*, *9*, 1-12.

Forrest, L. N., Grilo, C. M., & Udo, T. (2021). Suicide attempts among people with eating disorders and adverse childhood experiences: Results from a nationally representative sample of adults. *International Journal of Eating Disorders*, *54*(3), 326-335.

Forrest, L. N., Jacobucci, R. C., & Grilo, C. M. (2022). Empirically determined severity levels for binge-eating disorder outperform existing severity classification schemes. *Psychological Medicine*, *52*(4), 685-695.

Forrest, L. N., Jones, P. J., Ortiz, S. N., & Smith, A. R. (2018). Core psychopathology in anorexia nervosa and bulimia nervosa: A network analysis. *International Journal of Eating Disorders*, *51*(7), 668-679.

Frank, G. K., & Berner, L. A. (2020). *Binge eating: a transdiagnostic psychopathology*. Springer.

Franko, D. L., Tabri, N., Keshaviah, A., Murray, H. B., Herzog, D. B., Thomas, J. J., Coniglio, K., Keel, P. K., & Eddy, K. T. (2018). Predictors of long-term recovery in anorexia nervosa and bulimia nervosa: Data from a 22-year longitudinal study. *Journal of Psychiatric Research*, *96*, 183-188.

Frisell, T. (2021). Invited commentary: sibling-comparison designs, are they worth the effort? *American Journal of Epidemiology*, *190*(5), 738-741.

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000–2018 period: a systematic literature review. *The American Journal of Clinical Nutrition*, *109*(5), 1402-1413.

Garner, D. M. (1991a). Eating disorder inventory-2 (EDI-2). *Proffesional Manual. odessa: Psychological Assessment. Resources*.

Garner, D. M. (1991b). *Eating Disorder Inventory-2: EDI2; Professional Kit*. Psychological Assessment Resources Odessa, FL.

Garner, D. M. (2004). Eating disorder inventory-3 (EDI-3). *Professional manual. Odessa, FL: Psychological Assessment Resources*, *1*.

Garner, D. M., Olmsted, M. P., Polivy, J., & Garfinkel, P. E. (1984). Comparison between weight-preoccupied women and anorexia nervosa. *Psychosomatic Medicine*, *46*(3), 255-266.

Gianini, L., Roberto, C. A., Attia, E., Walsh, B. T., Thomas, J. J., Eddy, K. T., Grilo, C. M., Weigel, T., & Sysko, R. (2017). Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers. *International Journal of Eating Disorders*, *50*(8), 906-916.

Gibson, D., Workman, C., & Mehler, P. S. (2019). Medical complications of anorexia nervosa and bulimia nervosa. *Psychiatric Clinics*, *42*(2), 263-274.

Giel, K. E., Bulik, C. M., Fernandez-Aranda, F., Hay, P., Keski-Rahkonen, A., Schag, K., Schmidt, U., & Zipfel, S. (2022). Binge eating disorder. *Nature Reviews Disease Primers*, *8*(1), 16.

Golden, N. H., & Mehler, P. S. (2020). Atypical anorexia nervosa can be just as bad. *Clevel Clin J Med*, *87*(3), 172-174.

Gorrell, S., Byrne, C. E., Trojanowski, P. J., Fischer, S., & Le Grange, D. (2022). A scoping review of non-specific predictors, moderators, and mediators of family-based treatment for adolescent anorexia and bulimia nervosa: a summary of the current research findings. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *27*(6), 1971-1990.

Gorrell, S., Hail, L., Kinasz, K., Bruett, L., Forsberg, S., Delucchi, K., Lock, J., & Le Grange, D. (2019). A test of the DSM-5 severity specifier for bulimia nervosa in adolescents: Can we anticipate clinical treatment outcomes? *International Journal of Eating Disorders*, *52*(5), 586-590.

Gorrell, S., Reilly, E. E., Schaumberg, K., Anderson, L. M., & Donahue, J. M. (2019). Weight suppression and its relation to eating disorder and weight outcomes: a narrative review. *Eating Disorders*, *27*(1), 52-81.

Gregertsen, E. C., Mandy, W., Kanakam, N., Armstrong, S., & Serpell, L. (2019). Pre-treatment patient characteristics as predictors of drop-out and treatment outcome in individual and family therapy for adolescents and adults with anorexia nervosa: a systematic review and meta-analysis. *Psychiatry Research*, *271*, 484-501.

Grilo, C. M., Ivezaj, V., & White, M. A. (2015a). Evaluation of the DSM-5 severity indicator for binge eating disorder in a clinical sample. *Behaviour Research and Therapy*, *71*, 110-114.

Grilo, C. M., Ivezaj, V., & White, M. A. (2015b). Evaluation of the DSM-5 severity indicator for binge eating disorder in a community sample. *Behaviour Research and Therapy*, *66*, 72-76.

Guo, W., & Xiong, W. (2024). From gut microbiota to brain: implications on binge eating disorders. *Gut Microbes*, *16*(1), 2357177.

Hagan, K. E., Clark, K. E., & Forbush, K. T. (2017). Incremental validity of weight suppression in predicting clinical impairment in bulimic syndromes. *International Journal of Eating Disorders*, *50*(6), 672-678.

Hambleton, A., Pepin, G., Le, A., Maloney, D., Touyz, S., & Maguire, S. (2022). Psychiatric and medical comorbidities of eating disorders: findings from a rapid review of the literature. *Journal of Eating Disorders*, *10*(1), 132.

Hamilton, A., Mitchison, D., Basten, C., Byrne, S., Goldstein, M., Hay, P., Heruc, G., Thornton, C., & Touyz, S. (2022). Understanding treatment delay: perceived barriers preventing treatment-seeking for eating disorders. *Australian & New Zealand Journal of Psychiatry*, *56*(3), 248-259.

Hammerle, F., Huss, M., Ernst, V., & Bürger, A. (2016). Thinking dimensional: prevalence of DSM-5 early adolescent full syndrome, partial and subthreshold eating disorders in a cross-sectional survey in German schools. *BMJ open*, *6*(5), e010843.

Hanson, J. A., Phillips, L. N., Hughes, S. M., & Corson, K. (2020). Attention-deficit hyperactivity disorder symptomatology, binge eating disorder symptomatology, and body mass index among college students. *Journal of American College Health*, *68*(5), 543-549.

Harrop, E. N., Hutcheson, R., Harner, V., Mensinger, J. L., & Lindhorst, T. (2023). “You Don’t Look Anorexic”: Atypical anorexia patient experiences of weight stigma in medical care. *Body Image*, *46*, 48-61.

Hart, L. M., Granillo, M. T., Jorm, A. F., & Paxton, S. J. (2011). Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. *Clinical Psychology Review*, *31*(5), 727-735.

Hartmann, S., Timm, C., Barnow, S., Rubel, J. A., Lalk, C., & Pruessner, L. (2024). Web-based cognitive behavioral treatment for Bulimia Nervosa: a Randomized Clinical Trial. *JAMA Network Open*, *7*(7), e2419019-e2419019.

Hay, P., Aouad, P., Le, A., Marks, P., Maloney, D., Touyz, S., & Maguire, S. (2023). Epidemiology of eating disorders: population, prevalence, disease burden and quality of life informing public policy in Australia—a rapid review. *Journal of Eating Disorders*, *11*(1), 23.

Hay, P. J., Touyz, S., Claudino, A. M., Lujic, S., Smith, C. A., & Madden, S. (2019). Inpatient versus outpatient care, partial hospitalisation and waiting list for people with eating disorders. *Cochrane Database of Systematic Reviews*(1).

Haynos, A. F., Wall, M. M., Chen, C., Wang, S. B., Loth, K., & Neumark-Sztainer, D. (2018). Patterns of weight control behavior persisting beyond young adulthood: Results from a 15-year longitudinal study. *International Journal of Eating Disorders*, *51*(9), 1090-1097.

Hazzard, V. M., Loth, K. A., Hooper, L., & Becker, C. B. (2020). Food insecurity and eating disorders: A review of emerging evidence. *Current Psychiatry Reports*, *22*, 1-9.

Hebebrand, J., Gradl-Dietsch, G., Peters, T., Correll, C. U., & Haas, V. (2024). The diagnosis and treatment of anorexia nervosa in childhood and adolescence. *Deutsches Ärzteblatt International*, *121*(5), 164.

Hebebrand, J., Hildebrandt, T., Schlögl, H., Seitz, J., Denecke, S., Vieira, D., Gradl-Dietsch, G., Peters, T., Antel, J., & Lau, D. (2022). The role of hypoleptinemia in the psychological and behavioral adaptation to starvation: Implications for anorexia nervosa. *Neuroscience & Biobehavioral Reviews*, *141*, 104807.

Herzog, D. B., & Eddy, K. T. (2018). Comorbidity in eating disorders. *Annual Review of Eating Disorders*, 35-50.

Hilbert, A. (2019). Binge-eating disorder. *Psychiatric Clinics*, *42*(1), 33-43.

Hilbert, A., Bishop, M. E., Stein, R. I., Tanofsky-Kraff, M., Swenson, A. K., Welch, R. R., & Wilfley, D. E. (2012). Long-term efficacy of psychological treatments for binge eating disorder. *The British Journal of Psychiatry*, *200*(3), 232-237.

Hilbert, A., Hoek, H. W., & Schmidt, R. (2017). Evidence-based clinical guidelines for eating disorders: international comparison. *Current opinion in psychiatry*, *30*(6), 423-437.

Hilbert, A., Petroff, D., Herpertz, S., Pietrowsky, R., Tuschen-Caffier, B., Vocks, S., & Schmidt, R. (2020). Meta-analysis on the long-term effectiveness of psychological and medical treatments for binge-eating disorder. *International Journal of Eating Disorders*, *53*(9), 1353-1376.

Himmerich, H., Bentley, J., Kan, C., & Treasure, J. (2019). Genetic risk factors for eating disorders: an update and insights into pathophysiology. *Therapeutic Advances in Psychopharmacology*, *9*, 2045125318814734.

Himmerich, H., Hotopf, M., Shetty, H., Schmidt, U., Treasure, J., Hayes, R. D., Stewart, R., & Chang, C.-K. (2019a). Psychiatric comorbidity as a risk factor for mortality in people with anorexia nervosa. *European Archives of Psychiatry and Clinical Neuroscience*, *269*, 351-359.

Himmerich, H., Hotopf, M., Shetty, H., Schmidt, U., Treasure, J., Hayes, R. D., Stewart, R., & Chang, C.-K. (2019b). Psychiatric comorbidity as a risk factor for the mortality of people with bulimia nervosa. *Social Psychiatry and Psychiatric Epidemiology*, *54*, 813-821.

Hirtz, R., & Hinney, A. (2020). Genetic and epigenetic findings in anorexia nervosa. *Medizinische Genetik*, *32*(1), 25-29.

Hoek, H. W. (2006). Incidence, prevalence and mortality of anorexia nervosa and other eating disorders. *Current Opinion in Psychiatry*, *19*(4), 389-394.

Howard, M., Gregertsen, E. C., Hindocha, C., & Serpell, L. (2020). Impulsivity and compulsivity in anorexia and bulimia nervosa: a systematic review. *Psychiatry Research*, *293*, 113354.

Howard, T. L., Williams, M. O., Woodward, D., & Fox, J. R. (2023). The relationship between shame, perfectionism and Anorexia Nervosa: A grounded theory study. *Psychology and Psychotherapy: Theory, Research and Practice*, *96*(1), 40-55.

Hübel, C., Marzi, S. J., Breen, G., & Bulik, C. M. (2019). Epigenetics in eating disorders: a systematic review. *Molecular Psychiatry*, *24*(6), 901-915.

Hundemer, G. L., Clarke, A., Akbari, A., Bugeja, A., Massicotte-Azarniouch, D., Knoll, G., Myran, D. T., Tanuseputro, P., & Sood, M. M. (2022). Analysis of electrolyte abnormalities in adolescents and adults and subsequent diagnosis of an eating disorder. *JAMA Network Open*, *5*(11), e2240809-e2240809.

Ivancic, L., Maguire, S., Miskovic-Wheatley, J., Harrison, C., & Nassar, N. (2021). Prevalence and management of people with eating disorders presenting to primary care: a national study. *Australian & New Zealand Journal of Psychiatry*, *55*(11), 1089-1100.

Iwajomo, T., Bondy, S. J., de Oliveira, C., Colton, P., Trottier, K., & Kurdyak, P. (2021). Excess mortality associated with eating disorders: population-based cohort study. *The British Journal of Psychiatry*, *219*(3), 487-493.

Jacobi, C., Hayward, C., de Zwaan, M., Kraemer, H. C., & Agras, W. S. (2004). Coming to terms with risk factors for eating disorders: application of risk terminology and suggestions for a general taxonomy. *Psychological bulletin*, *130*(1), 19.

Jaeger, B., Ruggiero, G., Edlund, B., Gomez-Perretta, C., Lang, F., Mohammadkhani, P., Sahleen-Veasey, C., Schomer, H., & Lamprecht, F. (2002). Body dissatisfaction and its interrelations with other risk factors for bulimia nervosa in 12 countries. *Psychotherapy and Psychosomatics*, *71*(1), 54-61.

Jakovina, T., Crnković Batista, M., Ražić Pavičić, A., Žurić Jakovina, I., & Begovac, I. (2018). Emotional dysregulation and attachment dimensions in female patients with bulimia nervosa. *Psychiatria Danubina*, *30*(1), 72-78.

Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of basic and clinical pharmacy*, *5*(4), 87.

Jebeile, H., Lister, N. B., Baur, L. A., Garnett, S. P., & Paxton, S. J. (2021). Eating disorder risk in adolescents with obesity. *Obesity Reviews*, *22*(5), e13173.

Jenkins, P., Lebow, J., & Rienecke, R. D. (2018). Weight suppression as a predictor variable in the treatment of eating disorders: A systematic review. *Journal of Psychiatric and Mental Health Nursing*, *25*(5-6), 297-306.

Jérolon, A., Perduca, V., Delsedime, N., Abbate-Daga, G., & Marzola, E. (2022). Mediation models of anxiety and depression between temperament and drive for thinness and body dissatisfaction in anorexia nervosa. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *27*(7), 2569-2581.

Jones, E. J., Egan, S. J., Howell, J. A., Hoiles, K. J., & Mazzucchelli, T. G. (2020). An examination of the transdiagnostic cognitive-behavioural model of eating disorders in adolescents. *Eating Behaviors*, *39*, 101445.

Kambanis, P. E., Mancuso, C. J., Becker, K. R., Eddy, K. T., Thomas, J. J., & De Young, K. P. (2024). Course of avoidant/restrictive food intake disorder: Emergence of overvaluation of shape/weight. *Journal of Eating Disorders*, *12*(1), 54.

Kärkkäinen, U., Mustelin, L., Raevuori, A., Kaprio, J., & Keski-Rahkonen, A. (2018). Do disordered eating behaviours have long-term health-related consequences? *European Eating Disorders Review*, *26*(1), 22-28.

Karwautz, A., Rabe-Hesketh, S., Hu, X., Zhao, J., Sham, P., Collier, D., & Treasure, J. (2001). Individual-specific risk factors for anorexia nervosa: a pilot study using a discordant sister-pair design. *Psychological Medicine*, *31*(2), 317-329.

Kästner, D., Weigel, A., Buchholz, I., Voderholzer, U., Löwe, B., & Gumz, A. (2021). Facilitators and barriers in anorexia nervosa treatment initiation: a qualitative study on the perspectives of patients, carers and professionals. *Journal of Eating Disorders*, *9*, 1-11.

Käver, L., Hinney, A., Rajcsanyi, L. S., Maier, H. B., Frieling, H., Steiger, H., Voelz, C., Beyer, C., Trinh, S., & Seitz, J. (2024). Epigenetic alterations in patients with anorexia nervosa—a systematic review. *Molecular Psychiatry*, 1-15.

Keel, P. K., & Brown, T. A. (2010). Update on course and outcome in eating disorders. *International Journal of Eating Disorders*, *43*(3), 195-204.

Keel, P. K., & Heatherton, T. F. (2010). Weight suppression predicts maintenance and onset of bulimic syndromes at 10-year follow-up. *Journal of Abnormal Psychology*, *119*(2), 268.

Kenny, T. E., & Carter, J. C. (2018). I weigh therefore I am: Implications of using different criteria to define overvaluation of weight and shape in binge-eating disorder. *International Journal of Eating Disorders*, *51*(11), 1244-1251.

Kerr-Gaffney, J., Hayward, H., Jones, E. J., Halls, D., Murphy, D., & Tchanturia, K. (2021). Autism symptoms in anorexia nervosa: A comparative study with females with autism spectrum disorder. *Molecular Autism*, *12*(1), 47.

Keshishian, A. C., Tabri, N., Becker, K. R., Franko, D. L., Herzog, D. B., Thomas, J. J., & Eddy, K. T. (2021). Comorbid depression and substance use prospectively predict eating disorder persistence among women with anorexia nervosa and bulimia nervosa. *Journal of Behavioral and Cognitive Therapy*, *31*(4), 309-315.

Keski-Rahkonen, A. (2021). Epidemiology of binge eating disorder: prevalence, course, comorbidity, and risk factors. *Current Opinion in Psychiatry*, *34*(6), 525-531.

Keski-Rahkonen, A., Raevuori, A., Bulik, C. M., Hoek, H. W., Sihvola, E., Kaprio, J., & Rissanen, A. (2013). Depression and drive for thinness are associated with persistent bulimia nervosa in the community. *European Eating Disorders Review*, *21*(2), 121-129.

Khalil, R. B., Kassab, A., Richa, S., Seneque, M., Lefebvre, P., Sultan, A., Avignon, A., Maimoun, L., Renard, E., & Courtet, P. (2024). Weight suppression at lowest weight as an indicator of eating disorder clinical severity: A retrospective cohort study. *Eating Behaviors*, *53*, 101853.

Konstantakopoulos, G., Tchanturia, K., Surguladze, S., & David, A. (2011). Insight in eating disorders: clinical and cognitive correlates. *Psychological Medicine*, *41*(9), 1951-1961.

Krug, I., Binh Dang, A., Granero, R., Agüera, Z., Sánchez, I., Riesco, N., Jimenez-Murcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2021). Drive for thinness provides an alternative, more meaningful, severity indicator than the DSM-5 severity indices for eating disorders. *European Eating Disorders Review*, *29*(3), 482-498.

Krug, I., Casasnovas, C., Granero, R., Martinez, C., Jiménez-Murcia, S., Bulik, C., & Fernández-Aranda, F. (2008). Comparison study of full and subthreshold bulimia nervosa: Personality, clinical characteristics, and short-term response to therapy. *Psychotherapy Research*, *18*(1), 37-47.

Krug, I., Dang, A. B., & Hughes, E. K. (2024a). There is nothing as inconsistent as the OSFED diagnostic criteria. *Trends in Molecular Medicine*, *30*(4), 403-415. https://doi.org/10.1016/j.molmed.2024.01.006

Krug, I., Dang, A. B., & Hughes, E. K. (2024b). There is nothing as inconsistent as the OSFED diagnostic criteria. *Trends in Molecular Medicine*.

Krug, I., Dang, A. B., Sánchez, I., Granero, R., Agüera, Z., Gaspar-Perez, A., Jimenez-Murcia, S., & Fernandez-Aranda, F. (2024). How to assess eating disorder severity in males? The DSM-5 severity index versus severity based on drive for thinness. *Eating Disorders*, *32*(1), 81-97.

Kukk, K., & Akkermann, K. (2022). The interplay between binge eating risk factors: Toward an integrated model. *Journal of Health Psychology*, *27*(1), 92-102.

Laczkovics, C., Czernin, K., Carlitscheck, J., Zeiler, M., Schlund, P., Wunram, H. L., Lehmkuhl, G., & Krischer, M. (2023). Personality Disorder in Adolescent Patients with Anorexia Nervosa. *Psychopathology*, *56*(4), 268-275.

Lampard, A. M., Tasca, G. A., Balfour, L., & Bissada, H. (2013). An evaluation of the transdiagnostic cognitive-behavioural model of eating disorders. *European Eating Disorders Review*, *21*(2), 99-107.

Larsen, J. T., Bulik, C. M., Thornton, L. M., Koch, S. V., & Petersen, L. (2021). Prenatal and perinatal factors and risk of eating disorders. *Psychological Medicine*, *51*(5), 870-880.

Lavinia-Maria, P., Iorga, M., Anamika, J., & Petrariu, F. (2020). Adverse childhood experiences in relationship with eating disorders. *The Medical-Surgical Journal*, *124*(1), 123-133.

Le Grange, D., Lock, J., Accurso, E. C., Agras, W. S., Darcy, A., Forsberg, S., & Bryson, S. W. (2014). Relapse from remission at two-to four-year follow-up in two treatments for adolescent anorexia nervosa. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(11), 1162-1167.

Le Grange, D., Pradel, M., Pogos, D., Yeo, M., Hughes, E. K., Tompson, A., Court, A., Crosby, R. D., & Sawyer, S. M. (2021). Family-based treatment for adolescent anorexia nervosa: Outcomes of a stepped-care model. *International Journal of Eating Disorders*, *54*(11), 1989-1997.

Lieberman, A., Robison, M., Wonderlich, S. A., Crosby, R. D., Mitchell, J. E., Crow, S. J., Peterson, C. B., Le Grange, D., Bardone-Cone, A. M., & Kolden, G. (2023). Self-hate, dissociation, and suicidal behavior in bulimia nervosa. *Journal of Affective Disorders*, *335*, 44-48.

Linardon, J. (2018). Rates of abstinence following psychological or behavioral treatments for binge-eating disorder: Meta-analysis. *International Journal of Eating Disorders*, *51*(8), 785-797.

Linardon, J., Gleeson, J., Yap, K., Murphy, K., & Brennan, L. (2019). Meta-analysis of the effects of third-wave behavioural interventions on disordered eating and body image concerns: Implications for eating disorder prevention. *Cognitive behaviour therapy*, *48*(1), 15-38.

Linardon, J., & Wade, T. D. (2018). How many individuals achieve symptom abstinence following psychological treatments for bulimia nervosa? A meta-analytic review. *International Journal of Eating Disorders*, *51*(4), 287-294.

Linardon, J., Wade, T. D., De la Piedad Garcia, X., & Brennan, L. (2017). The efficacy of cognitive-behavioral therapy for eating disorders: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, *85*(11), 1080.

Liu, J., & Zhang, H. (2023). The Role of Negative Emotions in Binge Eating Disorder. *Journal of Education, Humanities and Social Sciences*, *8*, 1449-1453.

Lock, J., Le Grange, D., Agras, W. S., Moye, A., Bryson, S. W., & Jo, B. (2010). Randomized clinical trial comparing family-based treatment with adolescent-focused individual therapy for adolescents with anorexia nervosa. *Archives of general psychiatry*, *67*(10), 1025-1032.

Longo, P., Bevione, F., Amodeo, L., Martini, M., Panero, M., & Abbate-Daga, G. (2024). Perfectionism in anorexia nervosa: Associations with clinical picture and personality traits. *Clinical Psychology & Psychotherapy*, *31*(1), e2931.

Loth, K. A., MacLehose, R., Bucchianeri, M., Crow, S., & Neumark-Sztainer, D. (2014). Predictors of dieting and disordered eating behaviors from adolescence to young adulthood. *Journal of Adolescent Health*, *55*(5), 705-712.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335-343.

Lowe, M., Miller-Kovach, K., & Phelan, S. (2001). Weight-loss maintenance in overweight individuals one to five years following successful completion of a commercial weight loss program. *International Journal of Obesity*, *25*(3), 325-331.

Lowe, M. R. (1984). Dietary concern, weight fluctuation and weight status: Further explorations of the Restraint Scale. *Behaviour Research and Therapy*, *22*(3), 243-248.

Lowe, M. R. (1986). Dieting and binging: Some unanswered questions.

Lowe, M. R., Berner, L. A., Swanson, S. A., Clark, V. L., Eddy, K. T., Franko, D. L., Shaw, J. A., Ross, S., & Herzog, D. B. (2011). Weight suppression predicts time to remission from bulimia nervosa. *Journal of Consulting and Clinical Psychology*, *79*(6), 772.

Lowe, M. R., Davis, W., Lucks, D., Annunziato, R., & Butryn, M. (2006). Weight suppression predicts weight gain during inpatient treatment of bulimia nervosa. *Physiology & Behavior*, *87*(3), 487-492.

Lowe, M. R., Haller, L. L., Singh, S., & Chen, J. Y. (2020). Weight dysregulation, positive energy balance, and binge eating in eating disorders. *Binge eating: A transdiagnostic psychopathology*, 59-67.

Lowe, M. R., Piers, A. D., & Benson, L. (2018). Weight suppression in eating disorders: a research and conceptual update. *Current Psychiatry Reports*, *20*, 1-12.

Lucas, I., Jiménez-Murcia, S., & Fernández-Aranda, F. (2022). DSM–5 severity ratings for eating disorders: Some clinical implications and future directions.

Lydecker, J. A., & Grilo, C. M. (2022). Psychiatric comorbidity as predictor and moderator of binge-eating disorder treatment outcomes: an analysis of aggregated randomized controlled trials. *Psychological Medicine*, *52*(16), 4085-4093.

Madhavan Unny, N., Zarina, A., & Beena, V. (2023). Fluid and Electrolyte Balance. In *Textbook of Veterinary Physiology* (pp. 193-211). Springer.

Mairhofer, D., Zeiler, M., Philipp, J., Truttmann, S., Wittek, T., Skala, K., Mitterer, M., Schöfbeck, G., Laczkovics, C., & Schwarzenberg, J. (2021). Short-term outcome of inpatient treatment for adolescents with anorexia nervosa using DSM-5 remission criteria. *Journal of Clinical Medicine*, *10*(14), 3190.

Mallorquí-Bagué, N., Lozano-Madrid, M., Granero, R., Mestre-Bach, G., Vintró-Alcaraz, C., Sánchez, I., Jiménez-Murcia, S., & Fernández-Aranda, F. (2023). Cognitive and clinical gender-related differences among binge-spectrum eating disorders: Analysis of therapy response predictors. *European Eating Disorders Review*, *31*(3), 377-389.

Mandelli, L., Arminio, A., Atti, A.-R., & De Ronchi, D. (2019). Suicide attempts in eating disorder subtypes: a meta-analysis of the literature employing DSM-IV, DSM-5, or ICD-10 diagnostic criteria. *Psychological Medicine*, *49*(8), 1237-1249.

Mares, S. H., Burger, J., Lemmens, L. H., van Elburg, A. A., & Vroling, M. S. (2022). Evaluation of the cognitive behavioural theory of eating disorders: A network analysis investigation. *Eating Behaviors*, *44*, 101590.

Marzilli, E., Cerniglia, L., & Cimino, S. (2018). A narrative review of binge eating disorder in adolescence: prevalence, impact, and psychological treatment strategies. *Adolescent health, medicine and therapeutics*, 17-30.

Mashalpourfard, M. (2018). Relationship of social anxiety, body image perception and depression with bulimia nervosa and anorexia nervosa in youth. *Journal of Fundamentals of Mental Health*, *20*(2).

Mason, T. B., Do, B., Chu, D., Belcher, B. R., Dunton, G. F., & Lopez, N. V. (2022). Associations among affect, diet, and activity and binge-eating severity using ecological momentary assessment in a non-clinical sample of middle-aged fathers. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 1-9.

Mason, T. B., Smith, K. E., Crosby, R. D., Engel, S. G., & Wonderlich, S. A. (2021). Examination of momentary maintenance factors and eating disorder behaviors and cognitions using ecological momentary assessment. *Eating Disorders*, *29*(1), 42-55.

Matherne, C. E., Watson, H., Fassnacht, D. B., Ali, K., Zerwas, S., Peat, C., Runfola, C., Levine, M. D., Marcus, M. D., & Zimmer, B. (2022). An exploratory investigation of predictors of outcome in face-to-face and online cognitive-behavioural therapy for bulimia nervosa. *European Eating Disorders Review*, *30*(4), 373-387.

Maurel, L., MacKean, M., & Lacey, J. H. (2024). Factors predicting long-term weight maintenance in anorexia nervosa: a systematic review. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *29*(1), 24.

McGrath, J. J., Lim, C., Plana-Ripoll, O., Holtz, Y., Agerbo, E., Momen, N., Mortensen, P., Pedersen, C., Abdulmalik, J., & Aguilar-Gaxiola, S. (2020). Comorbidity within mental disorders: a comprehensive analysis based on 145 990 survey respondents from 27 countries. *Epidemiology and Psychiatric Sciences*, *29*, e153.

McIntosh, V. V., Jordan, J., Luty, S. E., Carter, F. A., McKenzie, J. M., Bulik, C. M., & Joyce, P. R. (2006). Specialist supportive clinical management for anorexia nervosa. *International Journal of Eating Disorders*, *39*(8), 625-632.

Meczekalski, B., Podfigurna-Stopa, A., & Katulski, K. (2013). Long-term consequences of anorexia nervosa. *Maturitas*, *75*(3), 215-220.

Mehler, P. S., Anderson, K., Bauschka, M., Cost, J., & Farooq, A. (2023). Emergency room presentations of people with anorexia nervosa. *Journal of Eating Disorders*, *11*(1), 16.

Mehler, P. S., Blalock, D. V., Walden, K., Kaur, S., McBride, J., Walsh, K., & Watts, J. (2018). Medical findings in 1,026 consecutive adult inpatient–residential eating disordered patients. *International Journal of Eating Disorders*, *51*(4), 305-313.

Mehler, P. S., & Rylander, M. (2015). Bulimia Nervosa–medical complications. *Journal of Eating Disorders*, *3*, 1-5.

Meneguzzo, P., Di Pascoli, L., Bindolo, M. C., Collantoni, E., Meregalli, V., Zanetti, T., Veronese, A., Tenconi, E., & Favaro, A. (2023). The psychological and medical landscape of anorexia nervosa over a decade: a retrospective study of possible physical and psychological shifts. *Journal of Clinical Medicine*, *12*(23), 7294.

Mestre-Bach, G., & Potenza, M. N. (2023). Potential biological markers and treatment implications for binge eating disorder and behavioral addictions. *Nutrients*, *15*(4), 827.

Micali, N., Martini, M. G., Thomas, J. J., Eddy, K. T., Kothari, R., Russell, E., Bulik, C. M., & Treasure, J. (2017). Lifetime and 12-month prevalence of eating disorders amongst women in mid-life: a population-based study of diagnoses and risk factors. *BMC Medicine*, *15*, 1-10.

Milano, W., Ambrosio, P., Carizzone, F., De Biasio, V., & Capasso, A. (2020). Other Specific Feeding or Eating Disorders. *Pharmacology OnLine*, *3*, 109-125.

Milano, W., Milano, L., & Capasso, A. (2018). Health consequences of bulimia nervosa. *Biomedical Research and Clinical Practice*, *3*(1), 1-5.

Milembamane, M., Moussa, N. M., Twynstra, J., & Seabrook, J. A. (2023). Maternal Eating Disorders and Adverse Birth Outcomes: A Systematic Review and Meta-Analysis. *Canadian Journal of Dietetic Practice and Research*, *85*(1), 45-53.

Miller, A. E., Racine, S. E., & Klonsky, E. D. (2021). Symptoms of anorexia nervosa and bulimia nervosa have differential relationships to borderline personality disorder symptoms. *Eating Disorders*, *29*(2), 161-174.

Mills, R., Hyam, L., & Schmidt, U. (2023). A narrative review of early intervention for eating disorders: barriers and facilitators. *Adolescent Health, Medicine and Therapeutics*, 217-235.

Miranda-Olivos, R., Agüera, Z., Granero, R., Jiménez-Murcia, S., Puig-Llobet, M., Lluch-Canut, M. T., Gearhardt, A. N., & Fernández-Aranda, F. (2023). The role of food addiction and lifetime substance use on eating disorder treatment outcomes. *Nutrients*, *15*(13), 2919.

Miskovic-Wheatley, J., Bryant, E., Ong, S. H., Vatter, S., Le, A., Touyz, S., & Maguire, S. (2023). Eating disorder outcomes: findings from a rapid review of over a decade of research. *Journal of Eating Disorders*, *11*(1), 85.

Mitchell, J., Huckstepp, T., Allen, A., Louis, P., Anijärv, T., & Hermens, D. (2024). Early adaptive schemas, emotional regulation, and cognitive flexibility in eating disorders: subtype specific predictors of eating disorder symptoms using hierarchical linear regression. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *29*(1), 54.

Mitchison, D., Hay, P., Slewa-Younan, S., & Mond, J. (2012). Time trends in population prevalence of eating disorder behaviors and their relationship to quality of life. *PloS one*, *7*(11), e48450.

Mitchison, D., Mond, J., Bussey, K., Griffiths, S., Trompeter, N., Lonergan, A., Pike, K. M., Murray, S. B., & Hay, P. (2020). DSM-5 full syndrome, other specified, and unspecified eating disorders in Australian adolescents: prevalence and clinical significance. *Psychological Medicine*, *50*(6), 981-990.

Momen, N. C., Plana-Ripoll, O., Yilmaz, Z., Thornton, L. M., McGrath, J. J., Bulik, C. M., & Petersen, L. V. (2022). Comorbidity between eating disorders and psychiatric disorders. *International Journal of Eating Disorders*, *55*(4), 505-517.

Monteleone, A. M., Monteleone, P., Serino, I., Scognamiglio, P., Di Genio, M., & Maj, M. (2015). Childhood trauma and cortisol awakening response in symptomatic patients with anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, *48*(6), 615-621.

Monteleone, A. M., Tzischinsky, O., Cascino, G., Alon, S., Pellegrino, F., Ruzzi, V., & Latzer, Y. (2022). The connection between childhood maltreatment and eating disorder psychopathology: a network analysis study in people with bulimia nervosa and with binge eating disorder. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 1-9.

Mora-Maltas, B., Lucas, I., Granero, R., Vintró-Alcaraz, C., Miranda-Olivos, R., Baenas, I., Sánchez, I., Jiménez-del Toro, J., Sánchez-González, J., & Krug, I. (2023). Cognitive flexibility and DSM-5 severity criteria for eating disorders: assessing drive for thinness and duration of illness as alternative severity variables. *Journal of Eating Disorders*, *11*(1), 155.

Moskowitz, L., & Weiselberg, E. (2017). Anorexia nervosa/atypical anorexia nervosa. *Current Problems in Pediatric and Adolescent Health Care*, *47*(4), 70-84.

Murray, S. B., & Anderson, L. K. (2015). Deconstructing “atypical” eating disorders: an overview of emerging eating disorder phenotypes. *Current Psychiatry Reports*, *17*, 1-7.

Nagata, J. M., Garber, A. K., Tabler, J. L., Murray, S. B., & Bibbins-Domingo, K. (2018). Prevalence and correlates of disordered eating behaviors among young adults with overweight or obesity. *Journal of General Internal Medicine*, *33*, 1337-1343.

Nagy, H., Paul, T., Jain, E., Loh, H., Kazmi, S. H., Dua, R., Rodriguez, R., Naqvi, S. A. A., & Bidika, E. (2023). A clinical overview of anorexia nervosa and overcoming treatment resistance. *Avicenna Journal of Medicine*, *13*(01), 003-014.

Nieves, J. W., Formica, C., Ruffing, J., Zion, M., Garrett, P., Lindsay, R., & Cosman, F. (2005). Males have larger skeletal size and bone mass than females, despite comparable body size. *Journal of Bone and Mineral Research*, *20*(3), 529-535.

Nordgaard, R. A. (2020). *Revisiting the significance of duration of illness and level of underweight in predicting severity and outcome: towards a staging model for anorexia nervosa*

Paeratakul, S., White, M. A., Williamson, D. A., Ryan, D. H., & Bray, G. A. (2002). Sex, race/ethnicity, socioeconomic status, and BMI in relation to self-perception of overweight. *Obesity Research*, *10*(5), 345-350.

Paolacci, S., Kiani, A. K., Manara, E., Beccari, T., Ceccarini, M. R., Stuppia, L., Chiurazzi, P., Dalla Ragione, L., & Bertelli, M. (2020). Genetic contributions to the etiology of anorexia nervosa: New perspectives in molecular diagnosis and treatment. *Molecular Genetics & Genomic Medicine*, *8*(7), e1244.

Patel, N., Tyagi, R., Biswas, D., Birjees, A., Rajesh, C., & Khan, S. (2023). Anorexia Nervosa: Evaluating Disparities in Places of Death in the United States Over 22 Years Using the CDC WONDER Database. *Cureus*, *15*(12).

Pawaskar, M., Witt, E. A., Supina, D., Herman, B. K., & Wadden, T. A. (2017). Impact of binge eating disorder on functional impairment and work productivity in an adult community sample in the United States. *International Journal of Clinical Practice*, *71*(7), e12970.

Peckmezian, T., & Hay, P. (2017). A systematic review and narrative synthesis of interventions for uncomplicated obesity: weight loss, well-being and impact on eating disorders. *Journal of Eating Disorders*, *5*, 1-15.

Penas-Lledo, E., Fernández-Aranda, F., Jiménez-Murcia, S., Granero, R., Penelo, E., Soto, A., Gunnard, K., & Menchon, J. (2009). Subtyping eating disordered patients along drive for thinness and depression. *Behaviour Research and Therapy*, *47*(6), 513-519.

Peñas-Lledó, E., Bulik, C. M., Lichtenstein, P., Larsson, H., & Baker, J. H. (2015). Risk for self-reported anorexia or bulimia nervosa based on drive for thinness and negative affect clusters/dimensions during adolescence: A three-year prospective study of the TC h AD cohort. *International Journal of Eating Disorders*, *48*(6), 692-699.

Pennesi, J.-L., & Wade, T. D. (2016). A systematic review of the existing models of disordered eating: Do they inform the development of effective interventions? *Clinical Psychology Review*, *43*, 175-192.

Pettersen, G., Rosenvinge, J. H., Bakland, M., Wynn, R., Mathisen, T. F., & Sundgot-Borgen, J. (2018). Patients' and therapists' experiences with a new treatment programme for eating disorders that combines physical exercise and dietary therapy: the PED-t trial. A qualitative study protocol. *BMJ open*, *8*(1), e018708.

Pike, K., Hilbert, A., Wilfley, D. E., Fairburn, C., Dohm, F.-A., Walsh, B., & Striegel-Moore, R. (2008). Toward an understanding of risk factors for anorexia nervosa: a case-control study. *Psychological Medicine*, *38*(10), 1443-1453.

Poulsen, S., Lunn, S., Daniel, S. I., Folke, S., Mathiesen, B. B., Katznelson, H., & Fairburn, C. G. (2014). A randomized controlled trial of psychoanalytic psychotherapy or cognitive-behavioral therapy for bulimia nervosa. *American Journal of Psychiatry*, *171*(1), 109-116.

Preti, A., Rocchi, M. B. L., Sisti, D., Camboni, M. V., & Miotto, P. (2011). A comprehensive meta-analysis of the risk of suicide in eating disorders. *Acta Psychiatrica Scandinavica*, *124*(1), 6-17.

Prost-Lehmann, C., Shankland, R., França, L. R., Laurent, A., & Flaudias, V. (2018). Symptomatology long-term evolution after hospitalization for anorexia nervosa: Drive for thinness to explain effects of body dissatisfaction on type of outcome. *Psychiatry Research*, *266*, 212-217.

Puccio, F., Fuller-Tyszkiewicz, M., Ong, D., & Krug, I. (2016). A systematic review and meta-analysis on the longitudinal relationship between eating pathology and depression. *International Journal of Eating Disorders*, *49*(5), 439-454.

Puckett, L. (2023). Renal and electrolyte complications in eating disorders: a comprehensive review. *Journal of Eating Disorders*, *11*(1), 26.

Puckett, L., Grayeb, D., Khatri, V., Cass, K., & Mehler, P. (2021). A comprehensive review of complications and new findings associated with anorexia nervosa. *Journal of Clinical Medicine*, *10*(12), 2555.

Qian, J., Wu, Y., Liu, F., Zhu, Y., Jin, H., Zhang, H., Wan, Y., Li, C., & Yu, D. (2021). An update on the prevalence of eating disorders in the general population: a systematic review and meta-analysis. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 1-14.

Quadflieg, N., & Fichter, M. (2024). Course of Bulimia Nervosa and Binge-Eating Disorder. In *Handbook of Eating Disorders and Obesity* (pp. 85-90). Springer.

Quadflieg, N., Naab, S., Voderholzer, U., & Fichter, M. M. (2022). Long-term outcome in males with anorexia nervosa: A prospective, sex-matched study. *International Journal of Eating Disorders*, *55*(3), 393-398.

Raevuori, A., Linna, M. S., & Keski-Rahkonen, A. (2014). Prenatal and perinatal factors in eating disorders: a descriptive review. *International Journal of Eating Disorders*, *47*(7), 676-685.

Ramacciotti, C. E., Dell'Osso, L., Paoli, R. A., Ciapparelli, A., Coli, E., Kaplan, A. S., & Garfinkel, P. E. (2002). Characteristics of eating disorder patients without a drive for thinness. *International Journal of Eating Disorders*, *32*(2), 206-212.

Ratković, D., Knežević, V., Dickov, A., Fedrigolli, E., & Čomić, M. (2023). Comparison of binge-eating disorder and food addiction. *Journal of International Medical Research*, *51*(4), 03000605231171016.

Raugh, I. M., Chapman, H. C., Bartolomeo, L. A., Gonzalez, C., & Strauss, G. P. (2019). A comprehensive review of psychophysiological applications for ecological momentary assessment in psychiatric populations. *Psychological Assessment*, *31*(3), 304.

Reas, D. L., Williamson, D. A., Martin, C. K., & Zucker, N. L. (2000). Duration of illness predicts outcome for bulimia nervosa: A long-term follow-up study. *International Journal of Eating Disorders*, *27*(4), 428-434.

Regier, D. A., Kuhl, E. A., & Kupfer, D. J. (2013). The DSM-5: Classification and criteria changes. *World Psychiatry*, *12*(2), 92-98.

Ricciardelli, L. A., & McCabe, M. P. (2001). Dietary restraint and negative affect as mediators of body dissatisfaction and bulimic behavior in adolescent girls and boys. *Behaviour Research and Therapy*, *39*(11), 1317-1328.

Ricciardelli, L. A., McCabe, M. P., Holt, K. E., & Finemore, J. (2003). A biopsychosocial model for understanding body image and body change strategies among children. *Journal of Applied Developmental Psychology*, *24*(4), 475-495.

Rieger, E., Touyz, S. W., Swain, T., & Beumont, P. J. (2001). Cross-cultural research on anorexia nervosa: Assumptions regarding the role of body weight. *International Journal of Eating Disorders*, *29*(2), 205-215.

Riquin, E., Raynal, A., Mattar, L., Lalanne, C., Hirot, F., Huas, C., Duclos, J., group, E., Berthoz, S., & Godart, N. (2021). Is the severity of the clinical expression of anorexia nervosa influenced by an anxiety, depressive, or obsessive-compulsive comorbidity over a lifetime? *Frontiers in Psychiatry*, *12*, 658416.

Riva, A., Brasola, E., Sforza, S. E., Marfone, M., Biso, F., & Nacinovich, R. (2024). Anorexia Nervosa in Comorbidity With Borderline Personality Disorder in Adolescence: A Specific Clinical Endophenotype? *European Eating Disorders Review*.

Robinson, L., Flynn, M., & Cooper, M. (2024). Individual differences in motivation to change in individuals with eating disorders: A systematic review. *International Journal of Eating Disorders*, *57*(5), 1069-1087.

Romano, K. A., Heron, K. E., Amerson, R., Howard, L. M., MacIntyre, R. I., & Mason, T. B. (2020). Changes in disordered eating behaviors over 10 or more years: A meta-analysis. *International Journal of Eating Disorders*, *53*(7), 1034-1055.

Romeo, M., Cavaliere, G., & Traina, G. (2024). Bulimia Nervosa and Depression, from the Brain to the Gut Microbiota and Back. *Frontiers in Bioscience-Landmark*, *29*(8), 277.

Rosenbaum, M., & Leibel, R. L. (2010). Adaptive thermogenesis in humans. *International Journal of Obesity*, *34*(1), S47-S55.

Rosling, A. M., Sparén, P., Norring, C., & von Knorring, A. L. (2011). Mortality of eating disorders: A follow-up study of treatment in a specialist unit 1974–2000. *International Journal of Eating Disorders*, *44*(4), 304-310.

Rozakou-Soumalia, N., Dârvariu, Ş., & Sjögren, J. M. (2021). Dialectical behaviour therapy improves emotion dysregulation mainly in binge eating disorder and bulimia nervosa: a systematic review and meta-analysis. *Journal of Personalized Medicine*, *11*(9), 931.

Rubin, A. G., Schvey, N. A., Shank, L. M., Altman, D. R., Swanson, T. N., Ramirez, E., Moore, N. A., Jaramillo, M., Ramirez, S., & Davis, E. K. (2021). Associations between weight-based teasing and disordered eating behaviors among youth. *Eating Behaviors*, *41*, 101504.

Ruggeri, M., Leese, M., Thornicroft, G., Bisoffi, G., & Tansella, M. (2000). Definition and prevalence of severe and persistent mental illness. *The British Journal of Psychiatry*, *177*(2), 149-155.

Safer, D. L., Telch, C. F., & Chen, E. Y. (2009). *Dialectical behavior therapy for binge eating and bulimia*. Guilford Press.

Sagar, A. (2005). Long term health risks due to impaired nutrition in women with a past history of bulimia nervosa. *Nutrition Noteworthy*, *7*(1).

Sala, M., Keshishian, A., Song, S., Moskowitz, R., Bulik, C. M., Roos, C. R., & Levinson, C. A. (2023). Predictors of relapse in eating disorders: A meta-analysis. *Journal of Psychiatric Research*, *158*, 281-299.

Sala, M., & Levinson, C. A. (2016). The longitudinal relationship between worry and disordered eating: Is worry a precursor or consequence of disordered eating? *Eating Behaviors*, *23*, 28-32.

Sand, L., Bøe, T., Shafran, R., Stormark, K. M., & Hysing, M. (2021). Perfectionism in adolescence: Associations with gender, age, and socioeconomic status in a Norwegian sample. *Frontiers in Public Health*, *9*, 688811.

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, *23*(3), 246-255.

Santomauro, D. F., Melen, S., Mitchison, D., Vos, T., Whiteford, H., & Ferrari, A. J. (2021). The hidden burden of eating disorders: an extension of estimates from the Global Burden of Disease Study 2019. *The Lancet Psychiatry*, *8*(4), 320-328.

Sardahaee, F. S., Holmen, T. L., Micali, N., & Kvaløy, K. (2017). Effects of single genetic variants and polygenic obesity risk scores on disordered eating in adolescents–The HUNT study. *Appetite*, *118*, 8-16.

Schaefer, L. M., Hazzard, V. M., Smith, K. E., Johnson, C. A., Cao, L., Crosby, R. D., Peterson, C. B., Crow, S. J., Bardone-Cone, A. M., & Joiner, T. E. (2021). Examining the roles of emotion dysregulation and impulsivity in the relationship between psychological trauma and substance abuse among women with bulimic-spectrum pathology. *Eating Disorders*, *29*(3), 276-291.

Schmidt, U., Adan, R., Böhm, I., Campbell, I. C., Dingemans, A., Ehrlich, S., Elzakkers, I., Favaro, A., Giel, K., & Harrison, A. (2016). Eating disorders: the big issue. *The Lancet Psychiatry*, *3*(4), 313-315.

Schnicker, K., Hiller, W., & Legenbauer, T. (2013). Drop-out and treatment outcome of outpatient cognitive–behavioral therapy for anorexia nervosa and bulimia nervosa. *Comprehensive Psychiatry*, *54*(7), 812-823.

Sharpe, D. (2015). Your chi-square test is statistically significant: now what? *Practical Assessment, Research & Evaluation*, *20*(8), n8.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(20), 22-33.

Skowron, K., Kurnik-Łucka, M., Dadański, E., Bętkowska-Korpała, B., & Gil, K. (2020). Backstage of eating disorder—About the biological mechanisms behind the symptoms of anorexia nervosa. *Nutrients*, *12*(9), 2604.

Smink, F. R., van Hoeken, D., Oldehinkel, A. J., & Hoek, H. W. (2014). Prevalence and severity of DSM-5 eating disorders in a community cohort of adolescents. *International Journal of Eating Disorders*, *47*(6), 610-619.

Søeby, M., Gribsholt, S. B., Clausen, L., & Richelsen, B. (2024). Overall and cause-specific mortality in anorexia nervosa; impact of psychiatric comorbidity and sex in a 40-year follow-up study. *International Journal of Eating Disorders*.

Solmi, M., Monaco, F., Højlund, M., Monteleone, A. M., Trott, M., Firth, J., Carfagno, M., Eaton, M., De Toffol, M., & Vergine, M. (2024). Outcomes in people with eating disorders: A transdiagnostic and disorder-specific systematic review, meta-analysis and multivariable meta-regression analysis. *World Psychiatry*, *23*(1), 124-138.

Solmi, M., Wade, T., Byrne, S., Del Giovane, C., Fairburn, C., Ostinelli, E., De Crescenzo, F., Johnson, C., Schmidt, U., & Treasure, J. (2021). Comparative efficacy and acceptability of psychological interventions for the treatment of adult outpatients with anorexia nervosa: a systematic review and network meta-analysis. *The Lancet Psychiatry*, *8*(3), 215-224.

Spindler, A., & Milos, G. (2007). Links between eating disorder symptom severity and psychiatric comorbidity. *Eating Behaviors*, *8*(3), 364-373.

Stackpole, R., Greene, D., Bills, E., & Egan, S. J. (2023). The association between eating disorders and perfectionism in adults: A systematic review and meta-analysis. *Eating Behaviors*, *50*, 101769.

Steiger, H., Gauvin, L., Joober, R., Israel, M., Badawi, G., Groleau, P., Bruce, K. R., Kin, N. N. Y., Sycz, L., & Ouelette, A. S. (2012). Interaction of the BcII glucocorticoid receptor polymorphism and childhood abuse in bulimia nervosa (BN): relationship to BN and to associated trait manifestations. *Journal of Psychiatric Research*, *46*(2), 152-158.

Steinhausen, H.-C. (2002). The outcome of anorexia nervosa in the 20th century. *American Journal of Psychiatry*, *159*(8), 1284-1293.

Steinhausen, H.-C. (2009). Outcome of eating disorders. *Child and adolescent psychiatric clinics of North America*, *18*(1), 225-242.

Steinhausen, H.-C. (2020). *Eating disorders in adolescence: Anorexia and bulimia nervosa* (Vol. 3). Walter de Gruyter GmbH & Co KG.

Steinhausen, H.-C., & Weber, S. (2009). The outcome of bulimia nervosa: findings from one-quarter century of research. *American Journal of Psychiatry*, *166*(12), 1331-1341.

Steward, T., Mestre-Bach, G., Agüera, Z., Granero, R., Martín-Romera, V., Sánchez, I., Riesco, N., Tolosa-Sola, I., Fernández-Formoso, J. A., & Fernández-García, J. C. (2016). Enduring changes in decision making in patients with full remission from anorexia nervosa. *European Eating Disorders Review*, *24*(6), 523-527.

Stewart, M. P., & Baumann, O. (2024). The Effectiveness of Adolescent-Focused Therapy and Family-Based Therapy for Anorexia Nervosa. *Psychological Reports*, 00332941241226687.

Stewart, T. M., Martin, C. K., & Williamson, D. A. (2022). The complicated relationship between dieting, dietary restraint, caloric restriction, and eating disorders: is a shift in public health messaging warranted? *International Journal of Environmental Research and Public Health*, *19*(1), 491.

Stice, E. (2001). A prospective test of the dual-pathway model of bulimic pathology: mediating effects of dieting and negative affect. *Journal of Abnormal Psychology*, *110*(1), 124.

Stice, E., Gau, J. M., Rohde, P., & Shaw, H. (2017). Risk factors that predict future onset of each DSM–5 eating disorder: Predictive specificity in high-risk adolescent females. *Journal of Abnormal Psychology*, *126*(1), 38.

Stice, E., Marti, C. N., & Rohde, P. (2013). Prevalence, incidence, impairment, and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. *Journal of Abnormal Psychology*, *122*(2), 445.

Stice, E., Onipede, Z. A., & Marti, C. N. (2021). A meta-analytic review of trials that tested whether eating disorder prevention programs prevent eating disorder onset. *Clinical Psychology Review*, *87*, 102046.

Stice, E., Rohde, P., Shaw, H., & Desjardins, C. (2020). Weight suppression increases odds for future onset of anorexia nervosa, bulimia nervosa, and purging disorder, but not binge eating disorder. *The American Journal of Clinical Nutrition*, *112*(4), 941-947.

Stice, E., Rohde, P., Shaw, H., & Gau, J. M. (2017). Clinician-led, peer-led, and internet-delivered dissonance-based eating disorder prevention programs: Acute effectiveness of these delivery modalities. *Journal of Consulting and Clinical Psychology*, *85*(9), 883.

Stice, E., Rohde, P., Shaw, H., & Gau, J. M. (2020). Clinician-led, peer-led, and internet-delivered dissonance-based eating disorder prevention programs: Effectiveness of these delivery modalities through 4-year follow-up. *Journal of Consulting and Clinical Psychology*, *88*(5), 481.

Stice, E., Shaw, H., & Nemeroff, C. (1998). Dual pathway model of bulimia nervosa: Longitudinal support for dietary restraint and affect-regulation mechanisms. *Journal of Social and Clinical Psychology*, *17*(2), 129-149.

Striegel-Moore, R. H., & Bulik, C. M. (2007). Risk factors for eating disorders. *American psychologist*, *62*(3), 181.

Stroebe, W. (2022). *Dieting, overweight and obesity: Self-regulation in a food-rich environment*. Routledge.

Sun, L., Sun, Z., Wu, L., Zhu, Z., Zhang, F., Shang, Z., Jia, Y., Gu, J., Zhou, Y., & Wang, Y. (2021). Prevalence and risk factors for acute posttraumatic stress disorder during the COVID-19 outbreak. *Journal of Affective Disorders*, *283*, 123-129.

Tan, E. J., Cistullo, L., Castle, D. J., Rossell, S. L., Jenkins, Z. M., & Phillipou, A. (2022). Depression, perceived disability and unemployment are associated with reduced life satisfaction in anorexia nervosa. *Eating Disorders*, *30*(3), 323-330.

Thornton, L. M., Mazzeo, S. E., & Bulik, C. M. (2011). The heritability of eating disorders: methods and current findings. *Behavioral Neurobiology of Eating Disorders*, 141-156.

Tian, C., Manfei, X., Justin, T., Hongyue, W., & Xiaohui, N. (2018). Relationship between Omnibus and Post-hoc Tests: An Investigation of performance of the F test in ANOVA. *Shanghai Archives of Psychiatry*, *30*(1), 60.

Tith, R. M., Paradis, G., Potter, B. J., Low, N., Healy-Profitós, J., He, S., & Auger, N. (2020). Association of bulimia nervosa with long-term risk of cardiovascular disease and mortality among women. *JAMA psychiatry*, *77*(1), 44-51.

Toppino, F., Longo, P., Martini, M., Abbate-Daga, G., & Marzola, E. (2022). Body mass index specifiers in anorexia nervosa: anything below the “extreme”? *Journal of Clinical Medicine*, *11*(3), 542.

Tseng, M. C. M., Chien, L. N., Tu, C. Y., & Liu, H. Y. (2023). Mortality in anorexia nervosa and bulimia nervosa: A population-based cohort study in Taiwan, 2002–2017. *International Journal of Eating Disorders*, *56*(6), 1135-1144.

Udo, T., Bitley, S., & Grilo, C. M. (2019). Suicide attempts in US adults with lifetime DSM-5 eating disorders. *BMC medicine*, *17*, 1-13.

Udo, T., & Grilo, C. M. (2018). Prevalence and correlates of DSM-5–defined eating disorders in a nationally representative sample of US adults. *Biological Psychiatry*, *84*(5), 345-354.

Udo, T., & Grilo, C. M. (2019). Psychiatric and medical correlates of DSM-5 eating disorders in a nationally representative sample of adults in the United States. *International Journal of Eating Disorders*, *52*(1), 42-50.

Ulfvebrand, S., Birgegård, A., Norring, C., Högdahl, L., & von Hausswolff-Juhlin, Y. (2015). Psychiatric comorbidity in women and men with eating disorders results from a large clinical database. *Psychiatry Research*, *230*(2), 294-299.

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders*, *48*(7), 946-971.

van Eeden, A. E., van Hoeken, D., Hendriksen, J. M., & Hoek, H. W. (2023). Increase in incidence of anorexia nervosa among 10-to 14-year-old girls: A nationwide study in the Netherlands over four decades. *International Journal of Eating Disorders*, *56*(12), 2295-2303.

Van Eeden, A. E., Van Hoeken, D., & Hoek, H. W. (2021). Incidence, prevalence and mortality of anorexia nervosa and bulimia nervosa. *Current Opinion In Psychiatry*, *34*(6), 515-524.

van Hoeken, D., & Hoek, H. W. (2020). Review of the burden of eating disorders: mortality, disability, costs, quality of life, and family burden. *Current Opinion in Psychiatry*, , *33*(6), 521-527.

Van Son, G. E., van der Meer, P. A., & Van Furth, E. F. (2013). Correlates and associations between weight suppression and binge eating symptomatology in a population-based sample. *Eating Behaviors*, *14*(2), 102-106.

Vansteelandt, K., Rijmen, F., Pieters, G., Probst, M., & Vanderlinden, J. (2007). Drive for thinness, affect regulation and physical activity in eating disorders: A daily life study. *Behaviour Research and Therapy*, *45*(8), 1717-1734.

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, *18*, 1-18.

Veldic, M., Jia, Y.-F., Choi, Y., Ayers-Ringler, J. R., Biernacka, J. M., Geske, J. R., McElroy, S., Frye, M., & Choi, D.-S. (2017). 450. In Bipolar Disorder, SLC1A2 Promoter Hypomethylation is Associated with Binge Eating Disorder and Nicotine Dependance. *Biological Psychiatry*, *81*(10), S183-S184.

Villa, F., Crippa, A., Rosi, E., Nobile, M., Brambilla, P., & Delvecchio, G. (2023). ADHD and eating disorders in childhood and adolescence: An updated minireview. *Journal of Affective Disorders*, *321*, 265-271.

Vogel, E. N., Singh, S., & Accurso, E. C. (2021). A systematic review of cognitive behavior therapy and dialectical behavior therapy for adolescent eating disorders. *Journal of Eating Disorders*, *9*, 1-38.

Wade, T. D., Treasure, J., & Schmidt, U. (2011). A case series evaluation of the Maudsley Model for treatment of adults with anorexia nervosa. *European Eating Disorders Review*, *19*(5), 382-389.

Wade, T. D., Wilksch, S. M., & Lee, C. (2012). A longitudinal investigation of the impact of disordered eating on young women's quality of life. *Health Psychology*, *31*(3), 352.

Waller, G., Micali, N., & James, A. (2014). General Practitioners are poor at identifying the eating disorders. *Advances in Eating Disorders: Theory, Research and Practice*, *2*(2), 146-157.

Walsh, B. T., Hagan, K. E., & Lockwood, C. (2023). A systematic review comparing atypical anorexia nervosa and anorexia nervosa. *International Journal of Eating Disorders*, *56*(4), 798-820.

Watson, H., & Bulik, C. (2013). Update on the treatment of anorexia nervosa: review of clinical trials, practice guidelines and emerging interventions. *Psychological Medicine*, *43*(12), 2477-2500.

Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R., Gaspar, H. A., Bryois, J., Hinney, A., Leppä, V. M., & Mattheisen, M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, *51*(8), 1207-1214.

Webb, H., Dalton, B., Irish, M., Mercado, D., McCombie, C., Peachey, G., Arcelus, J., Au, K., Himmerich, H., & Louise Johnston, A. (2022). Clinicians' perspectives on supporting individuals with severe anorexia nervosa in specialist eating disorder intensive treatment settings. *Journal of Eating Disorders*, *10*(1), 3.

Weigel, A., Löwe, B., & Kohlmann, S. (2019). Severity of somatic symptoms in outpatients with anorexia and bulimia nervosa. *European Eating Disorders Review*, *27*(2), 195-204.

Weissman, R. S. (2019). The role of sociocultural factors in the etiology of eating disorders. *Psychiatric Clinics*, *42*(1), 121-144.

Westermair, A. L., Weber, S., Westmoreland, P., Mehler, P. S., Elsner, F., & Trachsel, M. (2024). Scoping review of end-of-life care for persons with anorexia nervosa. *Annals of Palliative Medicine*, *13*(3), 68507-68707.

Westmoreland, P., Duffy, A., Rienecke, R., Le Grange, D., Joiner, T., Manwaring, J., Watters, A., & Mehler, P. (2022). Causes of death in patients with a history of severe anorexia nervosa. *Journal of Eating Disorders*, *10*(1), 200.

Whitelaw, M., Lee, K. J., Gilbertson, H., & Sawyer, S. M. (2018). Predictors of complications in anorexia nervosa and atypical anorexia nervosa: degree of underweight or extent and recency of weight loss? *Journal of Adolescent Health*, *63*(6), 717-723.

Widiger, T. A., & Crego, C. (2018). Mental disorders as discrete clinical conditions: Dimensional versus categorical classification. *Adult Psychopathology and Diagnosis, Eighth Edition*, 3-32.

Wild, B., Friederich, H.-C., Zipfel, S., Resmark, G., Giel, K., Teufel, M., Schellberg, D., Löwe, B., de Zwaan, M., & Zeeck, A. (2016). Predictors of outcomes in outpatients with anorexia nervosa–results from the ANTOP study. *Psychiatry Research*, *244*, 45-50.

Wilfley, D. E., Citrome, L., & Herman, B. K. (2016). Characteristics of binge eating disorder in relation to diagnostic criteria. *Neuropsychiatric Disease and Treatment*, 2213-2223.

Wilkop, M., Wade, T. D., Keegan, E., & Cohen-Woods, S. (2023). Impairments among DSM-5 eating disorders: A systematic review and multilevel meta-analysis. *Clinical Psychology Review*, *101*, 102267.

Withnell, S. J., Kinnear, A., Masson, P., & Bodell, L. P. (2022). How different are threshold and other specified feeding and eating disorders? Comparing severity and treatment outcome. *Frontiers in psychology*, *13*, 784512.

World Health Organization(WHO). (2019). *The ICD-11 classification of mental and behavioural disorders*. World Health Organization.

Wulan, S., Westerterp, K., & Plasqui, G. (2010). Ethnic differences in body composition and the associated metabolic profile: a comparative study between Asians and Caucasians. *Maturitas*, *65*(4), 315-319.

Yamamiya, Y., Desjardins, C. D., & Stice, E. (2023). Sequencing of symptom emergence in anorexia nervosa, bulimia nervosa, binge eating disorder, and purging disorder in adolescent girls and relations of prodromal symptoms to future onset of these eating disorders. *Psychological Medicine*, *53*(10), 4657-4665.

Yamamiya, Y., & Stice, E. (2024). Risk Factors That Predict Future Onset of Anorexia Nervosa, Bulimia Nervosa, Binge Eating Disorder, and Purging Disorder in Adolescent Girls. *Behavior Therapy*, *55*(4), 712-723. https://doi.org/https://doi.org/10.1016/j.beth.2023.10.002

Yao, S., Larsson, H., Norring, C., Birgegård, A., Lichtenstein, P., D'Onofrio, B. M., Almqvist, C., Thornton, L. M., Bulik, C. M., & Kuja-Halkola, R. (2021). Genetic and environmental contributions to diagnostic fluctuation in anorexia nervosa and bulimia nervosa. *Psychological Medicine*, *51*(1), 62-69.

Yilmaz, Z., Hardaway, J. A., & Bulik, C. M. (2015). Genetics and epigenetics of eating disorders. *Advances in Genomics and Genetics*, 131-150.

Yim, S. H., & Schmidt, U. (2019). Experiences of computer-based and conventional self-help interventions for eating disorders: A systematic review and meta-synthesis of qualitative research. *International Journal of Eating Disorders*, *52*(10), 1108-1124.

Yokokura, M., Terada, T., Bunai, T., Nakaizumi, K., Kato, Y., Yoshikawa, E., Futatsubashi, M., Suzuki, K., Yamasue, H., & Ouchi, Y. (2019). Alterations in serotonin transporter and body image-related cognition in anorexia nervosa. *NeuroImage: Clinical*, *23*, 101928.

Yoon, C., Mason, S. M., Hooper, L., Eisenberg, M. E., & Neumark-Sztainer, D. (2020). Disordered eating behaviors and 15-year trajectories in body mass index: findings from project eating and activity in teens and young adults (EAT). *Journal of Adolescent Health*, *66*(2), 181-188.

Yu, Z., & Tan, M. (2016). Disordered eating behaviors and food addiction among nutrition major college students. *Nutrients*, *8*(11), 673.

Zaccagnino, M., Cussino, M., Callerame, C., & Civilotti, C. (2018). The many faces of eating disorders: Anorexia nervosa (AN), bulimia nervosa (BN), binge eating disorder (BED), other specified feeding or eating disorder (OSFED), bulimarexia, and orthorexia. *Trauma-Informed Approaches to Eating Disorders*.

Zayas, L. V., Wang, S. B., Coniglio, K., Becker, K., Murray, H. B., Klosterman, E., Kay, B., Bean, P., Weltzin, T., & Franko, D. L. (2018). Gender differences in eating disorder psychopathology across DSM-5 severity categories of anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, *51*(9), 1098-1102.

Zeeck, A., Herpertz-Dahlmann, B., Friederich, H.-C., Brockmeyer, T., Resmark, G., Hagenah, U., Ehrlich, S., Cuntz, U., Zipfel, S., & Hartmann, A. (2018). Psychotherapeutic treatment for anorexia nervosa: a systematic review and network meta-analysis. *Frontiers in Psychiatry*, *9*, 158.

Zeeland, S.-V., Bloss, C., Tewhey, R., Bansal, V., Torkamani, A., Libiger, O., Duvvuri, V., Wineinger, N., Galvez, L., & Darst, B. (2014). Evidence for the role of EP HX2 gene variants in anorexia nervosa. *Molecular Psychiatry*, *19*(6), 724-732.

Zerwas, S., Lund, B. C., Von Holle, A., Thornton, L. M., Berrettini, W. H., Brandt, H., Crawford, S., Fichter, M. M., Halmi, K. A., & Johnson, C. (2013). Factors associated with recovery from anorexia nervosa. *Journal of Psychiatric Research*, *47*(7), 972-979.