Exploring the qualities of Electronic Health Record medical student documentation

Lisa Cheshire
ORCID Identifier 0000-0003-3935-7608

Submitted in fulfilment of the requirements of the degree of Masters of Education (Stream 150) Major Thesis

October 2016

Melbourne Graduate School of Education
Department of Medical Education
University of Melbourne
Abstract

Written communication within the health professions has been rapidly changing over the last decade. Implementation of Electronic Health Records (EHR) in health services is now widespread. Medical student teaching and learning of the skills specifically required for EHRs has lagged behind the implementation. Very few original studies have focused on EHR skills and there are no validated measures by which to assess any of the EHR skills students are expected to develop. Our study explored the attributes of quality EHR documentation recorded by medical students, with the purpose of the EHR documentation being the communication between health care professionals to share or transfer the clinical care of a patient.

Recently there have been published validated instruments for measuring quality in physician EHR documentation, one being Physician Documentation Quality Instrument (PDQI-9). The purpose of this study was to explore the attributes of quality of EHR documentation written by first-year clinical medical students by building upon existing literature. The PDQI-9 was used as a basis for defining the attributes of quality in EHR documentation as a foundation for assessing and providing feedback on the performance of documentation to medical students.

With the focus on assessment, and providing a content validated test domain for assessment in quality EHR documentation, we utilised Kane’s framework for validity to structure the study and a mixed method study design to achieve the depth of exploration required to examine the performance of quality documentation fully. The study was conducted in two stages. In the first stage of the study, an expert panel of assessors applied the PDQI-9 to existing EHR data recorded by first clinical year medical students in a graduate entry program. The assessors both scored the records and justified their grading. Descriptive statistics and thematic analysis were undertaken on the data collected, and the findings triangulated with the literature review. The second stage employed explanatory semi-structured interviews with the
expert assessors to better understand the findings of the first stage and reach consensus on a test domain for assessing quality documentation recorded by medical students.

Outcomes from our study indicated that the PDQI-9 in its current format was not valid in a medical student setting, however most of the attributes assessed by the PDQI-9 were deemed relevant and meaningful to assess if their interpretations were clarified. In addition, Professionalism of documentation was regarded as a quality attribute. Consensus was reached on modifications that have the potential to improve the validity of the assessment of quality documentation recorded by medical students. Further studies need to complete Kane’s framework of validity for an assessment instrument and collect evidence to broaden the validity of the scoring, the generalization of the assessment items, the extrapolation to the real world and the implications of this assessment for students and health services.
Declaration

This is to certify that

- the thesis comprises only my original work towards the masters except where indicated in the Preface
- due acknowledgment has been made in the text to all other material used
- the thesis is 25,000 in length exclusive of tables, maps, bibliographies and appendices, as approved by the Research Higher Degrees Committee

Signed: [Signature]

Printed Name: Lisa Cheshire

Date: 26 October 2016
Acknowledgements

Many people need to be thanked for helping me complete my Masters. First and foremost, my primary supervisor, Dr. Amy Gullickson walked with me through the whole process, providing timely and valuable feedback at all stages. She is an excellent educator who asks the right questions at just the right time, her help was invaluable. My secondary supervisor, Professor Geoff McColl was a great enabler of the whole process, and positively supported all proposals. My advisory committee Chair, Professor David Beckett was incredibly pragmatic and provided advice on all processes and ways forward throughout the whole process. Associate Professor Kris Elliott, a member of my advisory committee, was integral in starting the exploration on EHRs.

Professor Sue Finch provided excellent statistical advice, which had a strong influence on the study design. The Education Technology Team, in particular, Steve Martin, made it all possible. And finally, all my Department of Medical Education colleagues for endless pieces of advice on all manners of questions, it is a wonderfully supportive environment to study and work in.
Table of Contents

Abstract .......................................................................................................................... i
Declaration ..................................................................................................................... iii
Acknowledgements ........................................................................................................ iv
List of Tables ................................................................................................................... viii
List of Figures ................................................................................................................ ix
Glossary and Acronyms ................................................................................................. x

1 Introduction .................................................................................................................. 1
1.1 EHR documentation ............................................................................................. 1
1.2 EHR and medical education ................................................................................. 3
1.3 Purpose of the Study ............................................................................................ 4
1.4 Overview of Thesis .............................................................................................. 5

2 Literature Review ....................................................................................................... 6
2.1 Attributes of quality medical student EHR documentation ................................ 7
   2.1.1 Original studies ............................................................................................. 7
   2.1.2 Evaluations of simulated EHR programs ..................................................... 9
   2.1.3 Expert Opinion ........................................................................................... 11
2.2 Assessing quality of EHR documentation in physicians ...................................... 16
   2.2.1 Physician Document Quality Instrument .................................................. 17
   2.2.2 QNOTE ......................................................................................................... 20
   2.2.3 Summary ..................................................................................................... 23
2.3 Interpretations of quality ...................................................................................... 25
   2.3.1 Physician preference for quality in documentation .................................... 25
2.4 Medical Education Assessment Literature .......................................................... 27
2.5 Summary ............................................................................................................... 28

3 Methodology and Methods ......................................................................................... 30
3.1 Content-Related Validity Studies ......................................................................... 30
3.2 Kane’s Conceptual Framework to Validation ..................................................... 31
   3.2.1 Overview Kane’s Argument-Based Approach to Validation ....................... 31
3.3 Methodology .......................................................................................................... 35
3.4 Methods .................................................................................................................. 37
   3.4.1 Overview of the study design ................................................................. 37
3.5 Limitations of content-related validity studies .......................................................... 41
3.6 First Stage .................................................................................................................. 42
  3.6.1 Participants: Assessors ....................................................................................... 43
  3.6.2 Medical Student Documentation ....................................................................... 46
  3.6.3 Data Collection ................................................................................................. 49
  3.6.4 Procedure ........................................................................................................ 51
  3.6.5 Stage 1 Analysis ............................................................................................... 52
3.7 Second Stage .......................................................................................................... 53
  3.7.1 Data collection ................................................................................................. 54
  3.7.2 Stage 2 Analysis ............................................................................................... 57
  3.7.3 Summary ......................................................................................................... 57

4 Findings ....................................................................................................................... 59
  4.1 Overview of Analysis ............................................................................................ 59
  4.2 Overview of Stage 1 Descriptive Statistics ............................................................ 60
  4.3 Organising themes of analysis ............................................................................. 64
  4.4 Attributes that were similarly interpreted and meaningful in a medical student context
  67
    4.4.1 Thorough ......................................................................................................... 67
    4.4.2 Synthesised ..................................................................................................... 70
  4.5 Attributes that were interpreted in multiple ways and meaningfulness was
  inconclusive ................................................................................................................ 72
    4.5.1 Organised ......................................................................................................... 72
    4.5.2 Comprehensible .............................................................................................. 76
    4.5.3 Up to Date ........................................................................................................ 79
    4.5.4 Useful ................................................................................................................ 82
  4.6 Attributes that were difficult to interpret and meaningfulness was unlikely .......... 86
    4.6.1 Succinct ............................................................................................................ 86
    4.6.2 Internally Consistent ...................................................................................... 91
  4.7 Other issues arising from Stage 2 analysis ............................................................ 94
  4.8 Summary ............................................................................................................... 95

5 Discussion .................................................................................................................. 96
5.1 Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students? ................................................................. 98
5.2 The quality attributes and how they can be modified ...................................................... 100
  5.2.1 Level of training ........................................................................................................ 103
  5.2.2 Test domain for assessing quality EHR documentation in medical students ....... 104
5.3 Limitations and Implications for Future Study .......................................................... 111
5.4 Future Directions for Research .................................................................................. 112
5.5 Contribution ................................................................................................................. 114

References ......................................................................................................................... 114
Appendix A. PDQI-9 ........................................................................................................ 121
Appendix B. QNOTE ......................................................................................................... 122
Appendix C. Assessor consent, demographics, and logistics survey ......................... 123
Appendix D. Assessor Training Package ........................................................................ 128
Appendix E. Example of Interview Schedule ................................................................ 145
Appendix F. Proposed rubric for assessing quality of EHR documentation recorded by first year clinical medical students ................................................................. 165
# List of Tables

Table 2.1 Quality attributes referenced by original studies and simulated EHR reports ..............................................................11
Table 2.2 RIME framework and corresponding attributes of quality, EHR content and use ........................................................................13
Table 2.3 Quality attributes from expert opinion and impact studies ..........................................................16
Table 2.4 PDQI exploratory factor analysis .........................................................................................................................18
Table 2.5 PDQI-9 validated attributes and descriptions .................................................................................................19
Table 3.1 Kane's inferences (2006) ...............................................................................................................................32
Table 3.2 Kane’s inferences and the study research question .........................................................................................33
Table 3.3 Process of Kane's framework and our study design .........................................................................................34
Table 3.4 Relationship between research questions, Kane’s inferences, data collection and analysis ................................................................40
Table 3.5 Selection Criteria for assessors .......................................................................................................................44
Table 3.6 Selection criteria for medical student documentation ...................................................................................48
Table 3.7 Semi-structured Interview Schedule .............................................................................................................55
Table 4.1 Descriptive statistics from Stage 1 ...................................................................................................................61
Table 4.2 Descriptive statistics from scoring for Case 10 ..............................................................................................63
Table 4.3 Range of variance for each attribute ..............................................................................................................64
Table 4.4 The relationship between the research questions, Kane’s inferences and the presentation of findings ...........................................................................66
Table 4.5 Summary of findings ........................................................................................................................................96
Table 5.1 Kane's inferences of Generalisation and Interpretation ..................................................................................97
Table 5.2 Rules for criteria of merit modified from Scriven (1994) ..............................................................................98
Table 5.3 Overview of findings and modifications proposed .........................................................................................101
Table 5.4 Directions for future studies ..........................................................................................................................113
List of Figures

Figure 2.1 Number of supporting articles for the attributes of quality in the PDQI-9 and QNOTE..................................................................................................................24
Figure 3.1 Kane's framework adapted from Cook et al., (2015) ................................32
Figure 3.2 Process of study design ........................................................................39
Figure 3.3 Assessor medical education experience ..............................................45
Figure 3.4 Assessor specialty ..................................................................................45
Figure 3.5 Assessor years of experience with EHR ..............................................46
Figure 4.1 Overview of study design and analysis .................................................60
Figure 4.2 Scoring of Thorough across cases for each assessor ..........................61
Figure 4.3 Scoring of Up to Date across cases for each assessor .......................62
Figure 4.4 Assessor variance in scoring of Thorough across cases ...................67
Figure 4.5 Assessor variance in scoring of Synthesised across cases ..................70
Figure 4.6 Assessor variance in scoring Organised across cases .......................73
Figure 4.7 Assessor variance in scoring Comprehensible across cases ..............76
Figure 4.8 Assessor variance on scoring of Up to Date across cases ..................79
Figure 4.9 Assessor variance in scoring Useful across cases ...............................83
Figure 4.10 Assessor variance in scoring Succinct across cases ..........................87
Figure 4.11 Assessor variance of scoring Internally Consistent across cases ........92
### Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Admission Record</strong></td>
<td>A part of the medical record that documents the patient’s status, reasons why the patient is being admitted for inpatient care and the initial management plan.</td>
</tr>
<tr>
<td><strong>Clinical Phase</strong></td>
<td>Later stage of medical course and set within a health service.</td>
</tr>
<tr>
<td><strong>Content domain</strong></td>
<td>All observable attributes of a performance.</td>
</tr>
<tr>
<td><strong>Content Validity</strong></td>
<td>If the chosen observable attributes to assess, the test domain, are a fair representation of all observable attributes that could be assessed, the content domain.</td>
</tr>
<tr>
<td><strong>Content-related validity</strong></td>
<td>How the test and content domain are defined and in particular, the relationship between them. This encompasses both face and content validity.</td>
</tr>
<tr>
<td><strong>Criterion-related validity</strong></td>
<td>The performance of the assessment model based on the theory and definition of the competency, how well the assessment and test domain, lives up to its claims.</td>
</tr>
<tr>
<td><strong>Discharge Summary</strong></td>
<td>A synopsis of the patient’s admission to a hospital that provides pertinent information for the continuation of care following discharge.</td>
</tr>
<tr>
<td><strong>Electronic Health Record (EHR)</strong></td>
<td>An electronic version of a patient’s medical history that is maintained by health providers.</td>
</tr>
<tr>
<td><strong>Entry to Practice</strong></td>
<td>The educational, competency and experience-related qualifications required in a particular region before a person can practice a profession.</td>
</tr>
<tr>
<td><strong>Face Validity</strong></td>
<td>A personal judgment that a test appears to be a good measure of the phenomena to be assessed.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>First Year Clinical Medical Student</td>
<td>A medical student in their first year of placement in a health service, with learning occurring within the health care settings.</td>
</tr>
<tr>
<td>Medical Student</td>
<td>A student undertaking study that will allow them entry to practice medicine.</td>
</tr>
<tr>
<td>Patient Record Program</td>
<td>An online patient record program in Year 2 MD, University of Melbourne.</td>
</tr>
<tr>
<td>Physician (US)</td>
<td>A medical practitioner who has completed their vocational training and practices in specialised practice (US definition and used in this study).</td>
</tr>
<tr>
<td>Physician (Australian)</td>
<td>A medical practitioner who has completed their vocational training and practices in a specialised medical practice. This practice is within a medical discipline as opposed to, for example, surgery or primary health care.</td>
</tr>
<tr>
<td>Pre-clinical</td>
<td>Early stage of medical course focused on biomedical sciences with minimal patient contact; usually based at the University.</td>
</tr>
<tr>
<td>Pre-vocational training</td>
<td>The experience gained after completion of their basic medical qualification and prior to vocational training.</td>
</tr>
<tr>
<td>Progress Note</td>
<td>A part of the medical record that documents a patient’s current clinical status and any change in patient status since the last documentation, specifically indicating any significant and relevant changes and ongoing management plan.</td>
</tr>
<tr>
<td>Test domain</td>
<td>The chosen observable attributes to assess as a fair representation of all observable attributes that could be assessed on a performance.</td>
</tr>
<tr>
<td>Vocational medical training</td>
<td>The educational, competency and experience-related qualifications in a recognised medical specialty training program required in a particular region before a person can practice in a specialised medical practice.</td>
</tr>
</tbody>
</table>
# Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE</td>
<td>Alliance for Clinical Education</td>
</tr>
<tr>
<td>ETT</td>
<td>Educational Technology Team of University of Melbourne responsible for Patient Record Program</td>
</tr>
<tr>
<td>MD</td>
<td>Doctor of Medicine</td>
</tr>
<tr>
<td>PDQI</td>
<td>Physician Document Quality Instrument (22 item)</td>
</tr>
<tr>
<td>PDQI-9</td>
<td>Physician Document Quality Instrument (9 item)</td>
</tr>
<tr>
<td>RIME</td>
<td>Reporter-Interpreter-Manager-Expert framework</td>
</tr>
<tr>
<td>QNOTE</td>
<td>Quality of document instrument</td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
</tr>
</tbody>
</table>
1 Introduction

This thesis has explored the quality of EHR documentation of first-year clinical medical students. This exploration has served as a foundation for assessing the performance of EHR documentation in medical students. The first part of this chapter provides a context to EHRs and the importance of quality EHR documentation to patient care and safety, followed by a description of the purpose of the study and an overview of the thesis chapters.

1.1 EHR documentation

An EHR is an electronic version of a patient’s medical history that is maintained by health providers over time, and may include all of the key administrative clinical data relevant to that person’s care under a particular provider, including demographics, admission, progress and discharge notes, medications, investigations and all relevant documentation (Ellaway, Graves & Greene, 2013). In time it is envisioned that different EHRs platforms will be able to communicate between hospitals and primary and secondary healthcare. The EHR is a record held by the provider, not a patient-held record.

EHRs, and the documentation within it have multiple purposes. One of the key purposes of clinical documentation is to communicate between health care professionals to share or transfer the care and responsibility of a patient. The documentation provides a record of clinical decision making throughout the patient’s admission to hospital, and between different hospital admissions. Other purposes include providing easily accessible administrative data, health care statistics for research requirements and an information trail to evaluate and ensure best quality assurance practices (Häyrinen, Saranto & Nyänen, 2008). In medical training, EHRs have the potential to contribute to learning in a variety of methods. EHRs can play a role in logging learner encounters by automatically extracting data, providing clear demonstrations of both practitioners’ and students’ clinical decision making and for
research, auditing and reflective practice (Ellaway et al., 2013). This study specifically focused on the purpose of the documentation being for communication between health care professionals. However, any definition of quality in EHR documentation has to be mindful of the multiple purposes documentation is intended to meet; all definitions of quality must intersect and not be mutually exclusive.

With respect to clinical documentation for communication between health care professionals, this is a core skill of practice, and high-quality documentation is regarded as pivotal to improving patient outcomes (Häyrinen et al., 2008). All clinical documentation has the potential to be a high-risk scenario for patient safety (Cummings et al., 2010). Poor quality documentation directly impacts on patient care and safety, and can result in preventable adverse events (Bartlett, Blais, Tamblyn, Clermont & MacGibbon, 2008; Forster, Murff, Peterson, Gandhi & Bates, 2008; Witherington, Pirzada & Avery, 2008), poor continuity of care (Van Walraven, Oake, Jennings & Forster, 2008), avoidable re-admissions and inefficient health care practices (Kripalani et al., 2007).

EHRs have become an important feature of health care systems (Burke et al., 2015; Pageler, Friedman & Longhurst, 2013; Schiff & Bates, 2010; State of Victoria, Department of Health, 2014). The majority of health professionals internationally, and all within Australia will need to document within an EHR, or components of one, from the first day of their professional lives (Dattakumar, Gray, Henderson, Maeder & Chenery, 2012). The implementation of EHRs has been endorsed at multiple levels of jurisdiction, at the national level by health departments and locally by health services and practices, and continues to attract significant funding (Australian Government, DHS, 2009; Blumenthal & Tavenner, 2010; Elliott, Judd & McColl, 2011).

EHRs have the potential to enhance communication by addressing the deficits in handwritten documentation, in particular, legibility, that had adversely impact patient care (Embi et al. 2013; Shoolin, Ozeran, Hamann & Bria, 2013). Easier access to documentation, decreasing errors, particularly with medications, improving data
quality, legibility and timeliness of communication should all be facilitated by EHRs
(Kripalani et al., 2007; Van Walraven & Rokosh, 1999). EHRs are being linked to
improved caregiver decisions and patient outcomes (Blumenthal & Tavenner, 2010;
Burke et al., 2015; Pageler et al., 2013; State of Victoria, Department of Health, 2014).
However, to date, not all evaluations of EHR documentation have been favourable and
there continues to be concerns about the quality of documentation impacting on
patient care (Callen, McIntosh & Li, 2010; Jansen & Grant, 2003; O’Leary et al., 2009;
mean easy transition.’ (p.501)

Many factors may potentially have contributed to EHRs not living up to expectations.
These include the competing purposes of EHR documentation, information technology
platforms that are not intuitive or meeting the needs of the context they are intended
to serve, to insufficient training in EHR use, insufficient education of documentation
and understanding what constitutes high quality documentation (Callen, Alderton &
McIntosh, 2008; Jansen & Grant, 2003; Orrico, 2008; Uitvlugt, Siegert, Janssen, Nijpels
& Karapinar-Carkit, 2015). There is less emphasis on legibility, and new concerns
emerging with the capacity of a digital platform to allow ‘cut and paste’ and the easy
inclusion of a vast number of reports in documentation. This has led to concerns about
a lack of originality and increasing redundant information that decreases the
readability and functional usability of the document. These demands all need to be
balanced and prioritised when teaching and assessing quality EHR documentation.

1.2 EHR and medical education

All future doctors will have to document in EHRs. Despite the widespread
implementation of EHRs, and the concerns raised on the quality of EHRs and the
variable impact of EHR on patient safety, there are few studies that focus on the
learning or assessment of EHR documentation, or use, in medical education (Tierney,
Pageler, Kahana & Longhurst, 2013; Ellaway et al., 2013; Otto & Kushniruk, 2009).
Wald, George, Reis and Taylor (2014), Dattakumar et al. (2012), Otto and Kushniruk
Schenarts and Schenarts (2011) and Strauss (2010) all agree that health informatics skills should become a core competency of medical curricula; however, the pace of education in health care information technology is not aligned with the rapid expansion of EHR in health service environments (Wald et al., 2014; McGowan, Cusack & Poon, 2008). The Alliance of Clinical Educators (ACE) have made key recommendations that students’ use of EHR be restricted until they can demonstrate competence in the documentation of ‘essential elements’ (Hammoud et al., 2012) Given the significance EHRs will have in the future health care of patients, it is imperative that the learning and assessment of EHR skills in medical student education are addressed.

1.3 Purpose of the Study

The purpose of our study was to explore the attributes of quality of EHR documentation written by first-year clinical medical students. This exploration built on existing literature to define the attributes of quality in EHR documentation as a foundation for assessing and providing feedback on the performance of documentation to medical students. Assessment is central to medical education, and the valid creation of the assessment items for performance fundamental (Cook, Brydges, Ginsburg & Hatala, 2015). Kane’s conceptual framework for an argument-based approach to validity was utilised to provide a structure for this study and guide the collection and interpretation of validity evidence (Kane, 2006).

A recently validated, 9-item instrument for assessing the quality of physician EHR documentation, the Physician Document Quality Instrument (PDQI-9) was used as the platform to build an understanding of the attributes of medical student EHR documentation (Stetson, Bakken, Wrenn & Siegler, 2011). The findings of the literature review provided a justification for this selection, in terms of the attributes identified in the medical student literature, the primary focus, and purpose of the documentation and current medical education literature on assessment.
Building upon recently published validated instruments, the research questions for this study were:

1. Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students?
2. What are the quality attributes of EHR documentation by medical students in their first clinical year?
3. What modifications to the PDQI-9 could be made to improve validity in this setting?

The study design triangulated methods to provide a robust validity argument. By exploring the application of the PDQI-9 to EHR records written by first clinical year medical students, numerical, narrative and spoken data was obtained from assessors with credibility to provide the requested insights of quality in EHR documentation and detailed descriptions of observed behaviours, for the specific level of the student. Data was collected until thematic saturation and consensus was reached and there was a coherence of the final interpretations supporting a content basis for the assessment process in medical students.

1.4 Overview of Thesis

The thesis is divided into five chapters. The context and importance of the study and aim have been discussed in this initial chapter. Chapter Two provides a critical review of the key literature on the quality attributes of medical student EHR documentation. Chapter Three details the Methodology and the conceptual framework of the study, the method, and rationale for the design. Chapter Four presents the analysis of the data, integrating the findings from the study and the thesis concludes with Chapter Five and a discussion of the conclusion, the key findings, limitations of the study and recommendations for future directions in research.
2 Literature Review

While there is a large body of literature on both medical student and physician hand-written documentation skills and assessment, we made the decision to focus our review on studies specifically addressing electronic health record documentation to keep the analysis focused on issues specifically related to this emerging mode of communication. Therefore, this chapter examines the literature on the attributes of quality EHR documentation written by medical students, salient discourses that directly impact the assessment of quality in medical students and the associations between the two.

The chapter will be divided into the following sections.

- Attributes of quality medical student EHR documentation
- Assessing quality of documentation in physicians
- Contention between narrative expressivity, EHR platforms, and the impact this has on the definition of quality
- Medical education assessment literature that directly impacts on the process of assessment
- Based on these findings, the review concludes with a justification for the selection of the PDQI-9 as a platform for this study.

To attempt to identify all significant articles, and provide a comprehensive review of the literature, two searches were conducted using multiple terms. The first search attempted to retrieve articles with terms synonymous with electronic health records, written communication skills, documentation skills and medical student. The search was broadened to include written communication, electronic health records, assessment, instrument, and validation. The electronic databases CINAHL, Medline (Web of Knowledge), Scopus, Web of Science, PsycINFO, Medline (Ovid), Medline (ISI), EMBASE (Ovid), ERIC and proprietary search engines were searched. Articles applicable to this review frequently did not share keywords, and variations of keywords were
often identified in consensus and expert opinion articles. Hand searching was also performed by reviewing references of retrieved articles.

Nearly all the literature on EHR documentation is from the United States (US). It is assumed that the findings on attributes are generalisable internationally, although there will be different local logistical problems. The delay in addressing this growing area of medical practice and education has been noted by authors in both Australia and Canada (Dattakumar et al., 2012; Strauss, 2010).

## 2.1 Attributes of quality medical student EHR documentation

The first section of this chapter provides the analysis of the existing literature to discover what quality attributes are expected in medical student EHR documentation. This section steps through the discussion by examining the original studies in medical student EHR documentation that provided some insight into attributes of quality, followed by the findings inferred from the evaluation of innovative simulated EHR programs in medical courses, and reviews on the impact of EHRs on medical education. Relevant findings from published expert opinions on medical student EHR documentation conclude the section.

### 2.1.1 Original studies

Our literature search identified no studies that specifically addressed the quality of medical student EHR documentation, nor any readily identifiable validated measures of assessment in this context. We have summarised the five studies that made a reference to attributes of quality documentation in analysing or assessing medical student written content in an EHR. Either their method or findings implied expected qualities in student EHR documentation.

Heiman et al. (2014) surveyed 116 third year medical students about their EHR documentation practices. The purpose of this study was to report on the prevalence of copying and pasting into medical records from their own previous documents and those of others. The findings were that most students (95%) reported copying from
their own previous notes, and 22% reported copying from residents. 10% indicated that copying from others was an acceptable practice. Students expressed the wish to be taught responsible electronic documentation. In conclusion, as well as the need for students to receive feedback on responsible documentation, the authors stated that there should be a focus on documentation that was concise, well-reasoned and updated patient assessments, as well as an effort to minimise redundancy and emphasise originality and accuracy.

Rouf, Chumley and Dobbie (2008) surveyed 33 third year medical students on how the use of EHRs impacted on their learning. Considering qualities of documentation, the students liked that the EHRs assisted in the organisation of content (70%), and the ease of accessing essential information. Most students (69%) reported that the EHR improved their documentation and that they received more feedback on their documentation. There was little definition on what was meant by improving documentation, the few qualitative comments made used the terms Complete and Organised.

A frequently cited study by Szauter, Ainsworth, Holden and Mercado (2006) compared students undertaking examinations on video to the student documentation of the encounter. Only 4% of the documentation completely and accurately reported what was witnessed to occur in the student-patient encounter. The study concluded the need to teach and assess documentation with an emphasis on accuracy and honesty.

Usatine, Landlay and Posalski (2002) designed a small observational study to test the hypothesis that medical students’ chart documents were as good as physicians’. The main purpose of this study was to justify the inclusion of medical students’ documentation in hard copy patient records or EHRs. The researchers reviewed both electronic and written documentation. Patient encounters of both physicians and medical students and the subsequent EHR documentation were evaluated using explicit criteria developed by the study team. The attributes; Comprehensible, Complete, and Accurate were all deemed to be important by the researchers. The
criteria included identification of the diagnosis and a listed differential diagnosis, implying importance on clinical reasoning and synthesis of the patient encounter. Overall the students’ documents were more thorough and comprehensible than physician documentation but rated lower on synthesis.

Similar in purpose to the study by Usatine et al. (2002), Ferenchick et al. (2013) performed a study to determine medical students’ readiness to document into a real EHR-based. The criteria for readiness to document in the real setting was based on the number of errors the student made in documentation into a simulated EHR program. The study included 222 third year medical students on internal medicine rotations and their documentation into Michigan State University’s simulated EHR program. Thirty percent of the students made at least one error and 13 % had more than one error. The researchers believed that lack of attention to detail or lack of conscientiousness directly contributed to the rate of errors. Ferenchick et al. concluded that students’ readiness to accomplish tasks successfully within an EHR was undetermined. The relevance to this review was the importance placed on accurate documentation and the researchers’ interpretation that errors were likely to be made when documentation was not thorough and complete.

In summary, from original studies with a focus on an aspect of medical student EHR documentation, the attributes of Complete, Concise, Organised, Comprehensible, Synthesised or Well-reasoned, Updated, with minimal Redundancy, Original, and Accurate were referenced as being desirable attributes of quality medical student EHR documentation. Accurate was mentioned by four of the studies. Complete, or Thorough, was referred to by three of the studies and Organised, Synthesised documentation were criteria of assessment in at least two of the studies.

2.1.2 Evaluations of simulated EHR programs
Our literature search identified two papers on EHR curriculum initiatives that detailed attributes of quality medical student EHR documentation in an aspect of the
evaluation or review (Milano, Hardman, Plesiu, Rdesinski & Biagioli, 2014; Wagner, Roskos, Demuth & Mavis, 2010).

Wagner, Roskos, Demuth and Mavis (2010) described grading the documents submitted in a simulated EHR at Michigan State University. The researchers assessed the documents based on an explicit competency rubric, with grading attributes for each section of the record. The focus of the paper was the evaluation of the program and the assessment process was not detailed. From the elements of the rubric outlined in the report, important attributes included Thorough, Accurate and Timelines.

Milano, Hardman, Plesiu, Rdesinski and Biagioli (2014) reported on Oregon University’s simulated EHR program. A three-week simulated EHR program was part of the third year family medicine rotation curriculum. The focus of the article was the development and implementation of the program, rather than the assessment of the documentation. The researchers assessed students’ chart work using a rubric tied to the curriculum’s general and specific objectives, which were in line with the Accreditation Council for Graduate Medical Education (ACGME) competencies. The rubric assessed that the student demonstrated skills in record maintenance and demonstrated an ability to order prescriptions, laboratory tests, vaccines, imaging studies and document consultations. Assessment measures focused on Thorough, Complete, Accurate (or correction of inaccuracies), Evidence-based and Prioritised.

The inferences made from the reports of simulated EHR programs on the attributes of quality documentation support those from other studies. Accurate and Complete, or Thorough, documentation was regarded as important attributes by both reports. Organised documentation with respect to timelines, and Evidence-based and Prioritised were additional qualities referenced to. Many studies focused on the absence of omissions, inconsistencies, time delays, and errors. Nearly all rubrics in the above studies, focused on content rather than quality. The attributes and supporting original studies and simulated EHR reports are summarised in Table 2.1.
<table>
<thead>
<tr>
<th>Attribute of quality medical student EHR documentation</th>
<th>Supporting Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accurate</strong></td>
<td>Seven (7) supporting studies</td>
</tr>
<tr>
<td><strong>Complete or Thorough</strong></td>
<td>Five (5) supporting studies</td>
</tr>
<tr>
<td><strong>Synthesised or Well-reasoned</strong></td>
<td>Two (2) supporting studies</td>
</tr>
<tr>
<td></td>
<td>Usatine (2002), Heiman (2014)</td>
</tr>
<tr>
<td><strong>Comprehensible</strong></td>
<td>One (1) supporting study; Usatine (2002)</td>
</tr>
<tr>
<td><strong>Organised</strong></td>
<td>One (1) supporting study; Wagner (2010) {timelines}</td>
</tr>
<tr>
<td><strong>Concise</strong></td>
<td>One (1) supporting study; Heiman (2014)</td>
</tr>
<tr>
<td><strong>Updated</strong></td>
<td>One (1) supporting study; Heiman (2014)</td>
</tr>
<tr>
<td><strong>Evidence-based</strong></td>
<td>One (1) supporting study; Milano (2014)</td>
</tr>
<tr>
<td><strong>Prioritised</strong></td>
<td>One (1) supporting study; Milano (2014)</td>
</tr>
</tbody>
</table>

### 2.1.3 Expert Opinion

Our literature search identified three published expert opinions on EHRs and medical education that referenced the learning and assessment of documentation skills (Ellaway et al., 2013; Hammoud et al., 2012; Stephens, Gimbel, & Pangaro, 2011). An overview of these papers is provided in the following section. The impact of EHR on medical education was the focus of a further four opinion pieces identified. The authors did not specifically address quality EHR documentation or documentation skills, and any attributes that can be discerned from their deliberations are collectively discussed at the end of this section.
Expert opinion on EHR and medical education

Stephens, Gimbel and Pangaro (2011) demonstrated that their Reporter-Interpreter-Manager-Expert (RIME) framework could be systematically applied to teach medical students how to document, by aligning the clinical skills and EHR specific skills at each stage of development. They then outlined how this model could be used to track expectations at various stages of the course. Stephens et al. (2011) did not differentiate between recommended EHR content, attributes of quality documentation and the appropriate use of the EHR in his developmental framework. To make it clearer to differentiate between these three elements of the EHR, for each developmental stage, we have separated their expectations into attributes of quality, expected EHR content and use in Table 2.2 below.

The attributes valued were Accurate, Synthesised or Interpretation, Complete or Thorough, Originality, Comprehensible (semantic competence), Relevant, Clear, Logical and Prioritised. The priority for Stephens et al. (2011) was the development of clinical reasoning and interpretation of the patient encounter. In addition, it was recommended that any EHR assessment should include feedback on Professionalism. Students should demonstrate respect for the confidentiality of patient information in the context of publically situated computers and take ownership of the information they document. Their documentation should be original and independently acquired, and not ‘cut and pasted’ in.
Table 2.2 RIME framework and corresponding attributes of quality, EHR content and use

<table>
<thead>
<tr>
<th>Student Level of Development</th>
<th>EHR quality attribute</th>
<th>EHR use</th>
<th>EHR content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reporter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• (Approximately first to second clinical year)</td>
<td>• <em>Accurate</em></td>
<td>Data entry</td>
<td>• Basic information on history and exam</td>
</tr>
<tr>
<td></td>
<td>• <em>Clear</em></td>
<td></td>
<td>• Uses appropriate clinical language</td>
</tr>
<tr>
<td></td>
<td>• <em>Synthesised</em> - answers the what questions about patient care</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <em>Complete</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <em>Comprehensible</em> – semantic competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <em>Originality</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <em>Professionalism</em> – respect patient confidentiality and privacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interpreter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• (Approximately Second and third clinical year)</td>
<td>• <em>Prioritised</em></td>
<td>Date assessment</td>
<td>• Records clinical assessment</td>
</tr>
<tr>
<td></td>
<td>• <em>Synthesised or interprets</em> - addresses the why questions</td>
<td></td>
<td>• Constructs a differential diagnosis</td>
</tr>
<tr>
<td><strong>Manager</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• (Approximately third clinical year to second post graduate year)</td>
<td>• <em>Synthesised</em> - answers how do we solve this</td>
<td>Data assimilation</td>
<td>• Records management options</td>
</tr>
<tr>
<td></td>
<td>• <em>Logical</em></td>
<td></td>
<td>• Recommends appropriate diagnostic and treatment options</td>
</tr>
<tr>
<td></td>
<td>• <em>Comprehensible</em> - semantic competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Educator</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• (In practice)</td>
<td>• <em>Evidence-based</em>; modifies plan with evidence</td>
<td>Clinical Decision Support</td>
<td>• Records a patient-centered plan</td>
</tr>
<tr>
<td></td>
<td>• <em>Synthesised</em></td>
<td></td>
<td>1.</td>
</tr>
</tbody>
</table>

Ellaway et al. (2013) wrote a review on medical education with the increasing application of EHRs. They added their voices to the growing number of academics calling for learners to be provided with guidance in developing EHR skills and commented that this cannot be left to ‘ambient’ exposure. With respect to attributes
of quality documentation and the emphasis on learning, the authors supported Stephens’ et al. (2011) RIME framework.

The Alliance for Clinical Education (ACE) issued a collaborative statement on expectations of medical student documentation in EHRs for educators (Hammoud et al., 2012) This statement was issued in response to growing concerns about the lack of student access to EHRs in the clinical environment and the concern about medical students learning the skills of documentation with limited exposure. (Gliatto, Masters & Karani, 2009; Usatine et al., 2002) Given the change in documentation across the US, ACE were adamant that ‘schools must develop a set of medical competencies related to charting in the EHR and state how they would evaluate it’ (p.257) and that EHR documentation be integrated into medical student curriculum. ACE valued Accurate, Complete, Succinct and Organised documentation. Recommendations included the need to assess all types of documents, and that the format of the document as well as content needed to be evaluated. Similar to Stephens et al. (2011) and Heiman et al. (2014), ACE emphasised the need to reduce ‘cut and paste’, and recommended that documentation be used to evaluate professionalism, honesty, and respect for patients. The guide referred to RIME as a possible developmental framework for documentation skills.

*Expert opinion on the impact of EHR on medical education*

Five review articles reported on the impact of the EHR on education, both medical student and resident, and made comments on EHR documentation skills and qualities. (Gliatto et al., 2009; Mintz et al., 2009; Schenarts, P.J & Schenarts, 2012; Tierney et al., 2013; Wald et al., 2014). Uniformly they agreed that medical students were expected to develop effective written communication skills during undergraduate education and that EHRs had the potential to enhance and impair this development.

The literature on quality learning in the area of medical student documentation was sparse in all reports. Quality attributes that were mentioned included: Accurate, Comprehensive, Complete, Utility, Originality, Prioritised, Synthesised and Integrated
clinical information. Similar to the RIME model, there was a very strong emphasis on the written demonstration of clinical reasoning. Tierney et al. (2013), writing on the benefits, challenges and future directions of medical education in the EHR era, stated that one potential benefit of EHR documentation was that if effective, students should be able to spend less time gathering information and more time synthesising clinical data. Gliatto, Masters and Karani (2009) agreed that through the process of writing, students can organise their thinking, and express their reasoning processes. Interpretation of clinical information was seen as essential to the effective exchange of information to other health professionals that enabled and did not jeopardise the safe care of patients.

The reviewers expressed considerable concern about the negative impact EHR documentation may have on the originality of the document. Within the EHR system, it was possible to indiscriminately ‘copy and paste’ from any other documentation within the record with mindless repetition resulting in deteriorating quality (Shoolin et al., 2013). ‘Copy and paste’ habits needed to be specifically addressed and how this practice impacted on professionalism, respect, and honesty, as well as decreasing the accuracy and utility of the note. They all believed that there was the potential for EHR documentation to decrease the expression of clinical reasoning if care was not taken to stress its importance in training and reinforce this in any assessment of documentation.

In summary, the attributes supported by expert opinion on medical student documentation and reviews on the impact of EHR on medical education are collated and compared to the number of supporting studies in the original or evaluation studies in Table 2.3. Of particular emphases was the importance of clinical reasoning, the risk to this attribute by the functionality of the EHR system, and the importance of identifying ‘copy and paste’ habits, as well as commenting on the Professionalism of the documentation.
Table 2.3 Quality attributes from expert opinion and impact studies

<table>
<thead>
<tr>
<th>Attribute of Quality EHR documentation</th>
<th>Number of Supporting Expert Opinions and Impact Reviews</th>
<th>Number of Supporting Original studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurate</td>
<td>Six (6)</td>
<td>Seven (7)</td>
</tr>
<tr>
<td>Complete or Thorough</td>
<td>Six (6)</td>
<td>Six (6)</td>
</tr>
<tr>
<td>Originality or Redundancy (copy and paste)</td>
<td>Six (6)</td>
<td>One (1)</td>
</tr>
<tr>
<td>Professionalism</td>
<td>Six (6)</td>
<td>-</td>
</tr>
<tr>
<td>Prioritised</td>
<td>Five (5)</td>
<td>One (1)</td>
</tr>
<tr>
<td>Synthesised or well-reasoned</td>
<td>Five (5)</td>
<td>Two (2)</td>
</tr>
<tr>
<td>Utility</td>
<td>Four (4)</td>
<td>-</td>
</tr>
<tr>
<td>Succinct</td>
<td>One (1)</td>
<td>Concise – One (1)</td>
</tr>
<tr>
<td>Organised</td>
<td>One (1)</td>
<td>One (1)</td>
</tr>
<tr>
<td>Comprehensible</td>
<td>One (1)</td>
<td>One (1)</td>
</tr>
<tr>
<td>Relevant</td>
<td>One (1)</td>
<td>-</td>
</tr>
<tr>
<td>Evidence-based</td>
<td>One (1)</td>
<td>One (1)</td>
</tr>
<tr>
<td>Clear</td>
<td>One (1)</td>
<td>-</td>
</tr>
<tr>
<td>Logical</td>
<td>One (1) study</td>
<td>-</td>
</tr>
</tbody>
</table>

The literature on EHR and medical student documentation was limited. With inconclusive findings from the medical student literature, the search was broadened to original studies exploring the assessment of attributes of quality in physician EHR documentation. The findings are discussed in the next section.

2.2 Assessing quality of EHR documentation in physicians.

The findings from the literature review on the attributes of quality EHR documentation in medical student documentation has relied heavily on expert opinion and review articles. Only the two expert opinion pieces by Stephens et al. (2009) and ACE (Hammoud et al., 2012) directly commented on attributes and expectations of medical student EHR documentation. All other findings have been inferred.

Despite the importance and pivotal role of documentation, there is no agreed upon definition of quality documentation for physicians. Assessing the quality of documentation was an arduous task prior to EHR, requiring medical records to be manually sourced, with significant issues with legibility and labor-intensive review.
These factors may have all delayed research into this area (Embi et al., 2013). With the advent of EHRs, there has been increased interest in defining quality, with easier access to legible physician documentation. Two validated instruments for assessing quality physician EHR documentation were identified (Burke et al., 2014; Stetson et al., 2012) that endeavored to define quality EHR documentation in physicians and based on their findings, validate an instrument to assess quality EHR documentation. The studies and how the researchers endeavored to define quality physician EHR documentation, and from this determine the observable attributes to assess, namely the test domain, as a fair representation of all observable attributes that could be assessed in their definition, the content domain, and subsequently validation of the resulting instruments is described in the following section.

2.2.1 Physician Document Quality Instrument
Stetson, Morrison, Bakken and Johnson (2008) set about creating and validating an instrument to assess quality of the EHR documentation of physicians. This was a two stage process and the instrument was solely focused on assessing the quality of the documentation, not the content or the use of the EHR platform. The purpose of the documentation in this study was to support clinical communication with other health care providers engaged in the care of the patient. The researchers initially reported on the design and validation of the 22 item Physician Documentation Quality Instrument (PDQI) in 2008, and the refining and validating of a modified nine-item instrument in 2012, the PDQI-9 (Stetson et al., 2008; Stetson, Bakken, Wrenn & Siegler, 2012).

In the preliminary study, Stetson et al. (2008) analysed the findings from a literature review searching for adjectives describing quality in EHR documentation. Opinions were provided on the relevance of the resulting attributes from seven clinical experts. Twenty-five factors were identified from the literature review, and a further five were added by the clinical experts. In an exploratory factor analysis of the attributes, 63 medical professionals were surveyed and asked to grade each attribute on how well it described a high-quality document. Of the 30 attributes, an instrument with 22
adjectives was developed. After analysis, the adjectives aligned to four factors for documentation quality, outlined in Table 2.4.

Table 2.4 PDQI exploratory factor analysis

<table>
<thead>
<tr>
<th>Factor</th>
<th>Interpretation</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor I. Well-formed</td>
<td>The content is being presented in a way that contains a logical structure leading to a reasonable conclusion.</td>
<td>Clear, Uncluttered, Organised, structure, Non-redundant, Synthesised</td>
</tr>
<tr>
<td>Factor II. Comprehensible</td>
<td>The note transferred information with meaning and comprehension.</td>
<td>Legible, Coherent, Useful, Correct, Comprehensible and Consistent</td>
</tr>
<tr>
<td>Factor III. Accurate</td>
<td>The extent to which the document reflects the true state of the patient</td>
<td>Up to date, Complete, Accurate, Thorough, Current, Relevant</td>
</tr>
<tr>
<td>Factor IV. Compact</td>
<td>The note has appropriate density of information within the given content</td>
<td>Brief, Concise, Succinct, Focused</td>
</tr>
</tbody>
</table>

An initial 22-item assessment was created and in their initial study, Stetson et al. (2008) provided preliminary evidence of its content validity. The authors viewed the factors derived from this study to be inclusive of both the form and content of documents. Factor III. Accurate, focused on content; Factor I. Well-formed and Factor IV. Compact referred to how the note was composed, the form. Factor II. Comprehensible was both content and form-based. Stetson et al. (2008) interpreted this as meaning that factors describing both content and form were needed to make the distinction between the quality of data and the quality of documents, which are collections of data. They concluded that the results extended any previous work on data quality, ‘through validation of form as a key constituent of documentation quality’ (p.540).

In 2012, the researchers extended their prior work, by refining the PDQI and testing the validity of a 9 item version, the PDQI-9 (Stetson et al., 2012). In this study, the instrument was modified by further exploratory factor analysis to eliminate items that were not relevant to all EHR documents and when removed did not impact on the internal consistency of the original instrument, and were judged by an expert in
documentation quality to be redundant. Three sets of documents, including admission records, progress notes and discharge summaries, from EHRs were evaluated by two groups of physicians using the PDQI-9. Attributes were scored on a Likert scale, from 1 to 5, from not at all evident to extremely evident. Assessors were asked to make an overall assessment of the document in addition to the PDQI-9.

The results supported the validity of PDQI-9, a simpler-to-use tool, for rating the quality of EHR physician documents. The PDQI-9 score was able to discriminate between satisfactory and unsatisfactory as judged in the overall assessment of the document. From narrative comments, no assessor had difficulty using or interpreting the PDQI-9. The following nine quality attributes were validated: Up to Date, Accurate, Thorough, Useful, Organised, Comprehensible, Succinct, and Synthesised. The descriptions of the validated attributes are in Table 2. 5 and a copy of the PDQI-9 is Appendix. A.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description of Ideal Document</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to Date</td>
<td>The document contains the most recent results and recommendations</td>
</tr>
<tr>
<td>Accurate</td>
<td>The document is true. It is free of incorrect information</td>
</tr>
<tr>
<td>Thorough</td>
<td>The document is complete and documents all the issues of the importance to the patient</td>
</tr>
<tr>
<td>Useful</td>
<td>The document is extremely relevant, providing valuable information and/or analysis</td>
</tr>
<tr>
<td>Organised</td>
<td>The document is well-formed and structured in a way that helps the reader understand the patient’s clinical course</td>
</tr>
<tr>
<td>Comprehensible</td>
<td>The document is clear, without ambiguity or sections that are difficult to understand</td>
</tr>
<tr>
<td>Succinct</td>
<td>The document is brief, to the point, and without redundancy</td>
</tr>
<tr>
<td>Synthesised</td>
<td>The document reflects the author’s understanding of the patient’s status and ability to develop a plan of care</td>
</tr>
<tr>
<td>Internally consistent</td>
<td>No part of the document ignores or contradicts any other part.</td>
</tr>
</tbody>
</table>
The researchers commented that there was no item that correlated with the extent of ‘copy and paste’ and originality of the documentation that was a recurrent concern in findings from studies into EHR documentation. This was not picked up by the inclusion of Redundancy in the description of the attribute Succinct. They recommended that a separate Redundancy score, or an overall assessment on Originality, be made in conjunction with the PDQI-9.

In summary, the PDQI-9 is a validated global instrument that makes an assessment on the entire clinical document. It is intended to assess the quality of EHR physician documentation, it is not intended to assess data content and the presence or absence of specific document components. The authors stated one purpose of this instrument was to create a tool that could be used for teaching purposes with medical students and residents.

The attributes Synthesised, Accurate, Thorough, Comprehensible, Useful and Organised all correlate with attributes identified and supported in the previous literature. Addressing the format of the document, and the expression of clinical reasoning through the attribute Synthesised matches the emphasis that the expert opinions have placed on medical student documentation. Omitted attributes in the PDQI-9, that were identified in the medical education literature, are the lack of assessment of originality, as declared by the authors, and direct feedback to the student on the professionalism of the document.

2.2.2 QNOTE
In 2012, a second instrument was validated assessing quality in EHR documentation, with an alternative interpretation of quality, focusing on the document elements rather than the document (Burke et al., 2014). Similar to the PDQI-9, the resulting instrument, the QNOTE, was produced with a two stage process. A large qualitative study was undertaken to explore the quality of outpatient clinical notes from multiple stakeholders (Hanson, Stephens, Pangaro & Gimbel, 2012). With the findings from the exploration used as the content domain for the performance of quality
documentation, the QNOTE was developed and validated in the second study (Burke et al., 2014). The purpose of the documentation in this study was for communication and to support patient safety, medical education, medico-legal investigations, and reimbursement.

The aim of the initial qualitative exploration was to develop a comprehensive definition of quality in outpatient clinical notes, that encompassed both EHR and handwritten documentation (Hanson et al., 2012). There were 163 participants, across four stakeholder groups, including clinicians, nursing staff, patients and administration. The majority of participants were working with EHR systems. Qualitative methods included both focus groups and individual interviews, and the researchers reported their findings in three organising themes.

1. The characteristics, or attributes, of quality in clinical notes,
2. The desired elements within the clinical notes and
3. The system supports to improve the quality of notes.

With respect to the findings on attributes of quality in clinical documentation, the study found that the document needed to contain sufficient information for the reader to understand its rationale and tell a coherent and continuous story. Eleven codes were identified to describe the attributes of quality: Concise, Sufficient, Explanatory, Relevant, Continuity of Story, Clear, Prioritised, Readable, Organised, Current and Accurate, and Ease of Translation into codes. There was substantial overlap on the attributes of quality across all stakeholder groups. These were very similar attributes to the Stetson et al. (2008) study.

The preliminary study detailed the importance of the integration of insights and care across disciplines, and incorporation of patient priorities (Hanson et al., 2012). Within 11 attributes, the additional attributes of Continuity of Story and Explanatory, were included that encompassed this integrated view. The researchers related Continuity of Story to narrative expressivity, and the highest quality documents, in the context of a
complete patient record, were deemed as those that made clear the connections between different encounters. A single document represented only one chapter in an evolving health care story. Twenty codes were identified under the theme of desired elements, and these corresponded with document components, for example, patient complaint, problem list, medications and past medical history.

With a content domain derived from the Hanson et al.’s (2012) findings, Burke et al. (2014) developed and validated an instrument to measure the quality of EHR documentation, the QNOTE. The researchers differentiated the QNOTE from the PDQI-9 by creating a non-global instrument for measuring the quality of a clinical note. The components of the QNOTE related the themes of desired elements and attributes of quality from the preliminary study. The reporting was unclear about the method used to select the desired elements to assess, and the corresponding attributes.

The QNOTE is a checklist of 12 document elements from the 20 reported in the original study. Each document element is assessed on 3 or 4 of the attributes of quality, out of the original 11. The attributes that were assessed for each element included 3 of Clear, Complete, Concise, Current (Up to Date), Organised, Prioritised and Sufficient information. The document’s quality is a combination of its element scores. The resulting QNOTE is a checklist instrument that focuses on content with over 36 points of assessment. Despite the Continuity of the Story and Explanatory being related to the highest quality documents in the qualitative study, these attributes were not included in the QNOTE.

The QNOTE was validated by randomly assigning 100 ambulatory notes to eight general internal medicine and eight family medicine practicing physicians to assess, using a variety of assessment processes (Burke et al., 2014). Burke et al. (2014) believed strongly that it was the combination of the elements that formed the basis of quality. A high-quality document contained the necessary and significant elements required to achieve the main goal to support patient care and improve outcomes. Each clinical element had its own set of quality characteristics, and that it was not possible
to combine these to a global assessment. The researchers were directly contrasting the QNOTE to the PDQI-9 in this discussion. A copy of the QNOTE is in Appendix B.

2.2.3 Summary

The PDQI-9 and QNOTE provide validated measures of assessment of quality EHR documentation in physicians. There has been considerable work conducted to produce the QNOTE and PDQI-9, and define the content domain for quality in physician EHR documentation. There is also extensive overlap in the observable attributes that describe this performance.

The instruments have taken considerably different approaches in both the focus of their interpretation of quality and in the format of the rubric. The QNOTE focuses on data and document elements and the PDQI-9 assesses the entire document. The PDQI-9 is a global assessment of quality relying on the judgment of the assessors. The QNOTE is a 36 item checklist and is non-global. The authors of the QNOTE regard their instrument as being more objective, and the PDQI-9 as a subjective assessment.

Similar attributes, with significant overlap, were identified, with both studies finding that the synthesised document was essential to a high-quality document. Five of the nine attributes were synonymously described; Thorough and Complete, Succinct and Concise, Comprehensible and Understandable, Organised, and Up to Date. Figure 2.1 visually demonstrates the number of articles from the medical student literature review that support the attribute of the PDQI-9 and QNOTE.
There are some difficulties with the interpretation and definition of the concept of *Succinct*. Expressions used are succinct, concise and redundant information. In addition, the literature strongly supported any assessment of *Succinct*, and redundant information in documentation, to encompass feedback on the *Originality* of the documentation. Stetson et al. (2009) noted that the PDQI-9 did not successfully capture *Redundancy*, in terms of *Originality*, with this term.

Attributes in the medical student literature review that were not identified in either the PDQI-9 or QNOTE were *Originality*, *Professionalism*, and there was support from one source for both *Logical* and *Evidence based*. The PDQI-9 has one attribute, *Internally Consistent* that had no support from either QNOTE or the literature review.

The purpose of any study is to build upon existing knowledge. In considering the suitability of the PDQI-9 or QNOTE for assessing medical student EHR documentation the literature was reviewed on the tension between the different interpretations of
quality and the relationship the findings have with the medical student literature review, and secondly the current literature on medical education on performance-based assessment. The findings are outlined in the following sections.

2.3 Interpretations of quality

The competing emphases of the PDQI-9 and QNOTE illustrated alternative interpretations of quality, the form or the content, respectively (Burke et al., 2014; Stetson et al., 2012). With respect to the PDQI-9, this distinction was more obviously illustrated by the findings from a cross-sectional study assessing physician document quality using the PDQI-9 and comparing this to disease-specific clinical quality scores based on specific document elements (Edwards, Meri, Volk, Schiff & Bates, 2013). The researchers found that the PDQI-9, as a subjective assessment of document quality did not correlate with EHR extracted disease-specific quality indicators. They noted that documents with clear narrative that appeared organised and easy to read could often lack key quality data, while notes with automatically imported data and copy and paste elements contained more quality indicators but were often long, less organised, more difficult to read, leading to lower subjective scores. Documents that scored highly on *Useful*, *Synthesised* and *Organised* also scored high on an overall general impression scores. It was concluded that subjective note quality was significantly impacted on by exhaustive documentation of quality markers, and documents were considered of higher quality if there was more evidence of synthesis, organisation and clinical usefulness of the document. The presence of document elements alone was not sufficient to create a high-quality document, that health care professionals found useful.

2.3.1 Physician preference for quality in documentation

From the Edwards et al (2013) study and the findings from our literature review on quality medical student EHR documentation, the importance that physicians place on documentation being synthesised, and after reading the document, the author’s reasoning, thought processes and uncertainties are understood, was obvious
(Rosenbloom, Crow, Blackford & Johnson, 2007). The EHR has had unintended consequences on documentation. EHR platforms preference the entry of document elements rather than a cohesive collection of information and allow indiscriminate copying and pasting, and repetition of outdated or incorrect information. As mentioned in the introduction to this thesis, the evaluation studies of EHRs have produced mixed results, and one of the reasons proposed for these findings is the tension between the structure of EHR platforms favoring documentation of elements and the preferred documentation practices of physicians favoring form (Mamykina, Vawdrey, Stetson, Zheng & Hripcsak, 2012; Rosenbloom et al., 2011).

Rosenbloom, Crow, Blackford and Johnson (2007) conducted a qualitative study exploring cognitive factors that influence perceptions of computer-based clinical documentation tools. Based on in-depth interviews with 16 health care providers, including both physicians and nurses, five factors were identified that influenced satisfaction: document system efficiency, accessibility, expressivity, quality and patient care quality. Quality was defined similarly to our findings as Thorough, Accurate, Legible and Compliant with administrative documentation standards. Expressivity was a key factor impacting satisfaction that emerged from this study. Expressivity is the narrative character of the text in clinical documents, how well a document conveys the patient’s and provider’s impressions, reasoning and thought processes.

In their following viewpoint paper, Rosenbloom et al. (2011) referred to the large number of requirements that were now being placed on clinical documentation. Clinical documents had to be simultaneously accurate, detailed, readable and useful to the reader, as well as meeting requirements for quality clinical management of conditions, administrative requirements, and structure and standardisation of data that can be reused to meet this. Physicians, however, valued expressivity of reasoning and thought process; the level of concern and uncertainty. They also placed importance on the linguistic nuance necessary for describing aspects of the patient encounter using words or phrases that the health care provider deemed appropriate. The researchers concluded that doctors needed a system that allowed them to use the
words, grammar, and style that they believed necessary to convey the patient encounter accurately.

Rosenbloom et al. (2007, 2011) studies and opinion commenced a discourse on EHR structure in the literature, with calls to move away from EHR structure, with the sole focus on content, to documentation being an information synthesis activity rather than a composition activity (Han & Lopp, 2013; Mamykina et al., 2014; Shoolin et al., 2013). Edwards et al. (2014) concluded from their comparison of the PDQI-9 and quality indicators, that an EHR system should simultaneously support the creation of subjectively high-quality notes and help ensure performance and documentation of delivery of high-quality care. This allowed physicians to focus on synthesising complex data and writing high-quality notes that are easily understood by other clinicians.

Any assessment method has to have an agreement on the essential focus of the assessment (Crossley & Jolly, 2012). The debate in the literature on quality as defined by content or form has favoured narrative expressivity and synthesised documentation. This has been supported by the findings of our literature review, with a strong endorsement of prioritising clinical reasoning and synthesised documentation (Stephens et al., 2011; Hammoud et al., 2012). Creating quality synthesised documents is a high-level skill. Ensuring that document elements are complete requires attention to detail and, or conscientiousness but is not necessarily a complex (Ferenchick et al., 2013). This consensus on the attributes of quality documentation needed to be taken into account when considering the focus of assessment of medical student EHR documentation.

2.4 Medical Education Assessment Literature

International lead researchers in medical education assessment have supported a move away from reducing the assessment of complex performances to checklists in favour of global ratings that value assessor judgments (Crossley, Gavin, Booth & Wade, 2011; Govaerts & van der Vleuten, 2013; van der Vleuten & Schuwirth, 2005). This has resulted in assessing the task as a whole and viewing the ability to handle a complex
professional task as an integration of skills, or attributes, that can be observed. Medical assessment, in general, needed to be assessing higher cognitive factors, and competencies that reflected an increasing level of sophistication and independence (Crossley et al., 2011). Focusing on the completion of elements in a document does not respect the holistic and tacit nature of the expertise of documentation.

The concern with overly detailing an assessment method to gain objectivity was that it trivializes the assessed material and decreases validity (Van der Vleuten & Schuwirth, 2005). In the validation study of the QNOTE, Burke et al. (2014) commented directly on their instrument resulting in significantly higher quality scores and higher intra-rater agreement than a global subjective assessment. Van der Vleuten and Schuwirth (2005) argued that global rating, relying on the holistic professional judgment on the part of the assessor, provided a more faithful reflection of expertise than detailed checklists and respected the characteristic of authenticity. Subjective global assessments were an important part of any program of assessment if used sensibly and expertly. The challenge was to make the assessment rigorous without belittling the complexity of the performance for objectivity reasons. Assessing elements of the process may not necessarily equate to a competent performance of the whole task.

2.5 Summary

In summary, there was no single clear definition of quality in EHR documentation in either physician or medical student populations. There have been no specific studies addressing quality attributes in medical student EHR documentation. These studies did not specifically focus on the quality of documentation. There have been more rigorous studies with the direct purpose of defining quality in the physician group. The attributes extracted from the medical student literature and in the validated measures of quality of physician EHR documentation, PDQI-9, and QNOTE, largely overlap.

Narrative expressivity in the documentation that clearly conveys the author’s clinical reasoning was highly valued by experts in medical education and in the physician literature. The PDQI-9 assesses synthesised documentation, and in the qualitative
study that was the basis for the QNOTE instrument, ‘explanatory’ and continuity of study’ were key findings across four stakeholder groups, although not included in the final instrument (Hanson et al., 2012).

In addition, the current literature in competency based assessment in medical education supported assessing the clinical tasks as a whole, integrating the many attributes that impact on the performance. The assessment should acknowledge the increasing levels of sophistication and independence required for the task and rely heavily on the professional judgment of the assessor.

The PDQI-9 is a global assessment of quality EHR documentation that relies on the professional judgment of the assessor after reading the document as a whole. It has a strong emphasis on Synthesised as an attribute of quality documentation. Of the qualities extracted from the medical student literature, the only omissions are Prioritised, and assessment of Originality and Professionalism. The PDQI-9 included Internal Consistency which had no corresponding finding in the review. For these reasons, the PDQI-9 was chosen as a base to build upon further exploration of the attributes of quality for medical student EHR documentation.

The following chapter describes the conceptual framework, methodology, methods utilised by this study to provide evidence in line with our research question. This study built upon the PDQI-9, and established test domain in physicians, based on the correlations to medical student EHR documentation skills and assessment in the literature, to provide evidence for a test domain to assess the attributes of quality in medical student EHR documentation.
3 Methodology and Methods

This chapter initially describes content-related validity and the conceptual framework employed by this study to assist in the collection of evidence for a content and test domain assessing the quality of EHR documentation in medical students. This is followed by a description of the methodology, rationale of the study design and methods chosen to meet this objective.

3.1 Content-Related Validity Studies

Validating global assessments is challenging. Each attribute needs to contextualized, weighted and balanced in order to achieve real content-related validity as an assessment of a complex task in a real world setting (Crossley & Jolly, 2012). Medical documentation is a complex task, even more multifaceted with the advent of EHRs (Mamykina et al., 2014). When little is known about the performance of a task, as was the case with medical student EHR documentation, the first step in the process is to provide evidence for all observable attributes of the task, namely the content domain, and ensure that what results in being assessed, the test domain, reflected the key aspects of performance. The test domain must be a fair representation of all observable attributes that could be assessed and a reflection of relative importance and free from unrelated or irrelevant observations (Portney & Watkins, 2015). Kane (2006) speaks in terms of content-related validity, how the domains are defined and in particular, on the relationship between the content domain and the test domain.

Content-related validity is distinct from criterion-related validity. Criterion-related validity focuses on the performance of the assessment model based on the theory and definition of the competency; how well the assessment and test domain lives up to its claims (Lissitz, 2009; Trochim, 2006). All criterion-related validity studies rely on the attributes, or criterion, being fair and accurate descriptions of the performance being assessed. With this study, we have focused on providing well evidenced, content-related validity, as an essential prerequisite to further criterion-related validity studies (Cook & et al., 2015; Kane, 2009).
3.2 Kane’s Conceptual Framework to Validation

As a framework for systematically organising the rigorous collection of evidence to support findings on content-related validity, we chose to structure this study on Kane’s Argument-Based Approach to Validation (Cook et al. 2015; Kane, 2006, 2009). This section provides an overview of the framework and the rationale for its use in our study. The subsequent section structures the purpose, research questions, methodology and methods chosen for this study on Kane’s framework.

3.2.1 Overview Kane’s Argument-Based Approach to Validation.

Kane’s conceptual framework guides the researcher to focus on the main assumptions being made in the validity process of an assessment and prioritising the gathering of evidence to argue for and against this assumption (Lissitz, 2009). The focus is on the interpretation and uses of the test rather than the test itself. Kane has argued that validity can never be proven absolutely, the task of validation was to present a body of evidence that supported the proposed use of the test and empirical evidence that justified the interpretations that were made (Cook et al. 2015; Kane, 2006).

The key to the validity argument is prioritisation. Kane emphasized the importance of stating assumptions and prioritizing the collection of evidence to address the weakest assumption in the assessment process. The evidence is then synthesised into a comprehensive and coherent argument (Kane, 2009). The process is iterative, and with the evidence collected the instrument can be further modified, the assumptions can be re-visited and further evidence can be collected.

Identifying assumptions made in an assessment process is a difficult task. Kane broke down the main inferences made in an assessment task into four categories to assist researchers to recognise assumptions relevant to their assessment and prioritise the most important pieces of evidence to gather. In a practical guide to Kane’s framework for medical education, Cook, Brydges, Ginsburg and Hatala (2015) explained that the evaluation process was often performed in stages and rarely was it possible to address all inferences in the one study. The main inferences identified by Kane in the
assessment process are described in Table 3.1. Researchers need to consider each inference and determine the priority of addressing them in turn. In time, the complete validity argument for an instrument contains multiple sources of evidence that span several, if not all, inferences.

**Table 3.1 Kane’s inferences (2006)**

<table>
<thead>
<tr>
<th>Inference</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scoring</td>
<td>Translating an observation into one or more scores</td>
</tr>
<tr>
<td>Generalisation</td>
<td>Using the score as a reflection of performance in a test setting</td>
</tr>
<tr>
<td>Extrapolation</td>
<td>Using the score(s) as a reflection of real world performance</td>
</tr>
<tr>
<td>Implications</td>
<td>Applying the score to inform a decision or action</td>
</tr>
</tbody>
</table>

A summary of Kane’s framework is diagrammatically represented in Figure 3.1.

**Figure 3.1 Kane’s framework adapted from Cook et al., (2015)**

Multiple factors contributed to the selection of Kane’s framework for our study. Primarily it was chosen as a conceptual framework that helped structure a validity argument, given that the aim and research questions were focused on establishing the content-related validity of the assessment of quality in medical student EHR.
documentation. Kane’s conceptual framework was a compromise between sophisticated validity theory and reality (Cook et al. 2015; Kane, 2006).

Secondly, Kane provided a clear structure for considering inferences made in the assessment process and assisted us in identifying the specific inference(s) we were addressing in our study. The two inferences that directly applied to the research questions of this study were Scoring and Extrapolation. Table 3.2 outlines the association between the relevant inferences of scoring and extrapolation and our research question.

Table 3.2 Kane’s inferences and the study research question

<table>
<thead>
<tr>
<th>Study Research question</th>
<th>Kane’s Inference</th>
<th>Questions asked to explore this inference by Kane (Cook et al., 2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students?</td>
<td>Extrapolation</td>
<td>Is the test domain in physicians seen as relevant and important, and closely aligned to the test domain in a medical student setting? Is it meaningful?</td>
</tr>
<tr>
<td></td>
<td>Scoring</td>
<td>Is the current structure of the instrument and the test scoring likely to create fair and reproducible scores?</td>
</tr>
<tr>
<td>2. What are the quality attributes of EHR written documentation by medical students in their first clinical year?</td>
<td>Extrapolation</td>
<td>Is the domain content different in medical students? Are the attributes seen in physicians’ documents present, and are there medical student specific attributes that should also be assessed?</td>
</tr>
<tr>
<td></td>
<td>Scoring</td>
<td>What judgments were made in scoring? Are the attributes interpreted similarly? How are they defined in a medical student context?</td>
</tr>
<tr>
<td>3. What modifications to the PDQI-9 could be made to improve validity in this setting?</td>
<td>Scoring</td>
<td>Given the above evidence, what potential modifications could be made to make scoring more reliable and fair? What is the likely test domain? What are the likely criteria?</td>
</tr>
</tbody>
</table>
Thirdly, Kane’s framework is an extremely pragmatic and simple way to view the validity argument. The framework did not rely heavily on psychometrics and was equally suitable for qualitative ungraded competency based assessment. The complexities of validating a global assessment of a complex task have been previously outlined, and it was anticipated prior to data collection that the strength of this study would heavily rely on the explanatory qualitative data exploring the informed judgment of the expert assessors.

Kane’s framework follows four basic steps. The steps are summarised in Table 3.3 with the correlations to our study design. In subsequent sections, we discuss the steps of Kane’s framework as applied to our study.

Table 3.3 Process of Kane’s framework and our study design

<table>
<thead>
<tr>
<th>Step</th>
<th>Process</th>
<th>Study Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define the proposed use</td>
<td>The starting point for Kane’s framework is a clear statement of interpretation and use of the test.</td>
<td>Aims and research questions of the study</td>
</tr>
<tr>
<td>State the main assumptions</td>
<td>Identify the main assumptions inherent in the application of the assessment depending on the proposed use.</td>
<td>The analysis of the literature review</td>
</tr>
<tr>
<td>Collection of evidence</td>
<td>Collect evidence to support or refute the assumptions made –</td>
<td>Methodology and Methods</td>
</tr>
<tr>
<td>Evaluate the evidence</td>
<td>Evaluate the evidence and modify instrument or proposed use</td>
<td>The interpretation of the study findings</td>
</tr>
</tbody>
</table>

Step 1. Define the Proposed use – the Purpose

This initial step in Kane’s framework defines the proposed use of the assessment. Our introduction explained the context and issues with EHR documentation and the purpose of the study. To summarise, the purpose of this study was to explore the attributes of quality EHR documentation in medical students as a foundation for assessing and providing feedback on the performance of documentation to medical students. The proposed use of any resulting instrument from this study was to assess
the quality of EHR documentation in medical students, with the provision of detailed information on performance, and feedback.

**Step 2. State the main assumptions**

Limited evidence was available to guide the development of an assessment instrument to assess this performance in medical students and tension about the interpretation of quality in EHR documentation. The literature examined provided a strong argument for building upon the test domain of quality in EHR documentation in physicians. The main assumption being questioned by our study was that attributes regarded as contributing to the quality documentation in physicians were relevant and important in the development of skills in medical students, and an appropriate platform from which to develop.

As a result of building upon recently published validated instruments, the research questions were

1. Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students?
2. What are the quality attributes of EHR documentation by medical students in their first clinical year?
3. What modifications to the PDQI-9 could be made to improve validity in this setting?

**Step 3. Collection of Evidence. Methodology and Methods**

The third step in Kane’s framework focuses on the decisions made about the collection of evidence.

### 3.3 Methodology

Content-related validity is achieved through the systematic method of collecting expert judgment that should be based on research evidence, and, or professional and
legal standards, and is usually qualitative in nature (Kane, 2006). If an approved
description of performance or guidelines to apply are not available in the literature,
then a panel of experts usually make an informed judgment about what constitutes a
competent performance and expectations that will be acceptable (Norcini & Shea,
1997).

Our priority was to develop the most appropriate design to answer the research
questions, specifically addressing the outlined inferences and assumptions in Step 1
and 2. We made the decision to use a mixed methods design, employing the collection
and analysis of quantitative and qualitative data in a single study, to increase the range
and scope of inquiry to explore all dimensions of the performance that may emerge
(Lavelle, Vuk & Barber, 2013). We believed that choosing a single method alone to
determine the content and test domain of a currently poorly defined complex
performance may potentially provide an incomplete picture of the whole. It was
intended that the strengths of the individual methods produced a better overall
understanding of the dimensions than the individual strands alone (Hesse-Biber, 2011;
Teddle & Tashakorri, 2009).

Additionally, a mixed methods study design enabled the triangulation of evidence to
strengthen the findings from each data collection method of the study and allowed
sequential collection of data to improve construct validity of the design (Creswell &
Plano Clark, 2011; Hesse-Biber, 2011) It was also important that the study was
practical, valued the assessor’s judgments and maximized the expression of their
expertise in a realistic time frame that they could be expected to invest in the process.
The emphasis on pragmatism married well with both the underlying worldview for
using a mixed methods design and Kane’s conceptual framework (Creswell & Plano
Clark, 2011).

To complement the design, the analytic strategy was the integration of hypothesis
testing and hypothesis generation in a single study. Creswell and Plano Clark, (2011)
regarded this as the hallmark of mixed methods research. Grounded theory guided the
analysis to generate themes and the relationships between themes (Glaser & Strauss, 1967; Malterud, 2001). This was achieved by open, axial and selective coding. There was continuous interaction with the data, making comparisons and asking questions.

### 3.4 Methods

#### 3.4.1 Overview of the study design

The research design had to specifically maximize the strength of evidence collected with human judgment at its core, minimize risks of subjectivity and present a validity argument at a level of certainty to enable decision making on the competency of EHR documentation skills in medical students (Davidson, 2005; Kane, 2006).

Recognised methods for establishing content domain included review of past literature, interview or poll experts, asking experts to think aloud as they perform the task, or judge the task, and the use of another instrument that has achieved a degree of validation to assist in defining the domain content (Cook et al. 2015; Portney & Watkins, 2015). We decided to employ a number of these methods to triangulate the findings and strengthen the conclusions we made.

The study design was conducted in two stages. The first stage was the application of the PDQI-9 to medical student EHR documentation by expert assessors, followed by the second stage, semi-structured interviews with the expert assessors.

The first stage tested the hypothesis that attributes of quality in physician documentation were similar to those seen in medical students and a validated instrument for assessing this would be valid in a medical student context. Numerical data from scoring the documentation and narrative comments for the justification of the scores and the meaningfulness of the attribute to assess in a medical student context were collected concurrently. Data collection and analysis occurred sequentially in Stage 1. The descriptive statistics from the quantitative arm of the study were intended to be indicative, and the thematic analysis of the qualitative data helped explain the numerical data and why the hypothesis may, or may not be correct. After
collection, the integrated Stage 1 numerical and qualitative findings were triangulated with the literature review findings to provide a framework for the second stage of the study.

In Stage 2, the method used was semi-structured interviews with the expert assessors. The aim was to further the exploration of the attributes of medical student EHR documentation and reach consensus on both the findings from the first stage and a proposed test domain for assessing quality in medical student EHR documentation. The triangulation and integration of Stage 1 and literature review findings created an initial framework for the interview schedules. Data collection and analysis occurred concurrently in Stage 2. The interview schedules evolved and adapted to the data analysis of preceding interviews. Stage 2 data analysis was iterative moving between thematic analysis of Stage 1, codes and text to reach saturation and consensus.

Throughout both stages of the study, there was a strong emphasis on capturing the language of the expert assessors, and understanding their decision processes, to enable a complete picture of the content and test domain and the judgments made. Figure 3.2 summarises the process of the study design and Table 3.4 outlines how the data collection and analysis addresses the research questions and Kane’s inferences. Each stage and further rationales for the specific methods are outlined in more detail in following sections. The Human Research and Ethics Committee of the Melbourne Graduate School of Education, University of Melbourne, approved this study.
Figure 3.2 Process of study design

Study Design Process

Stage 1. Application of the PDQI-9:
- to medical student EHR documentation
- by expert assessors

Data collected concurrently
Produces
- Numerical Data
- Qualitative narrative data

Data analysed concurrently
Produces
Integration of Stage 1 analysis
Indicative descriptive statistics
Explanatory thematic analysis

Stage 1 Analysis triangulated with
literature review findings
Initial framework for interview schedule

Stage 2. Semi-structured interviews with
Expert assessors

Data collected and analysed
concurrently
Modifications of subsequent
interview schedules

Consensus on test domain for assessing quality
of EHR documentation in medical students
Table 3.4 Relationship between research questions, Kane’s inferences, data collection and analysis

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Data Collection product</th>
<th>Data Analysis product</th>
<th>Kane’s inference</th>
</tr>
</thead>
</table>
| Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students? | Stage 1: Numerical item scores  
Stage 1: Narrative justifications for scoring | Stage 1  
Descriptive statistics - indicative  
Coded text and explanatory thematic analysis | Extrapolation Scoring |
| What are the quality attributes of EHR documentation by medical students in their first clinical year? | Stage 1: Numerical item scores  
Stage 1: Narrative  
• Justifications for scoring  
• Meaningfulness of attribute  
• Additional observable attributes | Stage 2  
Interview transcripts  
• Interpretation of attribute  
• Meaningfulness of attribute in medical students | Extrapolation Scoring |
| What modifications to the PDQI-9 could be made to improve validity in this setting? | Stage 2: Qualitative transcripts:  
• Exploration of potential modifications based on all previous findings | Stage 2 analysis resulting in recommendations for test domain in medical students | Scoring |
3.5 Limitations of content-related validity studies

Care had to be taken with the study design to address the main criticisms of the methods and evidence from content related validity studies (Kane, 2009). Content-related validity has often been appraised as being weak and subjective given that it has relied largely based on human judgment (Davidson, 2005; Trochim, 2006). If the study processes are transparent, defensible and fair, and directly attend to concerns about subjectivity, the data should be objective and free of personal bias and at a level of certainty appropriate for decision-making in a given context (Davidson, 2005). Davidson (2005) clearly outlined three inherent risks associated with evidence derived from human judgment and how these could be addressed in study design to increase the robustness of the evidence collected. In the following paragraphs, we have outlined and explained how they were addressed in this study.

Subjectivity 1 is the risk to a study by the inappropriate application of personal or cultural preference and biases. This bias can be minimized by firstly, recognizing it as a threat to the study, and then taking measures to counter this. The literature review highlighted the tension between the different interpretations of quality in EHR documentation, and there were published advocates, in addition to the different validated instruments, for both approaches (Burke et al., 2015; Stetson et al., 2009). Personal orientation about the interpretation of quality, the purposes of the EHR documentation, expectations of medical students, expectations from different specialties all may have altered the judgment of the attributes in medical students and resulted in a poorly defined content domain for performance. The selection of the assessors strove to reduce these potential biases.

The assessors had to be predominantly medical educators of many years’ experience, with a clear understanding of student performance and expectations and the development of skills. It was necessary that they were familiar with EHR documentation but not experts with a potential different agenda to meet (Davidson, 2005). When assessing documentation, they had to be able to clearly define their
expectations of a medical student performance, and the general approach this must take with the development of skills, and divorce this from the expectations of their fully qualified clinical colleagues. Their use of EHRs was as a clinician sharing or transferring patient care and responsibility, not as a medical administrator or information technology expert. They were also fully informed of the decision process for the selection of the PDQI-9, and the current medical education literature on workplace based assessment to appreciate the process of a global assessment of quality rather than a checklist, and the value placed on the language used to make judgments.

Subjectivity 2 is the risk to the study by relying on informed judgment. The evidence to be provided in the study relied on the expert informed judgment to be valid. Kane (2009) countered the criticism on the subjectivity of informed judgment by arguing that these objections were overstated and that human judgment is required for all of the inferences. Crossley and Jolley (2012) agreed with this and maintained that clinical assessors if asked in ‘the right way’ can make highly reliable judgments on complex performances. To address this concern, the collection of evidence was triangulated in this study and consideration was given to adapting the PDQI-9 to gather the evidence in such a way that it directly addressed the development of the content domain and the test domain in medical students in both numerical and narrative form.

Subjectivity 3 is the risk to the study by the person’s inner Life. The third type of subjectivity refers to evidence collected about a person’s inner life or experiences that may be true to the individual but unable to be substantiated by external observation. This risk was very unlikely to be relevant to EHR documentation and its assessment.

3.6 First Stage

In the first stage, expert assessors applied the PDQI-9 to a sample of first clinical year medical student EHR documentation and produced both numerical scores and qualitative justifications for the scoring. The quantitative and qualitative data were collected simultaneously, and the point of mix was at analysis. Qualitative data was
also collected on the assessors’ judgment of the meaningfulness of the attribute in the medical students’ setting and any possible attributes that should also be included in this population. Findings were analysed with descriptive statistics and thematic analysis.

3.6.1 Participants: Assessors

The following section focuses on the participant details of the assessors: the required selection criteria needed to assess performance in this setting, the subsequent inclusion criteria, and sampling strategy and demographics are described.

Selection criteria

The selection of expert assessors has been partially addressed in the earlier section on the limitations of content-related validity studies. The concerns raised in addressing subjectivity bias were in keeping with the writings of Norcini & Shea (1997) on the selection of experts to reach an agreement on judging and rating performance. To summarise from the literature, the following should be inclusion criteria in order to reduce the risk of subjectivity and to provide meaningful results on the performance expected from students (Crossley, 2013; Norcini & Shea, 1997; Portney & Watkins, 2015)

- Experts should be knowledgeable about the competencies of the students during training and at its completion
- Experts have a similar purpose for the assessment and should be sensitive to this purpose
- Experts must be familiar with the content of the assessment
- Experts should appreciate the demands of practice but not be wedded too closely to a very specific skill. This can result in higher their expectations of students and competing agendas for performance.
- Selection of experts should ensure an appropriate mix of racial, ethnic and gender representation
Consistent with literature of the selection of experts, the selection criteria for expert assessors in our study is listed in Table 3.5

Table 3.5 Selection Criteria for assessors

<table>
<thead>
<tr>
<th>Selection Criteria for Assessors</th>
</tr>
</thead>
<tbody>
<tr>
<td>The assessors must be long-standing medical educators, with a clear understanding of the skill level and expectations of the students</td>
</tr>
<tr>
<td>Assessors had to be qualified medical doctors and users of EHR systems</td>
</tr>
<tr>
<td>Assessors had to use EHR systems for the purpose of patient health care and sharing and transferring responsibility for patient care. They were not to be medical administrators or IT experts with an alternate purpose for EHR documentation or a different set of student expectations</td>
</tr>
<tr>
<td>Assessors had to be able to understand the development of skills and they would preference the development of generic skills over and above the expectation of EHR documentation from fully qualified physicians.</td>
</tr>
<tr>
<td>Assessors were an appropriate mix of racial, ethnic and gender representation</td>
</tr>
</tbody>
</table>

**Sampling strategy**

Assessors were purposively sampled for the selection criteria, through the researchers’ personal knowledge regarding the expertise of the assessor, and through recommendations from the Head of the Medical Schools within Victoria, Australia.

Assessors meeting the selection criteria were invited via an email to participate in this study. Eight clinicians were invited, of which seven agreed to participate in the first stage of the study. A follow-up email invitation was extended to the assessors to participate in the second stage of the study. Of these seven, six participated in interviews to reach consensus on the findings.

**Demographics of Assessors**

Assessors who agreed to participate were sent a Google form ‘Assessor Consent, Demographics and Logistics’. A copy of the survey is in Appendix C. Data was collected on their specialty, years since graduation, number of years involved in medical student education and assessment, and number of years’ experience with EHR, as well as other
postgraduate activities or qualifications. The data is visually represented in Figures 3.3, 3.4 and 3.5

**Figure 3.3 Assessor medical education experience**

**Medical Education Experience**

- < 5 years
- 5 - 10 years
- 10 - 20 years
- > 20 years

**Figure 3.4 Assessor specialty**

**Speciality**

- General Practice
- Medical
- Other
All assessors held lead positions in medical education units at their respective Universities. They worked across the three medical schools in Victoria, and three held higher qualifications in education. They were distributed across metropolitan and regional environments; four were from inner metropolitan, two from outer metropolitan and one from a regional city.

### 3.6.2 Medical Student Documentation

In Stage 1 the assessors applied the PDQI-9 to a selection of medical student EHR documentation. This section describes the documentation, selection criteria and sampling strategy for the documentation.

The medical student EHR documents were selected from the Patient Record Program at the University of Melbourne. This program is part of the Year 2 Doctor of Medicine (MD) at the University of Melbourne. The MD is a four-year graduate program. Year 1 focuses on biomedical science. Year 2 is the first year they are placed in a clinical setting. The objectives of the Patient Record Program are to help students develop expertise in the practice of creating, storing and managing high-quality patient records. It is an online patient record system designed for clinical education. The system allows students to capture details of their clinical encounters and store de-identified patient information in a secure database. It also mediates online access to specialist medical databases such as Australian Medicines Handbook. The
documentation from the 2013 and 2014 Year 2 cohorts were selected as these students, at the time of the study, were still enrolled in the MD and this facilitated gaining their consent.

Given the complexities of validating a global assessment, a level of uncertainty remained about the utility of the PDQI-9 in a medical student setting. We made the decision, with advice from the University of Melbourne Statistics Department, to commence the study with approximately 10 documents. This number of assessed documents would allow us to make a judgement on whether there was general agreement, or disagreement, in scoring across cases and attributes. If the preliminary numerical data analysis supported that the PDQI-9 was similarly interpreted and rated over a small number of documents, the PDQI-9 could be trialed with a larger cohort of assessors and documents. If a significant amount of clarification of the PDQI-9 was required in this setting, the focus would be on the explanatory qualitative data generated from the application of the PDQI-9 and data collection directed at reaching consensus on the observable attributes, and a test domain, for this performance in medical students.

Selection criteria

The documents for the Stage 1 exploratory phase needed to vary in quality, and specifically on the attributes being assessed by the PDQI-9, to try to differentiate scoring and judgments on these attributes. Despite the small number of documents required in this stage of the study, there was also an attempt to sample across year groups, clinical setting contexts and at different stages of learning during their first year. The selection criteria for the EHR documentation is listed in Table 3.5.
### Sampling strategy

All students from Year 2 2013 and 2014 were invited to permit their documentation to be assessed in our study. The entire year groups were invited in order to avoid any bias secondary to specific clinical setting environments. Students were asked to opt out of the project if they did not want their documentation to be assessed. To preserve confidentiality, students advised a third party if they wished to opt out. The third party directly liaised with the information technology team responsible for maintaining the database. This team, referred to as the Educational Technology Team (ETT), removed the documentation from the opted out students from the database. The research team was unaware of who opted out; however, of over 700 students, across two-year cohorts, only two students opted out.

All permitted documentation was de-identified with respect to the student and patient. The ETT converted all documentation into PDFs. Each document was given a case identification number. This allowed the ETT team to link the document back to the database, and to provide the researchers with de-identified documents. All the de-identified student EHR documents were randomised. We were forwarded the first 100 randomised de-identified student EHR documents as a password protected file and

---

**Table 3.6 Selection criteria for medical student documentation**

<table>
<thead>
<tr>
<th>Selection criteria for student EHR documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attributes</strong></td>
</tr>
<tr>
<td><strong>Year Groups</strong></td>
</tr>
<tr>
<td><strong>Through the year</strong></td>
</tr>
<tr>
<td><strong>Clinical Placements</strong></td>
</tr>
<tr>
<td><strong>Word count</strong></td>
</tr>
</tbody>
</table>
continued to request documents until there was a sample of documents that met the selection criteria.

We made the decision on the documents to be included in the first stage of the study, applying the criteria in Table 3.6. Eleven documents were purposively selected.

3.6.3 Data Collection
This section details the process of data collection for the first stage of the study. Necessary adaptations of the PDQI-9 were required to ensure the data collected directly addressed the research questions. These adaptations are outlined next, followed by details of the process for assessment.

Adaptations of the PDQI-9
Attributes
The attributes and the one sentence descriptor of the attribute were unchanged from the original instrument. The only exception was the removal of the attribute Accurate. In this study, it was not possible to judge the accuracy of the submitted EHR documents, given they were de-identified. The assessment of Accurate in documentation has been noted to be difficult in a real world setting where the assessors of clinical documentation have often not witnessed the patient encounter that the documentation resulted from (Boulet, Rebbecchi, McKinley & Whelan, 2004). Creating assessment tasks that allow Accurate to be assessed reliably has been logistically challenging. There is widespread consensus that regardless of the level of training, the documentation should be an accurate and true record of the patient encounter, free of incorrect information. In any final rubric, if the assessment task can be designed in such a way that allows for the assessment of Accurate, this item should be included. The exclusion of Accurate reflected the discrepancies between the extrapolation inference, whereby the assessment of Accurate in the real world is regarded as a priority, and the inference of generalisation, that the assessment of Accurate documentation is difficult to capture.
Benchmark

The benchmark for the PDQI-9 was a gold standard physician document. There were no accepted standards for medical student EHR documentation. Given the documents used were from first-year clinical medical students we believed that to continue to have such a highly developed benchmark, as that of a fully qualified physician, would render the feedback from assessing the documents with the PDQI-9 of little use to the students. Three medical education experts within the Department of Medical Education, University of Melbourne, were selected to provide their opinion on what the benchmark should be for rating medical student documents in this project. They were asked to make an informed judgment about setting a benchmark that would be acceptable. They were provided with three de-identified student documents and the PDQI-9.

They agreed with the concerns regarding the gold standard physician document being the standard and were of the opinion that if the benchmark was the most proximate achievable goal for the student, this provided more useful feedback and guidance for the student. Furthermore, a level of competency of a graduating medical student, and the expectations of a doctor at the beginning of their first postgraduate year, was an easier level to define for the assessors, than an indeterminate point in their training. Given the variation in curriculum delivery across medical courses, there would be no other similarly defined point in their learning. Scoring five in any one quality equated to a competent performance for an end-of-course medical student. In summary, the benchmark chosen for this exercise was that of the standard expected from a graduating student, completing a medical degree, and at the beginning of their first post-graduate year.

Scoring

1. Numerical. The scoring remained as a five-point Likert scale. In the PDQI-9, the descriptors at either end of the scale were ‘Not at all’ and ‘Extremely’. To be
more explicit, these were changed to ‘Not at all evident’ and ‘Extensively demonstrated’.

2. Qualitative. Assessors were asked to provide a detailed justification for each score assigned.

3.6.4 Procedure
In the first stage, the seven assessors were provided with information to improve the consistency of assessing the documents, and ensure that the assessors were aware of the purpose of the EHR documentation and assessment and the importance of capturing the language of their decision making. The Assessor Training Package is in Appendix D.

The information included

- Assessor Training Package
- A link to the adapted PDQI-9 in a Google form titled the ‘Electronic Health Record Documentation Quality Instrument’
- Eleven student EHR documents in a password protected PDF. Each document was given a case identification number located at the top of the de-identified EHR student document

The Assessor Training Package included details of the process of the assessment, an overview of the student documentation, an overview of the PDQI-9 and the attributes to be scored and a PDF copy of the Google Form ‘Electronic Health Record Documentation Quality Instrument’. They were provided with the justification for the change in benchmarking and the purpose of the documents and focus of assessment. We asked assessors to think generally about medical student documentation quality when scoring and justifying their marking, rather than focus on one particular program that they may be familiar with.

Assessors were given their own unique assessor identifier. The assessors had to complete a separate PDQI-9 for each student document. They accessed the adapted
PDQI-9 by the link to the Google form. They were able to read the student document within the pdf and then score the document on the Google Form. At the top of the form, the assessor entered their unique assessor identifier and the Case identification number for the student document. For each document, a new form was to be completed.

The document was then rated in terms of complexity of the patient described in the record. For each attribute the assessor was asked, to the best of their ability to score the written document for the specific quality attribute. They were asked to explain why they attributed individual scores and for an overall comment on the meaningfulness of the specific quality attribute in assessing medical student documentation.

At the end of the form, the assessor was asked to make a global assessment of the record in its entirety and justify this, particularly highlighting if there was a significant aspect of this record that contributed to the global score, which may have seemed at odds with the scores from the individual attributes. Finally, they were asked to comment on the purpose of EHR written documentation for medical students, and any attributes that may be important and not included in the assessment.

3.6.5 Stage 1 Analysis

Findings from Stage 1 were analysed with descriptive statistics and thematic analysis. For the quantitative strand of Stage 1, descriptive statistics were used. Data was entered into Minitab and graphs of scoring attributes across cases were generated.

For the qualitative strand of Stage 1, we used grounded theory to advise our analysis. The overall main organizing theme for the dataset was the description of the attributes of quality in a medical student document. We applied open, axial and selective coding techniques to the written qualitative data of Stage 1 and interview transcripts of Stage 2 (Creswell, 2013).
The initial analysis was focused on the attributes in the PDQI-9 and included the justifications for scoring, the interpretation and the meaningfulness of the attribute. All narrative data was placed in excel spreadsheets and the data collected for each attribute was open coded, and then grouped into themes. The data was looked at from multiple perspectives. Relationships were pursued between the data and the coding, with the focus in turn on the attributes, the cases, the assessors and the numerical score and justifications to ensure that the themes generated were explanatory across the entire database. The data did not reach saturation in Stage 1.

After Stage 1 analysis of both quantitative and qualitative findings the attributes grouped into three main themes. These were:

A. Attributes that were similarly interpreted and meaningful in a medical student context
B. Attributes that were interpreted in multiple ways and meaningfulness was inconclusive
C. Attributes that were difficult to interpret and meaningfulness was unlikely

3.7 Second Stage

The purpose of the second stage of the study was to further explore the findings of the Stage 1 analysis, reach consensus on the interpretations of the attributes from the PDQI-9 in a medical student context, whether with agreed upon interpretations the attribute was meaningful to assess in a medical student context and if there were any necessary modifications of the PDQI-9 that could be made to improve validity of the PDQI-9 in this setting. Any observable and meaningful attributes in quality medical student EHR documentation, that were not included in the PDQI-9, were also explored to further define the content and test domain for this performance by the medical student. The instrument chosen for this stage of the study to reach consensus was the semi-structured interview. Semi-structured interviews enabled the collection of rich data that could be guided to direct specific issues as well as allowing for the exploration of an individual’s perspective (Creswell, 2008).
There were a number of reasons for the selection of the semi-structured interview as the data collection method for the second stage. The main strength of this method was its ability to be directed to provide in-depth exploration and clarity of emerging themes from both Stage 1 analysis and the literature. The assessors were ideal participants for interviewing. They were articulate, not hesitant to speak and willing to share ideas. The interview format allowed assessors to voice their opinions unconstrained by the perspectives of others, answer with their own words and draw from actual cases (Johnson & Turner, 2003). Pragmatics were a further reason for choosing semi-structured interviews. Even though the time for myself was increased in conducting, transcribing and analysing the interviews, the logistics of organising a focus group with this population were problematic. Finally, the semi-structured interview format continued to capture the language of the assessors in the description of the attributes of quality in medical student EHR documentation.

Stage 2 data collection focused on reaching consensus and saturation, whereby no new themes on the interpretation and meaningfulness for each of the attributes were emerging before we stopped collecting data (Creswell, 2008; Johnson and Turner, 2003; Teddlie & Tashakorri, 2009).

3.7.1 Data collection
The seven assessors were invited to participate in a one-hour semi-structured interview with the student researcher. Interviews were conducted with six out of seven of the assessors. Interviews were audiotaped and later transcribed by an independent research assistant. The transcripts were identified with the assessor’s unique identifier from Stage One.

The framework for the interviews was based on the integration and triangulation of findings from the descriptive statistics and thematic analysis from Stage 1, and the literature findings. The primary interview questions focused on the interpretation of the attribute, the meaningfulness of it in medical students and potential modifications for the PDQI-9 arising from this discussion. The general schedule of the semi-
structured interviews is described in Table 3.7. An example of interview schedule is in Appendix E. The interview process was iterative, with the interview schedules being adapted to incorporate the findings and analysis from previous interviews and extend the exploration of the attributes.

In summary, for each attribute, we captured the appropriate interpretation of the attribute in a medical student context, how the attribute would be observed in the performance of EHR documentation, if this observation was meaningful to assess, and modifications to assessment that would need to occur to achieve this.

Table 3.7 Semi-structured Interview Schedule

<table>
<thead>
<tr>
<th>Semi-Structured Interview Guide</th>
<th>Purpose and supplementary information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interview Guide</strong></td>
<td>• To reach consensus on purpose prior to ongoing discussion in line with the first step in Kane’s framework</td>
</tr>
<tr>
<td><strong>Purpose and supplementary information</strong></td>
<td></td>
</tr>
<tr>
<td><strong>What is the purpose of student documentation in an educational setting?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>The attributes fell into three categories after analysis:</strong></td>
<td></td>
</tr>
<tr>
<td>1. Attributes that were similarly interpreted and meaningful in a medical student context</td>
<td>• Explain Stage 1 thematic analysis</td>
</tr>
<tr>
<td>2. Attributes that were interpreted in multiple ways and meaningfulness was inconclusive</td>
<td>• Each theme and the attributes within it were then discussed in turn</td>
</tr>
<tr>
<td>3. Attributes that were difficult to assess with the documents</td>
<td>• Aim of interview is to reach consensus</td>
</tr>
<tr>
<td>1. Attributes that were similarly interpreted and meaningful in a medical student context.</td>
<td>Interview schedule with</td>
</tr>
<tr>
<td>For each attribute in the theme:</td>
<td>• Stage 1 cases, numerical data and referenced quotes from analysis</td>
</tr>
<tr>
<td>• Explore reasons for the difference in marks</td>
<td>• Literature findings and</td>
</tr>
<tr>
<td>o Clarify interpretation</td>
<td>• Any other pertinent findings of from Stage 1 analysis</td>
</tr>
<tr>
<td>o Impact of benchmarking, and assessor variability</td>
<td></td>
</tr>
<tr>
<td>Explore</td>
<td>Evolution of discussion based on concurrent Stage 2 analysis</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>• How it would be observed in medical student documentation at different stages</td>
</tr>
<tr>
<td></td>
<td>• Meaningfulness of attribute to assessment</td>
</tr>
<tr>
<td></td>
<td>• Modifications to assessment of the attribute</td>
</tr>
</tbody>
</table>

2. Attributes that were interpreted in multiple ways and meaningfulness was inconclusive

For each attribute in the theme:

• Clarify confirm the interpretations of Stage 1
• Explore
  • How it would be observed in medical student documentation at different stages
  • Meaningfulness in a medical student
  • Modifications to assessment of the attribute

3. Attributes that were difficult to interpret and meaningfulness was unlikely

For each attribute in the theme:

• Explore
  • Difficulties in assessing
  • Interpretation in medical students
  • How the attribute could possibly be observed in medical student documentation at different stages
  • Modifications to assessment of the attribute

Further exploration

• Any other observable attributes arising from Stage 1 that were not included in the PDQI-9
• Brief discussion on
  • Accurate as an attribute
  • Other problems observed with scoring, for example, multiple comments about low word counts
3.7.2 Stage 2 Analysis

The analytic strategy of Stage 2 built upon the Stage 1 thematic analysis. The main organising theme for the dataset was the description of the attributes of quality in a medical student document. The data from the interviews was coded and iterative comparison between Stage 1 data and analysis to deepen the understanding of the attributes within a medical student context, and the themes emerging. Stage 2 analysis was occurring concurrently with data collection, and the ongoing findings and analysis contributed to the evolving interview schedules of subsequent interviews, to explore each attribute in depth until no further themes were emerging.

3.7.3 Summary

In this study, we were seeking answers to questions on content-related validity. Content-based validity evidence was collected and analysed focusing on establishing an understanding of the attributes that most clearly define the performance of quality EHR documentation in medical students. Kane’s framework for a validity argument structured our study, and assisted in articulating the main assumptions we were making in the assessment of medical student EHR documentation, and prioritising the collection of evidence.

We chose a mixed methods study design to allow for a fuller understanding of the performance of EHR documentation in medical students and the triangulation of the data to strengthen the conclusions we made from our study. The qualitative data was analysed using grounded theory.

Our study gave serious consideration to the systematic collection of evidence to reduce the risk of biases and potential criticisms to subjectivity. It was conducted in two stages, the application of the PDQI-9 to medical student EHR documentation by expert assessors, and semi-structured interviews with the expert assessors in the second stage. Careful selection of expert assessors to provide the judgments necessary to determine the content and test domain for quality EHR documentation by medical students addressed the threats to the validity of this study. Throughout data collection
importance was placed on capturing the language of the assessors to better understand their judgment process. The following chapter presents the findings from our study.
4 Findings

Chapter 4 presents the findings from our study. The findings and conclusions, discussed in Chapter 5, address the final stage in Kane’s Framework, evaluating the evidence to support or reject the proposed use and assumptions of the assessment process, and propose a revision of the instrument or its use.

4.1 Overview of Analysis

After Stage 1 analysis the attributes were organised into three main themes. Stage 2 analysis considered each attribute within the themes, and sought to clarify the interpretation of the attribute in medical student EHR documentation, the meaningfulness of the attribute to be assessed, and potential modifications to the PDQI-9, and test domain.

A summary of Stage 1 numerical data is presented initially. Following this, the chapter is structured on the three main organising themes from the Stage 1 analysis. The discussion of the attributes, and the analysis from Stage 2 is presented within these themes. The chapter concludes with other pertinent findings from the data collected, that falls outside the three main organising themes. An overview of the process of collecting evidence and the analysis is presented in Figure 4.1.
4.2 Overview of Stage 1 Descriptive Statistics

When the numerical data was analysed there was a wide range of scores given for each attribute in a particular case, refer to Table 4.1. The scoring scale for the PDQI-9 was a Likert scale from 1 through to 5, with the attribute being scored from 1 ‘Not at all evident’ to 5 ‘Extensively demonstrated’.
Table 4.1 Descriptive statistics from Stage 1

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Number of Scored Cases</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thorough</td>
<td>77</td>
<td>2.88</td>
<td>1.21</td>
<td>1.47</td>
</tr>
<tr>
<td>Synthesised</td>
<td>77</td>
<td>2.62</td>
<td>1.19</td>
<td>1.42</td>
</tr>
<tr>
<td>Organised</td>
<td>77</td>
<td>3.41</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>Comprehensible</td>
<td>77</td>
<td>3.71</td>
<td>0.92</td>
<td>0.84</td>
</tr>
<tr>
<td>Up to Date</td>
<td>77</td>
<td>2.58</td>
<td>1.36</td>
<td>1.85</td>
</tr>
<tr>
<td>Useful</td>
<td>77</td>
<td>3.03</td>
<td>1.01</td>
<td>1.03</td>
</tr>
<tr>
<td>Succinct</td>
<td>77</td>
<td>3.57</td>
<td>0.89</td>
<td>0.80</td>
</tr>
<tr>
<td>Internally consistent</td>
<td>77</td>
<td>3.57</td>
<td>1.26</td>
<td>1.59</td>
</tr>
</tbody>
</table>

An example of the raw scores for the attributes, *Thorough* and *Up to Date* are graphically depicted below in Figures 4.2 and 4.3 respectively, to further illustrate the variation in scores across assessors on the same cases. Each individual graph represents the scoring of an individual assessor, for example, assessor AA, BB, ... for the specific attribute across Cases 1 to 11.

Figure 4.2 Scoring of Thorough across cases for each assessor
A different description of the data is tabulated in Table 4.2, with the raw scores for the scoring of Case 10 and the descriptive statistics derived from this.
Table 4.2 Descriptive statistics from scoring for Case 10

<table>
<thead>
<tr>
<th>Assessor</th>
<th>Thorough</th>
<th>Useful</th>
<th>Organised</th>
<th>Up to Date</th>
<th>Comprehensible</th>
<th>Succinct</th>
<th>Synthesised</th>
<th>Internally consistent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>BB</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>CC</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>DD</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>FF</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>GG</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>HH</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>22</td>
</tr>
</tbody>
</table>

Summary Statistics for Case 10

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thorough</td>
<td>1</td>
<td>4</td>
<td>2.29</td>
<td>1</td>
<td>2</td>
<td>1.24</td>
</tr>
<tr>
<td>Useful</td>
<td>1</td>
<td>4</td>
<td>2.29</td>
<td>1</td>
<td>2</td>
<td>1.24</td>
</tr>
<tr>
<td>Organised</td>
<td>4</td>
<td>4</td>
<td>3.00</td>
<td>4</td>
<td>3</td>
<td>1.00</td>
</tr>
<tr>
<td>Up to Date</td>
<td>4</td>
<td>4</td>
<td>2.29</td>
<td>4</td>
<td>4</td>
<td>1.57</td>
</tr>
<tr>
<td>Comprehensible</td>
<td>4</td>
<td>4</td>
<td>3.29</td>
<td>4</td>
<td>3</td>
<td>0.57</td>
</tr>
<tr>
<td>Succinct</td>
<td>3</td>
<td>5</td>
<td>3.57</td>
<td>3</td>
<td>3</td>
<td>0.62</td>
</tr>
<tr>
<td>Synthesised</td>
<td>2</td>
<td>4</td>
<td>2.14</td>
<td>3</td>
<td>3</td>
<td>1.48</td>
</tr>
<tr>
<td>Internally consistent</td>
<td>3</td>
<td>5</td>
<td>2.86</td>
<td>1</td>
<td>3</td>
<td>1.81</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>5</td>
<td>21.71</td>
<td>3</td>
<td>3</td>
<td>50.90</td>
</tr>
</tbody>
</table>

The obvious lack of agreement among raters in scoring across cases meant that the resulting Stage 1 quantitative data was indicative only. It also clearly demonstrated that the PDQI-9 in its current format was not valid in a medical student context, on numerical scoring alone, with a Likert scale of 1 to 5 for the presence of each attribute.

It was difficult to extract patterns and meaning from the raw numerical scoring of the attributes across assessors and cases. Given the aim of this study was to provide content validity for an assessment of EHR documentation, it was more relevant to understand why the assessors scored as they did, rather than studying the scoring of attributes across cases. The most helpful expression of the numerical data to assist in the thematic analysis of Stage 1 was the range of scoring as represented by variance, rather than the raw scores themselves. In a visual format, the numerical values of variance demonstrated this more noticeably than standard deviation. The variance was useful in clearly demonstrating the spread of scores from the mean for a given attribute for each case. The variance in scoring for each attribute, and the interpretation of this, is outlined in Table 4.3.
### Table 4.3 Range of variance for each attribute

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Number of cases</th>
<th>Variance ≤ 1.0</th>
<th>Variance 1.0 – 1.5</th>
<th>Variance ≥1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Useful</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thorough</td>
<td>8</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthesised</td>
<td>7</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succinct</td>
<td>7</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehensible</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Organised</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Up to Date</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Internally consistent</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

In the sections below for each attribute, the presentation begins with a graphical depiction of the variance in the scoring across assessors. However, Stage 1 qualitative data was strongly relied upon to understand the scoring of the attributes.

### 4.3 Organising themes of analysis

After Stage 1 analysis of both quantitative and qualitative findings the attributes grouped into three main themes. These were:

A. Attributes that were similarly interpreted and meaningful in a medical student context

B. Attributes that were interpreted in multiple ways and meaningfulness was inconclusive

C. Attributes that were difficult to interpret and meaningfulness was unlikely

In the following sections, each theme is discussed separately. For each attribute categorised in a theme, an interpretation of the indicative descriptive statistics. Stage 1 thematic analysis, the focus of the Stage 2 data collection and Stage 2 analysis are integrated to address the observed interpretation of the attribute in medical students,
the meaningfulness of assessing this attribute and proposed modifications for the assessment of the attribute for medical student EHR documentation.

How the presentation of the findings in this manner answers the research questions and is related to Kane’s inferences is outlined in Table 4.3.
<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Kane’s Inference</th>
<th>Presentation of Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students?</strong></td>
<td>Scoring (The observation generates a fair, accurate, reproducible quantitative score, or insightful narrative comment) &amp; Extrapolation (The score being a reflection of the real world performance)</td>
<td>If observed, is the attribute meaningful in medical student documentation? How meaningful is it for students to be scored, and/or provided feedback on, as part of the assessment on quality EHR documentation?</td>
</tr>
<tr>
<td><strong>What are the quality attributes of EHR written documentation by graduate medical students in their first clinical year?</strong></td>
<td>Scoring &amp; Extrapolation</td>
<td>Is this attribute observed in medical student documentation? How is this attribute interpreted? How do we accurately define this attribute to resonate with potential assessors and improve reproducibility and accuracy?</td>
</tr>
<tr>
<td><strong>What modifications to the PDQI-9 could be made to improve validity in this setting?</strong></td>
<td>Scoring</td>
<td>From consensus on interpretation and meaningfulness of the attribute, what potential modifications could be made to the PDQI-9 to make scoring more reliable and fair and improve the validity of assessment? What is the modified test domain?</td>
</tr>
</tbody>
</table>

Table 4.4 The relationship between the research questions, Kane’s inferences and the presentation of findings
4.4 Attributes that were similarly interpreted and meaningful in a medical student context

Introduction

Theme A includes attributes that were similarly interpreted and meaningful in a medical student context. The Stage 1 quantitative data grossly suggested similarities in scoring the attributes of Thorough and Synthesised across the seven assessors. The Stage 1 qualitative analysis demonstrated uniform agreement that these attributes were similarly interpreted and regarded as meaningful and relevant an attribute of high-quality medical student documentation. Thus, the Stage 2 Data collection was focused on confirming the findings from Stage 1, and explored potential modifications that could be made to the interpretation of the attribute in medical student documentation to improve the validity of the assessment.

4.4.1 Thorough

The original descriptor from the PDQI – 9 defined Thorough as being ‘the note is complete and documents all of the issues of importance to the patient.’

Figure 4.4 Assessor variance in scoring of Thorough across cases
The graph above depicts the variance of the raw scores from the application of the PDQI-9 to the documentation cases for the specific attribute, across the 11 cases. The variance in the scoring is plotted on the y-axis, labeled as ‘Assessor variance’, and the case numbers are the x-axis coordinates. For *Thorough*, the graph demonstrates that for all, save case 4, 10 and 11, the variance of scoring was under 1, and for the other three cases was less than 1.5. This broadly indicates there was a similarity in scoring *Thorough* across the seven assessors.

**Interpretation**

The Stage 1 qualitative analysis demonstrated that the interpretation of *Thorough* by the assessors was consistent with the PDQI-9 descriptor. Justifications for scoring were based on the completeness of the information provided in the document. Phrases commonly used when the scoring was low were limited or incomplete information, gaps in information, essential information missing. When the scoring was higher, justifications included a complete recording of some of the information and all necessary details included.

The focus in the document for the scoring of this attribute was the completeness of the history and examination sections of the presentation. The completeness of these sections was regarded as more important for medical students, and more emphasis was placed on these sections, in the development of this attribute.

Stage 2 data collection and analysis confirmed the consensus in interpretation.

**Meaningfulness**

The Stage 1 qualitative analysis supported *Thorough* as meaningful in a medical student setting and in fact assessors’ placed greater importance on thoroughness in medical student documentation compared to documentation recorded by graduated doctors.
Thoroughness is a meaningful attribute for assessing student records, as long as the students have been informed that this is one of the assessment criteria. This record had lots of opportunities for the student to enter the extra data that would be useful as record of the patient’s care and as a learning tool for his/herself (AA Case 008)

Stage 2 data collection and analysis confirmed that this attribute was extremely meaningful in the changed context.

Modifications

Given that there was consensus on the interpretation and meaningfulness of Thorough, Stage 2 data collection explored potential modifications that could be made to the interpretation of Thorough for medical student EHR documentation to improve the validity of the assessment.

Interviews continued the exploration of the theme from the Stage 1 data on the importance the assessors had placed on student documentation being thorough, over and above what would be expected of a physician. Stage 2 findings were that the assessors agreed that modifications should be made to the descriptor(s) of this attribute to focus on the importance of thorough documentation in the early years of training. They agreed it was important that the students in their early years of medical practice demonstrate their knowledge by being overly thorough on the history of the presenting complaint and examination. The students need to be advised of this emphasis and expectation, and the reasoning supporting this, in any resulting descriptors. Care would be needed in the description of Through, and expectations, as the student progresses through the course.

I think it is actually interesting because it is more meaningful the more junior they are (BB Interview)
.. we do say they need to be very thorough early on so we can see what they are thinking about and what information they have gathered. (FF Interview)

4.4.2 Synthesised

The original descriptor from the PDQI-9 defined Synthesised as being ‘the note reflects the author’s understanding of the patient’s status and ability to develop a care plan’.

Figure 4.5 Assessor variance in scoring of Synthesised across cases

Similar to Thorough, the Stage 1 quantitative data largely demonstrated minimal variance in scoring Synthesised. Seven cases had a variance of less than 1, and the remainder were less than 1.5.

Interpretation

In the Stage 1 qualitative analysis, the assessors were in uniform agreement with the interpretation of Synthesised and this corresponded with the first part of the PDQI-9 descriptor, ‘the note reflects the author’s understanding of the patient’s status’. Justifications often were phrased in terms of demonstration of clinical reasoning and interpretation of the presentation. Similar to Thorough, the focus was on the history and examination being written in such a way that clinical reasoning was evident, and this was often confirmed for the assessor by the inclusion of an appropriate
differential diagnosis. With reference to the PDQI-9 descriptor, the importance was in understanding the patient’s status, rather than developing a management plan.

Stage 2 data collection and analysis confirmed this interpretation.

Meaningfulness

In the Stage 1 qualitative analysis, this attribute was regarded as being extremely meaningful in a medical student setting and Stage 2 data collection and analysis confirmed the meaningfulness of Synthesised.

*Synthesis needs to be assessed; this is the only way we can see if the student has "got" what they are writing about (GG Case 008)*

*to actually interpret the information not just report but interpret and then convey something to the reader that says what is going to happen next not just what happened. (DD Interview)*

Modifications

Given that there was consensus on the interpretation and meaningfulness of Synthesised, Stage 2 data collection explored potential modifications that could be made to the interpretation of Synthesised for medical student EHR documentation to improve the validity of the assessment.

The discussion explored the emphasis placed on a synthesised history and examination and how this could be scored consistently. Assessors agreed that modifications should be made to reflect the importance of a synthesised history and examination in assessing this attribute and that the documents created were a written demonstration of the students’ clinical reasoning. Guidelines and descriptors for students needed to focus on how students could achieve both this objective and produce a high-quality document for use in a clinical setting.
a well-structured differential diagnosis is an example of synthesis so the differential that is appropriately ordered and targeted to the patient rather than just at textbook list is a synthesised differential diagnosis. A targeted history and examination is an example of synthesis as well (DD Interview)

4.5 Attributes that were interpreted in multiple ways and meaningfulness was inconclusive

Introduction

Theme B includes attributes that were interpreted in multiple ways and meaningfulness was inconclusive. There was a mixed variance for the attributes Organised, Comprehensible, Up to Date and Useful from the Stage 1 quantitative data. Organised, Comprehensible and Useful were not dissimilar in the level of variance in scoring to that of Synthesised and Thorough, seemingly indicating similarities in scoring among assessors. However, qualitative analysis of Stage 1 data demonstrated multiple interpretations of the attributes used in scoring and differing opinions of the meaningfulness of the attributes.

The Stage 2 data collection focused on clarifying the most appropriate interpretation of the attribute for medical student EHR documentation, and whether, with this agreed upon interpretation, it was meaningful to assess in medical student documentation. If there was consensus on meaningfulness, potential modifications to the PDQI-9 to improve the validity of the assessment were discussed.

4.5.1 Organised

The original descriptor from the PDQI-9 defined Organised as being ‘the note is well formed and structured in a way that helps the reader understand the patient’s clinical course’.
In the Stage 1 quantitative data, the scoring of *Organised* demonstrated six cases with a variance below 1.0, two cases greater than 1.0 and less than 1.5 and one case above 1.5. This data indicated an increased variance in scoring *Organised*.

**Interpretation**

The increase in variance in Stage 1 numerical data was partially explained by the interpretations assessors used to score *Organised*. *Organised* was interpreted in two different ways by the assessors in the Stage 1 qualitative analysis.

1. Organised use of the EHR. The correct information has been placed in the appropriate sections in the EHR.  
   
   Correctly uses the framework of the software program to enhance organisational structure/document form. Information is documented in the correct/expected sections (AA Case 003)

2. Organised presentation of the patient encounter. The student had presented the information logically and in an organised way. When this interpretation was used there was a key emphasis on the logical presentation of findings.
**sequential, logical presentation of the history (CC Case 005)**

Based on the dichotomy in Stage 1 qualitative findings, the objective of the Stage 2 data collection was to clarify the most appropriate interpretation of *Organised*.

It was agreed by the assessors that both interpretations were important to assess and provide the student with feedback. With regards to the quality of EHR documentation, there was a consensus that the interpretation of *Organised* should be focused on the organisation of content to be well formed and structured rather than to the use of a software program. There was agreement that students needed to have feedback on whether they demonstrated an organised use of the EHR platform, and that this would need to be addressed, but was not considered an attribute of high-quality EHR documentation.

*Meaningful*

In the Stage 1 qualitative analysis there were differing opinions on the meaningfulness of this attribute that arose from the problems with interpretation.

> *The basic organisational structure is forced by the EHR, this attribute has to be assess beyond this (BB Case 007)*
> *No, this is not very meaningful. Organised is not very relevant when students are presented with sections to populate (FF Case 009)*

In the Stage 2 data collection and analysis when *Organised* was defined in terms of a logical presentation of information there was a consensus that is was a meaningful attribute to assess.

> *I think Organised should reflect organisation of thinking not just the task or performed in a certain order (DD Interview)*
... what you are hoping is that how well organised the note is reflects how organised the student actually is in their performance of the task not just in how they write it down but that they can perform that task thoroughly and quickly (DD Interview)

There was discussion that an organised document was not necessarily a high-quality document, although a high quality document would have to be organised. The assessors agreed that the score on Organised may not correlate with an overall judgment of the quality of the document.

_to be well organised is a great starting point. It doesn’t necessarily mean the knowledge is there for a start and the content stems from the knowledge not from how well organised the presentation is_ (DD Interview)

_So yeah they could still put something in every box and they could make them sound quite logical but they may not actually be relevant, logical ..._ (CC Interview)

**Modifications**

Reaching consensus on the interpretation and meaningfulness was the focus of the Stage 2 data collection and analysis of the attribute Organised. To improve the validity of the assessment, there was a consensus that the PDQI-9 descriptor would have to be reworded to make the interpretation clearer to the students and the assessors. The preferred interpretation of Organised, was regarded as a higher level skill and developmental, and needed scoring or feedback along the continuum of attribute development. The organised use of the EHR was regarded as the student attention to detail and conscientious reporting, and not on a continuum. Providing the students with a comment that they had or had not organised their content in the appropriate fields in the EHR was recommended and agreed upon.
4.5.2 Comprehensible

The original descriptor from the PDQI-9 defined Comprehensible as being ‘the note is clear without ambiguity or sections that are difficult to understand’.

**Figure 4.7 Assessor variance in scoring Comprehensible across cases**

In the Stage 1 quantitative data, the variance in the scoring of Comprehensible was between 0.5 to 1.7, with seven cases below 1.0, two between 1.0 and 1.5, and two above 1.5. Numerical data indicated that there was an increased variance in scoring this attribute.

**Interpretation**

After Stage 1 qualitative analysis, Comprehensible was interpreted, and the scores justified by assessors, in two different ways.

1. Comprehensible in terms is what is written is understandable.

   *The little data that is entered is perfectly comprehensible (AA case 007 Scored 3)*

2. Comprehensible in terms of how what is written contributes to understanding.
I understood what happened (GG Case 001)

In the Stage 2 data collection and findings, the assessors confirmed that there were two interpretations used for this attribute when scoring the cases.

Comprehensible for me is I have read it and I understand it and the fact that I will be able to deliver excellent care having read the notes. And I wasn’t looking at the English or how the apostrophe or whether the spelling was right.... (HH Interview)

I treated it fairly simplistically actually whether I could understand it, just could I understand it? And I am not surprised that it didn’t correlate well with the rest because it could be a very, very superficial... (CC Interview)

As an attribute for high-quality documentation, there was a consensus that Comprehensible should mean that the reader understands the patient status.

But I also think that the purpose of the record is to comprehend what is going on in a patient and so it is of fundamental importance that after you read it you feel informed (BB Interview)

However, it was agreed that the students also need to be given feedback on their literacy skills, their use of abbreviations, English grammar, and spelling. There was particular concern about the overuse of acronyms.

If you know take for example how students record dosages of medications, we know that there have been awful errors where people have written ten units and it is people have thought that ‘u’ is a zero. And I think that you know there are certain things that can be misunderstood and so I think they do need feedback about this is appropriate terminology (BB Interview)

Meaningful
Given the difficulties in interpretation of this attribute, there was considerable discord about whether this was a meaningful attribute to assess in the Stage 1 qualitative analysis.

As a basis of something they are going to do a lot more in the future it is pretty important that the records are able to be understood by the various readers, so necessary for only agreed abbreviations to be used, and for the layout to be organised so that it is easy to follow. This I think should be one of the criteria that gets included in the set for assessment (AA Case 007)

As an attribute of high-quality documentation consensus was reached with Stage 2 data collection and findings that Comprehensible in terms of understanding the patient encounter was meaningful in a medical student context.

**Modifications**

Reaching consensus on the interpretation and meaningfulness was the focus of Stage 2 data collection. Similar to Organised, both interpretations, understanding the patient encounter and understanding the language used were regarded as relevant and important to provide feedback on. The interpretation on understanding the patient encounter was agreed upon as the focus for high-quality documentation, and there needed to be significant modification of the PDQI-9 descriptor to reflect this and improve consistency of scoring.

Although consensus on the interpretation of quality EHR documentation was reached, the assessors were insistent that there had to be some mechanism for providing feedback on the students’ use of the English language; whether it is correct, with few spelling errors and appropriate use of abbreviations, or shortening of words, and minimal jargon.

One of the themes for discussion for Comprehensible was the learning and development of the appropriate language to use when corresponding with other
health care professionals. There was agreement that part of *Comprehensible* was the development of medical technical language. This should be incorporated into descriptors for this attribute.

*I guess to use medical language maybe that is part of it. I am not talking even jargon but just the stylistically too, you know you are not writing a Jane Austen novel. You are writing something that is going to be punchy and to the point and therefore the reader can instantly get to grips with what you are trying to say...... Partly I think it is mastery of medical language... (DD Interview)*

4.5.3 *Up to Date*

The original descriptor from the PDQI-9 defined *Up to Date* as being ‘the note contains the most recent test results and recommendations’.

![Assessor variance on scoring of Up to Date across cases](image)

In the Stage 1 quantitative data, there was no relationship between the assessors in the scoring of *Up to Date*. The variance is extremely distorted and there was no pattern in the scoring of *Up to Date* across the cases. Four cases had a variance below 1.0, two between 1.0 and 1.5, and four cases over 1.5.
Interpretation

There was considerable variance in numerical scores for the judgment of *Up to Date*. This matched the analysis of the Stage 1 qualitative data. Assessors had difficulty interpreting this attribute. Written justifications were categorised into two interpretations.

1. *Up to date*, in terms of the management of the patient was in keeping with best evidence.

   *This record shows no evidence of being up to date- the only thing that would help might be a comment on the evidence for the possible options for treatment of sinusitis (AA Case 001)*

   *The file did not demonstrate appropriate recommendations or management issues relating to the R shoulder pain (HH Case 002)*

2. *Up to date*, in terms of timelines.

   *The record presents all relevant information up to the current time (FF Case 005)*

Stage 2 data collection and analysis was focused on gaining a consensus on the most appropriate interpretation of this attribute and whether this was meaningful to assess. The assessors confirmed that they had used two different interpretations of *Up to Date* in scoring documents.

   ... *I kind of looked at it as Up to Date in terms of for me to take over the management of the patient do I have the most recent information to help me ... (FF Interview)*

   ... *I am thinking chronologically most recent not most evidence based (DD Interview)*
The preferred interpretation was that pertaining to timelines and time course, and most recent results and investigations. Once again, the emphasis was placed on displaying the time course clearly of the presenting complaint.

(Up to Date) requires a student to consider everything to this point in time and the influence of everything to this point in time and what is going on at the moment (BB Interview)

Up to date in terms of an evidence-based management of the problem relied heavily on the management plan. It was agreed that this would not be as meaningful in students at an earlier stage of training, where the emphasis is on the clinical history and examination rather than management.

I think it is a step too far to go there. (In first-year clinical students) (FF Interview)

There was a debate about whether evidence-based recommendations were an attribute of high-quality documentation. This was not resolved.

Meaningful

In the Stage 1 qualitative analysis with the difficulties with interpretation most assessors found that this was not an attribute that held meaning for a medical student documentation.

not sure of its relevance at any stage in the case record process (GG Case 11)

Once the interpretation was clarified in Stage 2 data collection and analysis, there was uniform agreement that timelines were an important component of a quality documentation and that students should be assessed and given feedback on it.

Modifications
In Stage 2 data collection and analysis, it was agreed that clear timelines were an important attribute of high-quality documentation, and there was a consensus that the descriptor in the PDQI-9 was very narrow in its focus on test results and recommendations only. For medical students, the description of *Up to Date* should focus more broadly than this and encompass timelines in all parts of the document particularly the presentation and examination.

In addition, the preference was that this attribute should not stand alone as an attribute and be included in the description of another attribute, such as *Comprehensible, Organised or Thorough*. It was determined that *Up to Date* most closely linked to the organisation of content, and any definition of *Organised* should include reference to timelines. The theme of the correct placement of *Up to Date* was not exhausted by the interviews.

... if we are thinking *Up to Date* is about a timeline then I think that can easily be covered by your organisation (CC Interview)

4.5.4 *Useful*

The original descriptor from the PDQI-9 defined *Useful* as being ‘the note is extremely relevant providing valuable information and/or analysis’.
Stage 1 numerical data demonstrated minimal variance in the scoring across the cases for this attribute. All except for Case 010 had a variance under 1.

**Interpretation**

Similar to other attributes in this theme, the assessors utilised two different interpretations to score *Useful* in Stage 1. As a reader and assessor of the document, the interpretation of *Useful* relied heavily on what the decided purpose of the document was. Scoring and justifications were divided on useful to whom, an educator or a health professional caring for the patient. Although directly advised to consider it as the second, it was difficult for the assessors to entirely do this, given they were aware that the documents had been created in an educational setting.

*This record is not really useful for ongoing care and nor would it be much use for revision of cases for the student (AA Case 008)*

The little variance in numerical scores despite the two different interpretations, implies that if a document is useful in one setting, it is very likely to be useful in another. If the document is useful as an exemplar of a case for future personal reference, it is likely that the reader found it useful written communication of the
student’s interpretation of the case, and this would substantially increase its usefulness in the sharing and handover of patient care.

Justifications for scoring this attribute when the assessor was considering usefulness to the student in a learning context included the usefulness for recalling an encounter, and as an exemplar of the clinical reasoning.

*This record is useful to students as a record of the interaction, a clear demonstration of clinical reasoning, identification of learning goals that have arisen from this encounter and then how these will be or were met (FF Case 002)*

*Limited utility beyond self. Minimally useful to others, some important information is conveyed (BB Case 004)*

When the document was scored with the underlying criteria being usefulness to other health care providers, justifications were in terms of the need for further clarification prior to taking on the care of the patient.

*The document was useful to others, although a few issues would need further clarification by 'reader' (BB Case 002)*

*This record contains information that is adequate for the patients' ongoing management (AA Case 006)*

The objective of the Stage 2 data collection was to clarify the interpretation of Useful in a medical student setting. The assessors interviewed confirmed how they had interpreted Useful and that Useful was closely linked to the perceived purpose of the document.

*I was only thinking of it from the reader’s point of view (CC Interview)*
The usefulness was whether this record if you went back the next day and let us say you got hit in the head and you had amnesia you could read the notes and say, “Okay, I know what I need to do.” ... And it is useful then in a hospital setting in regards to transfer of care between providers and allied health specialists (IH Interview)

Given that students are being assessed on high-quality EHR documentation, there was agreement that this attribute predominantly needed to focus on usefulness to a health care provider, however. Ideally, it would be useful to the student if the document was useful for future studies in its own right. However, the consensus was that the usefulness to the students was the task and the feedback they received from this.

**Meaningful**

After Stage 1 qualitative analysis, regardless of the interpretation, this attribute was regarded as meaningful by all the assessors. After clarifying the interpretation, Stage 2 data collection and analysis confirmed the meaningfulness of *Useful*.

**Modifications**

Following on from the theme in the Stage 1 analysis of the purpose of the document, it was agreed in Stage 2 analysis that the purpose of the documentation had to be very clearly defined for the students and assessors and that there needed to be some consideration of the educational context.

With a focus on usefulness to other health care providers, and considering the stage of the students’ learning, the emphasis should be that the reader finds the account of the history and examination and subsequent differential diagnosis useful to them, with little need for clarification of these aspects for ongoing care.

*In (first year clinical) to make I think the minimum criteria for ... a document would be able for a fellow colleague to read and know the symptom or the*
condition and from that information examination findings, progress notes that they too would be able to agree or disagree with your differentials (HH Interview)

The theme of Useful to the student and how this could this could be described was explored. It was hoped, that part of the process of improving the validity of the assessment of high-quality documentation would be that the student was provided with authentic and relevant assessment and feedback that they would value and deem worthwhile. Other opinions on descriptors for this were in accord with Stage 1 findings, namely the usefulness of the document for students being as a recall of an interaction, the demonstration of their clinical reasoning and an exemplar of a presentation.

4.6 Attributes that were difficult to interpret and meaningfulness was unlikely

Introduction

The attributes in Theme C were difficult to interpret and meaningfulness was unlikely. Assessors had difficulties interpreting Succinct and Internally Consistent, revealed primarily by the Stage 1 qualitative data. Given the difficulties in scoring, there was resulting doubt about the meaningfulness of these attributes in assessing medical student documentation quality. Therefore, Stage 2 data collection focused on clarifying whether it was possible to reach agreement on interpretations for these attributes and whether it was possible and meaningful to assess this on medical student EHR documentation. If there was consensus on meaningfulness, potential modifications to the PDQI-9 to improve the validity of the assessment were discussed.

4.6.1 Succinct

The original descriptor from the PDQI-9 defined Succinct as being ‘the note is brief to the point, and without redundancy’.
In Stage 1, numerical data for *Succinct* indicated that there was less variance in scoring this attribute, similar to *Synthesised*. Seven cases had a variance less than 1 in the score, and the remainder were all under 1.5.

**Interpretation**

From Stage 1 qualitative analysis, all of the assessors found this attribute difficult to judge in medical student EHR documentation. There were a number of reasons for this.

1. The assessors wanted the students to be overly thorough in their documentation, as discussed previously. In scoring the document, assessors placed great importance on thoroughness early in clinical training as a demonstration of clinical reasoning and understanding. They would rather the documentation was overly inclusive. The assessors tended to score high on *Succinct* as their expectations of demonstrating this attribute, and the importance of it in student documentation were very low at this stage of learning.
It is a skill learned with practice, better too many words initially than too few; (GG Case 001)

2. To be able to judge that a document is appropriately succinct, and judgment has been made by the writer on what information needs to be included and can be excluded, the reader needs to make a decision on the validity of this judgment and the expertise of the writer. With a medical student, the assumption that this judgment has been made accurately is difficult to make.

Succinctness is a higher order behaviour/judgments that involves value attribution and is not brevity per se. Biggest issue is that it is a higher level function and not brevity (GG overall comment in Stage 1)

Given the difficulties, most of the assessors defaulted to defining the attribute as an assessment of brevity. Poor quality documents with minimal content scored high.

Most of the records, this one included, are far too succinct! (AA Case 8 – scored 5)

Not necessarily a positive attribute, but definitely succinct (DD Case 007 – scored 2)

Although stated in the descriptor of Succinct in the PDQI-9, very few focused on the redundancy of language.

Too many words; redundant use of the language rather than demonstration of higher order thinking (GG Case 003)

The difficulties with interpretation resulted in consistently high scoring. If the assessment was in terms of the word count, most of the cases scored high. If it was in terms of a value judgment for the inclusion and exclusion of information, the assessors tended to give the students the benefit and assumed that there was a deliberate
reason for not documenting relevant information. These factors explained the low variance in scores seen with the numerical data.

The objective of the Stage 2 data collection was to clarify the interpretation of *Succinct* and whether this was meaningful to assess in medical students. Few justifications of scoring considered redundancy of language, as part of the descriptor for this attribute in the PDQ1-9, and this was to be clarified in the interviews given the considerable concern about this in the literature.

Stage 2 analysis clarified that the assessors found this a difficult attribute to assess in a medical student setting.

So I found that hugely, hugely challenging is it succinct because they are just giving us the relevant information or is it succinct because they have missed a whole lot of other possibilities? And it is really hard to tell without knowing the patient (BB Interview)

*In order to assess Succinct, you have to have some idea of the expertise or knowledge of the individual (BB Interview)*

After discussion of the literature and a consideration of the development of this attribute and how it may be demonstrated in medical student documentation, a consensus was reached that the focus of the interpretation should be on the redundancy of language and to tailor the information written to the purpose of the document, rather than being concise and brief.

*If they are going to prepare useful medical documents later on then there has to be some attention paid to whether it is succinct or not because you lose the reader if you are putting together a 20-page discharge summary (DD Interview)*
There is a difference between thoroughness and redundancy. So what I suppose I would like to see evolving in Year 2. This is from the assessor, is an elimination of redundancy so that it is still overly thorough (BB Interview)

Meaningful

From the Stage 1 qualitative analysis, this attribute was not deemed to be meaningful in a medical student setting. After Stage 2 data collection and analysis, it was agreed that if the focus was on reducing redundant language and tailoring the information provided to the purpose of the document, this would be a meaningful attribute to assess in medical student documentation.

but I think we should be pushing that they get that by the end of (first year clinical). What is relevant and what is not? We push that all the time with all our case discussions, all of our feedback about their case presentation is this relevant or not? (FF Interview)

Modifications

Stage 2 data collection explored the theme of the relationship between the development of Succinct and Thorough as attributes. The consensus was that there was a continuum between the two attributes that continued to change through the entire training of a doctor, from medical student to a fully qualified physician. Superimposed on this were the expectations from the specialty in which the student, or doctor, was training. In order to assist students to understand this, and for assessors to score these attributes, expectations had to be clearly described at all levels of training. The interpretation agreed upon for this study was focused on the skill level expected for a first-year clinical student.

I would definitely encourage that in medical students as first be thorough and if that means that the result is you collect extraneous information that you might have to go back and weed out a little bit later fine that is okay. That is the
preferential over missing information because you are trying to be too succinct

(DD Interview)

There was a discussion about the concern from the authors of the PDQI-9 that the instrument did not pick up on the issue of ‘copy and paste’ from other documents and Originality. The assessors agreed that the issue of copy and pasting redundant information into notes, and the resulting ‘note bloat’, was a concern from their own clinical experiences.

Yeah. But what happens with discharge summaries is that every single result that was ever done in copied and pasted so I have been reading discharge summaries of 22 pages (AA Interview)

As it was not possible to identify this behavior accurately with the documents they were provided, the discussions were inconclusive on how this should be assessed and what educational program could be set up to easily monitor for this.

4.6.2 Internally Consistent

The original descriptor from the PDQI-9 defined Internally Consistent as being ‘no part of the note ignores or contradicts any other part’.
In the Stage 1 quantitative data, there was only one case that had a variance under 1.0. Six cases had a variance above 2.0, and three between 1.0 and 1.5. There was no pattern in scoring between assessors.

**Interpretation**

From the Stage 1 qualitative analysis, assessors found it very difficult to assess *Internally Consistent*. Most assessors justified the scores with descriptions of the information provided being consistent within the document, and that there were no direct inconsistencies. This attribute frequently scored highly and the less information provided by the student, the less likely the document was to contradict itself, and the higher the scoring.

*So little information it would be difficult not to be internally consistent (AA Case 007)*

*internal consistency is difficult to assess as a positive feature, and is much easier as a negative when detected. There was insufficient data from which to generate much incongruence (GG Case 009)*
One assessor interpreted *Internally Consistent* being consistent with coding and diagnosis and the use of the EHR.

*This criterion was overall poorly scored for the 11 records. Coding is vital in EHR and diagnosis means just that. e.g. Angina, Peptic Ulcer Disease, Pneumonia as opposed to a symptom - chest pain, epigastric pain or productive cough…. (HH Case 011)*

The difficulty in scoring this attribute led to a wide range of scores.

The objective of the Stage 2 data collection was to clarify the interpretation of *Internally Consistent* and whether this was meaningful to assess in medical students. Clarification was sought on whether it encompassed coding. This attribute continued to be difficult to reach a consensus on its interpretation.

*Didn’t feel that useful from what the tasks that I performed and the records that I looked at I don’t know if I ever marked Internally Consistent down (DD Interview)*

*Meaningful*

From the Stage 1 qualitative data analysis, given the confusion in interpretation, assessors did not think this was a meaningful attribute to assess.

After Stage 2 data collection and analysis, *Internally Consistent* was not deemed meaningful in a medical student context to stand alone as an attribute by the majority of assessors. If it was intended to identify inconsistencies within a document then it was agreed that this did not warrant an attribute in its own right, and could be included in the assessment of another attribute.

*... it is probably looked after a little bit by the Useful and certainly looked after by the organisation part I think. I mean again because that is about the logic (CC Interview)*
it fits to me with, fits with Organised, Comprehensible and Internally Consistent. I think they are altogether some sort of tick the box, yes or no (FF Interview)

One assessor maintained individual views on Internally Consistent. They interpreted Internally Consistent as the documentation being written with consistent use of the EHR, including the correct coding of the condition or diagnosis. The other assessors although agreeing that coding was important, were unable to reach consensus on this opinion.

It is that Internally Consistent with appropriate structure as a basis of an outstanding e-record.... While I look at it as also Internally Consistent is how all the records are standardised to a way of doing things (HH Interview)

The assessors were being asked to assess individual documents, outside the context of an entire health record. It was appreciated that this attribute might hold more importance if the document was being assessed in the context of an entire health record.

4.7 Other issues arising from Stage 2 analysis

Professionalism

There was concern that the attributes did not provide feedback or assessment specifically on Professionalism. There was a consensus that a high-quality document should meet professional standards, and that the student should be advised of this. If there were any concerns about professionalism this needed to be flagged.

Low word count

The assessors agreed that minimal documentation distorted the assessing of the qualities. In an educational setting, there needed to be the option to advise the students that there was insufficient information on which to make an assessment of performance and they needed to resubmit a revised document.
Accurate

As explained in the Methods chapter, Accurate could not be assessed in this study. The assessors agreed with the literature that this was definitely an attribute of high-quality documentation. They also concurred with the literature that it is a logistically difficult attribute to assess in an educational context, where the assessor often does not know the patient or have access to the health services’ EHR.

Overlap of Attributes

Throughout the analysis, there were many comments about the overlap between the attributes. The assessors did not identify consistently specific overlaps. Many assessors commented that the attributes could probably be reduced after further validation of any assessment rubric. Quality EHR documentation is an abstract concept, and it was not possible to get agreement amongst the assessors on what attributes would be reliably included or encompassed by other attributes. Many commented on the need for further validation studies to clarify this.

*I think the Comprehensible, Synthesised and Internally Consistent all appear to be a bit inter-related as well (BB Interview)*

*As stated in previous comment, this criteria (Comprehensible) is closely linked with Thorough, Useful and Organised (HH Case 011)*

4.8 Summary

From the numerical data from Stage 1 of this study, the PDQI-9 in its current format was not valid in to assess the quality of medical student EHR documentation. Thematic analysis of Stage 1 narrative justifications for scoring resulted in the attributes being categorised in three themes.
A. Attributes that were similarly interpreted and meaningful in a medical student context

B. Attributes that were interpreted in multiple ways and meaningfulness was inconclusive

C. Attributes that were difficult to interpret and meaningfulness was unlikely

Stage 2 data collection was based on these themes and analysis sought to deepen the understanding of the attributes of quality by finding relationships between Stage 1 analysis, literature findings, and Stage 2 data, and building upon the data collection of Stage 2 to reach consensus. The following chapter will discuss the conclusions drawn from the findings and the study in general. An overview of the findings is presented in Table 4.4.

Table 4.5 Summary of findings

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Relevant</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td>The purpose should be clear and focus on documentation skills.</td>
</tr>
<tr>
<td>Thorough</td>
<td>+</td>
<td>Clearly in and emphasised</td>
</tr>
<tr>
<td>Synthesised</td>
<td>+</td>
<td>Clearly in and emphasised</td>
</tr>
<tr>
<td>Succinct</td>
<td>+</td>
<td>Focus on lack of redundancy and tailoring information provided to purpose of the document</td>
</tr>
<tr>
<td>Useful</td>
<td>+</td>
<td>Clarification of useful to both reader and student</td>
</tr>
<tr>
<td>Organised</td>
<td>+</td>
<td>Clarification meaningful interpretation is organisation of content</td>
</tr>
<tr>
<td>Comprehensible</td>
<td>+</td>
<td>Clarification meaningful interpretation is understanding the patient encounter with an emphasis on semantic competence</td>
</tr>
<tr>
<td>Up to Date</td>
<td>+/-</td>
<td>Included within another attribute</td>
</tr>
<tr>
<td>Accurate</td>
<td>+</td>
<td>Included on basis of literature review</td>
</tr>
<tr>
<td>Professionalism</td>
<td>+</td>
<td>Respect for patients and the health care team and setting</td>
</tr>
<tr>
<td>Internally consistent</td>
<td>-</td>
<td>Not meaningful as an attribute in its own right</td>
</tr>
</tbody>
</table>

5 Discussion

This chapter is divided into a discussion of the results of the study, and how it addresses each of the research questions, in turn, limitations of the study and directions for future research. Many of the future directions for research from this
study pertain to completing the validity argument framed by Kane’s concepts, with the gathering of more evidence directed at evaluating scoring and extrapolation as well as Generalisation and Implications inferences. As a reminder, the latter two inferences are summarised in Table 5.1.

Table 5.1 Kane’s inferences of Generalisation and Interpretation

<table>
<thead>
<tr>
<th>Kane’s Inferences</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generalisation</strong></td>
<td>How well do the selected test items in our test domain represent all of the</td>
</tr>
<tr>
<td>deals with test world</td>
<td>theoretically possible items in the relevant assessment universe? Addresses</td>
</tr>
<tr>
<td>performance</td>
<td>reproducibility, reliability</td>
</tr>
<tr>
<td><strong>Implications</strong></td>
<td>How the interpretation of the assessment is used for a specific use, decision</td>
</tr>
<tr>
<td>deals with the</td>
<td>or action?</td>
</tr>
<tr>
<td>consequences of the</td>
<td></td>
</tr>
<tr>
<td>assessment</td>
<td></td>
</tr>
</tbody>
</table>

Observable attributes were judged to be important to include in a test domain assessing the quality of EHR documentation in medical students based on rules defined by Scriven (1994) on refining and extending the range of synthesis inferences to judge criteria of merit. His rules are outlined in Table 5.2.
#### Table 5.2 Rules for criteria of merit modified from Scriven (1994)

<table>
<thead>
<tr>
<th>Rules for Criteria of Merit</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensive</td>
<td>The list has to cover everything related to quality in EHR documentation</td>
</tr>
<tr>
<td>Non-overlapping</td>
<td>None of the items on the list can be counted twice, or overlapping</td>
</tr>
<tr>
<td>Criteria of Merit</td>
<td>Only properties that are meaningful to the concept of a quality should be included</td>
</tr>
<tr>
<td>Commensurate</td>
<td>All items are at the same level and breadth</td>
</tr>
<tr>
<td>Observable</td>
<td>The criteria should be described well enough to determine their presence</td>
</tr>
</tbody>
</table>

#### 5.1 Is the application of the PDQI-9 valid for the assessment of EHR documentation recorded by medical students?

From the Stage 1 quantitative data, the PDQI-9 was not valid for the assessment of EHR documentation recorded by medical students. Stage 2 data analysis produced evidence that supported the majority of the attributes included in the PDQI-9 as being relevant to medical student assessment.

**Scoring**

Despite the small number of assessors and cases, the numerical data indicated that there was significant variation in scoring the individual attributes, and highlighted areas that needed further clarification to determine the test domain in medical students. Notable variation in scoring attributes was demonstrated, even for attributes where there was uniform agreement of their interpretation. The variation in the numerical score of these attributes more than likely reflected issues with benchmarking rather than an issue with the interpretation of the attribute. The justifications for the given attribute score in specific student documents were similar, despite the number attributed. Even though a training package was developed, and assessors were selected for their expertise in medical student education and their
perceived knowledge of the standards and expectations of students at different levels of training, their expectations varied. Future studies are needed to further define the scoring inference, and to set standards and benchmark the performance in order to improve the validity of any scoring associated with the assessment, or support any decision to be made on competency (Norcini & Shea, 1997).

Extrapolation

Although the assessors’ initial scoring demonstrated significant variability, with further exploration in Stage 2, most of the PDQI-9 attributes were judged to be a fair representation of the attributes that would be expected and should be assessed, in the performance of EHR documentation by medical students. After Stage 2 analysis, the attributes were placed into three categories.

First, those that were a fair representation of the attributes that would be expected, and should be assessed, in the performance of EHR documentation by medical students. Most of the PDQI-9 attributes fell into this category: *Thorough, Synthesised, Organised, Comprehensible, Useful, and Succinct*. The assessors recommended modifications to improve the validity of the assessment of these attributes.

Second, two attributes of the PDQI-9 were considered less important or unrelated to the assessment of quality in medical student EHR documents: *Internally Consistent* and *Up to Date*. Finally, an attribute that encompassed *Professionalism* was observable, was not in the PDQI-9 and needed inclusion in a test domain to assess medical student quality EHR documentation. This finding was in keeping with the literature review and expert opinions of EHR documentation (Gliatto et al., 2009; Hammoud et al., 2012; Schenarts, P. J. & Schenarts, 2011; Stephens et al., 2011; Tierney et al., 2013; Usatine et al., 2002)

In summary, the PDQI-9 as written was not valid in a medical student context. However, with clearer interpretations of the attributes that recognised the level of training of the medical student and likely skill acquisition, the attributes broadly
defined the test domain in medical students with minor inclusions and exclusions. Clear interpretations of the attributes and modifications on the test domain are discussed in the following section.

5.2 The quality attributes and how they can be modified

This section of the discussion will address the second and third questions of the study. The second research question, and step in the validity argument addressed what the observable attributes were, how they are described, identified and assessed in student EHR documentation. For authenticity, our study attempted to capture the language of the assessor to fully describe the observable attributes. The third research question stemming from this were modifications to the PDQI-9 test domain that would improve validity in the assessment of medical students. An overview of the discussion is presented in Table 5.3.
Table 5.3 Overview of findings and modifications proposed

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Decision</th>
<th>Recommendations</th>
<th>Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td>Level of training: All attributes to focus on presentation, history, examination and differential diagnosis</td>
<td>In description of attribute, include focus</td>
</tr>
<tr>
<td>Thorough</td>
<td>Include</td>
<td>Emphasis on overly thorough</td>
<td></td>
</tr>
<tr>
<td>Synthesised</td>
<td>Include</td>
<td>Emphasise importance</td>
<td></td>
</tr>
<tr>
<td>Succinct</td>
<td>Include</td>
<td>Focus on lack of redundancy and tailoring information provided to purpose of the document</td>
<td>Clear description of attribute using language of assessors</td>
</tr>
<tr>
<td>Useful</td>
<td>Include</td>
<td>Clarification of useful to both reader and student</td>
<td></td>
</tr>
<tr>
<td>Organised</td>
<td>Include</td>
<td>• Clarification interpretation is organisation of content</td>
<td>• Clear description of attribute using language of assessors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The organised use of the EHR not criteria of merit for quality</td>
<td>• Unsure if commensurate with other attributes - potentially ungraded</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The organised use of the EHR not criteria of merit for quality</td>
<td>• Additional to the assessment of quality, students should receive feedback on organised use of EHR</td>
</tr>
<tr>
<td>Comprehensible</td>
<td>Include</td>
<td>• Clarification – understanding the patient encounter</td>
<td>• Clear description of attribute using language of assessors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Emphasis on semantic competence</td>
<td>• Additional to the assessment of quality, students should receive feedback on English language, acronyms, abbreviations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• English language, acronyms, abbreviations not a criteria of merit for quality</td>
<td></td>
</tr>
<tr>
<td>Up to Date</td>
<td>+/-</td>
<td>• Part of the description of Organised</td>
<td>• Clear description of attribute using language of assessors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Seek further clarification on whether evidence-based documentation is a criterion of merit</td>
</tr>
<tr>
<td>Accurate</td>
<td>Include</td>
<td>Included on basis of literature review</td>
<td>• Further exploration of Scoring inference of Attribute to develop clear description of attribute using language of assessors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Needs further scoring exploration</td>
</tr>
<tr>
<td><strong>Professionalism</strong></td>
<td>Include</td>
<td>Respect for patients and the health care team and setting</td>
<td>• Not commensurate to other attributes – important, potentially a hurdle</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------</td>
<td>-----------------------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Internally consistent</strong></td>
<td>Exclude</td>
<td>No part of the note ignores or contradicts any other part.</td>
<td>Not a criteria of merit</td>
</tr>
</tbody>
</table>
The descriptions of the attributes in the original PDQI-9 were brief, and as evidenced by this study, open to multiple interpretations. We anticipated the difficulties with achieving validity of a global assessment in the justifications for selection of the PDQI-9, as compared to a more objective measurement of quality (van der Vleuten & Schuwirth, 2005). In line with assessment literature, we agreed that a global assessment based on assessor judgment, was a more faithful reflection of expertise and this was worth the extra rigor required to capture assessor language and judgments, using multiple methods to obtain data and triangulate results.

In collecting the evidence, we were mindful that the validity of the final test – score interpretation in a modified test domain would be largely dependent on how the attributes are defined (Cook et al., 2015; Kane, 2005). In the proposal of modifications to an assessment instrument with the intended outcome being an improvement of the validity of an instrument assessing a complex performance, the attributes had to resonate with assessors as true descriptions of the behavior being scored and allow for their professional judgment (Crossley & Jolly, 2012). To describe in full, the language of the assessors, is beyond the scope of this thesis, however, how this has been practically applied to a modified assessment instrument for the Patient Record Program at the University of Melbourne at the end of the Thesis in Appendix F.

In general, there were three main considerations when clarifying the description of the attribute, namely the level of training of the student, the development of the attribute in documentation and as a result, the most relevant interpretation for the attribute.

5.2.1 Level of training

Richer, more resonant descriptors of the attributes required considered opinion on the development of the attribute in documentation and corresponding training level of the student. Very little information was available about the development of these attributes over the training life of a medical professional. Guidelines outlined some of the behaviors expected at different levels of development in the use of the EHR
(Hammoud et al., 2012; Stephens et al., 2011;). These expectations were considered in reaching consensus with interpretations.

Additionally, there was uniform agreement that for all relevant attributes it had to be clear that the focus and priority of assessment was demonstrating these attributes in the documentation of the presentation, examination and differential diagnosis in first clinical year medical students. Stephens et al. (2009) described the expectations of students at this level of this study as being between reporters and interpreters and placed the same emphasis in their recommendations.

5.2.2 Test domain for assessing quality EHR documentation in medical students

*Thorough* and *Synthesised*

The quality attributes of EHR written documentation that were agreed upon as important to assess were *Thorough* and *Synthesised*. The majority of related studies on EHRs, whether stated directly or implied, supported complete and thorough documentation, that demonstrated narrative expressivity of clinical reasoning and synthesis (Gliatto et al., 2009; Mamykina et al., 2012; Rosenbloom et al., 2011; Schenarts, P. J. & Schenarts, 2011; Stephens et al., 2011; Tierney et al., 2013; Usatine et al., 2002).

One other notable modification in the interpretation of *Thorough* in medical students, in addition to the focus on the presentation and diagnosis, was the emphasis the assessors placed on this attribute, over and above what they would expect for a physician document. This was in keeping with reviewed guidelines and other expert opinions (Gliatto et al., 2009; Hammoud et al, 2012; Schenarts, P. J. & Schenarts, 2011; Stephens et al., 2011; Tierney et al., 2013; Usatine et al., 2002). When the writer of a document is junior, often the reader is uncertain about their level of expertise. Part of the performance of EHR documentation for medical students was to demonstrate that all relevant information has been asked or examined, and not overlooked. Any
resulting descriptor had to make it obvious that the student documentation had to be overly thorough, and include all relevant information.

**Synthesised**

The critical importance of *Synthesised* as a criteria of merit in assessing quality EHR documentation in students was reinforced by this study. As presented in the literature review, there is a growing body of work supporting documentation being an information synthesis activity (Loop & Han, 2013; Mamykina 2014; Shoolin et al., 2013). Medical professionals value expressivity of reasoning and thought process to convey the writer’s impression, reasoning and thought processes (Rosenbloom et al., 2007). Similar to the expectation of medical students and *Thorough*, it was essential that medical students and junior doctors, displayed their clinical reasoning to senior medical professionals in their documentation, as this could not be assumed.

**Succinct**

After refinement, *Succinct* was determined to be a criteria of merit. Assessors found *Succinct* extraordinarily difficult to judge initially with the PDQI-9 descriptor. Similar to the findings from the validation of the PDQI-9, the assessors did not regard *Succinct*, in terms of redundancy (Stetson et al., 2012). The assessment in this study and of the PDQI-9, was in direct contrast to the concern in the literature about the loss of concise, relevant documents in EHRs and the increase in redundant language, bloated notes and ‘copying and pasting’ (Hammoud et al., 2012; Heiman et al., 2014; Stephens et al., 2011).

The development of *Succinct* was regarded as being incremental with the initial skill acquisition being the reduction of irrelevant and redundant language rather than conciseness and brevity. As the student was introduced to the different health care documents, tailoring the information provided in their documentation to the purpose of the document was the next stage in the early development of being succinct. The meaningful interpretation of *Succinct* for medical students was a lack of redundant
information and tailoring information provided to the purpose of the document. The relationship between the development of the attributes, *Succinct* and *Thorough*, needs further elaboration if the modified instrument is to be adapted to further training levels.

With respect to ‘copying and pasting’ large amounts of redundant text into documents, the assessors recognised this as a concern but were inconclusive about how a simulated program in an educational setting would be able to assess for this in any meaningful way. Copying and pasting was unable to be identified in the documents in this study. Stetson et al. (2012) recommended that a separate *Redundancy score, or Originality assessment*, was made in addition to the PDQI-9 to assess this feature. Further studies, focusing on extrapolation, will need to address how *Originality* can be assessed and the implication inferences of doing so.

*Useful*

An extension of the interpretation of *Useful* to incorporate the educational setting that the students were in was necessary to include *Useful* as a criteria of merit. Despite the focus on high-quality documentation, given that the task was in an educational context, *Useful* must also be seen in terms of usefulness for the writer, the student, as well as the reader. The fact that the assessment task has a defined purpose and use to the student contributes directly to the validity of the assessment (Cook et al., 2015).

*Organised*

*Organised* was agreed upon as a criteria of merit if its description focused on the organisation of the content in the EHR within the fields. Feedback on organised use of the EHR was important but not regarded as an attribute of high-quality documentation. A further finding was that *Organised* may not be commensurate with other attributes of quality. A well-organised document may still be of low quality, although a high quality document would have to be well organised. Studies focused on refining Scoring, and Generalisation inferences, are needed to test these assertions.
when seeking to broaden the validity argument for the instrument beyond content-related validity. It may eventuate that Organised is ungraded and a hurdle rather than contributing to an overall numerical score.

Comprehensible

Comprehensible was an agreed criteria of merit if the descriptor referred to understanding the documentation in terms of the status of the patient. This attribute included the development of semantic competence and the nuanced language learnt to communicate with other health care professionals in the descriptor. These findings echoed the discussions in the literature around narrative expressivity and the value medical professionals place on linguistic nuance and the appropriate description of aspects of the patient encounter (Rosenbloom et al., 2007, 2011). No studies could be identified exploring the development of semantic competence in training.

Similar to Organised, the alternate interpretation of Comprehensible was not regarded as a criteria of merit, namely the use of the English language, shortening of words and abbreviations. This alternate interpretation remained a concern, and the impact of poor language skills on the readability of documentation. Wanting the students to receive feedback directed on this aspect of documentation, outside the assessment of quality, was in keeping with the international concerns (Gliatto et al., 2009; Hammoud et al., 2012; Stephens et al., 2011).

Up to Date

Up to Date was a criterion of merit that was seen as overlapping, and contributing to the understanding of the performance of Organised. The agreed upon interpretation of Up to Date pertained to clear timelines and time course within the documentation.

The alternate interpretation of Up to Date was that the documentation was evidence based. This might be a criteria of merit but the findings were inconclusive and the literature was not supportive. Further exploration is needed on whether having evidence-based documentation constitutes an attribute of quality, or whether it
overlaps with other attributes, and if it does need assessment, how this would be performed, observed and assessed.

**Accurate**

With the overwhelming agreement in the literature of the importance of accurate documentation, *Accurate* can be regarded as a criteria of merit without needing substantiation from this study. However, it was a limitation of our study that *Accurate* could not be assessed, and its interpretation clarified. Further study needs to focus on a clear interpretation of the attribute, *Accurate*, to determine if what it describes is accurate documentation of what the student actually did or accurate identification of the features of the patient presentation. In addition, similar to the issues with *Originality*, extrapolation studies need to address the juxtaposition of the difficulties with assessing this attribute in the assessment world, and the importance placed on it in the real world (Boulet et al., 2004).

**Internally Consistent**

*Internally Consistent* was not listed as a criteria of merit. It was deemed to be unrelated and irrelevant to medical student assessment, and assessors were unable to reach an interpretation of this attribute that would make it meaningful to assess. In addition, we found no support in the literature for this attribute beyond the studies by the developers of the PDQI-9 (Stetson et al., 2011)

The potential reasons for *Internally Consistent* not being regarded as a criteria of merit based on our study were the format of the document and the selection of assessors. The documentation provided were single documents, without reference to the patient’s whole EHR. *Internally Consistent* may be more meaningful if the document was part of the whole story of the patient, and was assessed in relation to other elements and documents within the record. The assessors were specifically selected for their expertise in medical education, and their use of the EHR being as physicians. They had to be familiar but not expert in EHR documentation. This attribute may be
more clearly defined and thus given greater merit if the assessors had broader skills in EHR use. Further studies that address Extrapolation and Implication inferences of the assessment need to include the viewpoints of a wider range of stakeholders. At this stage of understanding, Internally Consistent should be excluded from a modified instrument.

**Professionalism**

An important observable attribute of a quality document, that the assessors did not judge as having been assessed with the PDQI-9, was Professionalism. In the assessment of quality, documentation should be professional, demonstrating honesty and respect for patients and the health care team and setting. This was in keeping with recommendations from the literature that documentation be used to evaluate Professionalism (Gliatto et al., 2008; Schenarts, P. J., 2011; Stephens et al., 2011; Tierney et al., 2013; Usatine et al., 2002; Wald et al., 2014). Further exploration on the Scoring inference of a test item for Professionalism is required. Assessors were undecided whether Professionalism would stand alone as an attribute of quality documentation or be ungraded with feedback. There was a consensus that it should be a hurdle, unprofessional documentation was immediately unsatisfactory, and required faculty attention.

**Overlapping attributes**

Nearly all assessors gave opinions on observable attributes overlapping. There was agreement that the attributes could be reduced in number, and still be inclusive of all the observations that constituted a quality performance. However, there was no consistency as to what attributes would be seen as contributing to another. This, once again, reflected the difficulty in defining a complex abstract task with a global assessment. In studies focused on Generalisation, the test item scores will need to be closely examined and correlated. At this stage of the validity argument, the attributes as modified and defined are all important criteria of merit to be listed.
Further modifications

Over and above what has already been discussed, the assessors recommended that in any rubric created the ability to request the student to resubmit the task based on minimal information provided should be included. This recommendation stemmed from the fact that with an assessment of documentation with a low word count, the observation of the attributes distorted, for example, brevity with succinct, comprehension of minimal information, the organisation given the content matter is minimal. Some of these concerns may be resolved by the clearer interpretations of the attributes, in particular, *Succinct*. Notwithstanding this, the recommendation was that if not enough information had been documented to make a judgment on the performance there should be a process of resubmission, rather than the assessor completing the rubric.

Summary

In accordance with Scriven’s (1994) rules on criteria of merit, we have refined and synthesised the findings from this study to provide a test domain for assessing the quality of EHR documentation in medical students. Assessors reached an accord that the PDQI-9 test domain was broadly valid if the interpretation of the attributes was made clearer and relevant to the level of training of the student. With minor exclusions and inclusions, the PDQI-9 test domain was judged to reflect the key aspects of real performance in medical students. The amended descriptors have married the level of training of the student with the expected development of the attribute during training.

With the evidence provided for the test domain, in keeping with the assessment literature on assessor language and professional judgments, the modified PDQI-9 should result in improved validity of assessment (Crossley & Jolly, 2012).
5.3 Limitations and Implications for Future Study

There were a number of limitations to this study. We have already discussed the inability of this study to judge Accurate, the other limitations are as followed.

Assessors

As described in the methods chapter, we made an informed decision on the selection of the assessors with respect to the purpose of the study. For determining content related validity the numbers were adequate (Norcini & Shea, 1997; Davidson, 2005). A larger number of assessors, with a broader representation nationally and internationally, would strengthen the generalisability of these findings. Furthering the validity argument will require broader sampling and triangulation with a variety of perspectives reflected in the data to be analysed. This will provide more evidence for the Scoring inferences to enable the move to generalisation and interpretation studies (Cook et al., 2015).

Benchmarking

Benchmarking and standards were a key issue in the attributes that were interpreted similarly and still had significant variation in their scoring. Consensus needs to be reached on the level at which students are performing. Benchmarking and standard setting have to be one of the initial tasks in the further validation of this assessment and to complete the evidence for Scoring inference. An approved standard, in terms of what a student can and cannot do, is essential for assessors to apply when judging the performance of a student (Norcini & Shea, 1997).

The Student Documents

Two potentially modifiable features of the student documentation may have impacted on the scoring in Stage 1; the variability in the quality of the documents provided and determining a clear purpose for the documentation. Assessors found it difficult to assess documents of poor quality and low word count. Following studies should aim for a greater number of variable quality documents. The second issue was determining
the purpose of the documentation. The assessors were aware that the documents were sourced from an educational program with a simulated EHR, and were unable to completely disregard this fact when scoring the documents. When the assessor was having difficulties with determining the purpose of the document, they were unable to define their expectations clearly and reference their scoring to this. With differing expectations, scoring between assessors became more variable. Care has to be taken with any simulated EHR program to ensure that both the students and the assessors are clear about the same purpose of any documentation and, as far as is feasible, this should reflect real world practice. The first step in Kane’s validity framework is to state clearly the proposed purpose to ensure that the argument is focused. From this study, we would recommend that the purpose needs to be continually revisited and reinforced in any assessment process and evaluation.

Analysis

Finally, only one researcher qualitatively analysed, and coded the qualitative data from Stage 1 and 2. It would add significant strength to the findings from this study if multiple researchers had conducted the analysis, independently and reached the same consensus on the themes and codes.

5.4 Future Directions for Research

This exploratory study has provided a body of evidence based on expert opinion and informed judgment for the content-related validity of a modified physician instrument to assess the quality of EHR documentation recorded by medical students. A panel of experts reached consensus on the test domain for the task of creating quality EHR documentation, agreeing on definitions and descriptors of the observable attributes. The modifications were in line with current research on medical education assessment.

Kane’s framework is iterative, and with the test domain robustly studied in terms of content-related validity, the weakest assumption needs to be redefined and tested. The assumption from this study is that the modified assessment rubric will improve the
validity of the assessment of quality in EHR documentation recorded by medical students. This needs to be challenged, and with further study, the instrument can be modified, the assumptions revisited and further evidence collected. (Cook et al., 2015; Kane, 2006) We have collated the recommendations for future studies made throughout the discussion in Table 5.4.

Table 5.4 Directions for future studies

<table>
<thead>
<tr>
<th>Inference</th>
<th>Recommendations for future studies</th>
</tr>
</thead>
</table>
| **Scoring**     | - Benchmarking and standard setting at levels of training  
| Refers to the observations of a performance | - Clarification of scoring  
|                 |   - Evidence-based documentation as a quality  
|                 |   - Interpretation of Accurate and Professionalism  
|                 | - Behavioral observations provide reproducible assessments                                                                                   |
| **Generalisation** refers to the sample of observations reflecting the test world performance | - Overlapping attributes – item correlation  
|                 | - Test performance of all attributes, in particular, Organised and Professionalism  
|                 | - Sampling strategy – assessors, students, tasks, occasions, items                                                                        |
| **Extrapolation** refers to the relationships between the test world and the real world performance | - Scope of assessment  
|                 | - Authenticity of item/scenario - how to assess Accurate, copy and paste and Originality                                                       |
| **Implications** refers to the consequences of the assessment | - Explore the consequences of assessment to learners, health services, and the community  
|                 | - Ungraded measure of competency versus graded  
|                 | - Planned actions - implications for real world use of EHRs, depending on assessment                                                        |
| **Overall**     | - Revisit purpose  
|                 | - Complete the validity argument for assessment of quality in EHR documentation recorded by medical students                              |
Additional studies on the assessment of the quality use of EHR by medical students would significantly complement and enhance educational programs focused on EHRs and medical students, and the assessment of the quality EHR documentation.

Assessment criteria outlining the quality use of the EHR, including the use of clinical decision support tools, medication reconciliations, and ordering investigations are also needed. An assessment instrument focused on the use of the EHR would complement the assessment of quality EHR documentation, and might assist in resolving outstanding issues with the interpretation of Internally Consistent and Organised.

5.5 Contribution

We have provided a body evidence that proves that the PDQI-9 as currently formatted was not valid for assessing quality in EHR documentation recorded by medical students. However, with clearer interpretations of the attributes, the findings strongly support the test domain of the PDQI-9 for assessing the quality of EHR documentation recorded by medical students and as a result, modifications to the test domain have been recommended. We have made significant inroads into the necessary validity argument for the assessment of EHR documentation recorded by medical students. Our findings will impact assessment, as well as teaching and learning, of EHR documentation and as a result, our graduates will be better prepared for the challenging world of communication they are expected to perform in, to ensure patient care.
References


Wagner, D. P., Roskos, S., Demuth, R., & Mavis, B. (2010). Development and evaluation of a health record online submission tool (HOST). *Medical Education Online, 1*, 5350. doi: 10.3402/meo.v15i0.5350

## Appendix A.

### Physician Documentation Quality Instrument (PDQI-9)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Score</th>
<th>Description of Ideal Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Up to date</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note contains the most recent test results and recommendations.</td>
</tr>
<tr>
<td>2. Accurate</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note is true. It is free of incorrect information.</td>
</tr>
<tr>
<td>3. Thorough</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note is complete and documents all of the issues of importance to the patient.</td>
</tr>
<tr>
<td>4. Useful</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note is extremely relevant, providing valuable information and/or analysis.</td>
</tr>
<tr>
<td>5. Organized</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note is well-formed and structured in a way that helps the reader understand the patient’s clinical course.</td>
</tr>
<tr>
<td>6. Comprehensible</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note is clear, without ambiguity or sections that are difficult to understand.</td>
</tr>
<tr>
<td>7. Succinct</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note is brief, to the point, and without redundancy.</td>
</tr>
<tr>
<td>8. Synthesized</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>The note reflects the author’s understanding of the patient’s status and ability to develop a plan of care.</td>
</tr>
<tr>
<td>9. Internally Consistent</td>
<td>Not at all</td>
<td>Extremely</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>No part of the note ignores or contradicts any other part.</td>
</tr>
</tbody>
</table>

QNOTE, an instrument that scores the note based on the evaluative criteria (components) of each element.

<table>
<thead>
<tr>
<th>Elements and their components</th>
<th>Scoring - acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. CHIEF COMPLAINT(S)</strong></td>
<td></td>
</tr>
<tr>
<td>A. Sufficient information</td>
<td>□ Missing □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(enough information to direct HPI; includes pertinent details; includes duration)</td>
<td></td>
</tr>
<tr>
<td>B. Concise</td>
<td>□ Missing □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td>C. Clear</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(understandable to providers and others)</td>
<td></td>
</tr>
<tr>
<td>D. Organized</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(properly grouped, chronological, can find important information easily)</td>
<td></td>
</tr>
<tr>
<td><strong>2. HISTORY OF PRESENT ILLNESS (HPI)</strong></td>
<td></td>
</tr>
<tr>
<td>A. Sufficient information</td>
<td>□ Missing □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(enough information for its purpose; includes pertinent details)</td>
<td></td>
</tr>
<tr>
<td>B. Concise</td>
<td>□ Missing □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td>C. Clear</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(understandable to providers and others)</td>
<td></td>
</tr>
<tr>
<td>D. Organized</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(properly grouped, chronological, can find important information easily)</td>
<td></td>
</tr>
<tr>
<td><strong>3. PROBLEM LIST</strong></td>
<td></td>
</tr>
<tr>
<td>A. Current</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(up-to-date)</td>
<td></td>
</tr>
<tr>
<td>B. Ordered</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(from most to least important)</td>
<td></td>
</tr>
<tr>
<td>C. Concise</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td>D. Complete</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(addresses all important problems)</td>
<td></td>
</tr>
<tr>
<td><strong>4. PAST MEDICAL HISTORY</strong></td>
<td></td>
</tr>
<tr>
<td>A. Complete</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(addresses all important past medical history)</td>
<td></td>
</tr>
<tr>
<td>B. Concise</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td>C. Clear</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(understandable to providers and others)</td>
<td></td>
</tr>
<tr>
<td>D. Organized</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(properly grouped, chronological, can find important information easily)</td>
<td></td>
</tr>
<tr>
<td><strong>5. MEDICATIONS LIST</strong></td>
<td></td>
</tr>
<tr>
<td>A. Current</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(up-to-date)</td>
<td></td>
</tr>
<tr>
<td>B. Complete</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(contains all the current medications including dosages)</td>
<td></td>
</tr>
<tr>
<td>C. Concise</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(no non-current medications)</td>
<td></td>
</tr>
<tr>
<td><strong>6. ADVERSE DRUG REACTIONS AND ALLERGIES</strong></td>
<td></td>
</tr>
<tr>
<td>A. Current</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(up-to-date)</td>
<td></td>
</tr>
<tr>
<td>B. Sufficient information</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(enough information for purpose; includes pertinent details)</td>
<td></td>
</tr>
<tr>
<td>C. Clear</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(understandable to providers and others)</td>
<td></td>
</tr>
<tr>
<td><strong>7. SOCIAL AND FAMILY HISTORY</strong></td>
<td></td>
</tr>
<tr>
<td>A. Current</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(up-to-date)</td>
<td></td>
</tr>
<tr>
<td>B. Sufficient information</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(enough information for purpose; includes pertinent details)</td>
<td></td>
</tr>
<tr>
<td><strong>8. REVIEW OF SYSTEMS</strong></td>
<td></td>
</tr>
<tr>
<td>A. Complete</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(addresses all pertinent positives and negatives)</td>
<td></td>
</tr>
<tr>
<td>B. Concise</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td><strong>9. PHYSICAL FINDINGS</strong></td>
<td></td>
</tr>
<tr>
<td>(includes vital signs)</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td><strong>10. ASSESSMENT</strong></td>
<td></td>
</tr>
<tr>
<td>(diagnosis; differential)</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>A. Prioritized</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(displayed in order of importance: signs, symptoms, tests, procedures organized properly, includes care plan)</td>
<td></td>
</tr>
<tr>
<td>B. Sufficient information</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(enough information for purposes; includes pertinent details)</td>
<td></td>
</tr>
<tr>
<td>C. Clear</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(understandable to providers and others)</td>
<td></td>
</tr>
<tr>
<td>D. Concise</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td><strong>11. PLAN OF CARE</strong></td>
<td></td>
</tr>
<tr>
<td>(with goals and objectives)</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>A. Prioritized</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(displayed in order of importance: signs, symptoms, tests, procedures organized properly, includes care plan)</td>
<td></td>
</tr>
<tr>
<td>B. Sufficient information</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(enough information for purposes; includes pertinent details)</td>
<td></td>
</tr>
<tr>
<td>C. Clear</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(understandable to providers and others)</td>
<td></td>
</tr>
<tr>
<td>D. Concise</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
<tr>
<td>(focused, brief, not redundant)</td>
<td></td>
</tr>
<tr>
<td><strong>12. FOLLOW-UP INFORMATION</strong></td>
<td></td>
</tr>
<tr>
<td>(instructions for the patient; consults, orders, prescriptions)</td>
<td>□ Missing □ Not applicable □ Fully □ Partially □ Unacceptable</td>
</tr>
</tbody>
</table>


Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to http://group.bmj.com/group/rights-licensing/permissions
Appendix C.
Assessor Consent, Demographics and Logistics Survey

Assessor Consent, Demographics and Logistics

Thankyou for agreeing to participate in the study, ‘Exploring the qualities of Electronic Health Record written documentation by medical students in their first clinical year’.

This survey will take only a few minutes of your time and collect the following information:

- Your consent
- Confirm your credentials as an expert for this study
- Ensure correct contact details

Kind regards

Dr Lisa Cheshire
Dr Amy Gullickson
Professor Geoff McColl

Melbourne Graduate School of Education
The University of Melbourne Victoria 3010 Australia
T: +61 3 8344 8285  F: +61 3 8344 8529  W: www.edfac.unimelb.edu.au

Melbourne Medical School
The University of Melbourne Victoria 3010 Australia
T: +61 3 8344 9794  F: +61 3 8344 0188 W: www.medicine.unimelb.edu.au
Assessor Consent, Demographics and Logistics

* Required

CONSENT

Project Title: Exploring the qualities of Electronic Health Record written documentation by medical students in their first clinical year.

Investigators:
Dr Lisa Cheshire (Masters Student)
Dr Amy Gullickson (Principal Investigator)
Prof Geoff McColl (Co-investigator)

The Melbourne Graduate School of Education, in conjunction with the Department of Medical Education, Melbourne Medical School, University of Melbourne is conducting the project.

1. I consent to participate in the project named above, the particulars of which have been explained to me. A written copy of the information has been given to me to keep.
2. I authorise the researcher or assistant to use for this purpose the results of the assessed documentation
3. I acknowledge that:
   (a) The possible effects of the assessment process have been explained to me to my satisfaction;
   (b) I have been informed that I am free to withdraw from the project at any time without explanation or prejudice and to withdraw any unprocessed data previously supplied;
   (c) The project is for the purpose of research
   (d) I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements

Name: *

By typing your name into the space below you are indicating your consent

« Back  Continue »

50% completed
Assessor Consent, Demographics and Logistics

* Required

Credentials
It is important to be able to justify the selection of the panel experts. In order to do this, please fill out the details below

Your medical specialty: *

Year of Graduation from Medical School: *

If applicable, number of years since completion of postgraduate training:

Academic Institution affiliated with: *

Number of years involved in medical student education and assessment: *

- < 5 years
- 5 - 10 years
- 10 - 15 years
- Other: 

Number of years experience with electronic medical documentation: *

- < 5 years
- 5 - 10 years
- > 10 years

If applicable, other postgraduate activities or qualifications in the above areas of expertise:
Description that best describes the location of your main site of practice: *

75% completed
Assessor Consent, Demographics and Logistics

* Required

Logistical Information

Please provide your preferred email for documents to be sent to: *

Please provide an alternative contact: *
Due to the confidential nature of the documents you will be reviewing, we will send you a password for accessing encrypted files. For security purposes, this password needs to be sent to a different account than the email account you've listed above. Please provide a second account, either a mobile phone or alternate email where we can contact you with passwords

Never submit passwords through Google Forms.

100%: You made it.
Assessor Training Package

Study: Exploring the qualities of Electronic Health Record (EHR) documentation written by medical students in their first clinical year.

Please read the content of this training package before you to start to assess the student records.

Examiner name

Unique Assessor Identification: XX

---

**Melbourne Graduate School of Education**
The University of Melbourne Victoria 3010 Australia
T: +61 3 8344 8285  F: +61 3 8344 8529  W: [www.edfac.unimelb.edu.au](http://www.edfac.unimelb.edu.au)

In conjunction with

**Melbourne Medical School**
The University of Melbourne Victoria 3010 Australia
T: +61 3 8344 9794  F: +61 3 8344 0188 W: [wwwmedicine.unimelb.edu.au](http://wwwmedicine.unimelb.edu.au)
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key information for completing first stage</td>
<td>3</td>
</tr>
<tr>
<td>• Step by step instructions to complete the scoring of the de-identified student records</td>
<td></td>
</tr>
<tr>
<td>Important considerations for marking records</td>
<td>4</td>
</tr>
<tr>
<td>Overview of Medical Student Documentation</td>
<td>6</td>
</tr>
<tr>
<td>• Overview of the program the student electronic documentation has derived from and an explanation of the layout of the documents</td>
<td></td>
</tr>
<tr>
<td>Understanding the quality attributes you are scoring</td>
<td>9</td>
</tr>
<tr>
<td>Copy of the original PDQI -9</td>
<td>10</td>
</tr>
<tr>
<td>Copy of the Google Form ‘Electronic Health Record Documentation Quality Instrument’</td>
<td>11</td>
</tr>
<tr>
<td>Extra References and contact details of researchers</td>
<td>17</td>
</tr>
</tbody>
</table>
Process

Process – For each form http://goo.gl/forms/KzBf2Inc4I

1. Enter your unique assessor identifier XX
2. Enter the Case Identification Number located at the top of the de-identified EHR student document
3. Rate the case in terms of the complexity of the patient described in the record (not the student’s writing)
4. For each attribute, to the best of your ability:
   a. Score the written document for the specific quality attribute
   b. Provide your rationale for the score you have given in a sentence or two.
   c. Comment on this record with regard to the meaningfulness of the specific quality attribute.
5. At the end of the form you are asked to make a global assessment of the record in its entirety and justify this. Particularly highlight if there is a significant aspect of this record that has contributed to your global score, which may seem at odds with the scores from the individual attributes.
6. Finally comment on the purpose of EHR written documentation for medical students, and any attributes that maybe important and not included
7. Use a separate (new) form for each record
8. When you have completed scoring all 11 records please contact Lisa Cheshire to advise her that you will not be making any further entries.
Important Considerations for scoring records

1. **The purpose of the documentation and relevant quality attributes.** The Physician Document Quality Instrument (PDQI-9) was developed as a simple to use, validated, reliable way to measure the quality of physician documentation. The main underlying purpose of the physician documentation was to support clinical communication with other providers engaged in the care of the patient. The benchmark for validation of the PDQI-9 was that of a ‘gold standard’ physician document.

   The purpose for medical students’ written documentation will be slightly different than that for a fully qualified physician. Attributes regarded as indicating quality in physician documentation may or may not be meaningful to a medical student cohort. Medical student activities are embedded in an educational setting, and students are trying to demonstrate knowledge and learning, in addition to communication. Whilst filling out the forms please consider carefully the purpose EHR documentation serves in a medical student setting and attributes you think are relevant but not being assessed in the PDQI-9 to this context.

2. **Benchmark.** The benchmark for this exercise is that of the gold standard you would expect from a graduating student, completing a medical degree, and at the beginning of their Post Graduate Year 1 (PGY1), that is internship year. Scoring five in any one quality should equate to an excellent, fully competent performance for an end of course medical student. A contribution of this study will be an expert consensus on the quality of documentation a graduating student should have attained as they commence their PGY1.

3. **Generalisable findings and products.** Although many of you may be familiar with a written documentation program within your University, think generally about medical student documentation quality when scoring and justifying your marking. The results from this study are intended to be applicable to quality attributes for medical student EHR written documentation in any medical school.
Electronic Health Record student documentation

To help you understand the layout of the student documents you have received, below is a brief summary of the program and how students enter patient data.

The Electronic Health Record program is part of the Year 2 Doctor of Medicine (MD) program at the University of Melbourne. The MD is a four-year graduate program. Year 1 focuses on biomedical science. Year 2 is their first year in a clinical setting.

The EHR program was developed to help students develop expertise in the practice of creating, storing and managing high quality patient records. It is an online patient record system designed for clinical education. The system allows students to capture details of their clinical encounters and store de-identified patient information in a secure database. It also mediates online access to specialist medical databases such as Australian Medicines Handbook.

The students enter data in selected fields. The majority of the entry fields are free text boxes, see figure 2. The exceptions are the ‘Basic Details’ (figure 1) and ‘Medications’ (figure 3).
Figure 1. Basic Details Entry Page

Figure 2. Example of a ‘free text’ box. Presenting problems
When the student documentation is downloaded as a pdf the details will appear under the heading of the relevant text box that the data has been entered into. Please note, it is currently not possible to make the medication list more compact, and it is presented in the documents as it is entered in the record (see Figure 3).

As part of the educational program, there is a final section for the students to reflect on their learning from the case recorded.
Understanding the quality attributes you are scoring

Quality is an attribute that is judged by the individual who is reading a note, and this reflects that person’s needs, values and expectations. Note quality has remained an elusive concept.

The PDQI-9 was developed by factor analysis. Four sets of attributes consistently factored together across all note types (admission, progress and discharge). Factors were eliminated that were not relevant to all three-note types of interest. The bold attributes are in the PDQI-9 rating form.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Attribute</th>
<th>Description</th>
<th>Adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Well-formed</td>
<td>The information is being presented in a way that contains a logical structure leading to a reasonable conclusion.</td>
<td>Clear, uncluttered, organized, structured, non-redundant, synthesized</td>
</tr>
<tr>
<td>II</td>
<td>Comprehensible</td>
<td>The transfer of information from the note to the reader occurs with meaning and comprehension</td>
<td>Legible, coherent, useful, correct, comprehensible, consistent</td>
</tr>
<tr>
<td>III</td>
<td>Accurate</td>
<td>The extent to which the content of the document reflects the true state of the patient</td>
<td>Up to date, complete, accurate, thorough, current, relevant</td>
</tr>
<tr>
<td>IV</td>
<td>Compact</td>
<td>The note has the appropriate density of information given the content</td>
<td>Brief, concise, succinct and focused</td>
</tr>
</tbody>
</table>
**Physician Documentation Quality Instrument (PDQI-9)**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Score</th>
<th>Description of Ideal Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Up to date</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note contains the most recent test results and recommendations.</td>
</tr>
<tr>
<td>2. Accurate</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note is true. It is free of incorrect information.</td>
</tr>
<tr>
<td>3. Thorough</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note is complete and documents all of the issues of importance to the patient.</td>
</tr>
<tr>
<td>4. Useful</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note is extremely relevant, providing valuable information and/or analysis.</td>
</tr>
<tr>
<td>5. Organized</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note is well-formed and structured in a way that helps the reader understand the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>patient's clinical course.</td>
</tr>
<tr>
<td>6. Comprehensible</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note is clear, without ambiguity or sections that are difficult to understand.</td>
</tr>
<tr>
<td>7. Succinct</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note is brief, to the point, and without redundancy.</td>
</tr>
<tr>
<td>8. Synthesized</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The note reflects the author’s understanding of the patient’s status and ability to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>develop a plan of care.</td>
</tr>
<tr>
<td>9. Internally Consistent</td>
<td>Not at all</td>
<td>Extremely 1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No part of the note ignores or contradicts any other part.</td>
</tr>
</tbody>
</table>

**Note:**
For this study, you will be using an adapted version of the PDQI-9 form to rate EHR entries. There is widespread consensus in the literature, that regardless of the level of training, the documentation should be an accurate and true record, free of incorrect information. However, there is no way to judge the accuracy of the student documents you have received because you will not have access to the patient. Therefore, in the version you will use, ‘Accuracy’ has been removed as an item for calibration.
Electronic Health Record Documentation Quality Instrument

This EHR Documentation Quality Instrument is based on Physician Documentation Quality Instrument (PDQI-9)

A NEW form must be completed for each student document

On this page please:

1. Enter your unique assessor identification
2. Enter the Case Identification Number located at the top of the each EHR student document
3. Rate the case in terms of the complexity of the patient being described in the record (not the student's writing)

* Required

Assessor Identification Number: *

Case Identification Number: *

Complexity of Case *

- Low
- Medium
- High

Documentation Quality Instrument

For each of the following attributes, to the best of your ability:

a. Score the written document for the specific quality attribute

b. Provide your rationale for the score you have given in a sentence or two.

c. Comment on this record with regards to the meaningfulness of this specific quality attribute meaningfulness of this specific quality attribute.
Of note:
The benchmark for this exercise is that of the gold standard you would expect from a graduating student, completing a medical degree, and at the beginning of their Post Graduate Year 1 (PGY1), that is internship year. Scoring five in any one quality should equate to a competent performance for an end of course medical student.

Although you may be familiar with a written documentation program within your University, think generally about medical student documentation quality when scoring and justifying your marking.

Based on your global assessment of the written record please score the following attribute:

**Attribute: Thorough***
The note is complete and documents all of the issues of importance to the patient

1 2 3 4 5

Not at all evident  ○ ○ ○ ○ ○ Extensively demonstrated

**Explain why you attributed this score to Thorough***

Please make a comment about this record and the meaningfulness of Thorough

**Attribute: Useful***
The note is extremely relevant, providing valuable information and/or analysis

1 2 3 4 5

Not at all evident  ○ ○ ○ ○ ○ Extensively demonstrated
Explain why you attributed this score to Useful *

Please make a comment about this record and the meaningfulness of Useful

Attribute: Organised *
This note is well-formed and structured in such a way that helps the reader understand the patient's clinical course

1 2 3 4 5

Not at all evident    Extensively demonstrated

Explain why you attributed this score to Organised *

Please make a comment about this record and the meaningfulness of Organised

Attribute: Up to Date *
The note contains the most recent test results and recommendations.

1  2  3  4  5

Not at all evident  O  O  O  O  Extensively demonstrated

Explain why you attributed this score to Up to Date *

Please make a comment about this record and the meaningfulness of Up to Date

Attribute: Comprehensible *
The note is clear, without ambiguity or sections that are difficult to understand.

1  2  3  4  5

Not at all evident  O  O  O  O  Extensively demonstrated

Explain why you attributed this score to Comprehensible *

Please make a comment about this record and the meaningfulness of Comprehensible
Attribute: **Succinct** *

The note is brief, to the point and without redundancy.

Not at all evident 0 0 0 0  Extensively demonstrated

**Explain why you attributed this score to Succinct** *

Please make a comment about this record and the meaningfulness of Succinct


Attribute: **Synthesised** *

The note reflects the author's understanding of the patient's status and ability to develop a plan of care

Not at all evident 0 0 0 0  Extensively demonstrated

**Explain why you attributed this score to Synthesised** *
Please make a comment about this record and the meaningfulness of Synthesised

Attribute: Internally consistent *
No part of the note contradicts or ignores any other part

1 2 3 4 5

Not at all evident  ○ ○ ○ ○ Extensively demonstrated

Explain why you attributed this score to Internally Consistent *

Please make a comment about this record and the meaningfulness of Internally Consistent

Global Assessment
Please make a global assessment of this record in its entirety and justify this.

**Overall, the standard of this written record is** *

- Excellent
- Good Pass
- Pass
- Borderline
- Fail

**Explain why you attributed this global score to this record** *
Particularly highlight if there is a significant aspect of this record that has contributed to this global score, which may seem at odds with the scores from the individual attributes.

---

**Final Page**

Please comment on

1. The purpose of EHR written documentation for medical students

2. Any attributes that maybe important and not included, in the context of the documentation being written by medical students in an educational setting

You may wish to make comments as you go that are generated from the record you have just marked, and/or a summary comment towards the end of your marking.

**Comment:**

---

Submit

*Never submit passwords through Google Forms.*

100%: You made it.
References of importance to this study:

Crossley, J., & Jolly, B. (2012). *Making sense of work-based assessment: Ask the right questions, in the right way, about the right things, of the right people*


If you would like a copy of any of the references please let Lisa Cheshire know and these can be provided.

Contacts

- Dr Lisa Cheshire  
  Masters Student  
  Year 2 MD Academic Coordinator  
  Department of Medical Education  
  Email: cheshire@unimelb.edu.au

- Dr Amy Gullickson  
  Supervisor  
  Senior Lecturer & Academic Coordinator - Centre for Program Evaluation  
  Melbourne Graduate School of Education  
  Email: amy.gullickson@unimelb.edu.au

- Professor Geoff McColl  
  Supervisor  
  Head of the Melbourne Medical School  
  Professor Medical Education and Training  
  Email: gjmccoll@unimelb.edu.au
Appendix E.

Example of Interview Schedule

Exploring the qualities of Electronic Health Record medical student documentation

Interviewee: BB
Date: 7 October 2015
Time: 10 am
Place: DME
Interviewer: Lisa Cheshire

Thanks for participating in this project.

The first stage of the project wielded a large amount of data on the qualities of EHR medical student documentation. The interview today will focus on gaining consensus on issues raised in the first stage of the project to answer the questions.

Today we will be looking at analysis about the meaningfulness of the attribute in the context of the medical students, and secondly, whether there are modifications that may be made to help improve the validity of the PDQI-9 in this setting.

With permission, this interview will be audiotaped. Prior to data analysis, a transcript of the interview will be made available to you for clarification, expansion or modification, if you wish. I have a copy of the Plain Language Statement here if you wish to re-read it.

Here is a copy of the consent form please read it and if in agreement, sign.

You may withdraw consent at any time during the interview and up until data is de-identified. (Consent form read and signed).
Purpose

First discussion point is one of consensus on the purpose of student documentation. What is the purpose of student documentation in an educational setting?

Findings from the preliminary analysis:
The overall objectives are twofold:
- To develop expertise in the practice of creating, storing and managing high quality electronic patient records
- To create quality documents that demonstrates clinical reasoning, reflection and learning.

What are your opinions on this?

Attributes of Quality

The attributes fell into three categories after analysis:
1. Attributes that were interpreted by all similarly and the difference in marks probably reflects benchmarking, and assessor variability, rather than different interpretations of the attribute.
2. Attributes that were interpreted in different ways and marking reflected more the meaning of the attribute that was used to assess the record
3. Attributes that were difficult to assess with the documents
Attributes with similar interpretations

Overall aims

- To confirm interpretation and meaningfulness
- To consider how the performance is observed in medical students to improve interpretation
- Any modifications to improve validity

Synthesis

Original descriptor. The note reflects the author’s understanding of the patient’s status and ability to develop a plan of care.

Specific talking points:

Uniformly agreed that this attribute is vitally important. See own comment.

*Synthesis needs to be assessed; this is the only way we can see if the student has "got" what they are writing about’ AA Case 8*

Any further clarification of interpretation and meaningfulness.

This attribute correlated most closely to overall standard, it was hard to be excellent without a 4/5 in this section; and it invariably scored 1 - 2 if failed. There is something inherent about synthesis that we factor in our overall standard.

The difficulty with this attribute was one of benchmarking and another attribute that seemed to have a large focus on investigations and management in assessment. Can we see an evolution of synthesis throughout the medical course? Can the document demonstrate synthesis without investigations and a management plan?

Explore modifications to increase validity
**Thorough**

Original descriptor. The note is complete and documents all of the issues of importance to the patient.

Specific talking points:

On the whole Thorough has been used to indicate how many sections of the record students have entered information into and its completeness. It was the ‘checklist’ attribute. For a high score the section not only had to be addressed but it also had to be complete.

Would you agree with the interpretation as a ‘checklist’ attribute?

Uniformly this was seen as being meaningful in a medical student context.

Would you agree that Thorough is meaningful in a medical student context?

*Lots of missing information.*

*I think thorough is an important attribute for medical students at the end of the course. BB Case 1 (1)*

The main issue with the assessment of thorough was benchmarking. Except for a few exceptions justifications for scoring in this attribute were similar, the actual score awarded varied. What are your expectations?

What modifications could be made for this attribute to be assessed more reliably?

One issue is how to assess thorough when some parts are thorough and others are not. Comment.
Attributes with more than one interpretation for assessment

Overall aim

- If possible consensus on interpretation
- If agreed interpretation is this meaningful, and how is it observed
- Any modifications to improve validity

Organised

Original descriptor. The note is well-formed and structured in a way that helps the reader understand the patient’s clinical course.

Specific talking points:
Interpretation:

‘The basic organisational structure is forced by the EHR, this attribute has to be assess beyond this’. BB Case 5

Can you clarify what you meant by this?

In analysis, organised was interpreted in two different ways. The organisation of the content with respect to the software programme, and the organisation of content within sections.

Can you expand on this and give your opinion on what understanding you think a graduating student should have?

Keyword to emerge in justification of organisation was LOGICAL. Is this what we are trying to assess with this attribute?

Meaningfulness of Organised.
There were mixed opinions on whether this was meaningful or not, mainly due to the software program limitations.

What are your opinions on the meaningfulness of Organised?

Do you think that students need to be given feedback on both organisation of the EHR and within the sections?

Do they need to be assessed separately?
Do both interpretations develop during the medical course?

One suggestion is that organisation to the software program as a yes/no with feedback. What is your opinion on this?
**Useful**

Original descriptor. The note is extremely relevant, providing valuable information and/or analysis.

**Interpretation:**

Justifications for the assessment of useful were divided between usefulness to student and usefulness to audience. This largely explained the variance in marking of this attribute.

The PDQI -9 was assessing documentation. If the overarching purpose of this is maintained, how do we make it useful?

Clarify interpretation and meaningfulness
Explore Use for reader and use for student.
Explore modifications
Explore development

Crossover with synthesis and thoroughness – lack of synthesis and detail led to lack of usefulness. Do we agree that this is still an attribute that needs assessing in its own right and would give ‘useful’ feedback to the student?
**Up to Date**

Original descriptor. The note contains the most recent test results and recommendations.

Specific talking points.

Interpretation:
Up to Date was broadly seen as encompassing two aspects in justifications. Up to date equated with up to date management as in, in line with current best practice and Up to date equated with most recent findings. Others interpreted this as up to date management? What are your opinions on this interpretation?

Meaningful:
Do you think it should be assessed independently or part of the criteria for organised?
Should students be given feedback on this? If you were writing guidelines would you include this?
How would this develop – or difference in second and fourth year?

Interpretation of evidence based. Scoring was largely based on the absence or otherwise of management and investigations. How should this evolve in medical student course? What would we expect from a graduating student?

Considering EHR, emphasis on clinical decision support tools and using these appropriately, and recognising the need to use them. Is this something that would impact on this attribute and should it?

Once again raises the issue of missing data and students. If the data is missing, do you assume it is not contributing to the attribute - e.g a lack of investigations.
Comprehensible

Original descriptor - The note is clear, without ambiguity or sections that are difficult to understand

Specific talking points:

Correlated poorly with overall standard.

Interpretation:
You have interpreted this as understanding the document, as well as comments on abbreviations. Please clarify

While I can understand what the student has written, their thinking appears quite difficult to understand. If they are worried about an AAA (differential diagnosis) then why is this not being investigated with an US. Why are they doing suggesting a lipase be ordered, but there is no differential of pancreatitis? Why do they specifically mention CLD but there is no PMH suggesting this should be considered. I find the whole record hard to understand

BB Case 10 (2)

Others have interpreted it as the basics of English Language. Is what I am reading comprehensible? Explore

The majority of justifications focused on understanding what was written rather than understanding in the context of the presentation. Therefore, Link to overall standard - A large number of failing documents scored 4 or 5 in comprehensible
Do we need to give feedback on both?

There are comments on the development of medical professional language as part of what is encompassed by comprehensible. Explore semantic competence, how it develops.

In the context of Electronic Heath Record, is understanding of the record covered by another attribute – do we need to assess on whether it is clearly understood? Does another attribute advise the student of this? Does the student need feedback on this?
Are they different enough to warrant being assessed individually?
How would you see this attribute developing over the course?
Difficult attributes to assess

Overall aim

- Is there an interpretation that can be agreed upon?
- Is this meaningful
- How is it observed

Succinct

Original descriptor. The note is brief, to the point and without redundancy

Specific talking points:
This attribute was obviously difficult to mark with very few lower marks, and questionable correlations to other attributes. Regardless of the actual mark (anywhere between 1 – 5) there were similar comments and justifications at every level.

In grading these records succinctness was often substituted for brevity, a word count, and records that failed could score high on succinct.

It is a skill learned with practice, better too many words initially than too few;

Succinctness is a higher order behaviour/judgements that involves value attribution and is not brevity per se. Biggest issue is that it is a higher-level function and not brevity. GG Case 1

What is your interpretation of succinct?
What do you think of this comment?

Should succinct be included as an attribute in medical student documentation? Is this a meaningful attribute? Is this an attribute that is likely to develop within the medical course?

Compare a second year and a fourth year and the concept of succinct – could it be worded differently?
Could it be modified to be more meaningful?

A number commented on the balance between thorough and succinct and in medical students the former was a priority.

The issue with redundancy and ‘copy and paste’. Originality is not well captured by redundancy attribute – Succinct. Is this important and where should it be assessed?
What attribute would it fall under? PDQI-9 authors suggest that this is commented on in addition as such a big problem in the hospitals?
**Internally consistent**

Original descriptor. No part of the note ignores or contradicts any other part.

Specific talking points:

There was a lot of confusion about assessing this attribute and not many assessors regarded it as meaningful.

Interpretation of the attribute. The majority interpreted ‘internally consistent’ by identifying if there were any direct contradictions in their record. How would you interpret internally consistent?

Is it meaningful in a medical student context?

Does it need to be assessed?

Could it be modified to be more meaningful?

Can it be subsumed in organization or comprehensible?

Reference to coding: One interesting point raised that cannot be neglected when focusing on EHRs was

This criteria was overall poorly scored for the 11 records. Coding is vital in EHR and diagnosis means just that. e.g. Angina, Peptic Ulcer Disease, Pneumonia as opposed to a symptom - chest pain, epigastric pain or productive cough. Knowledge of this differentiation may occur with further education through the MD program. HH

Overall
It is vitally important and needs to be included, is it internally consistent? If not, where does it lie?
**Inclusion of accurate**

Descriptor: The note is true. It is free of incorrect information.

The general consensus in the literature is that this is important and meaningful to medical student documentation. It was difficult, if not impossible, to assess in the documents given in this study for accuracy, however should it be included as part of a modified instrument that is generalizable, but modifiable to different programs?

Do you agree that accurate should be an attribute?

Does this evolve during the medical course?

Have you any suggestions on how to assess this?

**Electronic Health Record requirements**

Consider issues specific for EHR if not already covered)

1. Clinical Decision Support

2. Originality

3. Order entry/medication reconciliation.

**Other attributes**

Any not included

Professionalism suggested – explore.

**Universal concerns**

Lack of information.
Supplementary Information for interviewer if required for interview with BB

Included

- Information in this document
- All the cases
- Descriptive statistics

Purpose

Perhaps part of the problem with our EHR documentation is that we’ve not made the purpose explicit. Perhaps when entering the record, we should allow students to record the kind of record they will make. Perhaps they could choose from a drop down list of clinical notes, discharge summary, referral letter, summary of case etc and then that would give much more direction and relevance to the task (and subsequently allow tutors to provide feedback likely to be of use in the future). AA

There are always concerns regarding students writing of materials that have no corresponding version in their clinical world. This does not mean that one cannot tag on a reflective commentary at the end but the main activity should resemble the real world for what is it you are hoping to transpose from one world to the next? GG

Perhaps part of the problem with ... EHR documentation is that we've not made the purpose explicit. Perhaps when entering the record we should allow students to record the kind of record they will make. Perhaps they could choose from a drop down list of clinical notes, discharge summary, referral letter, summary of case etc and then that would give much more direction and relevance to the task (and subsequently allow tutors to provide feedback likely to be of use in the future). BB

EHR documentation should be a learning tool for future reference / pursuing informal and formal feedback, as well as a means to practice the actual process of written professional communication in order to ensure this is competent / effective from the time of graduation onwards. DD
Attributes of similar interpretations

**Synthesis**

Synthesis needs to be assessed; this is the only way we can see if the student has "got" what they are writing about’ AA Case 8

‘from a patient centered perspective its one of the most important and difficult things GPs do’ GG Case 1

**Benchmarking**

again it’s the value attribution inherent within synthesis and the ability to weigh the importance of various findings and to be clear as to a management path. The student seems lost here, but then again, so might the resident (intern) also be GG overview

‘Most students had a basic understanding of this criteria by highlighting the major presenting complaint, other relevant information, +/- exam findings and a possible diagnosis whether it was in the correct section or not. However there were no outstanding responses and probably it reflects the junior level of the medical students inputting the data in the EHR….’ HH

**Thorough**

Lots of missing information.

I think thorough is an important attribute for medical students at the end of the course. BB Case 1 (1)

**Variance in marking**

Some marked on the information given and not on what was missing (this was uncommon) Case 4

Case 10 divided - unsure why, perhaps style.

Case 10. Highest variance = scored 4 (highest – two scored 1, 2 scored 2, 2 scored 3)

History is brief but to the point with a targeted examination and planned investigations and possible diagnoses AA Case 10 (4)

This presentation is not at all clear given the significant HOPC information missing (and no Medical History, Medications etc BB Case 10 (1)

Hard to say when it is so bad. Did a student really write this as it is very poor and therefore seems unrealistically bad? GG Case 10 (1)

A key area that is lacking - only a vague sense of what type of pain presentation this is, yet the differential diagnoses are of significant and life-threatening conditions. DD (2)
Attributes with multiple interpretations

Organised

Organised and up to date also appear to be related and again purpose dependent. The organisation of information in a letter to the GP (usually less PMH detail and more management/medication changes and ongoing issues) is completely different. Likewise the chronology of a hospital admission may be very relevant to the GP but less relevant to the specialist being asked to review a specific aspect of the case. AA

Comment

Meaningful

Organised seems a sensible assessment criterion however the electronic records all have their own formatting so the students don’t really get to demonstrate much in the way of organisation AA overall comment

No, this is not very meaningful. Organised is not very relevant when students are presented with sections to populate FF Overall comment. Although in a highly organised document said it was meaningful. it is part of utility GG Overall comment

for a complicated case with a number of factors organisation is clinically necessary GG overall Case 5

it is a variable necessity during consultation, the more complex and medically unexplained symptoms require it more GG overall Case 1

Two interpretations.

‘The basic organisational structure is forced by the EHR, this attribute has to be assessed beyond this’. BB Case 5

Everyone has agreed that that the case is superficially organised but not in a way that contributes to the reader's understanding. Case 3

Case 1, Case 7, Case 8 All failed

This is only organised because of the formatting of the document AA 3 Only via the EHR fields. Not enough content to be organised BB1

Can we agree to organised? Is it possible to come to consensus or is it a question of style and preference?

Case 2 illustration - GG score 2, CC scored 5.

Clearly organised; Could have been put in dot form or with minor headings to make it more easily obvious; Organised to the quick reader would have the history and examination in dot points rather than narrative AA Case 2 (3)

very well organised with relevant negatives in the history and examination. CC Case 2 (5)

it is verbatim recorded, like he undertook an OSCE out loud, rather than a patient focused problem analysis GG Case 5 (2)

Case 10 half commented poorly formed, and half comment well - organised.

it is disorganised mainly though lack of detail GG Case 10 (2)

Organised in the set format and organised within that AA Case 2 (4)

Case 11 divided assessors across 2 - 5, not the same assessors. Can we make
recommendations to students that are agreed upon, but also allow for individuality?

Even though there is a template this record is very well organised, includes the narrative bits clearly as well as well as the dot points and the AA Case 11 (5) Agreed CC and GG

Organised in terms of the EHR fields. Lack of organisation in terms of clear timeline, the various complaints on presentation and an investigation plan (and rationale). BB Case 11 (2) Agreed DD

Can organised only be assessed on what is written? Has it a very literal definition? Does brevity and incompleteness of content in the correct sections mean that the note is organised? Is organised independent of overall standard? Can you have a well-organised document that overall is of poor quality? Case 1, Case 7, Case 8 All failed

Case 1 Fail

The structure of the record keeps the documentation organised. The information entered in presenting complaint is roughly correlated with an 8 point history. BB Case 1 (3)

Useful

Own comment

I guess this makes me think about what use is going to be made of the student record. In the introductory content it was stated that the purpose of the record was to help students develop expertise in "creating, storing and managing high quality patient records". If student wants to recall a particular patient encounter for their own reference this might be fine, but if they are to recall this in order to review learning points, review pathophysiology etc then this record is not particularly helpful. I think this points to one of the challenges of the current system. Is the system being used to prove that you've seen a patient? If so then this record is useful. Or is the patient record entered to provide information history taking, examination findings, diagnostic thinking and identification of learning goals? In this case this record is not particularly useful. Case 1 BB Comment

Highest variance was Case 10

This would be useful for someone taking over the patient’s care, though a bit more info about past history and medications would be helpful in this case. AA Case 10 (4)

I’m not sure this one is useful at all - not enough information, jumbled differentials (included in the investigation section), and no clear plan for the patient Bb Case 10 (1)

Would be more useful with the right amount of detail and some critical analysis. DD Case 10 (2)

some important findings are missing and it contains no clear idea about what the patient's management will be GG Case 10 (1)
Up to Date

Complementary Literature Analysis

Discusses both time and evidence based medicine

1. Most recent assessment (interpreter), most recent requests (manager) – Stephens
2. For an educator - Actively search the medical literature; Uses embedded clinical support tools to access current evidence related to patient care and modifies plan (Stephens). This is focused on using the EHR, is it a part of quality of an EHR document?

Meaningfulness

Perhaps 'up to date' is generally not useful because it is not usually apparent what stage the patient’s at in terms of admission - should this be something that students are advised to clarify rather than disregard BB Case 11

Raises the issue of missing data and students.
Case 3 and CC, Case 5, Case 7
I can't comment on this AA Case 3 (1)
contains a reference to test results but these were not available at the time of wiring. CC Case 3 (4)
No evidence and in fact some wrong statements e.g patient will need to take warfarin for rest of life AA Case 5 (1)
included CTPA result, O2 reading CC Case 5 (5)
There are no investigation findings reported, nor respiratory exam findings nor the plan of management. FF Case 5 (2)
seemed to demonstrate most current management though did not include mention of direct thrombin inhibitors GG Case 5 (4)

Comprehensible

Interpretation.

To highlight definition of comprehensible in this context.
'The little data that is entered is perfectly comprehensible' AA Case 7 (3)
Its importance
‘This criterion becomes more important as Electronic medical records take over. The temptation to use lots of jargon and abbreviations may increase and the recent papers on this show that the many abbreviations are not understood by GPs - or often by doctors in other hospital specialties - but no one has been measuring this.’ AA overview

Complementary Literature Analysis

1. Clearly communicates clinical facts; uses appropriate clinical language (semantic competence)
2. Manager – clearly articulates a logical and semantically competent therapeutic plan (Stephens)
**Attributes difficult to interpret**

**Succinct**
Not sure of the value of rating scales that can directly contradict each other (eg Succinct vs Thorough), when perhaps a more generic rating could encompass both domains simultaneously. DD Overall comment

Thorough v’s Succinct is potentially going to be at odds - and dependent on the purpose (eg script for a specialty review v’s a discharge summary going to a GP). So I think the purpose or use of the content is fundamental to really be able to determine if it is thorough or succinct enough. Useful is another attribute that is extremely purpose dependent. BB Comment

**Internally consistent**
The difficulty being when the whole patient chart is not available and the record is of a poor standard. Similar difficulty with minimal data. If taken very literally, score 5 as minimal data recorded did not contradict. Refer to variance of Case 7

---

**So little information it would be difficult not to be internally consistent AA Case 7 (4) Agree FF, CC, DD**

**Too little information to make a judgment on this BB Case 7 (1)**

not that it was not consistent but rather then was little to be inconsistent about GG Case 7 (1)

‘again of limited utility when there is limited data within the case record’ GG Case 10

Possibly important to assess but not vital, better to pick the occasional off ones and give them feedback AA Overall

**Other issues**

**Accurate**

Complementary Lit Review Findings

R: The student honestly and accurately documents patient findings (Stephens); accuracy is freedom of errors but avoiding cutting and pasting/plagiarising; cut and pasting confined to objective findings such as reports radiology or pathology (ACE).

I: The student honestly and accurately documents patient findings and (Stephens) accurately creates a differential diagnosis and updated problem list; accuracy is freedom of errors but avoiding cutting and pasting/plagiarising; cut and pasting confined to objective findings such as reports radiology or pathology. (ACE)

M: In addition the student honestly and accurately documents patient findings and (Stephens) accurately creates a differential diagnosis and updated problem list, and plan; accuracy is freedom of errors but avoiding cutting and pasting/plagiarising; cut and pasting confined to objective findings such as reports radiology or pathology. (ACE)

Specific sections of the patient's clinical assessment should reflect the work product of
the final author and NOT be carried forward. This includes HPI, review of systems, physical examination, assessment, plan (Directors)

Lack of information
Possibly as a result of the EHR program. Blank information assumed that the student did not complete. In an ideal world, possibly have the student key in “not performed” or "negative" (infers information asked but negative result e.g. medical history, social history. This would provide some context on the EHR of the patient and quality of information. This applies to all records reviewed for stage 1 - files 1-11. BB
Appendix F
Proposed rubric for assessing quality of EHR documentation recorded by first year clinical medical students

<table>
<thead>
<tr>
<th>Insufficient skills for quality documentation</th>
<th>Emerging Skills for quality documentation</th>
<th>Appropriate skills for quality documentation</th>
<th>Well-developed skills for quality documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is insufficient documentation in this note to be able to make a judgment and provide worthwhile feedback. Please resubmit.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please select all the descriptors that you have observed in the student document. At least one descriptor needs to be selected for each attribute.

**Thorough:** The document is complete and documents all of the issues of importance to the patient’s care with respect to the history, examination and differential diagnosis and an outline of ongoing care.

- Your document has limited and/or incomplete information.
- Your document requires more depth and breadth of information.
- Your document is missing essential information.
- There is minimal consideration of the differential diagnosis in your document.
- Your document contains the minimum required information.
- Your document requires more depth and breadth of information.
- There are elements missing in your document that might help support the correct diagnosis.
- Your document has most sections thoroughly recorded.
- Most of the important aspects of the presenting problem, examination and differential diagnosis are thoroughly recorded.
- Some sections are incomplete or have missing information.
- There is a basic management plan.

**Organised:** The document is well formed in a way that helps the reader understand the patient’s clinical course.

- The basic content of your document was poorly organised.
- The clinical course needs to be made clear to the reader.
- Your information needs to be recorded in a logical manner.
- Timelines and dates have been omitted or are not consistent.
- The organisation of the presentation partially assists in understanding the presenting complaint.
- There is an attempt at timelines and progress of the presentation although this is still difficult to follow.
- Some sections are poorly structured.
- The logic used to display the information is not clear.
- The organisation of the presenting complaint assists the reader’s understanding.
- Your document includes reference to timelines but some details remain unclear.
- The logic used to display the information is understandable.
- All sections of your document are well organised.
- Your document is clearly referenced to timelines.
- The organisation of findings significantly assists the reader’s understanding of the case.
- The presentation of findings is clear and in a logical sequence.
**Succinct:** The document is concise, and without redundancy.

- There is redundant, or excessive, use of language and/or irrelevant material included.
- Your document appears in places to be a verbatim record of patient presentation with little attempt to eliminate redundancy.
- The note is too brief to make comment on this attribute.
- There has been an attempt to make the patient’s account more concise than the spoken word.
- There is redundant, or excessive, use of language and irrelevant material included.
- The information could be given more clearly with far fewer words.
- The information provided has not been tailored to meet the purpose of the document.
- Most of the information recorded is necessary.
- There is some redundant and irrelevant material included.
- The information could be given more clearly with fewer words.
- There is consideration of the information required to meet the purpose of an document.
- Your document is a concise account of the presentation.
- The information was clear and to the point, with minimal redundancy.
- All information presented was necessary.
- You have tailored your information to the purpose of the document.

**Synthesised:** The document reflects the author’s understanding of the patient’s status and ability to develop a differential diagnosis.

- Your document provides little evidence that you understand what is happening with the patient.
- There is no synthesis, or links, between the history, examination and differential diagnosis.
- There is little or no meaningful interpretation of the admission.
- Some aspects of the document indicate that you understand what is happening with the patient, but this does not continue throughout the record.
- There is minimal synthesis between the history, examination and differential diagnosis.
- There is no discussion of what needs to happen next.
- Your document demonstrates that you have a reasonable understanding of what is happening with the patient.
- You link presenting symptoms and examination to an appropriate differential diagnosis.
- There is some consideration of what needs to happen next consistent with the presentation provided.
- There is clear and consistent synthesis of the entire record.
- It is clear that you understand what is happening to the patient, and what needs to happen next.

**Comprehensible:** The reader easily understands the document.

- I had difficulty understanding the patient presentation and progress.
- The language used is essentially lay language and inappropriate for communication between health professionals.
- I was able to understand the basics of presentation and admission but would need significant clarification to continue to manage the patient.
- The language used needs to be more professional and technically precise.
- I was able to understand what has occurred to this patient and would only have to clarify a few issues in the history and examination.
- There is a clear use of the appropriate language for documentation read by health care professionals.
- I was able to understand fully what has occurred to this patient and would not require any further clarification.
- Language used is clear and appropriate for documentation between health care professionals.

**Useful:** The note is extremely relevant and provides valuable information and or/analysis.

- Your document would not be useful for the health care team providing.
- Your document would have limited usefulness to other health care.
- Your document is a useful document to other health care.
- Your document is a useful admission note, containing relevant.

<table>
<thead>
<tr>
<th>Succinct: The document is concise, and without redundancy.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• There is redundant, or excessive, use of language and/or irrelevant material included.</td>
</tr>
<tr>
<td>• Your document appears in places to be a verbatim record of patient presentation with little attempt to eliminate redundancy.</td>
</tr>
<tr>
<td>• The note is too brief to make comment on this attribute.</td>
</tr>
<tr>
<td>• There has been an attempt to make the patient’s account more concise than the spoken word.</td>
</tr>
<tr>
<td>• There is redundant, or excessive, use of language and irrelevant material included.</td>
</tr>
<tr>
<td>• The information could be given more clearly with far fewer words.</td>
</tr>
<tr>
<td>• The information provided has not been tailored to meet the purpose of the document.</td>
</tr>
<tr>
<td>• Most of the information recorded is necessary.</td>
</tr>
<tr>
<td>• There is some redundant and irrelevant material included.</td>
</tr>
<tr>
<td>• The information could be given more clearly with fewer words.</td>
</tr>
<tr>
<td>• There is consideration of the information required to meet the purpose of an document.</td>
</tr>
<tr>
<td>• Your document is a concise account of the presentation.</td>
</tr>
<tr>
<td>• The information was clear and to the point, with minimal redundancy.</td>
</tr>
<tr>
<td>• All information presented was necessary.</td>
</tr>
<tr>
<td>• You have tailored your information to the purpose of the document.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synthesised: The document reflects the author’s understanding of the patient’s status and ability to develop a differential diagnosis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Your document provides little evidence that you understand what is happening with the patient.</td>
</tr>
<tr>
<td>• There is no synthesis, or links, between the history, examination and differential diagnosis.</td>
</tr>
<tr>
<td>• There is little or no meaningful interpretation of the admission.</td>
</tr>
<tr>
<td>• Some aspects of the document indicate that you understand what is happening with the patient, but this does not continue throughout the record.</td>
</tr>
<tr>
<td>• There is minimal synthesis between the history, examination and differential diagnosis.</td>
</tr>
<tr>
<td>• There is no discussion of what needs to happen next.</td>
</tr>
<tr>
<td>• Your document demonstrates that you have a reasonable understanding of what is happening with the patient.</td>
</tr>
<tr>
<td>• You link presenting symptoms and examination to an appropriate differential diagnosis.</td>
</tr>
<tr>
<td>• There is some consideration of what needs to happen next consistent with the presentation provided.</td>
</tr>
<tr>
<td>• There is clear and consistent synthesis of the entire record.</td>
</tr>
<tr>
<td>• It is clear that you understand what is happening to the patient, and what needs to happen next.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comprehensible: The reader easily understands the document</th>
</tr>
</thead>
<tbody>
<tr>
<td>• I had difficulty understanding the patient presentation and progress.</td>
</tr>
<tr>
<td>• The language used is essentially lay language and inappropriate for communication between health professionals.</td>
</tr>
<tr>
<td>• I was able to understand the basics of presentation and admission but would need significant clarification to continue to manage the patient.</td>
</tr>
<tr>
<td>• The language used needs to be more professional and technically precise.</td>
</tr>
<tr>
<td>• I was able to understand what has occurred to this patient and would only have to clarify a few issues in the history and examination.</td>
</tr>
<tr>
<td>• There is a clear use of the appropriate language for documentation read by health care professionals.</td>
</tr>
<tr>
<td>• I was able to understand fully what has occurred to this patient and would not require any further clarification.</td>
</tr>
<tr>
<td>• Language used is clear and appropriate for documentation between health care professionals.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Useful: The note is extremely relevant and provides valuable information and or/analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Your document would not be useful for the health care team providing.</td>
</tr>
<tr>
<td>• Your document would have limited usefulness to other health care.</td>
</tr>
<tr>
<td>• Your document is a useful document to other health care.</td>
</tr>
<tr>
<td>• Your document is a useful admission note, containing relevant.</td>
</tr>
<tr>
<td>care for this patient</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>• Your document would be minimally useful as a record for yourself of the case</td>
</tr>
</tbody>
</table>

Please select what you have observed in the student documentation for the following:

1. **Your use of the English language:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>is correct, with few spelling errors. It contains correct grammar and appropriate use of abbreviations or shortening of words and minimal jargon needs attention. Be mindful of correct spelling, try to minimize the use of abbreviations and jargon, and adhere to the basics of English grammar.</td>
<td></td>
</tr>
</tbody>
</table>

2. **Your use of the Electronic Health Record:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>is organised, with appropriate use of the EHR program and its fields. When information was placed in a different section the reason for this was understandable and referenced elsewhere. Coding of the presentation was accurate requires more attention to the details. Documentation needs to be within the appropriate sections, unless an obvious explanation for choosing to document elsewhere is given. Coding of the presentation was not adequate for use by other professionals, including administration.</td>
<td></td>
</tr>
</tbody>
</table>

3. **Professionalism in documentation:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>This is a professional document with safe, respectful use of language and expressions There are concerns that the language and expressions used in this document are unprofessional. Please discuss this further with your clinical tutor</td>
<td></td>
</tr>
</tbody>
</table>

If you would like to make more specific feedback: