Author/s:
Russell, N;Allebone, J;Dandash, O;Hoermann, R;Cheung, AS;Zajac, JD;Handelsman, DJ;Kanaan, RA;Grossmann, M

Title:
Effect of estradiol on cognition in men undergoing androgen deprivation therapy: A randomized placebo-controlled trial

Date:
2022-11-01

Citation:
Russell, N., Allebone, J., Dandash, O., Hoermann, R., Cheung, A. S., Zajac, J. D., Handelsman, D. J., Kanaan, R. A. & Grossmann, M. (2022). Effect of estradiol on cognition in men undergoing androgen deprivation therapy: A randomized placebo-controlled trial. Clinical Endocrinology, 97 (5), pp.622-633. https://doi.org/10.1111/cen.14689.

Nicholas Russell ORCID iD: 0000-0001-5596-6601
1. Orwa Dandash ORCID iD: 0000-0001-8845-7280
2. Ada Cheung ORCID iD: 0000-0001-5257-5525
3. Mathis Grossmann ORCID iD: 0000-0001-8261-3457

4. Title Page

Title

Effect of estradiol on cognition in men undergoing androgen deprivation therapy: a randomized placebo-controlled trial.

Short title

Estradiol and cognition in men

Authors

Nicholas Russell[a,b], James Allebone[c], Orwa Dandash[c], Rudolf Hoermann[a], Ada S. Cheung[a,b], Jeffrey D Zajac[a,b], David J Handelsman[d], Richard A Kanaan[c*], Mathis Grossmann[a,b*]

*equal senior authors

[a]Department of Medicine (Austin Health), The University of Melbourne, Heidelberg, Australia

[b]Department of Endocrinology, Austin Health, Heidelberg, Australia

[c]Department of Psychiatry, The University of Melbourne, Austin Health, Heidelberg, Australia

[d]ANZAC Research Institute, University of Sydney and Department of Andrology, Concord Hospital, Sydney, New South Wales, Australia.

Correspondence

Dr Nicholas Russell, Department of Medicine Austin Health, The University of Melbourne, 145 Studley Road, Heidelberg, Victoria, 3084, Australia. Tel + 613 9496 5000; Fax + 613 9496 3365; Email nicholas.russell@austin.org.au

Acknowledgements

This is the author manuscript accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cen.14689.

The Cogstate battery was developed by and licensed from Cogstate Limited.

Conflict of Interest Statement

Mathis Grossmann has received research funding from Bayer Pharma, Novartis, Weight Watchers, Lilly and speaker's honoraria from Besins Healthcare. Ada S Cheung has received funding from Besins Healthcare for investigator-initiated studies utilizing estradiol and progesterone. David J Handelsman has received institutional funding for investigator-initiated testosterone pharmacology research from Besins Healthcare and Lawley and has provided expert testimony in anti-doping and professional standards tribunals and testosterone litigation. Nicholas Russell, James Allebone, Richard A Kanaan, Rudolf Hoermann, and Jeffrey D Zajac declare that they have no conflict of interest.

Funding statement

This work was supported by the National Health and Medical Research Council (NHMRC) (project grant number: 1099173). Ada Cheung is supported by a NHMRC Early Career Fellowship (#1143333). NR was supported by an Australian Postgraduate Award.

Data availability

Data available from the corresponding author upon reasonable request

Structured summary

Objective: Roles for estradiol in modulating cognition in men remain uncertain. We assessed the isolated effects of estradiol on cognition in men in the absence of testosterone.

Design: Randomised trial of transdermal estradiol 0.9mg daily, or matched placebo, for 6 months, hypothesizing that estradiol would improve verbal learning, verbal memory and spatial problem solving over time.

Patients: Men receiving androgen deprivation therapy (ADT) for prostate cancer.

Measurements: Cognition was assessed by a tablet-based cognitive battery (Cogstate) at baseline, month 1, month 3, and month 6. Anxiety and depression symptoms were assessed using the Hospital Anxiety and Depression Scale.

Results: 78 participants were randomised. Baseline mean scores were 21.0 (sd 4.1) for the International Shopping List test (ISL), assessing verbal learning and memory (higher scores better), and 60.4 (sd 19.5) for the Groton Maze Learning test (GML), assessing spatial problem solving (lower scores better). There was no significant difference in performance over time for the estradiol group versus the placebo group for the ISL, mean adjusted difference (MAD) 0.7 (95% CI -1.2 – 2.5), p=0.36, or the GML, MAD -3.2 (95% CI -12.0 – 5.6), p-0.53. There was no significant difference between groups over time in performance in any other cognitive domain, or on depression or anxiety scores.

Conclusions: We found no major effects of estradiol on cognition in men with castrate testosterone concentrations. Although the cognitive effects of ADT are debated, this study suggests that any such effects are unlikely to be prevented by administration of estradiol.

Keywords

Androgens

Cognition

Estradiol

Hypogonadism

Men

Prostatic Neoplasms

Testosterone

1. Introduction

The role of sex steroids in modulating cognition in men remains uncertain [1]. The predominant male sex steroid, testosterone (T), acts as an agonist of the androgen receptor, via conversion to dihydrotestosterone (DHT), a more potent androgen receptor agonist, and via aromatisation to estradiol (E2) to act on estrogen receptors. In non-human mammals, the androgen receptor, and all species of estrogen receptor, are expressed in the male brain, including in regions important for memory and cognition such as hippocampus, prefrontal cortex, and amygdala [2]. The human male brain is also a site of local production of E2 from its substrate T, due to high

expression of aromatase, including in the thalamus, hypothalamus, amygdala, and medulla [3]. Actions of T in the male brain could therefore be mediated directly, via DHT, and/or via E2. For eugonadal men, different laboratories, using sensitive and sophisticated test paradigms, have reported causal evidence for modification of reward processing by short term high-dose exogenous E2 [4] and correlational evidence for endogenous T concentrations in modulation of parochial altruism [5] and behavioural flexibility [6]. Some animal and human studies suggest that E2 treatment might improve aspects of cognition in androgen deprived males [7]. However, in RCTs assessing broad domains of cognition with standard tests, no effects of T treatment on cognition in older men with low or low-to-normal serum T have been identified [8–10].

It has been reported that 69% of men receiving androgen deprivation therapy (ADT) for prostate cancer decline in at least one cognitive domain [11], although not all studies agree [12,13]. In a meta-analysis of 14 observational studies, ADT, typically induced with gonadotrophin-releasing hormone (GnRH) analogues, was associated with worse visuomotor task performance, but not with decrements in attention/working memory, executive functioning, language, verbal memory, visual memory, or visuospatial ability [14]. A subsequent meta-analysis of 2 prospective and 4 retrospective observational studies was inconclusive with respect to an association between ADT and overall cognitive impairment (defined as decline of at least 1.5 standard deviations on at least 2 tasks) [15]. These and other syntheses of the literature have been hampered by different approaches to analysis, including focus on individual domains or overall cognition, a diversity of cognitive tests employed, and a lack of prospective studies, such that cognitive effects of ADT remain poorly defined [13,15].

There is suggestive evidence from a small observational study [16] and a randomized controlled trial (RCT) [17] that putative verbal memory and visual attention effects of ADT might improve with E2 administration. In premenopausal women, 8 weeks of conjugated equine estrogen treatment reversed deficits in verbal memory induced by 12 weeks of GnRH agonist treatment [18]. In aged male rats, E2 administration has been reported to improve spatial

memory [19]. A recent review called for additional studies of the effects of exogenous E2 on cognition in androgen-deprived males [7].

ADT represents the only situation in which an ethical requirement for T replacement is absent for prolonged periods, allowing any effects of exogenous E2 on cognition to be observed in the absence of T. We conducted a randomised placebo-controlled trial of transdermal E2 in men undergoing ADT, designed to assess the effect on fat mass and bone microarchitecture over 6 months [20]. This experimental design permitted us to observe the isolated effects of E2 on pre-defined domains of cognition, in the absence of T. We hypothesised that men randomised to receive E2, compared to those randomised to receive placebo, would have improved verbal learning and memory, and spatial problem solving over time.

2. Materials and Methods

We conducted a 6-month, randomised, double-blinded, placebo-controlled, parallel-group trial at Austin Health, a tertiary referral hospital affiliated with The University of Melbourne. Methods have been reported previously [20]. Briefly, participants were recruited from outpatient clinics from November 2017 to February 2020. Recruitment was then ceased prematurely due to the COVID-19 pandemic. Men were eligible for the study if they had been receiving GnRH agonists or antagonists for prostate cancer for a minimum of 4 weeks to ensure nadir sex steroid concentrations at baseline, with that therapy intended to continue for at least a further 6 months. Exclusion criteria included: impaired performance status (Eastern Cooperative Oncology Group Performance Status (ECOG) > 2); history of venous thromboembolism (VTE); breast cancer; systolic blood pressure > 160 or diastolic blood pressure > 100; New York Heart Association class 3 or 4 angina or heart failure; stroke, transient ischaemic attack, myocardial infarction, or angina within 12 months; current oral glucocorticoid treatment; prior chemotherapy; alcohol or illicit drug abuse.

The trial protocol was approved by the Austin Health Human Research Ethics Committee (HREC/16/Austin/98) and each participant provided written informed consent. The trial was

preregistered with the Australian New Zealand Clinical Trials Registry (identifier 12614000689673). We followed the CONSORT checklist in reporting this randomised trial (Supplementary Table).

Participants were randomly allocated, in a concealed fashion, to one of two intervention groups: E2 gel 1 mL (0.9 mg) per day or matching placebo gel 1 mL per day. We previously observed that E2 gel 0.9 mg daily increased the minimum serum E2 of men undergoing ADT into the reference range for healthy older men [21]. Randomisation occurred as follows: first, participants were stratified by ADT duration (≤3 or > 3 months) and then by eligibility to undergo brain MRI scanning (data to be reported separately); second, participants were allocated by restricted randomisation, using a computer-generated randomisation scheme in blocks of size 4, to E2 or placebo in a ratio of 1:1. This step was administered by clinical pharmacists independent of trial investigators.

E2 gel was Sandrena™ 1 mg/g estradiol (Aspen Pharmacare, St Leonards, NSW Australia). Placebo gel was a-gel™ (Fresenius Kabi, North Ryde, NSW Australia) and matched the E2 gel for colour, smell, and consistency. E2 and placebo were re-packaged into identical 10mL syringes by pharmacy, with instructions to apply 1mL each morning to the skin of the upper arms or abdomen. The concealed treatment allocation maintained the blinding of participants, investigators, and clinicians during treatment, with blinding maintained until data analysis, after the database had been cleaned and locked.

Study visits were conducted at baseline, month 1, month 3, and month 6 of the intervention. Gel syringes were collected each visit and residual volume recorded to assess adherence. Adverse events were graded according to Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 [22]. Handedness was recorded based on the self-reported preferred hand used for writing.

2.1 Cognitive battery

At each study visit, cognitive scores in 6 domains were measured by a short cognitive battery (Cogstate, New Haven, CT) delivered via tablet computer which automatically captured reaction time and accuracy of performance. The battery took 20-30 minutes to complete, depending on participant response times. The instructions for each task were read aloud using a standard script by a single trained assessor (NR) who remained present during the assessment. For each task, a familiarisation practice trial was presented before the scored task. An entire practice battery was completed by each participant once, immediately prior to their baseline assessment. The Cogstate battery does not demonstrate practice effects after the second assessment [23,24]. For each task, standard completion and data integrity criteria were applied. Tests that failed to meet completion criteria were excluded from the analysis.

The tasks, in order of testing, were as follows [25,26]:

1. International Shopping List (ISL), a test of verbal learning, memory, and recall. In each of three rounds, 12 supermarket items were read aloud, and the participant was given 1 minute to recall them. The outcome measure was the total number of correct responses, with a higher score indicating better performance (possible range 0-36).
2. Groton Maze Learning (GML), a test of executive function and spatial problem solving. The participant was asked to find the hidden route through a maze formed by a 10 by 10 grid of tiles. The correct pathway from the top left tile to the bottom right tile contained 28 steps. Eligible moves were up, down, left, or right. A correct move was indicated by a green tick, an incorrect move by a red cross. Following an incorrect move, the participant had to press the previous correct tile, before making an alternative move. There were five rounds. The outcome measure was the total number of errors made, with a lower score indicating better performance.
3. Detection Test (DET), a test of processing speed, visual attention, and psychomotor function. A playing card was displayed face-down in the centre of the screen. The card turned face-up at irregular intervals with the participant asked to press the 'yes' key on screen as soon as they detected the card had turned face-up. An error sound was played if

the ‘yes’ key was pressed before the card had turned over. The test terminated after 35 correct responses. The outcome measure was the speed of performance, reported as the mean of the $\log_{10}$ transformed reaction times for correct responses. A lower score indicates better performance.

4. Identification Test (IDN), a test of attention, psychomotor function, and information processing speed. A playing card was displayed face-down in the centre of the screen. Each time the card turned face-up it revealed either a black or red joker card. The participant was asked to press the ‘yes’ key for a red card and the ‘no’ key for a black card. An error sound played when an error was made. The task terminated after 30 responses. The outcome measure was the speed of performance, measured as the mean of the $\log_{10}$ transformed reaction times for correct responses. A lower score indicates better performance.
5. One Card Learning (OCL), a test of visual memory. A series of playing cards were displayed, one at a time, in the centre of the screen. The participant was asked to choose the ‘yes’ key or the ‘no’ key on screen to indicate whether each card had been displayed previously. The participant was alerted to an error by an error sound. The outcome was the accuracy of performance, measured as the arcsine square root of the proportion of correct responses. A higher score indicates better performance.
6. One Back (ONB), a test of working memory and attention. A series of playing cards was displayed, one at a time, in the centre of the screen. The participant was asked to choose ‘yes’ if the card displayed was the same as the previous card, or ‘no’ if the card displayed was different to the previous card. The participant was alerted to an error by an error sound. The test terminated after 31 correct responses. The outcome measure was speed of performance, measured as the mean of the $\log_{10}$ transformed reaction times for correct responses. A lower score indicates better performance.
7. Groton Maze Learning – Delayed Recall (GMR), a test of memory. The participant was asked to complete the 28-step maze a final time at the end of the battery. The outcome measure was the total number of errors with a lower score indicating better performance.

8. International Shopping List – Delayed Recall (ISLR), a test of memory. The participant was asked to recall the 12-item shopping at the end of the battery, on this occasion without the items being read aloud. The outcome measure was the number of correct items with a higher score indicating better performance.

2.2 Anxiety and Depression

At every study visit, anxiety and depression symptoms were assessed using the Hospital Anxiety and Depression Scale (HADS) inventory [27]. The HADS asks participants to rate 7 anxiety symptoms and 7 depression symptoms on a 0-3 Likert scale, based on how they have been feeling over the previous week. Higher scores correspond to higher symptom intensity. Anxiety and depression scores are separately totalled. Scores of 0-7 are considered normal, 8-10 borderline abnormal, and 11-21 abnormal. Use of the HADS inventory has been validated in medical patients as a cross-sectional screening tool and for tracking longitudinal changes in anxiety and depression [28]. The HADS inventory has been commonly used in studies of prostate cancer [29].

2.3 Blood samples

Fasting morning pre-dose blood samples were drawn at each visit. Serum was stored at -80 degrees Celsius for batched analysis of sex steroid profile by liquid chromatography mass spectrometry (LCMS/MS) [30]. Limits of detection (LOD) and quantification (LOQ) were, respectively, 11 pmol/L and 18 pmol/L for E2, and 0.03 nmol/L and 0.09 nmol/L for T. CVs were 5-10% and 2-8% respectively, for within-run reproducibility at three levels of quality control samples.

2.4 Study design and pre-specified outcomes

The pre-specified primary outcomes of this trial were total fat mass[20] and total volumetric bone density at the distal tibia (reported separately). Herein we report results for the pre-specified secondary endpoints of task performance on the Cogstate battery and on depression and anxiety symptoms assessed by the HADS.

2.5 Power Analyses

Power calculations were based on the primary endpoints and required 43 participants per group [20]. We aimed to recruit 54 participants per group, allowing for 20% attrition.

2.6 Statistical Analysis

Cogstate test battery scores were used on their transformed normative scales and are reported as means (with standard deviations, sd). Welch's t-test was used for between-group comparisons of the baseline scores and other normally distributed data. For non-normally distributed data, we used two-samples Wilcoxon's test, and for proportions, chi-squared test, or Fisher's exact test with low numbers. In the RCT, we assessed the treatment effect of E2 administration over placebo over 6 months on cognition and mood. The primary analysis was intention to treat including all subjects that were randomised as per their assigned group. Per protocol analysis was additionally done as a sensitivity analysis including only subjects that completed the trial and adhered to the protocol. Trial outcomes are reported as the mean adjusted difference (MAD) between the E2 group and placebo group, surrounded by a profiled 95% confidence interval, over the course of the trial. The significance level was tested as a single p value over all time points. MADs and p values were determined with the use of repeated measures mixed effects models including interaction of time point with treatment group (the treatment effect), treatment group, time points, strata of ADT duration ($\leq$3 or > 3 months) and baseline values of the respective variable as fixed effects, and subject as a random effect [31]. For sex steroid analysis, values below the limit of detection (LOD) were assigned the value of the LOD/$\sqrt{2}$ [32]. A two-sided p-value of <0.05 was considered indicative of statistical significance. Because the various outcomes were inter-correlated, no adjustments for multiple testing were made. Analyses were performed with the R statistical base package (version 4.0.4 for Mac) and the added packages lme4 1.1-26 and effects 4.1-4 [31,33,34].

3. Results

255 men had preliminary assessments for eligibility and 39 were randomised to each intervention. The flow of participants assigned to E2 or placebo through each phase of the study, reasons for non-randomisation, and protocol violations are shown in Figure 1. The

final screened participant was not randomised because the human research ethics committee ordered cessation of new clinical trial visits due to the COVID-19 pandemic. In the placebo group one participant withdrew from the trial, due to anxiety about attending study visits in the context of the pandemic, 38 completed the trial, and 37 completed the Cogstate battery at the 6-month visit. In the E2 group, one man withdrew (failed to attend study visits) and another became unwell due to new onset brain metastases, such that 37 men completed follow up.

Baseline characteristics by group are shown in Table 1. Overall, median duration of ADT at baseline was 4.4 months (IQR 2-10), median baseline testosterone was 0.3 nmol/L (interquartile range 0.2; 0.5) and median baseline E2 was 82 pmol/L (interquartile range 62; 101). Serum E2 increased in the E2 group over 6 months compared to the placebo group, MAD 207 pmol/L (95% CI 123 – 292), p<0.001 (Figure 2).

3.1 Cognitive battery

Baseline performance was similar in the E2 and placebo groups (Table 2). The main outcome measures for this study were total number of correct words recalled across 3 trials on the ISL test and total number of errors made across 5 trials on the GML test. In the overall cohort, baseline mean scores were 21.0 words (sd 4.1) for ISL and 60.4 errors (sd 19.5) for GML. In the RCT, there was no significant difference in performance over time for the E2 group versus the placebo group for ISL, assessing verbal learning and memory, MAD 0.7 words (95% CI -1.2 – 2.5), p=0.36, or for GML, assessing spatial problem solving, MAD -3.2 errors (95% CI -12.0 – 5.6), p-0.53. There was no significant difference between groups in performance on any of the other cognitive tests (Table 2).

For ISL, in the overall cohort, we observed higher (better) recalled word count over time (p<0.001), at 6 months by 2.6 words (95% CI 1.3 – 3.9). For GML, in the overall cohort, we observed lower (better) total error count over time (p=0.02), at 6 months -5.6 errors (95% CI -11.8 – 0.63), compared to baseline.

There was no significant interaction of treatment group with ADT duration on cognitive test performance.

3.2 Hospital Anxiety and Depression Scale

In the overall cohort, mean baseline depression score was 3.0 (sd 2.3) and mean baseline anxiety score was 3.9 (sd 3.4). In the RCT, there was no significant difference in depression score over time between E2 and placebo groups, MAD -0.1 (95% CI -1.1 – 1.0), p=0.98. Similarly, there was no significant difference in anxiety score over time between E2 and placebo groups, MAD -0.3 (-1.4 – 0.7), p=0.25 (Table 3). In the overall cohort, we observed a small increase in depression score over time (p=0.004), at 6 months MAD 0.8 (95% CI 0.1 – 1.6), compared to baseline. For anxiety, we observed no time-related effect in the overall cohort.

3.3 Per protocol analysis

In a sensitivity per-protocol analysis excluding men with protocol violations (Figure 1), cognitive and mood outcomes were similar (data not shown).

3.4 Adverse events

There were no serious adverse events detected other than prostate cancer progression which occurred in two participants in the placebo arm and one participant in the E2 arm. Two participants in the E2 arm ceased the intervention because of adverse effects on the breast. Incident CTCAE grade 1 or grade 2 gynaecomastia was twice as common in the E2 arm (44% versus 21%).

4. Discussion

In this trial, 6-months of transdermal E2 gel 0.9mg/d for men with castrate circulating T concentrations had no effect, compared to placebo, on verbal learning and memory or on spatial problem solving, or on any other domain of cognition, as assessed across 8 tasks as part of a validated cognitive battery. Participants in this trial were not selected for subjective or objective neuropsychological symptoms and performed comparably to age-matched

norms (Table 2) [35]. For ISL and GML, we observed scores that were marginally (approximately 10%) better at 6 months than at baseline. This was unexpected because the Cogstate battery is reported to be free of learning effects, and we were anticipating deterioration in performance over time on ADT. It is unknown whether these observed overall cohort differences represent an effect of or adaption to T depletion over time, another effect of prostate cancer or its management, or regression to the mean.

Although clinically meaningful changes in ISL and GML test scores have not been clearly defined, the changes in the overall cohort observed over 6 months are very small. By comparison, the baseline mean ISL test scores in well-characterised cohorts of healthy older adults and age-matched individuals with amnestic mild cognitive impairment and Alzheimer's disease were, respectively, 25.1 (sd 4.5) words, 17.6 (5.3) words, and 11.0 (5.2) words[24]. In healthy individuals, mean ISL score is 27.0 (sd 5.3) words in the 18-34 year age group and 23.0 (sd 4.7) words in the 70-79 year age group. In a cross-sectional comparison of 393 cognitively normal late-middle aged individuals and 78 demographically matched individuals with clinical or pre-clinical mild cognitive impairment from the Wisconsin Registry for Alzheimer's Prevention, GML test total errors were 51.8 (sd 16.3) and 62.7 (sd 17.0) respectively[36]. Normative data for the GML test show that there is an increase in total errors from 40.9 (sd 15.3) in the 18-34 year age group to 64.4 (26.2) in the 70-79 year age group[35]. The results of our study do not exclude benefits of E2 on cognition of small to medium effect size, but for comparison, the differences between these published cohorts of healthy younger and healthy older individuals and between healthy late middle-aged or older adults and demographically matched cohorts with mild cognitive impairment, are of medium to large effect size.

Few other experiments have been designed to allow assessment of the effect of E2 on cognition in men in the absence of T. A small study (n=18) observed men with castrate-resistant prostate cancer switching from a GnRH agonist to second-line experimental

therapy with high dose transdermal E2 (six 100mcg/24h patches worn concurrently), sufficient to maintain castrate T concentrations [16]. These men were compared over 4 weeks to healthy aged-matched controls and controls with prostate cancer receiving standard ADT. Verbal memory was worse in men receiving ADT compared to controls and improved only in the group switched to E2. However, a 12-week randomized, controlled trial (RCT) (n=21) of low dose oral E2 (1mg/d) in men receiving ADT did not support the hypothesis of improved verbal memory with this intervention[37]. In a 9-week RCT (n=23) of oral E2 (1mg/d) in men receiving a GnRH agonist for prostate cancer, the E2 group improved on a test of visual attention but not in 14 of 16 other neurocognitive measures, some testing overlapping cognitive processes [17]. Observational studies in transfeminine individuals have not observed a consistent cognitive effect of E2 treatment [38].

Other experiments in humans, while not designed to allow direct observation of E2 effects in the absence of T, have been consistent with a lack of short-term cognitive effects of changes in sex steroid concentrations. In a 6-week placebo-controlled experiment, in which groups of healthy younger and older men were randomised to ADT for 6 weeks with or without T replacement, with or without aromatase inhibition, there were no effects on cognition from global castration (T-/E2-) or isolated E2 deficiency (T+/E2-) [10]. In another 8-week study of experimentally induced hypogonadism in healthy men and women, the expected sex difference in spatial task performance persisted during hypogonadism, suggesting that it is unlikely to be influenced by short term changes in any sex steroid [39]. This study did not reproduce the female advantage in verbal fluency at baseline making it difficult to interpret the lack of sex difference during conditions of hypogonadism. Notably, major randomised trials of T supplementation in men with age-related low or low-normal T but not organic hypogonadism have not demonstrated improvements in cognitive function [8,9,40]. These data suggest that any effects of sex steroids in men are likely to be small and might only be detected in circumstances of reduced cognitive reserve. Our study, the design of which

allowed direct observation of the isolated of E2 effects on cognition, is consistent with these results.

The absence of effects of E2 observed here does not preclude important enduring effects of E2 on brain development in men, nor was this trial designed to evaluate proposed neuroprotective effects of E2. Pre-clinical data have suggested a potential role for E2 in preventing dementia [41], but clinical evidence has not aligned with this. In a UK cohort of over 30,000 men with over 133,000 person-years of follow-up, ADT exposure was not associated with an increased future risk of dementia [42], although other large observational analyses have reported this association [13].

4.1 Strengths and limitations of the study

This was a rigorously conducted randomised double-blinded controlled trial. We used a cognitive test battery that has been validated for longitudinal assessment and is free of significant practice effects [23]. However, our study was not powered to detect effects of E2 of small to moderate effect size on the cognitive tasks assessed, and it made no assessment of sexual cognition (sexual fantasies or thoughts) which is reportedly modified by E2 in men [7]. We selected an E2 dose that was previously shown in men receiving ADT to produce minimum serum E2 concentrations within the reference range for healthy older men [21]. In this trial, achieved E2 concentrations in the E2 group at 6 months were similar to those in our previous study, median 265 pmol/L (interquartile range 188 – 385). Smoothed age-specific reference ranges for Australian men, using the same E2 assay, have been reported as 37-195 pmol/L (men < 65 years), 22-165 pmol/L (75-85 years) and 22-173 pmol/L (>85 years) [43]. We have therefore achieved trough serum E2 concentrations at or above those seen in healthy eugonadal older men. Peak serum E2 concentrations, not measured in this study, could have been substantially higher. However, the relationship between serum E2 concentrations and estrogen effects in the brain in men with castrate T, is unknown. Given that 50-75% of E2 in men is derived from extragonadal aromatisation of T, serum E2

concentrations in eugonadal men reflects spill over from aromatase-produced E2 in peripheral tissues [44]. The intervention in this study may therefore represent a sub-physiological E2 replacement in tissues with high aromatase activity such as the brain. Moreover, as has been observed with transdermal E2 preparations in women, serum E2 concentrations show variability, as previously observed in our pilot study [21]. Our study was not designed to determine reasons for such variability, but differences in transcutaneous absorption and metabolism of E2 are likely to play a role. Of note, a study of T gel has concluded that baseline participant characteristics account for only a small fraction of the substantial variance in on-treatment T concentrations[45], although whether the same applies to E2 is not known. However, clear separation of E2 concentrations between groups, and reductions in bone turnover markers and higher incidence of nipple tenderness in the E2 group are consistent with estrogen effects in the men assigned to E2.

4.2 Significance and conclusions

We used an experimental paradigm that allowed effects of E2 on cognition in the absence of T in older men to be investigated over 6 months. Despite restoring circulating E2 to concentrations at least as high as those seen in eugonadal men, we detected no effect of E2 on a cognitive task performance battery over 6 months. These results align with the only two previous RCTs examining this intervention, both of which were smaller and of shorter duration. Taken in the context of the wider literature, it seems likely that short term effects of E2 on cognition in men, if present, are subtle. Although cognitive effects of ADT are debated, this study suggests that any such effects are unlikely to be prevented by a strategy of E2 administration.

5. References

1. Kaufman JM, Lapauw B, Mahmoud A, T'Sjoen G, Huhtaniemi IT. Aging and the Male Reproductive System. *Endocr Rev*. 2019;40(4):906-972. doi:10.1210/er.2018-00178

2. Cooke PS, Nanjappa MK, Ko C, Prins GS, Hess RA. Estrogens in male physiology. *Physiological reviews*. 2017;97(3):995-1043. doi:10.1152/physrev.00018.2016

3. Takahashi K, Hosoya T, Onoe K, et al. Association between aromatase in human brains and personality traits. *Sci Rep-uk*. 2018;8(1):16841. doi:10.1038/s41598-018-35065-4

4. Joue G, Chakroun K, Bayer J, et al. Sex Differences and Exogenous Estrogen Influence Learning and Brain Responses to Prediction Errors. *Cereb Cortex*. 2021. doi:10.1093/cercor/bhab334

5. Reimers L, Kappo E, Stadler L, Yaqubi M, Diekhof EK. Endogenous testosterone correlates with parochial altruism in relation to costly punishment in different social settings. *Peerj*. 2019;7:e7537. doi:10.7717/peerj.7537

6. Diekhof EK, Kraft S. The association between endogenous testosterone level and behavioral flexibility in young men – Evidence from stimulus-outcome reversal learning. *Horm Behav*. 2017;89:193-200. doi:10.1016/j.yhbeh.2017.02.006

7. Wibowo E. Cognitive Impacts of Estrogen Treatment in Androgen-Deprived Males: What Needs to be Resolved. *Current neuropharmacology*. 2017;15(999):1-1. doi:10.2174/1570159x15666170313122555

8. Resnick SM, Matsumoto AM, Stephens-Shields AJ, et al. Testosterone treatment and cognitive function in older men with low testosterone and age-associated memory impairment. *Journal of the American Medical Association*. 2017;317(7):717-727. doi:10.1001/jama.2016.21044

9. Huang G, Wharton W, Bhasin S, et al. Effects of long-term testosterone administration on cognition in older men with low or low-to-normal testosterone concentrations: a prespecified secondary analysis of data from the randomised, double-blind, placebo-controlled TEAAM trial. *Lancet Diabetes Endocrinol*. 2016;4(8):657-665. doi:10.1016/s2213-8587(16)30102-4

10. Young LA, Neiss MB, Samuels MH, Roselli CE, Janowsky JS. Cognition is not modified by large but temporary changes in sex hormones in men. *The Journal of clinical endocrinology and metabolism*. 2010;95(1):280-288. doi:10.1210/jc.2009-1346

11. Nelson CJ, Lee JS, Gamboa MC, Roth AJ. Cognitive effects of hormone therapy in men with prostate cancer: a review. *Cancer*. 2008;113(5):1097-1106. doi:10.1002/cncr.23658

12. Jamadar RJ, Winters MJ, Maki PM. Cognitive changes associated with ADT: a review of the literature. *Asian Journal of Andrology*. 2012;14(2):232-238. doi:10.1038/aja.2011.107

13. Kluger J, Roy A, Chao HH. Androgen Deprivation Therapy and Cognitive Function in Prostate Cancer. *Current oncology reports*. 2020;22(3):24. doi:10.1007/s11912-020-0884-1

14. McGinty HL, Phillips KM, Jim HSL, et al. Cognitive functioning in men receiving androgen deprivation therapy for prostate cancer: a systematic review and meta-analysis. *Supportive care in cancer: : official journal of the Multinational Association of Supportive Care in Cancer*. 2014;22(8):2271-2280. doi:10.1007/s00520-014-2285-1

15. Sun M, Cole AP, Hanna N, et al. Cognitive Impairment in Men with Prostate Cancer Treated with Androgen Deprivation Therapy: A Systematic Review and Meta-Analysis. *The Journal of Urology*. 2018;199(6):1417-1425. doi:10.1016/j.juro.2017.11.136

16. Beer TM, Bland LB, Bussiere JR, et al. Testosterone loss and estradiol administration modify memory in men. *The Journal of Urology*. 2006;175(1):130-135. doi:10.1016/s0022-5347(05)00049-2

17. Taxel P, Stevens MC, Trahiotis M, Zimmerman J, Kaplan RF. The effect of short-term estradiol therapy on cognitive function in older men receiving hormonal suppression therapy for prostate cancer. *Journal of the American Geriatrics Society*. 2004;52(2):269-273.

18. Sherwin BB, Tulandi T. "Add-back" estrogen reverses cognitive deficits induced by a gonadotropin-releasing hormone agonist in women with leiomyomata uteri. *J Clin Endocrinol Metabolism*. 1996;81(7):2545-2549. doi:10.1210/jcem.81.7.8675575

19. Luine V, Rodriguez M. Effects of estradiol on radial arm maze performance of young and aged rats. *Behav Neural Biol*. 1994;62(3):230-236. doi:10.1016/s0163-1047(05)80021-4

20. Russell N, Hoermann R, Cheung AS, Zajac JD, Handelsman DJ, Grossmann M. Effects of estradiol on fat in men undergoing androgen deprivation therapy: a randomized trial. *Eur J Endocrinol*. October 2021. doi:10.1530/eje-21-0663

21. Russell N, Hoermann R, Cheung AS, et al. Short-term effects of transdermal estradiol in men undergoing androgen deprivation therapy for prostate cancer: a randomized placebo-controlled trial. *European Journal of Endocrinology*. 2018;178(5):565-576. doi:10.1530/eje-17-1072

22. NIH National Cancer Institute. Common Terminology Criteria for Adverse Events version 4.03. https://nciterms.nci.nih.gov/ncitbrowser/pages/advanced_search.jsf?dictionary=CTCAE&version=4.03. Published April 29, 2010. Accessed July 4, 2017.

23. Falleti MG, Maruff P, Collie A, Darby DG. Practice Effects Associated with the Repeated Assessment of Cognitive Function Using the CogState Battery at 10-minute, One Week and One Month Test-retest Intervals. *J Clin Exp Neuropsyc*. 2006;28(7):1095-1112. doi:10.1080/13803390500205718

24. Lim YY, Jaeger J, Harrington K, et al. Three-Month Stability of the CogState Brief Battery in Healthy Older Adults, Mild Cognitive Impairment, and Alzheimer's Disease: Results from the Australian Imaging, Biomarkers, and Lifestyle-Rate of Change Substudy (AIBL-ROCS). *Arch Clin Neuropsych*. 2013;28(4):320-330. doi:10.1093/arclin/act021

25. Fredrickson J, Maruff P, Woodward M, et al. Evaluation of the Usability of a Brief Computerized Cognitive Screening Test in Older People for Epidemiological Studies. *Neuroepidemiology*. 2010;34(2):65-75. doi:10.1159/000264823

26. Koyama AK, Hagan KA, Okereke OI, Weisskopf MG, Rosner B, Grodstein F. Evaluation of a Self-Administered Computerized Cognitive Battery in an Older Population. *Neuroepidemiology*. 2015;45(4):264-272. doi:10.1159/000439592

27. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiat Scand*. 1983;67(6):361-370. doi:10.1111/j.1600-0447.1983.tb09716.x

28. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale - A review of validation data and clinical results. *J Psychosom Res*. 1997;42(1):17-41.

29. Nead KT, Sinha S, Yang DD, Nguyen PL. Association of androgen deprivation therapy and depression in the treatment of prostate cancer: A systematic review and meta-analysis. *Urologic oncology*. 2017;35(11):664.e1-664.e9. doi:10.1016/j.urolonc.2017.07.016

30. Harwood DT, Handelsman DJ. Development and validation of a sensitive liquid chromatography–tandem mass spectrometry assay to simultaneously measure androgens and estrogens in serum without derivatization. *Clinica Chimica Acta*. 2009;409(1-2):78-84. doi:10.1016/j.cca.2009.09.003

31. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*. 2015;67(1):1-48. doi:10.18637/jss.v067.i01

32. Handelsman DJ, Ly LP. An Accurate Substitution Method To Minimize Left Censoring Bias in Serum Steroid Measurements. *Endocrinology*. 2019;160(10):2395-2400. doi:10.1210/en.2019-00340

33. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2021. https://www.R-project.org/.

34. Fox J, Weisberg S. Visualizing Fit and Lack of Fit in Complex Regression Models with Predictor Effect Plots and Partial Residuals. *J Stat Softw*. 2018;87(9):1-27. doi:10.18637/jss.v087.i09

35. *Cogstate Pediatric and Adult Normative Data*. Melbourne, Australia: Cogstate Ltd; 2018.

36. Racine AM, Clark LR, Berman SE, et al. Associations between Performance on an Abbreviated CogState Battery, Other Measures of Cognitive Function, and Biomarkers in People at Risk for Alzheimer's Disease. *J Alzheimer's Dis*. 2016;Preprint(Preprint):1-14. doi:10.3233/jad-160528

37. Matousek RH, Sherwin BB. A randomized controlled trial of add-back estrogen or placebo on cognition in men with prostate cancer receiving an antiandrogen and a gonadotropin-releasing hormone analog. *Psychoneuroendocrinology*. 2010;35(2):215-225. doi:10.1016/j.psyneuen.2009.06.012

38. Miles C, Green R, Hines M. Estrogen treatment effects on cognition, memory and mood in male-to-female transsexuals. *Horm Behav*. 2006;50(5):708-717. doi:10.1016/j.yhbeh.2006.06.008

39. Guerrieri GM, Wakim PG, Keenan PA, et al. Sex differences in visuospatial abilities persist during induced hypogonadism. *Neuropsychologia*. 2016;81:219-229. doi:10.1016/j.neuropsychologia.2015.12.021

40. Emmelot-Vonk MH, Verhaar HJJ, Pour HRN, et al. Effect of Testosterone Supplementation on Functional Mobility, Cognition, and Other Parameters in Older Men: A Randomized Controlled Trial. *JAMA*. 2008;299(1):39-52.

41. Sahab-Negah S, Hajali V, Moradi HR, Gorji A. The Impact of Estradiol on Neurogenesis and Cognitive Functions in Alzheimer's Disease. *Cell Mol Neurobiol*. 2020;40(3):283-299. doi:10.1007/s10571-019-00733-0

42. Khosrow-Khavar F, Rej S, Yin H, Aprikian A, Azoulay L. Androgen Deprivation Therapy and the Risk of Dementia in Patients With Prostate Cancer. *J Clin Oncol*. 2017;35(2):201-207. doi:10.1200/jco.2016.69.6203

43. Handelsman DJ, Yeap B, Flicker L, Martin S, Wittert GA, Ly LP. Age-specific population centiles for androgen status in men. *European Journal of Endocrinology*. 2015;173(6):809-817. doi:10.1530/eje-15-0380

44. Russell N, Cheung A, Grossmann M. Estradiol for the mitigation of adverse effects of androgen deprivation therapy. *Endocrine-Related Cancer*. 2017;24(8):R297-R313. doi:10.1530/erc-17-0153

45. Bhasin S, Travison TG, O'Brien L, et al. Contributors to the substantial variation in on-treatment testosterone levels in men receiving transdermal testosterone gels in randomized trials. *Andrology-us*. 2018;6(1):151-157. doi:10.1111/andr.12428

6. Tables

Table 1: Baseline characteristics

	Placebo group (n=39)	**Estradiol group (n=39)**	**P Value**
Age (years) *mean (sd)*	72.1 (+/-6.7)	71.7 (+/-7.8)	0.82
Non-white race *n (%)*	0	2 (5%)	0.49
Left-handed *n (%)*	4 (10%)	2 (5%)	0.67
ECOG *n (%)*			0.73
0	35 (90%)	35 (90%)	
1	3 (8%)	3 (8%)	
2	1 (3%)	1 (3%)	
Educational (years) *mean (sd)*	14.4 (+/- 4.8)	14.4 (4.7)	0.91
Body mass index (kg/m2) *median (IQR)*	29.2 (27.4;32.0)	28.2 (26.3;32.7)	0.79
Prostate cancer stage *n (%)*			0.82
Localised	13 (33%)	10 (26%)	
Locally advanced	6 (15%)	10 (26%)	
Local recurrence	1 (3%)	1 (3%)	
M0	9 (23%)	7 (18%)	
Metastatic CSPC	10 (26%)	10 (26%)	
Metastatic CRPC	0 (0%)	1 (3%)	
PSA *median (IQR)*	0.17 (0.04;0.89)	0.24 (0.04;2.03)	0.49
Prostatectomy *n (%)*	18 (46%)	13 (13%)	0.36
Radiotherapy (previous or concurrent) *n (%)*	34 (87%)	31 (80%)	0.54
Prior course of ADT *n (%)*	5 (13%)	3 (8%)	0.71
ADT duration at baseline (months) *median (IQR)*	4.7 (2.1;9.9)	4.0 (2.0;8.8)	0.59
Planned ADT duration *n (%)*			0.84
6 months	5 (13%)	3 (8%)	
12 months	0	1 (3%)	

18 months	8 (21%)	6 (15%)	
24 months	8 (21%)	14 (36%)	
36 months	2 (5%)	1 (3%)	
Indefinite	15 (38%)	13 (33%)	
Undefined	1 (3%)	1 (3%)	
Smoking status *n (%)*			
Current *n (%)*	2 (5%)	0	0.49
Ex-smoker *n (%)*	19 (49%)	20 (51%)	1.0
Never smoked *n (%)*	18 (46%)	19 (49%)	0.82
Pack years *median* (IQR)	6.0 (0.0;17.5)	5.0 (0.0;16.5)	0.85
Alcohol (standard drinks/week) *median (IQR)*	3 (1;7)	5 (1;8)	0.60
Medical comorbidities *n (%)*			
Diabetes	7 (18%)	5 (13%)	0.75
Hypertension	31 (79%)	25 (64%)	0. 44
Chronic liver disease	1 (3%)	1 (3%)	1
Stroke	0	1 (3%)	1
IHD	4 (10%)	9 (23%)	0.22
Heart Failure	1 (3%)	0	1
CKD 3	5 (13%)	3 (8%)	0.46
CKD > 3	0	0	1
Medications *n (%)*			
GnRH agonist	35 (90%)	36 (92%)	1
GnRH antagonist	5 (13%)	3 (8%)	1
Anti-androgen	2 (5%)	2 (5%)	1
Testosterone (nmol/L) *median [IQR]*	0.3 [0.2;0.4]	0.3 [0.2;0.5]	0.76
Estradiol (pmol/L) *median [IQR]*	83 [61;100]	82 [67;102]	0.56

ADT, androgen deprivation therapy
CKD, chronic kidney disease
CSPC, castrate sensitive prostate cancer
CRPC, castrate resistant prostate cancer
ECOG, Eastern Cooperative Oncology Group Performance Status
GnRH, gonadotrophin-releasing hormone
IHD. Ischaemic heart disease
PSA, prostate specific antigen

Table 2: Cognitive battery results

TEST	**Normative Mean (sd)**[1]	**Placebo group (n=39)** Mean (sd)	**E2 group (n=39)** Mean (sd)	**MAD (E2 vs Placebo)** (95% CI)	**Overall P value**
HIGHER SCORE BETTER					
International Shopping List (total correct)	23.0 (4.7)	21.6 (3.4) 22.9 (4.0)	20.5 (4.7) 23.1 (4.9)	1.3 (-0.6 – 3.1) 1.6 (-0.3 – 3.5)	
0 months		22.7 (4.9)	23.2 (4.1)	0.7 (-1.2 – 2.5)	0.36
1 month		24.1 (5.2)	23.8 (4.4)		
3 months					
6 months					
International Shopping List Recall (total correct)	7.5 (2.4)	6.1 (2.4)	6.2 (2.5)	0.3 (-0.7 – 1.3)	
0 months		7.8 (2.3)	8.1 (2.4)	-0.1 (-1.1 – 0.9)	0.85
1 month		7.7 (2.4)	7.6 (2.5)	0.3 (-0.7 – 1.3)	
3 months		7.9 (2.8)	8.2 (2.2)		
6 months					
One Card Learning (accuracy[2])	0.95 (0.13)	0.91 (0.11)	0.94	0.01 (-0.04 –	
0 months		0.98 (0.10)	(0.09)	0.05)	0.98
1 month		0.96 (0.12)	1.01	0.01 (-0.04 –	
3 months		0.97 (0.11)	(0.10)	0.06)	

6 months			0.99 (0.11) 1.00 (0.11)	0.01 (-0.04 – 0.06)	
LOWER SCORE BETTER					
Groton Maze Learning (total errors) 0 months 1 month 3 months 6 months	64.4 (26.2)	56.2 (18.7) 52.3 (15.2) 57.9 (21.2) 53.1 (18.8)	64.6 (19.7) 57.3 (17.1) 61.0 (19.5) 60.7 (22.8)	-4.4 (-13.2 – 4.4) -6.7 (-15.6 – 2.2) -3.2 (-12.0 – 5.6)	0.53
Groton Maze Learning Recall (total errors) 0 months 1 month 3 months 6 months	13.4 (6.3)	9.30 (6.24) 9.68 (6.19) 8.71 (5.50) 8.60 (4.62)	11.7 (7.73) 10.5 (4.21) 10.5 (4.64) 9.52 (5.14)	-2.3 (-5.1 – 0.5) -1.0 (-3.8 – 1.8) -1.2 (-4.0 – 1.6)	0.46
Identification (speed[3]) 0 months 1 month 3 months 6 months	2.75 (0.09)	2.74 (0.06) 2.74 (0.07) 2.75 (0.06) 2.75 (0.06)	2.76 (0.06) 2.76 (0.09) 2.75 (0.04) 2.76 (0.07)	-0.00 (-0.02 – 0.02) -0.01 (-0.04 – 0.01) 0.00 (-0.02 – 0.02)	0.78
Detection (speed[3]) 0 months 1 month 3 months 6 months	2.58 (0.11)	2.58 (0.08) 2.62 (0.11) 2.62 (0.13) 2.63 (0.10)	2.63 (0.11) 2.63 (0.10) 2.64 (0.13) 2.64 (0.13)	-0.04 (-0.08 – 0.01) -0.03 (-0.07 – 0.02) -0.03 (-0.08 – 0.01)	0.36
One Back (speed[3]) 0 months 1 month 3 months 6 months	2.94 (0.10)	2.96 (0.10) 2.96 (0.09) 2.94 (0.08) 2.95 (0.10)	2.97 (0.10) 2.96 (0.11) 2.95 (0.08) 2.95 (0.09)	-0.01 (-0.04 – 0.02) -0.01 (-0.04 – 0.02) -0.00 (-0.03 – 0.02)	0.86

[1]*Normative mean and standard deviation are derived from the Cogstate normative dataset*[35]*. They are presented here for illustration. Formal statistical comparisons were not made with the normative dataset. The normative dataset is aggregated from multiple studies of healthy adults performed in 5 continents. This dataset is divided into age groups. Data for 70-79 year-olds is presented, which, for each test, is very similar for the data for the 60-69 and 80-89 year age groups.*

[2]accuracy of performance measured as the arcsine square root of proportion of correct responses

[3]speed of performance measured as the mean of the $\log_{10}$ of the reaction time for correct responses in milliseconds

Abbreviations:

MAD, mean adjusted difference

Table 3: Hospital Anxiety and Depression Scale results

	Placebo group (n=39) Mean (SD)	**E2 group (n=39)** Mean (SD)	**Mean adjusted difference (E2 vs Placebo)** (95% CI)	**Overall P value**
Depression Score				
0 months	3.0 (2.4)	3.0 (2.2)	-0.2 (-1.2 – 0.8)	
1 month	3.0 (2.7)	2.6 (2.2)	-0.2 (-1.2 – 0.9)	0.98
3 months	3.4 (3.2)	3.0 (2.8)	-0.1 (-1.1 – 1.0)	
6 months	3.7 (3.6)	3.4 (3.5)		
Anxiety Score				
0 months	3.1 (3.0)	4.7 (3.5)	-1.0 (-2.02 to -0.0)	
1 month	3.6 (3.6)	4.0 (3.3)	-0.5 (-1.46 – 0.1)	0.25
3 months	3.4 (3.1)	4.5 (3.9)	-0.3 (-1.35 – 0.7)	
6 months	3.4 (3.4)	4.6 (4.0)		

7. Figure legends

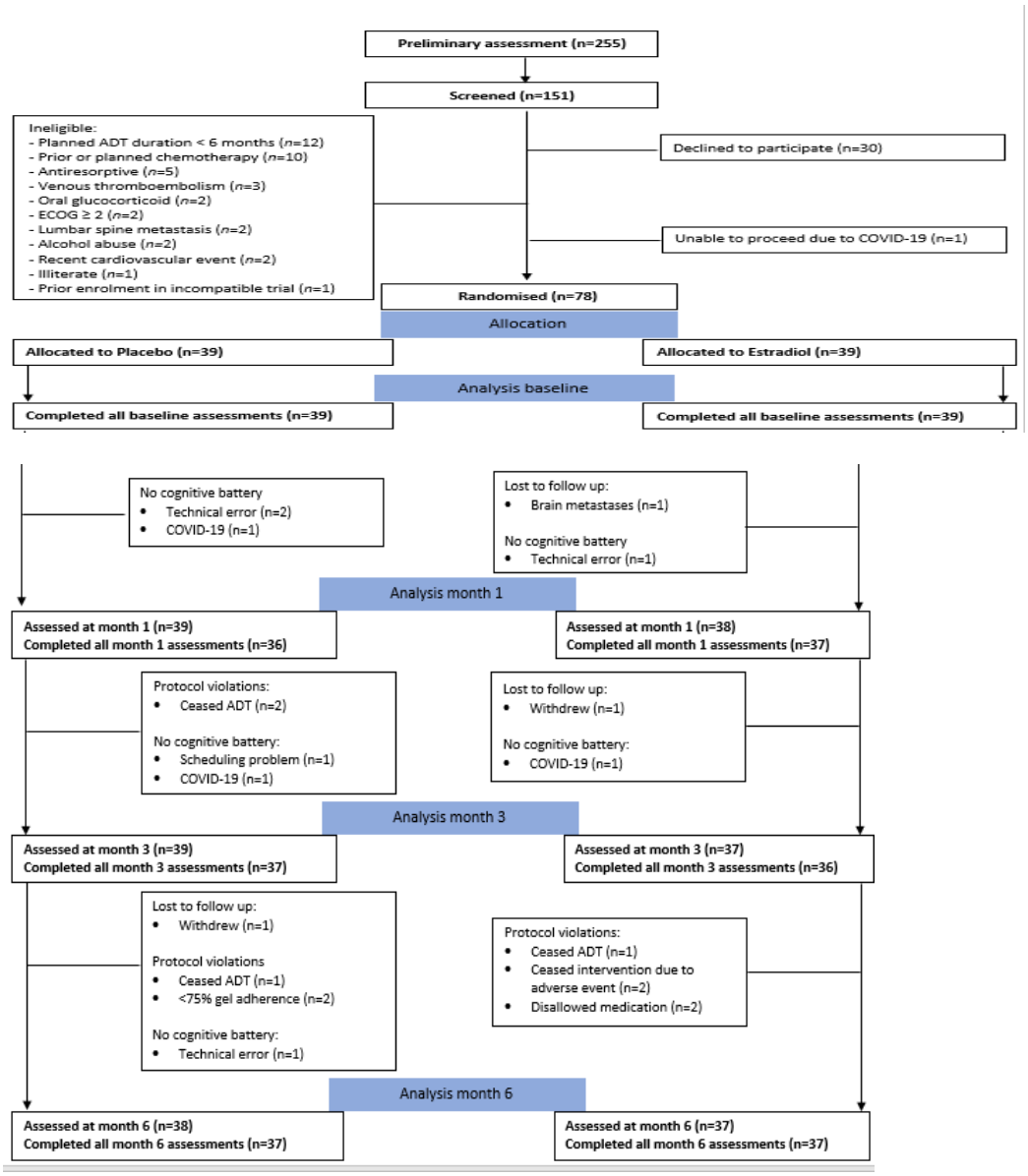

Figure 1: CONSORT flow diagram

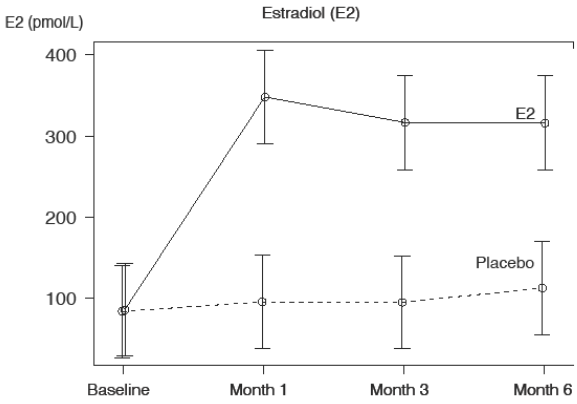

Figure 2: Adjusted mean and 95% confidence interval of serum estradiol concentration (pmol/L) by group and study visit. Mean adjusted difference, E2 group versus placebo group, 207 pmol/L (95% CI 123 – 292), p<0.001.