

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## Systematic review: the Androgens In Men Study.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-048013
Article Type:	Original research
Date Submitted by the Author:	11-Jan-2021
Complete List of Authors:	Marriott, Ross; The University of Western Australia, School of Population and Global Health Harse, Janis; The University of Western Australia, School of Population and Global Health Murray, Kevin; The University of Western Australia, School of Population and Global Health Yeap, Bu; The University of Western Australia, Medical School; Harry Perkins Institute of Medical Research
Keywords:	EPIDEMIOLOGY, Sex steroids & HRT < DIABETES & ENDOCRINOLOGY, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1 Systematic review: the Androgens In Men Study.

2

3 Corresponding author:

4 Ross J. Marriott PhD, School of Population and Global Health, Faculty of Health and  
5 Medical Sciences, The University of Western Australia, Clifton Street Building, Clifton  
6 Street, Nedlands 6009, Western Australia, Australia.

7 Email: [ross.marriott@uwa.edu.au](mailto:ross.marriott@uwa.edu.au); Telephone: +61 8 6488 1299

8

9 Authors:

10 Ross J. Marriott, School of Population and Global Health, The University of Western  
11 Australia, Nedlands 6009, Australia.

12 Janis Harse, School of Population and Global Health, The University of Western Australia,  
13 Nedlands 6009, Australia.

14 Kevin Murray, School of Population and Global Health, The University of Western Australia,  
15 Nedlands 6009, Australia.

16 Bu B Yeap, Medical School, The University of Western Australia, Perth 6009, Australia.

17

18 Short title:

19 Systematic review: Androgens In Men Study

20

21 Keywords:

22 Testosterone, Individual Participant Data, Cardiovascular Disease, Cancer, Mortality,  
23 Dementia, Meta-analysis.

24

1  
2  
3 25 Word count, excluding title page, abstract, strengths and limitations, references,  
4  
5 26 acknowledgements, contributions, figures and tables: 2,724 words.  
6  
7  
8 27

9  
10 28 **ABSTRACT**

11  
12 29 Objectives

13  
14 30 The overall study aim is to clarify the relation of endogenous sex hormones (primarily  
15  
16 31 testosterone) with major health outcomes in men.

17  
18 32 Setting

19  
20 33 Community-dwelling men.

21  
22 34 Participants

23  
24 35 20,180 adult males participated in the final set of studies identified and selected from a  
25  
26 36 systematic review. Eligible studies included prospective cohort studies with plasma or serum  
27  
28 37 testosterone concentrations measured for adult males using mass spectrometry with at least 5  
29  
30 38 years of follow-up data, with incident cardiovascular, cancer, mortality, dementia or  
31  
32 39 cognitive events recorded. Only published or grey literature items written in English were  
33  
34 40 considered.

35  
36 41 Primary and secondary outcome measures

37  
38 42 Planned prospective outcome measures: cardiovascular disease (CVD) events, CVD deaths,  
39  
40 43 all-cause mortality, cancer deaths, cancer diagnoses, cognitive decline, dementia. Outcome  
41  
42 44 measures analysed in this paper were of the published estimates most frequently reported in  
43  
44 45 selected studies: CVD deaths, all-cause mortality. All planned outcomes will be investigated  
45  
46 46 for the selected studies as a separate series of individual participant data (IPD) meta-analyses.

47  
48 47 Results

49  
50 48 Screening of 1,994 de-duplicated items identified 9 suitable studies, with an additional two  
51  
52 49 identified by colleagues (11 in total). Summary estimates of mean testosterone concentration  
53  
54 50

1  
2  
3 50 and age at recruitment for 20,180 adult males were  $15.4 \pm 0.7$  nmol/L and  $64.9 \pm 3.3$  yr. Despite  
4  
5 51 considerable variation in mean testosterone, a meta-regression estimated no significant  
6  
7 52 dependence on mean age at recruitment among studies (Slope = -0.03, 95% CI -0.11 – 0.06).  
8  
9  
10 53 Meta-analyses demonstrated no significant effect of a 5 nmol/L increase in testosterone on  
11  
12 54 the risk of all-cause mortality (hazard ratio, HR = 0.96, 95% CI 0.89 – 1.03) or death from  
13  
14 55 CVD (HR = 0.95, 95% CI 0.83 – 1.08).  
15  
16

## 17 56 Conclusions

18  
19 57 Analyses of published estimates did not demonstrate associations of endogenous testosterone  
20  
21 58 with CVD deaths or with all-cause mortality. Suggested further research includes the planned  
22  
23 59 IPD meta-analyses for selected studies, including scope for investigating non-linear effects.  
24  
25

## 26 60 Registration

27  
28 61 PROSPERO: CRD42019139668.  
29  
30 62

## 31 32 63 **STRENGTHS AND LIMITATIONS OF THIS STUDY**

- 33  
34 64 • This is the first systematic review on this topic to restrict selections to prospective cohort  
35  
36 65 studies of community-dwelling men with testosterone measured using mass spectrometry:  
37  
38 66 the “gold standard” method.  
39  
40 67 • Systematic searches were made of both the published and grey literature using online  
41  
42 68 search tools.  
43  
44 69 • Meta-analyses used estimates obtained from studies with at least five years of follow-up  
45  
46 70 data and from fitted models which controlled for (at least) the age, smoking status, and  
47  
48 71 body mass index or waist circumference of participants.  
49  
50 72 • Meta-analyses of published estimates were limited to assuming linear relationships,  
51  
52 73 however subsequent IPD meta-analyses planned to arise from this work will look to  
53  
54 74 explore non-linear associations.  
55  
56  
57  
58  
59  
60

- 1  
2  
3 75 • Analyses are of observational data, and so summary estimates will not fully eliminate the  
4  
5 76 possibility of confounding arising from unadjusted effects.  
6  
7  
8 77

9  
10 78 **1. INTRODUCTION**

11  
12 79 What does a low testosterone level mean for a man's health? In men, levels of testosterone,  
13  
14 80 the key male sex hormone (androgen), decline with increasing age, yet the basis for and  
15  
16 81 health consequences of this phenomenon remain unclear.[1-5] Many middle- and older-aged  
17  
18 82 men are told their levels are "too low", explaining the 12-fold increase in global testosterone  
19  
20 83 prescriptions over 2000-2011, costing \$1.8 billion.[6] The Androgens In Men Study (AIMS)  
21  
22 84 will seek to clarify the associations of androgens (primarily testosterone) with key health  
23  
24 85 outcomes in men (mortality, cardiovascular disease, cancer, cognitive decline and dementia).  
25  
26 86 The AIMS will conduct a systematic review and a series of individual participant data meta-  
27  
28 87 analyses to address these questions. In this paper we present the systematic review and meta-  
29  
30 88 analyses using published estimates from prospective cohort studies with at least 5 years of  
31  
32 89 follow-up data and testosterone measured using only mass spectrometry, the most reliable  
33  
34 90 method.[7]  
35  
36  
37  
38  
39  
40  
41

42 92 **2. METHODS**

43  
44 93 This systematic review, conducted 14 June—31 December 2019, was of "etiology and/or risk  
45  
46 94 type" studies.[8, 9] The pre-specified purpose of the systematic review was to identify studies  
47  
48 95 with suitable individual participant-level data (IPD) for collaborating with on a series of IPD  
49  
50 96 meta-analyses. The PEO (Population, Exposure, Outcomes) characteristics included: adult  
51  
52 97 men in the general community; endogenous circulating sex hormone concentration (primarily  
53  
54 98 testosterone); incident cardiovascular disease (CVD), mortality, cancers, cognitive decline,  
55  
56 99 dementia. Subgroup IPD meta-analyses are also planned for heart failure, myocardial  
57  
58  
59  
60

1  
2  
3 100 infarction, stroke; colorectal cancer, lung cancer, prostate cancer. A protocol was submitted  
4  
5 101 to PROSPERO on 23 July 2019 and registered on 20 November 2019 (registration number  
6  
7 102 CRD42019139668) and a protocol article has been published.[10]  
8  
9

10  
11

## 104 **2.1. Literature search and screening**

105 Four online search tools were used to identify available published (MEDLINE, EMBASE)  
106 and grey literature (OpenGrey, Mednar) *items* (journal articles, reports, theses, webpage  
107 articles) reporting on suitable prospective cohort *studies* (the underlying unique sources of  
108 data). Two reviewers (RJM, JH) independently screened the de-duplicated items against pre-  
109 specified criteria using Rayyan.[11] To optimise efficiency, title and abstract screenings were  
110 initially conducted (Step 1), followed by full text screenings of the selected abstracts (Step 2).  
111 Disagreements were resolved through subsequent discussions between reviewers and  
112 agreement quantified using Cohen's Kappa and percent agreement. Only items reporting on  
113 prospective population-based cohort studies of adults (combined sexes or of men alone) with  
114 mass spectrometry measurements of testosterone and at least five years of subsequent follow-  
115 up data on incident CVD events, cancer or dementia diagnoses, cognition assessments, or on  
116 all-cause, CVD, or cancer deaths were selected. The Newcastle-Ottawa Quality Assessment  
117 Scale for Cohort Studies (NOS) was used to assess quality of the selected items.[12] The  
118 terms and full criteria used for the MEDLINE search, PRISMA checklist, NOS star ratings  
119 and additional methods details are included in Supplementary Material.

120

## 121 **2.2. Meta-analyses of published estimates**

122 Published estimates (author names, publication year, cohort study name, number of  
123 participants analysed, model covariates, testosterone statistics (overall and for individual  
124 exposure levels), participant age statistics, numbers of outcome events, follow-up time,



1  
2  
3 125 hazard ratios (HRs) and 95% CIs of the most fully-adjusted model) were extracted from  
4  
5 126 selected articles by the first author (RJM). Testosterone statistics were converted into  
6  
7 127 standard units (nmol/L) and values representing categorical ranges were determined  
8  
9  
10 128 following Wang et al.[13] If not reported, the numbers of participants and events within  
11  
12 129 categories of testosterone, and the means of participant ages and testosterone concentrations  
13  
14 130 at baseline, were calculated. The numbers of participants within quartile or quintile categories  
15  
16 131 were calculated by dividing the total sample size by four or five. The numbers of events  
17  
18 132 within categories were solved using Newton's method by applying the algorithm of  
19  
20 133 Greenland and Longnecker.[14] Means and standard deviations for testosterone and age were  
21  
22 134 calculated from presented quartile estimates using the Box-Cox method of McGrath et al.,  
23  
24 135 which does not make distributional assumptions.[15]  
25  
26  
27  
28  
29

30  
31 137 A random effects meta-regression of mean baseline testosterone concentration on the mean  
32  
33 138 participant age at baseline was conducted using published estimates from: (i) only those items  
34  
35 139 identified in systematic searches; and (ii) all suitable articles, including those found outside of  
36  
37 140 systematic searches. A t-test of the meta-regression slope coefficient's departure from zero  
38  
39 141 was done after applying the Knapp and Hartung adjustment.  
40  
41  
42  
43

44 143 Dose-response random effects meta-analyses (DR-MAs) were conducted to summarise  
45  
46 144 published HR estimates for the associations of baseline testosterone concentrations with  
47  
48 145 incident all-cause deaths and with CVD cause-specific deaths. Estimates from an additional  
49  
50 146 article that had not been selected from systematic searches (Yeap et al[16]) were also used  
51  
52 147 because it reported suitable estimates from one of the selected studies, and had been  
53  
54 148 published within the literature search period. Contour-enhanced funnel plots were inspected  
55  
56 149 for publication bias and patterns in heterogeneity and Cochran Q tests for heterogeneity ( $I^2$ ),  
57  
58  
59  
60

1  
2  
3 150 as well as regression tests for funnel plot asymmetry,[17] were done.  
4  
5  
6 151

7  
8 152 The “metafor” package was used for meta-regressions, forest plots and funnel plots, the  
9  
10 153 “doseresmeta” package for DR-MAs, and the “estmeansd” package for calculating study  
11  
12 154 means and standard deviations from published quartile statistics in R version 4.0.2.[18-21]  
13  
14  
15 155

### 16 156 **2.3. Patient and public involvement**

17  
18  
19 157 This work uses existing published data. Patients and public were not involved in the design,  
20  
21  
22 158 conduct, reporting, or dissemination plans of the systematic review or meta-analyses.  
23  
24 159

## 25 26 160 **3. RESULTS**

### 27 28 161 29 30 31 162 **3.1. Literature search and study selection**

32  
33 163 The literature search returned 2,177 items (1,738 published and 439 from grey literature),  
34  
35 164 with 1,994 items remaining after duplicates had been removed, and after excluding two  
36  
37 165 Mednar items that had insufficient information available to review (Fig. 1). These included  
38  
39 166 1,764 journal articles, 111 webpage articles, 81 theses, and 38 unpublished reports/other  
40  
41 167 documents. Systematic screening of the returned, deduplicated items excluded 1,968,  
42  
43 168 classified five as “Maybe”, and selected 20 as suitable. Most (92.1%) items were excluded  
44  
45 169 from screening titles and abstracts at Step 1, with a much smaller percentage (6.6%) excluded  
46  
47 170 from screening the 157 full text items in Step 2. One item could not be screened in Step 2  
48  
49 171 because the full text was not available. Inter-reviewer agreement was a Cohen’s Kappa  
50  
51 172  $\kappa = 0.69$  (or 96.0 percent agreement) for Step 1 and  $\kappa = 0.82$  (or 98.1 percent agreement) for  
52  
53 173 Step 2.  
54  
55  
56  
57  
58 174  
59  
60

1  
2  
3 175 The 20 selected items collectively reported on eight prospective cohort studies: three from  
4  
5 176 Australia (Busselton Health Study *BHS*, The Concord Health and Ageing in Men Project  
6  
7 177 *CHAMP*, The Health In Men Study *HIMS*); three from Europe (European Male Ageing Study  
8  
9 178 *EMAS*, The *MrOS* Osteoporotic Fractures in Men study in Sweden, Study of Health in  
10  
11 179 Pomerania *SHIP*); and two from the USA (Atherosclerosis Risk in Communities *ARIC*,  
12  
13 180 Cardiovascular Health Study *CHS*). Two of the five items classified as “Maybe” reported on  
14  
15 181 the *MrOS USA* study, which were found, after further investigation, to be suitable for  
16  
17 182 selection. Two additional studies were identified as suitable based on information external to  
18  
19 183 the systematic searches and screenings: one from Australia (The Men Androgen  
20  
21 184 Inflammation Lifestyle Environment and Stress study *MAILES*); and one from the USA (the  
22  
23 185 Framingham Heart Study *FHS*). This is 11 cohort studies identified, in total. Additional details  
24  
25 186 on returned and screened items, and selected article attributes are provided in Supplementary  
26  
27 187 Material.  
28  
29  
30  
31  
32

188

### 189 **3.2. Meta-analysis and summary of selected articles.**

190 The quality of selected articles ranged from six to nine (out of nine) stars on the Newcastle-  
191 Ottawa Scale. Relatively high scores reflected that all articles: were of population-based  
192 studies; accurately measured the exposure (baseline testosterone concentration); included  
193 multivariable models adjusting for participant age and other risk factors; had outcomes  
194 measured or collected from record linkage, with or without expert adjudication; and had  
195 sufficient follow-up (Tables S5-S7, Supplementary Material). Relevant outcomes included  
196 CVD deaths (n=7 articles); all-cause deaths (n=6); strokes or cerebrovascular disease (n=6);  
197 cognitive function or cognitive decline (n=5); coronary heart disease (n=4); CVD events  
198 (n=4); cancer diagnoses (n=3); myocardial infarction (n=2); heart failure (n=1); and dementia

199 (n=1). All were published between 2010 and 2018, reflecting the relatively recent adoption of  
200 mass spectrometry as the “gold standard” for measuring endogenous testosterone levels.[7]

201  
202 The mean age of men at baseline ranged from middle-aged (49-54yr: BHS, FHS, MAILES,  
203 SHIP)[22-27] to elderly (72-77yr: CHAMP, CHS, HIMS, MrOS Sweden, MrOS USA).[28-  
204 36] Across the 11 studies, summary estimates for 20,180 adult males at baseline were  
205  $64.9 \pm 3.3$ yr for mean age and  $15.4 \pm 0.7$ nmol/L for mean testosterone. Although there appeared  
206 to be a slight declining trend in mean testosterone with mean age among studies (Meta-  
207 regression Slope= -0.07, 95% CI -0.21 – 0.07), this estimate was not significantly different  
208 from zero (P=0.27; Fig. 2a). However, the distribution of model residuals demonstrated  
209 significant heterogeneity (P<0.001) and funnel plot asymmetry (P=0.02). Additional  
210 diagnostics highlighted a relatively high mean testosterone estimate from Pencina et al.[37]  
211 (FHS) and a low mean testosterone estimate (relative to mean age) from Chan et al.[24]  
212 (BHS), as compared to the other studies (Supplementary Material). When restricted to  
213 systematically selected items (reporting on ARIC, BHS, CHAMP, CHS, EMAS, HIMS,  
214 MAILES, MrOS Sweden, SHIP studies), tests of residual heterogeneity were significant  
215 (P<0.001), funnel plot asymmetry (P=0.91) was non-significant, and the slope estimate  
216 (Meta-regression Slope= -0.03, 95% CI -0.11 – 0.06) was not significantly different from  
217 zero (P=0.50; Fig. 2b). These results demonstrate that varying distributions of participant age  
218 (likely reflecting differences in study-specific objectives and recruitment methods) did not  
219 explain the observed heterogeneity in published estimates of testosterone among the studies.

220  
221 Hazard ratios (HRs) for all-cause mortality were calculated from values in four of the  
222 selected articles (ARIC[38], BHS[24], CHS[39], EMAS[40]) and from one that was not  
223 selected, but had reported on the HIMS study during the literature search period.[16] All HRs

1  
2  
3 224 were adjusted for the age, smoking status, and body mass index (BMI) or waist  
4  
5 225 circumference of participants. A DR-MA estimated a summary HR of 0.96 (95% CI 0.89-  
6  
7 226 1.03) per 5nmol/L increase in testosterone (Fig. 3). The summary estimate was similar when  
8  
9 227 calculated using an alternative estimate from Yeap et al[16] (HR=0.97, 95% CI 0.92-1.03).  
10  
11 228 For both analyses, tests for residual heterogeneity ( $I^2=23.6\%$ ,  $P=0.26$ ;  $I^2=0.0\%$ ,  $P=0.76$ ) and  
12  
13 229 funnel plot asymmetry ( $P=0.09$ ;  $P=0.39$ ) were non-significant. A comparable HR was  
14  
15 230 calculated from a CHAMP study article[30] for inclusion in the forest plot but not in the DR-  
16  
17 231 MA, because a corresponding estimate of variance per 5nmol/L increase in testosterone could  
18  
19 232 not be calculated. Additional funnel plots, which included HR estimates from this CHAMP  
20  
21 233 article[30] (per 1 standard deviation decrease in testosterone, as reported in that article), also  
22  
23 234 demonstrated no significant asymmetry (Fig. S2c,d, Supplementary Material). These results  
24  
25 235 demonstrate no overall effect of baseline testosterone concentration on the relative hazard of  
26  
27 236 death from any cause after adjusting for factors including age, smoking status, and BMI or  
28  
29 237 waist circumference.  
30  
31  
32  
33  
34  
35 238  
36  
37 239 HRs for death caused by CVD demonstrated similar findings. A DR-MA using estimates  
38  
39 240 from the same five articles estimated a summary HR of 0.95 (95% CI 0.83-1.08) per 5nmol/L  
40  
41 241 increase in testosterone, with no significant residual heterogeneity ( $I^2=28.3\%$ ,  $P=0.23$ ) or  
42  
43 242 funnel plot asymmetry ( $P=0.20$ ; Fig. 4). Again, all HRs were adjusted for the age, smoking  
44  
45 243 status, and BMI or waist circumference. The DR-MA repeated using an alternative estimate  
46  
47 244 from Chasland et al.[25] for the BHS gave similar results (summary HR=0.93, 95% CI 0.83-  
48  
49 245 1.03; heterogeneity  $I^2=17.5\%$ ,  $P=0.30$ ; funnel plot asymmetry  $P=0.17$ ). These results  
50  
51 246 demonstrate no overall effect of baseline testosterone concentration on the relative hazard of  
52  
53 247 death from CVD after adjusting for factors including age, smoking status, and BMI or waist  
54  
55 248 circumference.  
56  
57  
58  
59  
60

1  
2  
3 2494 250 **4. DISCUSSION**

7 251 The systematic review identified nine studies, and when combined with an additional two  
8 252 identified by colleagues, comprises 11 in total, with data for over 20,000 men from Australia,  
9  
10 253 Europe, USA and the United Kingdom. Meta-regressions revealed significant heterogeneity  
11  
12 254 in testosterone measurements at baseline, which was not explained by the mean age of  
13  
14 255 participants among studies. However, DR-MA summary estimates demonstrated no  
15  
16 256 significant effects of baseline testosterone on the relative hazard of death from any cause or  
17  
18 257 from CVD, with negligible heterogeneity present. The DR-MAs, which suitably accounted  
19  
20 258 for correlations between estimates for different exposure categories within studies, were of  
21  
22 259 published estimates that had been adjusted for age, smoking status, and BMI or waist  
23  
24 260 circumference. Furthermore, only published estimates from prospective cohort studies of  
25  
26 261 community-dwelling men that had measured testosterone accurately using mass spectrometry  
27  
28 262 and had observed at least five years of follow-up data were used. Despite some of these  
29  
30 263 studies having reported an association between testosterone and mortality,[16, 30] the  
31  
32 264 collective body of evidence demonstrated no overall associations of endogenous testosterone  
33  
34 265 concentration with mortality or CVD mortality.  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 267 Previous meta-analyses investigating associations of endogenous testosterone with the health  
45  
46 268 outcomes of interest looked at CVD outcomes[41-43], all-cause mortality[41], and prostate  
47  
48 269 cancer[44]. Boyle et al.[44] and Holmegard et al.[42] both reported negligible heterogeneity  
49  
50 270 in their estimates. Boyle et al. found no significant association of a 5nmol/L increase in  
51  
52 271 testosterone with prostate cancer and Holmegard et al. estimated a 43% increase in risk of  
53  
54 272 ischemic stroke for men with testosterone levels below the 10<sup>th</sup> percentile, as compared to  
55  
56 273 men in the 11<sup>th</sup>-90<sup>th</sup> percentile range, from a meta-analysis of four articles.[42, 44] Ruige et  
57  
58  
59  
60

1  
2  
3 274 al. estimated an 11% decrease in risk of a CVD event from a standard deviation increase in  
4  
5 275 testosterone, and reported that significant heterogeneity was explained by larger effect sizes  
6  
7  
8 276 estimated for studies that recruited older men and for more recent articles.[43] Araujo et al  
9  
10 277 estimated a 35% increase in risk of all-cause mortality and a non-significant effect on CVD  
11  
12 278 mortality from a 2.18 standard deviation decrease in testosterone, although reported  
13  
14 279 significant heterogeneity, and suggested that effects were driven by differences between the  
15  
16 280 cohorts, such as underlying health status.[41] Two of these meta-analyses did not restrict  
17  
18 281 selections to prospective cohort studies[41, 44] and none restricted selections based on  
19  
20 282 testosterone assay method, although Ruige et al.[43] did find that assay method did not  
21  
22 283 explain heterogeneity in that study.  
23  
24  
25  
26  
27 284

28 285 The presented meta-analyses are the first to restrict selections to items of prospective cohort  
29  
30 286 studies of community-dwelling men with testosterone measured using mass spectrometry,  
31  
32 287 which is widely regarded as the reference method,[7] and with at least five years of follow-up  
33  
34 288 data. Accordingly, the presented summary estimates could arguably be viewed as the most  
35  
36 289 reliable to date. However, summary estimates represented associations that were assumed to  
37  
38 290 be linear at the scale of log hazards, which was a key limitation of the analyses and likely to  
39  
40 291 result in an oversimplification of true effects. For instance, although the 95% CI for the Pye  
41  
42 292 et al[40] study (calculated from HR estimates for quintile categories of testosterone)  
43  
44 293 overlapped one, an alternative set of estimates in that article (which could not be included in  
45  
46 294 the DR-MAs) reported a two-fold increase in the risk of all-cause mortality for men with very  
47  
48 295 low testosterone (<8nmol/L), as compared to “eugonadal” men (>11nmol/L). Pye et al[40]  
49  
50 296 postulated that their reported differences in estimates might be reflective of a nonlinear  
51  
52 297 association that emerges only when endogenous testosterone declines into the lower part of  
53  
54 298 the range (<8nmol/L). Furthermore, Yeap et al.[16] estimated an “U”-shaped association  
55  
56  
57  
58  
59  
60



1  
2  
3 299 between endogenous testosterone and all-cause mortality, as consistent with a lower relative  
4  
5 300 risk of health impacts for adult males with mid-range levels of testosterone. However, Shores  
6  
7 301 et al.[39] also used non-linear modelling but did not find any significant associations of  
8  
9 302 testosterone with all-cause or CVD mortality. Clearly, the investigation of non-linear  
10  
11 303 associations is required to more comprehensively investigate the associations of testosterone  
12  
13 304 concentrations with health outcomes in men.  
14  
15  
16

17 305  
18  
19  
20 306 Individual participant data (IPD) meta-analyses that incorporate flexible non-linear modelling  
21  
22 307 techniques will provide improved scope to clarify the nature of such associations. The ability  
23  
24 308 to apply a consistent statistical model to all studies, incorporate a more extended set of  
25  
26 309 covariates than may have been included at the individual study level, and to estimate effects  
27  
28 310 with increased statistical power, should result in more reliable summary estimates with  
29  
30 311 reduced bias. Furthermore, other hitherto unpublished variables may be available for sharing  
31  
32 312 by the collaborating studies to use in IPD meta-analyses, which could be useful for  
33  
34 313 constructing analysis covariates or outcome variables. For instance, articles from the ARIC  
35  
36 314 study that were identified from the systematic review reported on incident CVD event and  
37  
38 315 death outcomes, but documentation on the ARIC study website shows that data on other  
39  
40 316 prospective health outcomes, including cause-specific deaths and dementia diagnoses, are  
41  
42 317 also available upon request.[45] Although there have been recent advances with non-linear  
43  
44 318 modelling methods for the meta-analyses of published estimates,[18, 46] sufficient  
45  
46 319 information in the published articles, as is required for implementing these methods, was not  
47  
48 320 available. In future work, estimates from analyses of the IPD-level data will be used to  
49  
50 321 estimate and plot non-linear summary effects, and so will provide further improvements to  
51  
52 322 estimates of associations between androgen levels and health outcomes in men.  
53  
54  
55  
56  
57  
58  
59  
60 323



1  
2  
3 324 **ACKNOWLEDGEMENTS**  
4

5 325 We thank Terena Solomons for valuable advice and guidance with conducting the literature  
6  
7  
8 326 search and screening steps of the systematic review.  
9

10  
11 327

12  
13  
14 328 **AUTHORS' CONTRIBUTIONS**  
15

16 329 BBY, KM, RJM, JH contributed to the design of the systematic review. RJM conducted the  
17  
18 330 literature search and RJM, JH independently screened the returned items. RJM, KM  
19  
20 331 conducted the statistical analyses. All authors were involved in manuscript preparation and  
21  
22  
23 332 subsequent revisions, and approved this submission.  
24

25 333

26  
27 334 **COMPETING INTERESTS**  
28

29  
30 335 None declared.  
31

32 336  
33

34 337 **FUNDING**  
35

36 338 This work was supported by: (i) Western Australian Health Translation Network Medical  
37  
38  
39 339 Research Future Fund Rapid Applied Translation Grant (2018), Grant number N/A; (ii)  
40  
41 340 Lawley Pharmaceuticals, Western Australia, Grant number N/A.  
42

43  
44 341

45  
46 342 **DATA SHARING STATEMENT**  
47

48 343 No additional data available.  
49

50 344  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 345 **PATIENT CONSENT**  
4

5 346 This manuscript does not contain patient personal data.  
6  
7

8 347  
9

10 348 **ETHICS APPROVAL**  
11

12 349 The AIMS study has been assessed as exempt from ethics review by the Human Ethics office  
13  
14 350 at the University of Western Australia (file reference number RA/4/20/5014).  
15  
16

17 351  
18

19 352 **PROVENANCE AND PEER REVIEW**  
20

21 353 Not commissioned; externally peer reviewed.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

354 **REFERENCES CITED**

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

1. Ahern T, Swiecicka A, Eendebak RJA, et al. Natural history, risk factors and clinical features of primary hypogonadism in ageing men: Longitudinal Data from the European Male Ageing Study. *Clin Endocrinol (Oxf)* 2016;85:891-901. doi: 10.1111/cen.13152
2. Feldman HA, Longcope C, Derby CA, et al. Age Trends in the Level of Serum Testosterone and Other Hormones in Middle-Aged Men: Longitudinal Results from the Massachusetts Male Aging Study. *J Clin Endocrinol Metab* 2002;87(2):589-98.
3. Handelsman DJ, Yeap BB, Flicker L, et al. Age-specific population centiles for androgen status in men. *Eur J Endocrinol* 2015;173:809-17. doi: 10.1530/EJE-15-0380
4. Hsu B, Cumming RG, Hirani V, et al. Temporal Trend in Androgen Status and Androgen-Sensitive Outcomes in Older Men. *J Clin Endocrinol Metab* 2016;101(4):1836-46. doi: 10.1210/jc.2015-3810
5. Yeap BB, Manning L, Chubb SAP, et al. Progressive impairment of testicular endocrine function in ageing men: Testosterone and dihydrotestosterone decrease, and luteinizing hormone increases, in men transitioning from the 8th to 9th decades of life. *Clin Endocrinol (Oxf)* 2018;88:88-95. doi: 10.1111/cen.13484
6. Handelsman DJ. Global trends in testosterone prescribing, 2000–2011: expanding the spectrum of prescription drug misuse. *Med J Aust* 2013;199(8):548-51. doi: 10.5694/mja13.10111
7. Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the Journal of Clinical Endocrinology and Metabolism. *J Clin Endocrinol Metab* 2013;98(10):3971-73.
8. Munn Z, Stern C, Aromataris E, et al. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Med Res Methodol* 2018; 18.
9. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: A proposal for reporting. *JAMA* 2000;283(15):2008-12.
10. Yeap BB, Marriott RJ, Adams RJ, et al. Androgens In Men Study (AIMS): protocol for meta-analyses of individual participant data investigating associations of androgens with health outcomes in men. *BMJ Open* 2020; 10. <https://bmjopen.bmj.com/content/10/5/e034777> (accessed 19 Nov 2020).
11. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan - a web and mobile app for systematic reviews. *Syst Rev* 2016; 5.
12. Wells GA, Shea B, O'Connell DO, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses 2020. [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp) (accessed 10 Nov 2020).
13. Wang X, Ouyang Y, Liu J, et al. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* 2014; 349. <https://doi.org/10.1136/bmj.g4490> (accessed 19 Nov 2020).
14. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135(11):1301-09.
15. McGrath S, Zhao X, Steele R, et al. Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Stat Methods Med Res* 2020;29(9):2520-37. doi: 10.1177/0962280219889080
16. Yeap BB, Alfonso H, Chubb SAP, et al. In older men an optimal plasma testosterone is associated with reduced all-cause mortality and higher dihydrotestosterone with reduced ischemic heart disease mortality, while estradiol levels do not predict mortality. *J Clin Endocrinol Metab* 2014;99(1):E9-E18.
17. Sterne JAC, Egger M. Regression methods to detect publication and other bias in meta-analysis. In: Rothstein HR, Sutton AJ, Borenstein M, eds. *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. Chichester, England: Wiley 2005:99-110.

- 1  
2  
3 404 18. Crippa A, Orsini N. Multivariate dose-response meta-analysis: The dosresmeta R package. *J Stat*  
4 405 *Softw* 2016; 72.
- 5 406 19. estmeansd: Estimating the Sample Mean and Standard Deviation from Commonly Reported  
6 407 Quantiles in Meta-Analysis. R package version 0.2.1, 2020
- 7 408 20. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for  
8 409 Statistical Computing, 2020. <https://www.R-project.org/>
- 9 410 21. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*  
10 411 2010;36(3):1-48.
- 11 412 22. Bhasin S, Pencina MJ, Kaur Jasuja G, et al. Reference ranges for testosterone in men generated  
12 413 using liquid chromatography tandem mass spectrometry in a community-base sample of healthy  
13 414 nonobese young men in the Framingham Heart Study and applied to three geographically distinct  
14 415 cohorts. *J Clin Endocrinol Metab* 2011;96(8):2430-39.
- 15 416 23. Chan YX, Knuiaman MW, Divitini ML, et al. Lower Circulating Androgens Are Associated with  
16 417 Overall Cancer Risk and Prostate Cancer Risk in Men Aged 25-84 Years from the Busselton Health  
17 418 Study. *Horm Cancer* 2018;9(6):391-98. doi: <https://dx.doi.org/10.1007/s12672-018-0346-5>
- 18 419 24. Chan YX, Knuiaman MW, Hung J, et al. Neutral associations of testosterone, dihydrotestosterone  
19 420 and estradiol with fatal and non-fatal cardiovascular events, and mortality in men aged 17-97  
20 421 years. *Clin Endocrinol (Oxf)* 2016;85(4):575-82. doi: <https://dx.doi.org/10.1111/cen.13089>
- 21 422 25. Chasland LC, Knuiaman MW, Divitini ML, et al. Greater physical activity and higher androgen  
22 423 concentrations are independently associated with lower cardiometabolic risk in men. *Clin*  
23 424 *Endocrinol (Oxf)* 2017;87(5):466-74. doi: <https://dx.doi.org/10.1111/cen.13407>
- 24 425 26. Kische H, Gross S, Wallaschofski H, et al. Associations of androgens with depressive symptoms  
25 426 and cognitive status in the general population. *PLoS ONE* 2017;12(5):e0177272. doi:  
26 427 <https://dx.doi.org/10.1371/journal.pone.0177272>
- 27 428 27. Li JJ, Wittert GA, Vincent A, et al. Muscle grip strength predicts incident type 2 diabetes:  
28 429 population-based cohort study. *Metab Clin Exp* 2016;65:883-92.
- 29 430 28. Chan YX, Alfonso H, Chubb SA, et al. Higher Dihydrotestosterone Is Associated with the Incidence  
30 431 of Lung Cancer in Older Men. *Horm Cancer* 2017;8(2):119-26. doi:  
31 432 <https://dx.doi.org/10.1007/s12672-017-0287-4>
- 32 433 29. Hsu B, Cumming RG, Blyth FM, et al. Evaluating Calculated Free Testosterone as a Predictor of  
33 434 Morbidity and Mortality Independent of Testosterone for Cross-sectional and 5-Year Longitudinal  
34 435 Health Outcomes in Older Men: The Concord Health and Ageing in Men Project. *J Gerontol A Biol*  
35 436 *Sci Med Sci* 2018;73(6):729-36. doi: <https://dx.doi.org/10.1093/gerona/glx170>
- 36 437 30. Hsu B, Cumming RG, Naganathan V, et al. Temporal Changes in Androgens and Estrogens Are  
37 438 Associated With All-Cause and Cause-Specific Mortality in Older Men. *J Clin Endocrinol Metab*  
38 439 2016;101(5):2201-10. doi: <https://dx.doi.org/10.1210/jc.2016-1025>
- 39 440 31. Ohlsson C, Barrett-Connor E, Bhasin S, et al. High serum testosterone is associated with reduced  
40 441 risk of cardiovascular events in elderly men. The MrOS (Osteoporotic Fractures in Men) study in  
41 442 Sweden. *J Am Coll Cardiol* 2011;58(16):1674-81. doi:  
42 443 <https://dx.doi.org/10.1016/j.jacc.2011.07.019>
- 43 444 32. Ohlsson C, Labrie F, Barrett-Connor E, et al. Low serum levels of dehydroepiandrosterone sulfate  
44 445 predict all-cause and cardiovascular mortality in elderly Swedish men. *J Clin Endocrinol Metab*  
45 446 2010;95(9):4406-14. doi: <https://dx.doi.org/10.1210/jc.2010-0760>
- 46 447 33. Rosenberg MA, Shores MM, Matsumoto AM, et al. Serum androgens and risk of atrial fibrillation  
47 448 in older men: The Cardiovascular Health Study. *Clin Cardiol* 2018;41(6):830-36. doi:  
48 449 <https://dx.doi.org/10.1002/clc.22965>
- 49 450 34. Shores MM, Arnold AM, Biggs ML, et al. Testosterone and dihydrotestosterone and incident  
50 451 ischaemic stroke in men in the Cardiovascular Health Study. *Clin Endocrinol (Oxf)* 2014;81(5):746-  
51 452 53. doi: <https://dx.doi.org/10.1111/cen.12452>
- 52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 453 35. Sueoka KT, Ewing MS, Ensrud KE, et al. Higher endogenous testosterone levels associated with  
4 454 increased risk of coronary heart disease in elderly men: a prospective study. *Endocr Rev*  
5 455 2010;31(3):S858.  
6 456 36. Tivesten Å, Vandenput L, Carlzon D, et al. Dehydroepiandrosterone and its sulfate predict the 5-  
7 457 year risk of coronary heart disease events in elderly men. *J Am Coll Cardiol* 2014;64(17):1801-10.  
8 458 doi: <https://dx.doi.org/10.1016/j.jacc.2014.05.076>  
9 459 37. Pencina KM, Travison TG, Bhasin S, et al. Endogenous circulating testosterone and sex hormone-  
10 460 binding globulin levels and measures of myocardial structure and function: the Framingham  
11 461 Heart Study. *Andrology* 2019;7:307-14.  
12 462 38. Srinath R, Hill Golden S, Carson KA. Endogenous Testosterone and its Relationship to Preclinical  
13 463 and Clinical Measures of Cardiovascular Disease in the Atherosclerosis Risk in Communities Study.  
14 464 *J Clin Endocrinol Metab* 2015;100(4):1602-02. doi: 10.1210/jc.2014-3934  
15 465 39. Shores MM, Biggs ML, Arnold AM, et al. Testosterone, dihydrotestosterone, and incident  
16 466 cardiovascular disease and mortality in the cardiovascular health study. *J Clin Endocrinol Metab*  
17 467 2014;99(6):2061-8. doi: <https://dx.doi.org/10.1210/jc.2013-3576>  
18 468 40. Pye SR, Huhtaniemi IT, Finn JD, et al. Late-onset hypogonadism and mortality in aging men. *J Clin*  
19 469 *Endocrinol Metab* 2014;99(4):1357-66. doi: <https://dx.doi.org/10.1210/jc.2013-2052>  
20 470 41. Araujo AB, Dixon JM, Suarez EA, et al. Endogenous testosterone and mortality in men: a  
21 471 systematic review and meta-analysis. *J Clin Endocrinol Metab* 2011;96(10):3007-19.  
22 472 42. Holmegard HN, Nordestgaard BG, Jensen GB, et al. Sex hormones and ischemic stroke: a  
23 473 prospective cohort study and meta-analyses. *J Clin Endocrinol Metab* 2016;101(1):69-78.  
24 474 43. Ruige JB, Mahmoud AM, De Bacquer D, et al. Endogenous testosterone and cardiovascular  
25 475 disease in healthy men: a meta-analysis. *Heart* 2011;97:870-75.  
26 476 44. Boyle P, Koechlin A, Bota M, et al. Endogenous and exogenous testosterone and the risk of  
27 477 prostate cancer and increased prostate-specific antigen (PSA) level: a meta-analysis. *BJU Int*  
28 478 2016;118:731-41.  
29 479 45. Surveillance Dictionaries | Atherosclerosis Risk In Communities 2020.  
30 480 <https://sites.csc.unc.edu/aric/surveillance-dictionaries> (accessed 6 July 2020).  
31 481 46. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized  
32 482 dose-response data. *Stata J* 2006;6(1):40-57.  
33 483 47. Lee DM, Pye SR, Tajar A, et al. Cohort profile: the European Male Ageing Study. *Int J Epidemiol*  
34 484 2013;42(2):391-401. doi: <https://dx.doi.org/10.1093/ije/dyr234>  
35  
36  
37  
38  
39  
40 485  
41 486  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 487 **FIGURE LEGENDS**  
4

5 488  
6  
7 489 Figure 1. Studies returned from systematic review of the published and grey literature. Step 1  
8  
9 490 involved screening of titles and abstracts only and Step 2 the screening of full text items not  
10  
11 491 excluded at Step 1 (see Tables 1, 2). “Items” are individual articles or reports, with multiple  
12  
13 492 items returned for some studies (the purpose was to identify studies with suitable IPD-level  
14  
15 493 data). \* = Mednar items with insufficient information available to review; \*\* = Additional  
16  
17 494 studies identified through known contacts; \*\*\* = Screening criteria for five items selected as  
18  
19 495 “Maybe” in Step 2 were further investigated using information external to systematic  
20  
21 496 searches and screenings, resulting in the identification of one additional study with suitable  
22  
23 497 IPD-level data.  
24  
25  
26  
27  
28

29 498  
30 499 Figure 2. Meta-regression of mean testosterone on mean age for (a) all 11 cohort studies and  
31  
32 500 (b) 9 studies with articles that were selected by systematic literature searches and screening.  
33  
34 501 The size of plotted points refers are proportional to the inverse of the corresponding standard  
35  
36 502 errors (indicative of relative weightings), with lines demonstrating the fitted model and 95%  
37  
38 503 CIs. Plotted estimates are numbered as from the following articles (cohort studies):  
39  
40 504 1= Srinath et al.[38] (ARIC); 2= Chan et al.[24] (BHS); 3= Hsu et al.[30] (CHAMP); 4=  
41  
42 505 Shores et al.[34] (CHS); 5= Lee et al.[47] (EMAS); 6= Chan et al.[28] (HIMS); 7= Ohlsson  
43  
44 506 et al.[32] (MrOS Sweden); 8= Kische et al.[26] (SHIP); 9= Sueoka et al.[35] (MrOS USA);  
45  
46 507 10= Pencina et al.[37] (FHS); 11= Li et al.[27] (MAILES). \* = includes articles from two  
47  
48 508 additional studies (FHS, MAILES) that were not identified from systematic searches but by  
49  
50 509 colleagues.  
51  
52  
53  
54  
55

56  
57 511 Figure 3. Forest plot of a meta-analysis of published estimates: association of testosterone  
58  
59 512 with all-cause mortality. Plotted values are the estimated hazard ratios (HR) for death from  
60

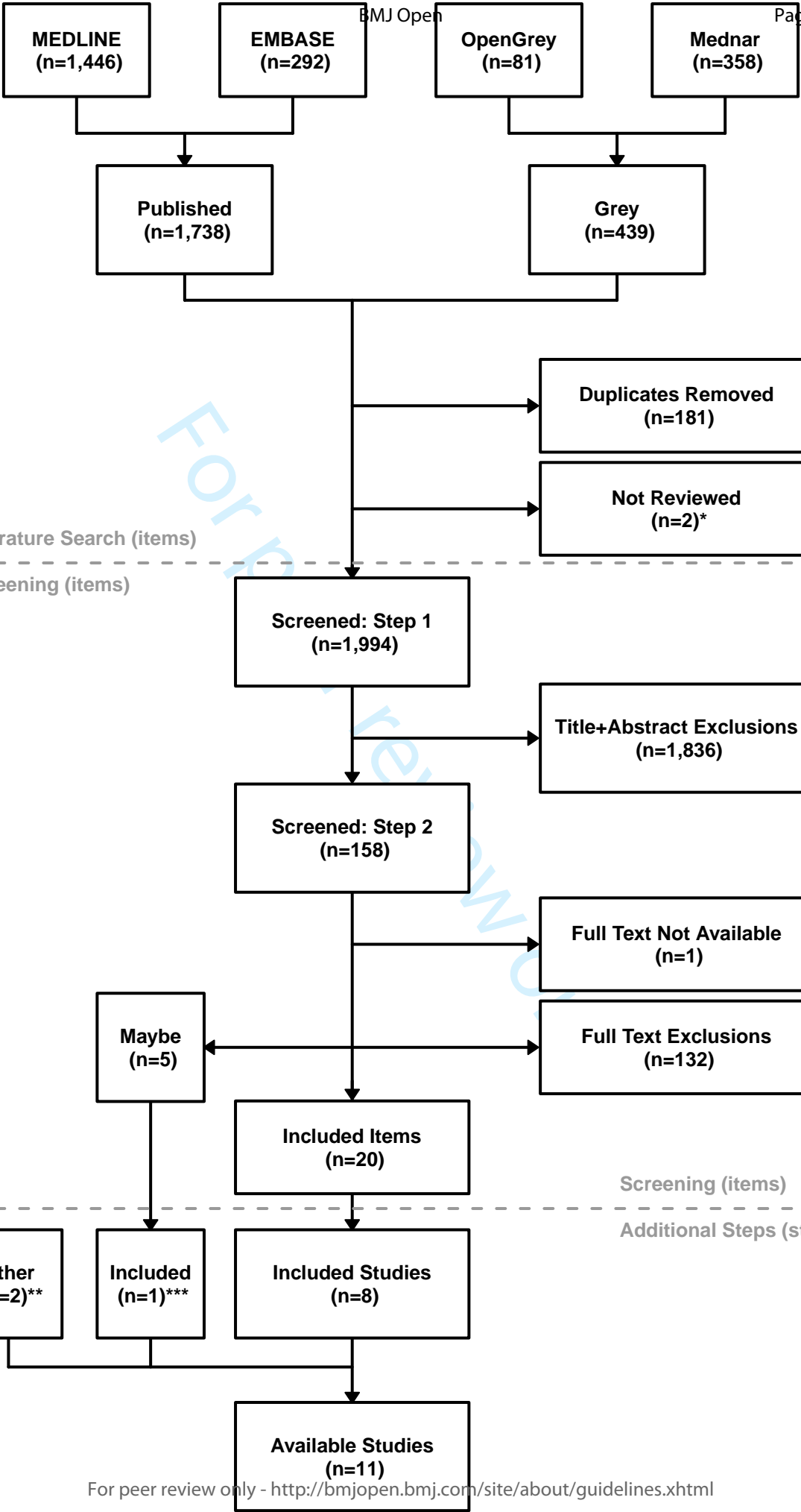
1  
2  
3 513 any cause, as attributed to an increase in endogenous testosterone concentration by 5 nmol/L.  
4  
5 514 The vertical reference line is HR=1. Study-specific estimates are presented for six of the  
6  
7 515 selected studies: BHS (Chan, 2016)[24]; EMAS (Pye, 2014)[40]; ARIC (Srinath, 2015)[38];  
8  
9 516 CHS (Shores, 2014b)[39]; HIMS (Yeap, 2014b)[16]; CHAMP (Hsu, 2016).[30] Summary  
10  
11 517 estimates are colour-coded as calculated using either the estimates from Yeap et al.[16]  
12  
13 518 calculated from the model including SHBG (black) or from the model including LH (grey). \*  
14  
15 519 This estimate from Hsu et al.[30] could not be used to calculate the summary estimate  
16  
17 520 because a variance estimate was not calculable for a 5nmol/L change in testosterone using the  
18  
19 521 published information.  
20  
21  
22  
23  
24  
25

26 523 Figure 4. Forest plot of a meta-analysis of published estimates: association of testosterone  
27  
28 524 with mortality caused by cardiovascular disease. Plotted values are the estimated hazard  
29  
30 525 ratios (HR) for death from any cause, as attributed to an increase in endogenous testosterone  
31  
32 526 concentration by 5 nmol/L. The vertical reference line is HR=1. Study-specific estimates are  
33  
34 527 presented for six of the selected studies: BHS (Chan, 2016; Chasland, 2017)[24, 25]; EMAS  
35  
36 528 (Pye, 2014)[40]; ARIC (Srinath, 2015)[38]; CHS (Shores, 2014b)[39]; HIMS (Yeap,  
37  
38 529 2014b)[16]; CHAMP (Hsu, 2016).[30] Summary estimates are colour-coded as calculated  
39  
40 530 using either the estimates from Chan et al.[24] (black) or Chasland et al.[25] (grey) for the  
41  
42 531 BHS. \* This estimate from Hsu et al.[30] could not be used to calculate the summary estimate  
43  
44 532 because a variance estimate was not calculable for a 5nmol/L change in testosterone using the  
45  
46 533 published information.  
47  
48  
49  
50

51 534

52 535





1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Literature Search (items)

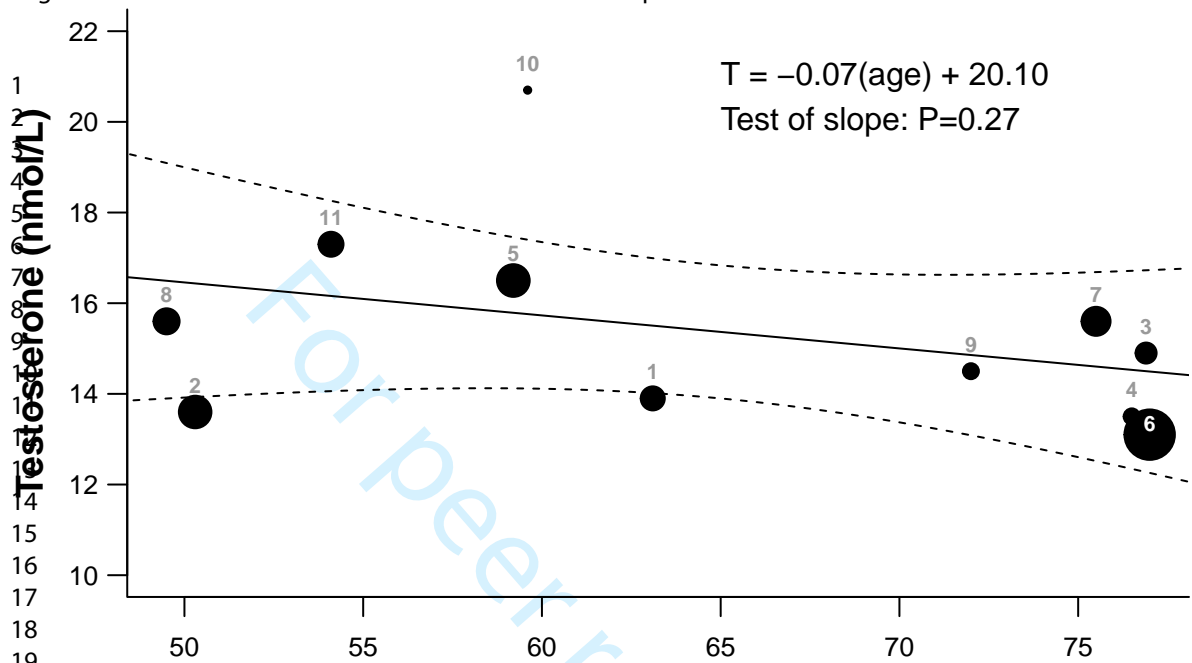
Screening (items)

Screening (items)

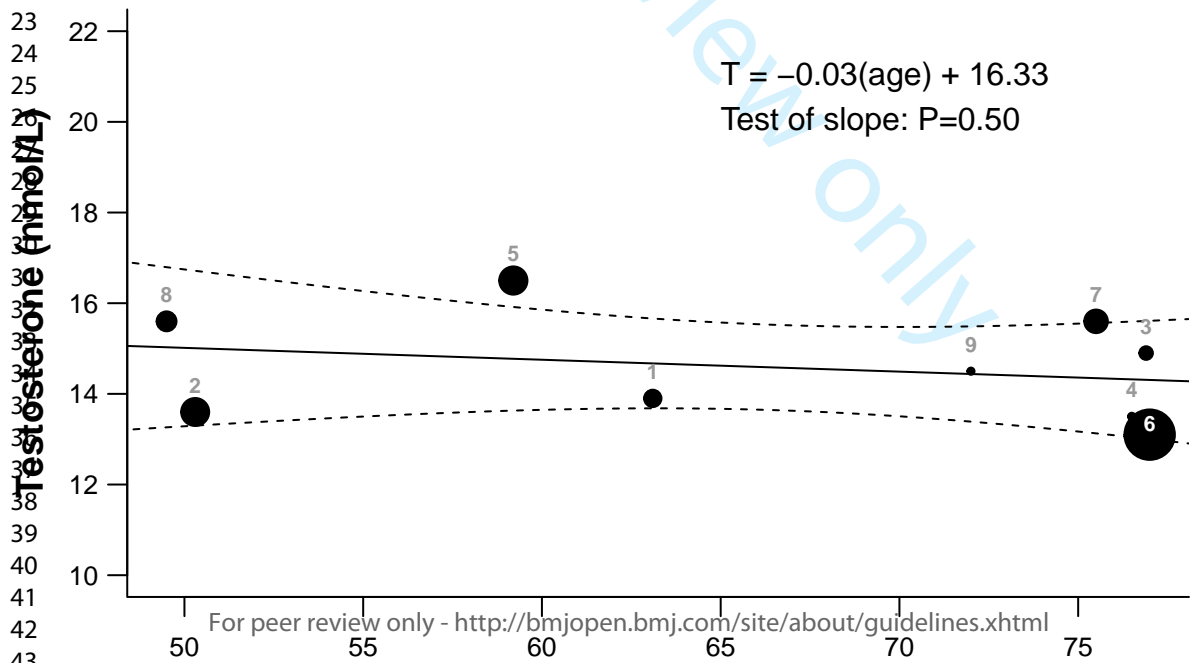
Additional Steps (studies)

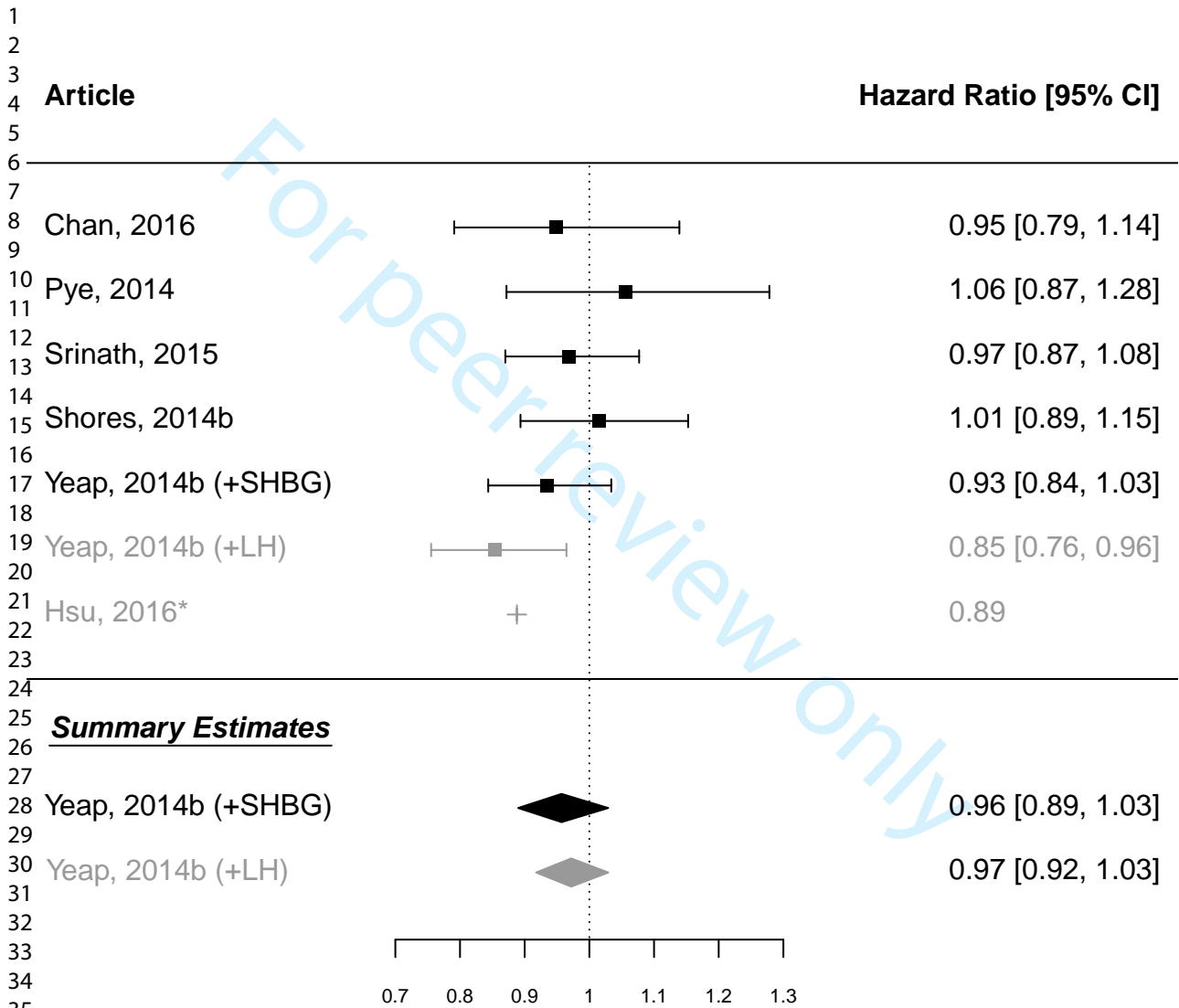


### a) All cohorts\*



### b) Selected items

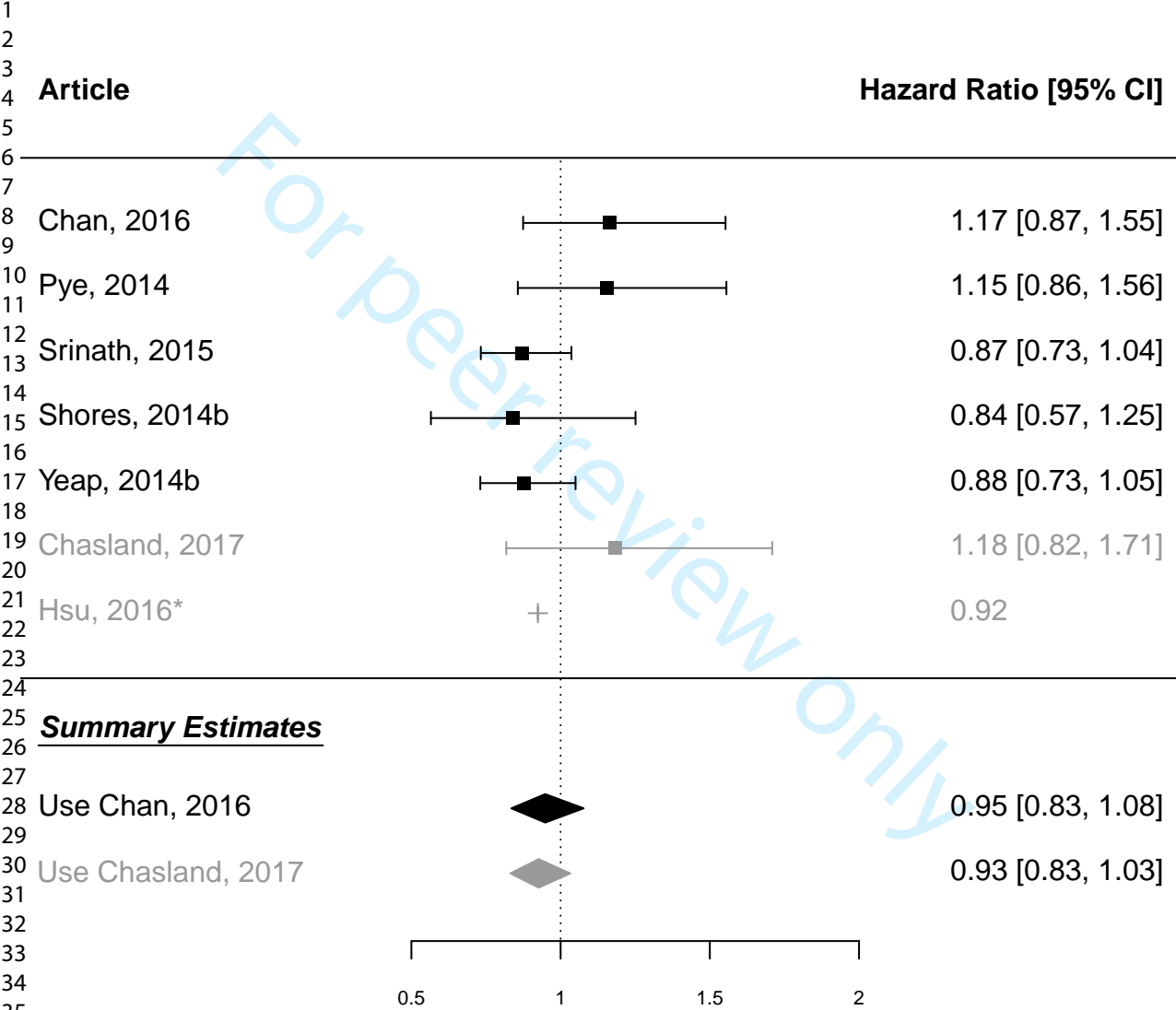




For peer review only - <http://bmjopen.bmj.com/site/about/guidelines.xhtml>

**HR: Increase in 5 nmol/L Testosterone**

36  
37  
38



*Supplementary Material: Systematic Review for the Androgens In Men Study.*

1		
2		
3		
4		
5	1	Table of Contents
6	2	Supplementary Material: Additional details on systematic searches and screening.....2
7	3	Supplementary Material: Tables .....7
8	4	Table S1. Full electronic search strategy used for MEDLINE database. ....7
9	5	Table S2: Selection criteria for screening items returned from the literature search.....8
10	6	Table S3: Adaptation of screening rules for different types of published and unpublished
11	7	items.....9
12	8	Table S4. Words mentioned in the titles or abstracts of reviewed items. ....10
13	9	Table S5. Attributes of selected items .....11
14	10	Table S6. Exposure levels, outcome assessment, covariates. ....15
15	11	Table S7. Newcastle-Ottawa Quality Assessment Scale for Cohort Studies: selected
16	12	articles. ....18
17	13	Table S8. PRISMA Checklist for Systematic Review: the Androgens In Men Study. ....19
18	14	Table S9. Extracted hazard ratio data for dose-response meta-analyses (DR-MAs).....22
19	15	Supplementary Material: Figures.....24
20	16	Figure S1. Meta-regression diagnostics.....24
21	17	Figure S2. Funnel plots for dose-response meta-analyses. ....26
22	18	References cited. ....27
23		
24		
25		
26		
27		
28		
29		
30		
31		
32	19	
33	20	
34		
35	21	
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		

*Supplementary Material: Systematic Review for the Androgens In Men Study.***22 Supplementary Material: Additional details on systematic searches and screening**

23  
24 We used online search tools to identify available published (MEDLINE, EMBASE) and grey  
25 literature (OpenGrey, Mednar) *items* (journal articles, reports, theses, webpage articles)  
26 reporting on suitable prospective cohort *studies* (the underlying unique sources of data). We  
27 used OpenGrey and Mednar because both were free search tools that we considered likely to  
28 identify additional grey literature items and studies in an expanded search beyond the  
29 mainstream publications. Mednar is a medically-focussed search engine of public and deep  
30 web resources, excluding subscription services.[1] OpenGrey is a searchable database  
31 containing citations for items including technical or research reports, theses, conference  
32 papers, and other types of grey literature.[2] Literature searches were conducted on 18-22  
33 July 2019, with no date restrictions set.

34  
35 Where possible (as functionality varied among the different tools), we placed the following  
36 restrictions on the search: items reporting on the results of a research study, longitudinal or  
37 prospective cohort studies, not of hormone therapy or deprivation treatments. Due to study  
38 timeframe and language translation limitations, we opted to search for only those items that  
39 were reported in the English language. The terms and full criteria used for the MEDLINE  
40 search are provided in Table S1, and the PRISMA checklist as Table S8.

41  
42 Selection criteria were set as applicable to the planned sets of IPD meta-analyses  
43 (Table S2).[3] Only items reporting on prospective population-based cohort studies, adults of  
44 combined sexes or of men alone, with individuals free of the disease at baseline, were sought.  
45 Items reporting a different design for the analysis of longitudinal data, such as nested case-  
46 control or case-cohort design, were also considered acceptable. A minimum of five years  
47 follow-up was selected, to ensure a sufficient number of incident events for statistical

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

1  
2  
3 48 modelling. We excluded items that did not measure testosterone using mass spectrometry,  
4  
5 49 which is regarded to be the ‘gold standard’ method,[4] although testosterone was not required  
6  
7 50 to be mentioned in the title or abstract, nor modelled as the primary exposure variable.

8  
9  
10 51 Selected items were to be studies of humans, reported in English, and reporting on analyses  
11  
12 52 of at least one of the AIMS outcomes.

13  
14 53  
15  
16  
17 54 Two reviewers (RJM, JH) independently screened the de-duplicated items against these pre-  
18  
19 55 specified criteria. To optimise efficiency, the selection of items proceeded in two steps. Title  
20  
21 56 and abstract screenings (Step 1) were followed by full text screening of items selected in Step  
22  
23 57 1 (Step 2). If an item was selected for exclusion, then the main reason for that decision was  
24  
25 58 recorded. If there was uncertainty in the decision to exclude, in Step 1 the reviewer selected  
26  
27 59 “include” (in Step 1) or “maybe” (in Step 2). At the end of each step, the two reviewers  
28  
29 60 sought to achieve consensus, through discussion, for each item that did not achieve  
30  
31 61 agreement. Exclusion reasons were used to inform discussions for achieving consensus. Items  
32  
33 62 with a consensus decision of “maybe” were further investigated by Reviewer 1 (RJM) using  
34  
35 63 information external to the systematic searches and screenings (reading further details of  
36  
37 64 methods used in cited articles, and from correspondence with authors or other researchers  
38  
39 65 currently working on the research study).

40  
41 66  
42  
43  
44  
45  
46  
47 67 This screening procedure was adjusted to accommodate the different types of items reviewed  
48  
49 68 (published articles, theses, webpage articles, unpublished reports; Table S3). A pilot set of  
50  
51 69 title-only screenings for 30 randomly chosen articles suggested that sufficient information  
52  
53 70 was contained within the titles alone for the purpose of Step 1 screenings.<sup>a</sup> Therefore, in cases

---

54  
55  
56  
57  
58  
59  
60  
<sup>a</sup> 30 titles were initially screened at random. 18 were flagged as not suitable, leaving 12 as potentially suitable. Subsequent Step 1 screening of titles with abstracts selected 25 of these articles for exclusion, with 5 retained for Step 2 (full text screening). All 5 were flagged as being potentially suitable in the pilot set of title-only screenings.

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

1  
2  
3 71 when an abstract was not available, only the titles were screened. Website items identified by  
4  
5 72 the Mednar search tool were the type of item that most often did not have abstract or  
6  
7 73 summary text, and in these cases the webpage text was reviewed in place of an abstract  
8  
9  
10 74 (Table S3).

11  
12 75  
13  
14 76 Endnote X8[5] was used for collating and storing the citations returned from literature  
15  
16 77 searches, and for de-duplicating and storing the selected references. The full citations,  
17  
18 78 including abstracts, were exported from Endnote for uploading into Rayyan[6], which is a  
19  
20 79 free web tool that was used for screening, recording exclusion decisions, and downloading  
21  
22 80 selection results.  
23  
24  
25

26 81  
27  
28 82 The literature search identified 2,177 items (1,738 published and 439 from grey literature),  
29  
30 83 with 1,994 items remaining after duplicates had been removed, and after excluding two  
31  
32 84 Mednar items that had insufficient information available to review (Fig. 1). Table S4 shows  
33  
34 85 the frequencies of returned items by search terms present in the titles and abstracts. Most  
35  
36 86 (72.7%) had the word “cancer”, and 1,107 (55.5%) of these had the word “prostate cancer”,  
37  
38 87 in the title or abstract. This, combined with frequent mentions of “androgen deprivation”  
39  
40 88 (29.2%), “radiotherapy” (18.6%), and “brachytherapy” (8.3%), show that items reporting  
41  
42 89 aspects of testosterone deprivation or suppression for treating prostate cancer were a  
43  
44 90 predominant feature of the returned items. Different types of returned items included 1,764  
45  
46 91 published articles, 111 webpage articles, 81 theses, and 38 unpublished reports/other  
47  
48 92 documents, and the percentages without abstract or webpage text screened in Step 1 were  
49  
50 93 2.6%, 1.8%, 24.7%, 65.8%, respectively (i.e., 4.7% overall).  
51  
52  
53  
54  
55

56 94  
57  
58  
59  
60

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

1  
2  
3 95 One thousand nine hundred sixty-eight items were excluded, five items were classified as  
4  
5 96 “Maybe”, and one item could not be screened because the full text version was not available,  
6  
7  
8 97 leaving  $n = 20$  suitable items selected (Fig. 1). Most (92.1%) of the exclusions were made  
9  
10 98 from reviewing titles and abstracts at Step 1, with a further 6.6% excluded from screening of  
11  
12 99 the 157 full text items in Step 2. Inter-reader agreement was a Cohen’s Kappa  $\kappa = 0.69$   
13  
14  
15 100 (or 96.0 percent agreement) for Step 1 and  $\kappa = 0.82$  (or 98.1 percent agreement) for Step 2.  
16  
17 101 Percentages of items with search terms (AIMS outcomes) in the title or abstract increased  
18  
19 102 after Step 1 in most cases except for “cancer” and “prostate cancer” (Table S4). This reflects  
20  
21 103 many exclusions in Step 1 that were of items reporting research on testosterone deprivation or  
22  
23 104 suppression treatments for prostate cancer.  
24  
25  
26 105  
27  
28 106 The systematic approach to literature searching and screening is widely held to be beneficial  
29  
30 107 to identifying studies that otherwise may not have been considered for inclusion, and thus to  
31  
32 108 minimise the prospect for reviewer biases affecting study selections and summary results.[7]  
33  
34  
35 109 This process is not perfect though, and in our case it did not identify two prospective cohort  
36  
37 110 studies that were known to be suitable, prior to commencing this review (FHS, MAILES).[3]  
38  
39 111 In the case of MAILES, this was one of the more recently commenced of the selected studies,  
40  
41 112 with its cohort profile article published in 2014,[8] and accordingly has had a comparatively  
42  
43 113 short timeframe within which to analyse and publish suitable findings. In the case of FHS,  
44  
45 114 associations of endogenous testosterone with male health outcomes had previously been  
46  
47 115 investigated and published, but not using mass spectrometry for measuring testosterone.[9,  
48  
49 116 10] Those articles were identified in the literature search but had been excluded on account of  
50  
51 117 assay method. Only relatively recently have testosterone measures been re-assayed for FHS  
52  
53 118 participants using mass spectrometry methods.[11] One article by Pencina et al[12] was  
54  
55 119 possibly within scope but not identified because it had not been entered into the MEDLINE  
56  
57  
58  
59  
60



*Supplementary Material: Systematic Review for the Androgens In Men Study.*

1  
2  
3 120 database prior to the literature search (article entry date = 14 May 2020). Furthermore, an  
4  
5 121 article that presented suitable estimates from one of the selected studies by Yeap et al[13]  
6  
7 122 was not identified from the literature search because it did not have “prospective”, “follow-  
8  
9 123 up”, “cohort study” or “longitudinal study” terms in its title or abstract, nor any of the  
10  
11 124 corresponding MeSH terms listed (refer to Table S1 for search terms used).  
12  
13  
14  
15 125

16  
17 126 In expanding our literature search to unpublished grey literature, it successfully located one  
18  
19 127 suitable item, which was a link to a Web MD webpage article, with further details published  
20  
21 128 in a conference abstract by Sueoka et al[14] that would otherwise have not been returned  
22  
23 129 from searching only the MEDLINE and EMBASE databases.  
24  
25  
26 130

*Supplementary Material: Systematic Review for the Androgens In Men Study.***131 Supplementary Material: Tables****132 Table S1. Full electronic search strategy used for MEDLINE database.**

134 The following is the search that was conducted on 18 July 2019 using MEDLINE.

- 135
- 136 1. Testosterone/ or Androgens/
- 137 2. (testosterone or androgen\* or sex hormone\* or sex steroid\*).ti.
- 138 3. (testosterone or androgen\*).ab.
- 139 4. cardiovascular diseases/ or heart diseases/ or heart failure/ or vascular diseases/ or stroke/
- 140 or myocardial infarction/ or coronary disease/ or cerebrovascular disorders/
- 141 5. (cardiovascular or stroke or myocardial infarction or heart failure).ti.
- 142 6. neoplasms/ or colorectal neoplasms/ or lung neoplasms/ or prostatic neoplasms/
- 143 7. cancer.ti.
- 144 8. mortality/ or mortality.ti.
- 145 9. dementia/ or cognition/ or dementia.ti. or cognit\*.ti.
- 146 10. Aging/psychology or Neuropsychological Tests/
- 147 11. 1 or 2 or 3
- 148 12. 4 or 5 or 6 or 7 or 8 or 9 or 10
- 149 13. 11 and 12
- 150 14. longitudinal studies/ or prospective studies/ or follow-up studies/ or cohort studies/
- 151 15. (prospective or follow-up or cohort study or longitudinal study).ti,ab.
- 152 16. 14 or 15
- 153 17. 13 and 16
- 154 18. (exogenous or replacement or therapy or hormone treatment).ti.
- 155 19. Hormone Replacement Therapy/
- 156 20. 18 or 19
- 157 21. 17 not 20
- 158 22. limit 21 to humans
- 159 23. limit 22 to english language
- 160 24. limit 23 to (adaptive clinical trial or address or autobiography or bibliography or
- 161 biography or case reports or clinical trial, all or clinical trial, phase i or clinical trial, phase ii
- 162 or clinical trial, phase iii or clinical trial, phase iv or clinical trial, veterinary or clinical trials,
- 163 veterinary as topic or clinical trial protocol or clinical trial or comment or controlled clinical
- 164 trial or dictionary or editorial or lecture or legislation or meta analysis or practice guideline or
- 165 pragmatic clinical trial or published erratum or randomized controlled trial or retracted
- 166 publication or "retraction of publication" or "review" or "scientific integrity review" or
- 167 "systematic review")
- 168 25. Retrospective Studies/ or Case-Control Studies/ or (retrospective analysis or case-
- 169 control).ti.
- 170 26. 24 or 25
- 171 27. 23 not 26

**173 Notes:**

174

175 Terms with a trailing "/" are MeSH terms and those with a trailing "\*" are truncated search

176 strings. Beforehand, a search of PROSPERO was conducted for another suitable strategy but

177 none were found. However, the above strategy is based upon one that has been used for a

178 similar study.[15] This search strategy is also published in the protocol article for the

179 Androgens In Men Study.[3]

180 **Table S2: Selection criteria for screening items returned from the literature search.** If neither Include nor Exclude could be selected for Step 1, then  
 181 reviewer selected "Include".

	Exclude	Include	Rationale	Used in Step 1		Used in Step 2
				Title & Abstract		Full-text
				Title only (no abstract)	Title & Abstract	
<b>Article type:</b>	Reviews, comments/opinion pieces, systematic reviews, dictionary, fact sheet, website information about diseases, fact sheets, etc.	Research study article / report, or an article that specifically refers to the results of one (e.g., a webpage referring to unpublished data).	These searches were of both published and unpublished scientific literature for the purpose of identifying prospective cohort studies that are likely to have the relevant data for planned IPD meta-analyses	Yes	Yes	Yes
<b>Study type:</b>	Retrospective or cross-sectional designs, case studies, case-control, surveys, RCTs or other trials, experiments, evaluation of androgen / testosterone therapy / deprivation / HRT or the effectiveness of any other type of intervention / surgery / treatment, genetics, etc.	Prospective cohort study.	A prospective cohort study design is of incident health outcomes for investigating etiology or disease risk for a cohort free of disease at baseline, and ideally should be representative of the local population, but may or may not be some demographic subset: e.g., age range, sex, ethnicity type.  Further details in Table 2.	Yes	Yes	Yes
<b>Population</b> (at baseline/date of recruitment to study)	Studies of juveniles only Studies of females only Individuals with some specific health condition/characteristic or following surgery / other medical treatment for specific illness	Adults (18 yr or older) Not females only Community-dwelling men	The study is of community-dwelling men.	Yes	Yes	Yes
<b>Exposure</b> (at baseline)	Do not exclude studies that do not model testosterone as the exposure: although it should be shown that it was measured for participants. If not mentioned in Step 2 then Exclude.	Endogenous testosterone	This will be the focal exposure for all IPD meta-analyses. However, as we are focussing on the identification of only those studies who have suitable androgen measurements available in IPD data, then testosterone does not necessarily need to be modelled as the focal exposure in included items. It is likely that details on the methods will be available only from full-text review.	Only if available	Only if available	Yes
	Testosterone not measured using mass spectrometry	Testosterone assay of serum or plasma sample using mass spectrometry (lc-ms or gc-ms)		Only if available	Only if available	Yes
<b>Outcome</b> (at follow-up)	Incident outcome not one of those type of events specified for inclusion.	Diagnosis/event of: cardiovascular disease (any); cancer (any); dementia. Deaths (any cause); deaths due to any type of cardiovascular disease; deaths due to any type of cancer. Cognition change / outcome	These are the outcomes for the planned IPD meta-analyses so it is important to seek IPD datasets from those studies who have already modelled these outcomes.  We refer to these as the "AIMS outcomes".	Yes	Yes	Yes
	Less than 5 years of follow-up data	Five or more years of follow-up data, with outcomes identified using systematic follow-up or data linkage.	As consistent across all included studies for IPD meta-analyses and set <i>a priori</i> . Likely that this will be available only from full-text review so not included Step 1.	No	No	Yes
<b>Language</b>	Title and/or abstract of article not in English	Title and/or abstract of article in English	As limited by the timeframe of this study and the native language of reviewers (a practicality).	Yes	Yes	Yes
<b>Species</b>	Studies not of humans	Studies of humans	We are studying humans.	Yes	Yes	Yes

182

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

183 Table S3: Adaptation of screening rules for different types of published and unpublished  
 184 items.

185

Item Type	Step 1	Step 2
Published article	Screen title (and abstract <sup>a</sup> )	Screen full text article
Thesis	Screen title (and abstract <sup>a</sup> )	Screen full thesis
Unpublished report / other document	Screen title (and abstract <sup>a,b</sup> )	Screen full document
Webpage	Screen title and webpage <sup>c</sup>	Screen full text article/document as identified from the webpage, or from a google search of information provided about the article, from the webpage.

186

187 <sup>a</sup> = when an abstract was available, otherwise title-only decisions were made (see Table 1).

188 <sup>b</sup> = or, if not an abstract, other suitable document summary, as returned by the search tool.

189 <sup>c</sup> = for webpage articles, the webpage text served as the proxy for an abstract, with the  
 190 proviso that the reviewer did not navigate to additional webpages during Step 1.

191

## Supplementary Material: Systematic Review for the Androgens In Men Study.

192 Table S4. Words mentioned in the titles or abstracts of reviewed items.<sup>a</sup>

Word(s)	Step 1 items (n=1,994)	Step 2 items (n=158)	Selected items (n=20)
<i>Search terms (AIMS outcomes)</i>			
cancer	1,449 (72.7)	72 (45.6)	6 (30.0)
colorectal cancer	9 (0.5)	4 (2.5)	2 (10.0)
lung cancer	10 (0.5)	6 (3.8)	2 (10.0)
prostate cancer	1,107 (55.5)	40 (25.3)	2 (10.0)
cardiovascular	219 (11.0)	49 (31)	15 (75.0)
heart failure	29 (1.5)	2 (1.3)	1 (5.0)
stroke	31 (1.6)	12 (7.6)	4 (20.0)
myocardial infarction	33 (1.7)	7 (4.4)	1 (5.0)
mortality	232 (11.6)	45 (28.5)	9 (45.0)
dementia	22 (1.1)	8 (5.1)	2 (10.0)
cognit*	87 (4.4)	20 (12.7)	4 (20.0)
<i>Other frequently observed (not search terms)</i>			
androgen deprivation	583 (29.2)	2 (1.3)	0 (0.0)
androgen receptor	235 (11.8)	10 (6.3)	0 (0.0)
brachytherapy	165 (8.3)	0 (0.0)	0 (0.0)
breast cancer	153 (7.7)	9 (5.7)	0 (0.0)
radiotherapy	371 (18.6)	0 (0.0)	0 (0.0)

193 <sup>a</sup> = Items summarised as numbers (percentages); \*= wildcard character designating truncation

## Supplementary Material: Systematic Review for the Androgens In Men Study.

Table S5. Attributes of selected items.

Item	Article	Country	Study name <sup>§</sup>	Baseline**				Follow-up (relevant outcomes)	
				No. adult males	Baseline period	Age (yr) Mean (sd)	T (nmol/L) Mean (sd)	Length of follow-up (yr) (person-years) <sup>¶</sup>	AIMS Longitudinal Outcomes (no. of events analysed)
<i>Selected from systematic review</i>									
1	Srinath, 2015[16]	USA	ARIC	1,558	1996-98	63.1 (5.6)	13.9 (5.7)	Md=12.8 (CHD); Md=13.1 (HF) (25,374; HF)	Coronary Heart Disease (CHD; 287) Heart Failure (HF; 104) CHD deaths (29) All-cause deaths (347)
2	Srinath, 2016[17]	USA	ARIC	1,558	1996-98	63.1 (5.6)	13.9 (5.7)	Md=14.1 (27,311)	Ischemic Stroke (79)
3	Chan, 2016[18]	Australia	BHS	1,804	1994-95	50.3 (16.8)	13.6 (4.9)	Mn=14.9 (31,930)	CVD events (234; 399)*** CVD deaths (71; 141)*** All-cause deaths (191; 319)***
4	Chasland, 2017[19]	Australia	BHS	1,649	1994-95	49.8 (15.3)	13.7 (4.9)	Tot=20	CVD events (415) CVD deaths (127)
5	Chan, 2018[20]	Australia	BHS	1,574	1994-95	51.1 (14.7)	13.5 (4.8)	Tot=20	Prostate cancer (116) Lung cancer (22) Colorectal cancer (48) Cancer (any; 289)
6	Hsu, 2015[21]	Australia	CHAMP	853	2005-07	76.9 (5.5)	14.6 (6.2)	Tot=5	Cognitive decline (95)
7	Hsu, 2016[22]	Australia	CHAMP	1,705	2005-07	76.9 (5.5)	14.9 (6.6)	Md=6.9; Tot=10 (11,764)	Cancer deaths (151) CVD deaths (185) Other deaths (174) All-cause deaths (510)
8	Hsu, 2018[23]	Australia	CHAMP	1,651	2005-07	76.9 (5.5)	14.7 (6.4)	Tot=5	All-cause deaths (382) CVD deaths (cases not reported) Cancer deaths (cases not reported) Other deaths (cases not reported) Change in: MMSE, SF-12 (Mental).

Supplementary Material: Systematic Review for the Androgens In Men Study.

Item	Article	Country	Study name <sup>s</sup>	Baseline**				Follow-up (relevant outcomes)	
				No. adult males	Baseline period	Age (yr) Mean (sd)	T (nmol/L) Mean (sd)	Length of follow-up (yr) (person-years) <sup>¶</sup>	AIMS Longitudinal Outcomes (no. of events analysed)
9	Rosenberg, 2018[24]	USA	CHS	1,019	1994	76.3 (4.9)	13.2 (6.2)	Md=9.5 (10,716)	Atrial Fibrillation (304)
10	Shores, 2014a[25]	USA	CHS	1,032	1994	76.5 (5.2)	13.5 (6.1)	Md=10; Tot=16 (19,220)	Ischemic stroke (114)
11	Shores, 2014b[26]	USA	CHS	NR	1994	NR	NR	Md=8.9 (CVD events) Md=10.8 yr (All-cause deaths). (9,184; CVD events)	CVD events (436) CVD deaths (157) All-cause deaths (777)
12	Lee, 2013[27]	Europe <sup>§§</sup>	EMAS	2,736	2003-05	59.2 (10.7)	16.5 (6)	Md=4.3; Tot=5 (14,486)	Cancer (any) Myocardial Infarction (MI) Heart Failure, Other heart conditions Stroke Cognitive function All-cause deaths (193)
13	Pye, 2014[28]	Europe <sup>§§</sup>	EMAS	2,599	NR	60 (11)	NR	Md=4.3; Tot=5 (11,140)	Cancer deaths (60) CVD deaths (56) All-cause deaths (147)
14	Chan, 2017[29]	Australia	HIMS	3,690	2001-04	77 (3.6)	13.1 (4.9)	Md=9.1, 9.2; Tot=11 (38,665)	Prostate cancer (348) Lung cancer (107) Colorectal cancer (137)
15	Ford, 2018[30]	Australia	HIMS	4,069	2001-04	NR	NR	Md=10.5; Tot=12 (44,404)	Dementia (499)
16	Yeap, 2014[31]	Australia	HIMS	3,690	2001-04	NR	NR	Mn=6.6 (2.3 sd) (28,036)	MI (344) Stroke (300)

## Supplementary Material: Systematic Review for the Androgens In Men Study.

Item	Article	Country	Study name <sup>§</sup>	Baseline**				Follow-up (relevant outcomes)	
				No. adult males	Baseline period	Age (yr) Mean (sd)	T (nmol/L) Mean (sd)	Length of follow-up (yr) (person-years) <sup>¶</sup>	AIMS Longitudinal Outcomes (no. of events analysed)
17	Ohlsson, 2010[32]	Sweden	MrOS	2,644	2001-04	75.5 (3.2)	15.6 (6.5)	Mn=4.5† (11,880)	CVD deaths (123) Cancer deaths (127) All-cause deaths (328)
18	Ohlsson, 2011[33]	Sweden	MrOS	2,416	2001-04	75.4 (3.2)	15.7 (6)	Md=5.1 (11,605)	CVD events (485) Chronic Heart Disease events (302) Cerebrovascular events (225)
19	Tivesten, 2014[34]	Sweden	MrOS	2,416	2001-04	75.4 (3.2)	15.7 (6)	Md=5.2 (12,070; CHD) (12,137; CBD)	Chronic Heart Disease (302; CHD) Cerebrovascular Disease (225; CBD)
20	Kische, 2017[35]	Germany	SHIP	1,962	1997-01	49.5 (16.3)	15.6 (6.1)	Tot=10	Change in cognitive status
<b>Decision = "Maybe". Item selected based on additional information</b>									
21	LeBlanc, 2010[36]	USA	MrOS	1,602	NR	NR	NR	Mn=4.5†† (26,977)	Cognitive function (and change in) Cognitive decline
22*	Sueoka, 2010[14]	USA	MrOS	697	2000-05	72 (5.5)	14.5 (5.1)	Av=3.9†† (6,247)	Coronary Heart Disease events (100)
<b>Other. Additional studies selected based on information external to the systematic review</b>									
	No articles were selected.	USA	FHS	3,352[12]	1998-05	59.6 (9.1)[12] 49.4 (13.8)[11]	20.7 (8.0)[12]	Tot=10 (for Atrial Fibrillation)[37]	Cardiovascular outcomes[37, 38] Deaths[37] Cause-specific deaths[38] Cancer[39]
	No articles were selected.	Australia	MAILES	1,632[40]	2002-06[8]	54.1 (11.4)[40]	17.3 (5.7)[40]	Md=4.95; IQR=4.35- 5.00[40] (12,686)	CVD events Deaths (99)[8] Cause-specific deaths[8]

§ Study name abbreviations: 'ARIC'= Atherosclerosis Risk in Communities; 'BHS'=Busselton Health Study; 'CHAMP'=The Concord Health and Ageing in Men Project; 'CHS'= Cardiovascular Health Study; 'FHS'= the Framingham Heart Study; 'HIMS'=The Health In Men Study; 'EMAS'=European Male Ageing Study; 'MAILES'= The Men Androgen Inflammation Lifestyle Environment and Stress study; 'MrOS Sweden'=The MrOS Osteoporotic Fractures in Men study in Sweden; 'MrOS USA' = The MrOS Osteoporotic Fractures in Men study USA; 'SHIP'=Study of Health in Pomerania SHIP.

§§ = UK, Italy, Belgium, Poland, Sweden, Spain, Hungary, Estonia



*Supplementary Material: Systematic Review for the Androgens In Men Study.*

¶¶ ‘Md’=median; ‘Mn’=mean; ‘Av’=average; ‘Tot’=total follow-up for the cohort (i.e., maximum, rounded down to nearest whole year); ‘IQR’=interquartile range. Unless provided in text, person-years was calculated by multiplying the median, mean, or average length of follow-up by the total number of adult male participants.

\* = Note that this is a published conference abstract so is not technically a “Full Text” item.

\*\* = Baseline statistics reported for whole cohort; ‘NR’ = statistics not reported for whole cohort; Means and standard deviations calculated by firstly transforming into standard units (for T: nmol/L) and then, where required, transforming from quartile statistics using the Box-Cox method of McGrath et al.[41]

\*\*\* = First number is for individuals without CVD at baseline.

† = Total follow-up exceeded 5 years, from baseline visit (2001-04) to end of mortality data collection (March 1, 2008).

†† = Note that since there was no published follow-up estimate exceeding 5 years (a requirement for selection) and it was not clear, based on the article information alone, whether the total follow-up was at least 5 years, these items were initially classified as “Maybe”. The length of follow-up for collection of AIMS outcome data was determined to be satisfactory from subsequent correspondence with MrOS USA researchers.

## Supplementary Material: Systematic Review for the Androgens In Men Study.

Table S6. Exposure levels, outcome assessment, covariates.

Study	Article	Longitudinal measure of association	Exposure* (testosterone)	Outcome ascertainment	Covariates
ARIC	Srinath, 2015[16]	HR	T quartiles	CVD events and deaths identified by annual questionnaires and continuous surveillance, independent from hospital admissions data (ICD codes). Cause of death from death certificates.	Age, race/centre, BMI, waist circumference, smoking status, diabetes mellitus, hypertension, LDL, HDL.
	Srinath, 2016[17]	HR	T tertiles	Definite or probable stroke events identified from hospital admissions, annual phone calls, study examinations adjudicated by a physician, with secondary physician adjudication if it disagreed with a computer algorithm.	Age, race, centre, BMI, waist circumference, smoking status, diabetes mellitus, hypertension, LDL, HDL.
BHS	Chan, 2016[18]	HR	T quartiles (results not shown), Continuous T.	Linked hospital admissions and deaths records (ICD codes)	Age, smoking, vigorous exercise, alcohol, BMI, diabetes, CVD, COPD, non-skin cancer, systolic blood pressure, hypertension, lipid lowering therapy, cholesterol, HDL, triglycerides, C-reactive protein, creatinine
	Chasland, 2017[19]	HR	Categories: Low (L) v High (H) T, physical activity(PA) LT+LPA, LT+HPA, HT+LPA, HT+HPA	Linked hospital admissions and deaths records (ICD codes)	Age, prevalent CVD, smoking, waist circumference, cholesterol, HDL, lipids medication, diabetes, systolic blood pressure, hypertension medication
	Chan, 2018[20]	HR	T quartiles, Continuous T.	Linked cancer and death registry records (ICD codes)	Age, marital status, occupation, smoking, alcohol consumption, leisure time physical activity, BMI, diabetes
CHAMP	Hsu, 2015[21]	Slope estimate (change in MMSE on baseline hormone level or longitudinal change in hormone level)	Continuous T, cFT	Clinic assessment: MMSE, Informant Questionnaire on Cognitive Decline as initial screen, followed by clinical assessment to diagnosis categories: normal cognition, MCI, dementia. During follow-up: A decline in MMSE by $\geq 3$ points	Age, BMI, smoking status, years of education, depression score (GDS)
	Hsu, 2016[22]	RR	Continuous T, cFT	Deaths identified from 4-monthly phone calls or deaths registry. Cause of death identified on death certificates independently by 2 medical practitioners.	Age, BMI, smoking status, comorbidity score
	Hsu, 2018[23]	HR, RR (Death outcomes); Slope estimates (MMSE, SF-12 Mental)	Categories: Low (<20 <sup>th</sup> centile) v Normal T combinations with Low (<20 <sup>th</sup> centile) v Normal cFT	Cause of death identified on death certificates independently by 2 medical practitioners.	Age, BMI, smoking status, comorbidity score
CHS	Rosenberg, 2018[24]	HR	Continuous T and cFT, T and cFT quintiles	Independently verified from ECGs taken annually for participants and from hospital discharge diagnoses	Age (stratified), race, education, income, clinic, smoking status, diabetes mellitus, BMI, loop diuretics, height, hypertension, depressed left ventricular ejection fraction, kidney function, systolic blood pressure, SHBG

Supplementary Material: Systematic Review for the Androgens In Men Study.

Study	Article	Longitudinal measure of association	Exposure* (testosterone)	Outcome ascertainment	Covariates
	Shores, 2014a[25]	HR	Continuous T, cFT (linear & non-linear), T categories	Medicare data, hospital records, imaging studies, autopsy results, death certificates, physician interviews data used for adjudications by committee, which included a neurologist.	Age, systolic blood pressure, anti-hypertensive medications, atrial fibrillation, diabetes, smoking, lipid-lowering drugs, HDL, cholesterol, creatinine, fasting glucose, diabetes medications.
	Shores, 2014b[26]	HR	Continuous T, cFT (linear or non-linear) categories: Q1, Q2-4	Medicare data, hospital records, imaging studies, autopsy results, death certificates, physician interviews data used for adjudications by committee, which included a neurologist.	Age, race, site, smoking status, alcohol consumption, hypertensive use, HDL, BMI, waist circumference, diabetes, SHBG.
EMAS	Lee, 2013[27]	N/A	No modelling of longitudinal outcomes reported	MI, heart failure, other heart conditions, cancers, stroke identified from postal questionnaire, MMSE for participants ≥65 yr old from clinic assessments Variable methods for data capture + validation among centres.	No modelling of longitudinal outcomes reported
	Pye, 2014[28]	HR	T, free T categories: quintiles, low v eugonadal T, LOH status.	Deaths identified from follow-up postal questionnaire or enquiry if no reply received, with 89% of deaths verified from death certificates, death registers, or medical/hospital records.	Age, site, BMI, smoking status, general health.
HIMS	Chan, 2017[29]	SHR	Continuous T, cFT.	Linked hospital admissions, death and cancer registry records (ICD, ICD-O-3 codes).	Age, BMI, smoking status, physical activity, alcohol consumption, diabetes mellitus, HDL, triglycerides, prior cancer diagnosis.
	Ford, 2018[30]	HR	Continuous T, free T Quartile categories of T, free T	Linked data (ICD codes) from inpatient and outpatient mental health services, hospital admissions, community aged care services, cancer and death registries.	Age, baseline cognitive function, depression, BMI, hypertension, CVD, plasma homocysteine.
	Yeap, 2014[31]	HR	T, free T as quartile categories	Linked hospital admissions, death and cancer registry records (ICD codes).	Age, education, smoking status, BMI, waist to hip ratio, hypertension, dyslipidemia, diabetes, creatinine, prior cancer or existing CVD. Also SHBG for models with T.
MrOS Europe	Ohlsson, 2010[32]	HR (in relation to DHEA, DHEA-S)	No: T modelled as a covariate only	Linked data (death and hospital discharge registries), death certificates.	Age, site, BMI, C-reactive protein, ApoB/A1, smoking status, diabetes, hypertension, prior CVD, prior cancer, low testosterone (in lowest quartile), low estradiol
	Ohlsson, 2011[33]	HR	T, free T as quartile categories, T as binary categories.	Linked data (death and hospital discharge registries), death certificates.	Age, morning sample, site, BMI, ApoB/A1, physical activity, smoking status, diabetes, hypertension
	Tivesten, 2014[34]	HR (in relation to DHEA, DHEA-S)	No: T modelled as a covariate only	Linked data (death and hospital discharge registries), death certificates.	Age, morning sample, site, BMI, ApoB/A1, C-reactive protein, estradiol, testosterone (i.e., continuous T), SHBG, eGFR, smoking status, diabetes, hypertension.
SHIP	Kische, 2017[35]	Slope estimate (change in MMSE on baseline hormone)	T, free T as continuous and as 10-year age group quartile categories.	MMSE score.	Age, BMI, smoking status, alcohol consumption, physical activity, hypertension, occupational status, education level, civil status, baseline MMSE.

## Supplementary Material: Systematic Review for the Androgens In Men Study.

Study	Article	Longitudinal measure of association	Exposure* (testosterone)	Outcome ascertainment	Covariates
MrOS USA	LeBlanc, 2010[36]	Change in mean score RR of clinically important decline	Free T quartiles and continuous free T	Cognitive tests at the baseline and follow-up visit from Part B of the Trail Making Test (Trails B) and the Modified Mental State Examination (3MS). Calculated from pre-defined drop in scores.	Age group, education level, race, general health, alcohol consumption, clinic, physical and mental health, physical activity, medications used at baseline, other sex steroids, SHBG.
	Sueoka, 2010[14]	HR	T quartiles	CHD events identified from 3-monthly contacts with participants. Incident events were reviewed and adjudicated by cardiologist using clinical records.	Age, clinic, BMI, blood pressure, lipid levels, smoking, hypertension, diabetes, use of lipid-lowering agents
FHS	N/A – no items were selected.	N/A	N/A	AF measured and adjudicated by cardiologists. Mortality data from death certificates, hospital or institutional records, obituaries, or direct notification[37] Medical records of CVD events reviewed by panel of experienced investigators. A heart study neurologist examined most participants with suspected stroke[38] Medical records of cancer diagnoses reviewed by two independent reviewers, with majority confirmed by pathology reports.[39]	N/A
MAILES	N/A – no items selected.	N/A	N/A	Self-reported and clinical follow-up data, death registry (linked data)[8]	N/A

\* T = total testosterone; cFT = calculated free testosterone; Q1=quartile 1; Q2-4=quartiles 2 to 4 combined.

Table S7. Newcastle-Ottawa Quality Assessment Scale for Cohort Studies: selected articles.

Article	Study	Selection (4 stars)	Comparability (2 stars)	Outcome (3 stars)	Notes on Selection	Notes on Outcome
Srinath 2015[16]	ARIC	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Srinath 2016[17]	ARIC	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Chan 2016[18]	BHS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Chasland 2017[19]	BHS	***	**	**	Prevalent cases not excluded <sup>a</sup>	Losses to f/u not mentioned; linked data <sup>b</sup>
Chan 2018[20]	BHS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Hsu 2015[21]	CHAMP	****	**	***		
Hsu 2016[22]	CHAMP	***	**	***	Prevalent cases not excluded <sup>c</sup>	
Hsu 2018[23]	CHAMP	***	**	***	Prevalent cases not excluded <sup>c</sup>	
Rosenberg 2018[24]	CHS	****	**	**		Losses to f/u not mentioned; linked data <sup>d</sup>
Shores 2014a[25]	CHS	****	**	**		Losses to f/u not mentioned; linked data <sup>d</sup>
Shores 2014b[26]	CHS	****	**	**		Losses to f/u not mentioned; linked data <sup>d</sup>
Lee 2013[27]	EMAS	NA	NA	NA	No modelling of longitudinal outcomes reported	
Pye 2014[28]	EMAS	***	**	***	Prevalent cases not excluded <sup>a</sup>	
Chan 2017[29]	HIMS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Ford 2018[30]	HIMS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Yeap 2014[31]	HIMS	***	**	**	Prevalent cases not excluded <sup>a</sup>	Losses to f/u not mentioned; linked data <sup>b</sup>
Ohlsson 2010[32]	MrOS Sw.	NA	NA	NA	Testosterone was not the exposure variable in this article	
Ohlsson 2011[33]	MrOS Sw.	****	**	***		
Tivesten 2014[34]	MrOS Sw.	NA	NA	NA	Testosterone was not the exposure variable in this article	
Kische 2017[35]	SHIP	***	**	***	Prevalent cases not excluded <sup>e</sup>	
LeBlanc 2010[36]	MrOS USA	****	**	*		Bias from loss to f/u; F/u OK: additional steps <sup>f</sup>
Sueoka 2010[14]	MrOS USA	***	**	*	Prevalent cases not excluded <sup>a</sup>	F/u OK: additional steps <sup>f</sup>
<b>Additional item not selected but included in DR-MA:</b>						
Yeap 2014b[13]	HIMS	***	**	**	Prevalent cases not excluded <sup>a</sup>	Losses to f/u not mentioned; linked data <sup>b</sup>

‘NA’ = Not applicable (see Notes); ‘f/u’ = follow-up (of incident events); ‘DR-MA’ = dose-response meta-analyses of published estimates.  
<sup>a</sup> = The influence of prevalent cases was statistically adjusted by including prevalent status as a model predictor.  
<sup>b</sup> = Follow-up of cases was assumed to be almost complete because analyses were of linked administrative data.  
<sup>c</sup> = The influence of prevalent cases was statistically adjusted by incorporating into a comorbidity status model predictor.  
<sup>d</sup> = Follow-up of cases was assumed to be almost complete because analyses were of linked administrative data (with expert adjudications).  
<sup>e</sup> = Outcome was change in cognition score, with baseline score (prevalent status) included as a model predictor.  
<sup>f</sup> = Total length of follow-up period was not reported but determined to be satisfactory from correspondence with MrOS USA researchers.

## Supplementary Material: Systematic Review for the Androgens In Men Study.

Table S8. PRISMA Checklist for Systematic Review: the Androgens In Men Study.

Section/topic	#	Checklist item	Reported on page #
<b>TITLE</b>			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
<b>ABSTRACT</b>			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5. For this type of review it is PEO instead.
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5, Table S2
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Table S1

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6, Tables S2-S3.
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6, Suppl. Data.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6, Suppl. Data.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6, Table S7.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I <sup>2</sup> ) for each meta-analysis.	7

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7
<b>RESULTS</b>			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Tables S5-6, S9
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9, Table S7
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9-11, Figs. 2-4

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-11, Figs. 3-4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10-11, Fig. S2
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9-10, Fig. S1
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	11-14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	11-14
<b>FUNDING</b>			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	15

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097



## Supplementary Material: Systematic Review for the Androgens In Men Study.

Table S9. Extracted hazard ratio data for dose-response meta-analyses (DR-MAs).

Article	Study	Outcome	Testosterone	Units	HR	95% CI	Notes
Chan 2016[18]	BHS	All-cause mortality	<10.20	nmol/L	ref.		
Chan 2016	BHS	All-cause mortality	10.20 - <13.04	nmol/L	0.84	(0.62-1.14)	
Chan 2016	BHS	All-cause mortality	13.04 - <16.58	nmol/L	0.86	(0.62-1.19)	
Chan 2016	BHS	All-cause mortality	≥16.58	nmol/L	0.9	(0.62-1.3)	
Pye 2014[28]	EMAS	All-cause mortality	<11.65	nmol/L	1.1	(0.6-1.8)	
Pye 2014	EMAS	All-cause mortality	11.65-14.61	nmol/L	0.7	(0.4-1.3)	
Pye 2014	EMAS	All-cause mortality	14.61-17.28	nmol/L	0.7	(0.4-1.3)	
Pye 2014	EMAS	All-cause mortality	17.28-21.20	nmol/L	1.2	(0.7-2)	
Pye 2014	EMAS	All-cause mortality	>21.20	nmol/L	ref.		
Srinath 2015[16]	ARIC	All-cause mortality	≤288.4	ng/dL	0.96	(0.7-1.34)	
Srinath 2015	ARIC	All-cause mortality	288.5-377.6	ng/dL	0.99	(0.72-1.35)	
Srinath 2015	ARIC	All-cause mortality	377.7-480.1	ng/dL	1	(0.74-1.35)	
Srinath 2015	ARIC	All-cause mortality	≥480.2	ng/dL	ref.		
Shores 2014b[26]	CHS	All-cause mortality	<278	ng/dL	1.06	(0.88-1.29)	
Shores 2014b	CHS	All-cause mortality	≥278	ng/dL	ref.		
Yeap 2014b[13]	HIMS	All-cause mortality	0.25-9.82	nmol/L	ref.		Fully-adjusted model + SHBG
Yeap 2014b	HIMS	All-cause mortality	9.82-12.53	nmol/L	0.81	(0.68-0.98)	
Yeap 2014b	HIMS	All-cause mortality	12.56-15.75	nmol/L	0.75	(0.61-0.92)	
Yeap 2014b	HIMS	All-cause mortality	15.79-46.50	nmol/L	0.77	(0.61-0.97)	
Yeap 2014b	HIMS	All-cause mortality	0.25-9.82	nmol/L	ref.		Fully-adjusted model + LH
Yeap 2014b	HIMS	All-cause mortality	9.82-12.53	nmol/L	0.84	(0.7-1.01)	
Yeap 2014b	HIMS	All-cause mortality	12.56-15.75	nmol/L	0.81	(0.67-0.97)	
Yeap 2014b	HIMS	All-cause mortality	15.79-46.50	nmol/L	0.89	(0.73-1.07)	
Chan 2016	BHS	CVD mortality	<10.20	nmol/L	ref.		
Chan 2016	BHS	CVD mortality	10.20 - <13.04	nmol/L	1.12	(0.7-1.78)	
Chan 2016	BHS	CVD mortality	13.04 - <16.58	nmol/L	1.39	(0.86-2.25)	
Chan 2016	BHS	CVD mortality	≥16.58	nmol/L	1.25	(0.69-2.25)	
Chasland 2017[19]	BHS	CVD mortality	<13.1	nmol/L	ref.		Total PA, "Low" PA, NS PA x T: these estimates were used
	BHS	CVD mortality	≥13.1	nmol/L	1.25	(0.77-2.03)	
Chasland 2017	BHS	CVD mortality	<13.1	nmol/L	0.69	(0.4-1.2)	Total PA, "High" PA, NS PA x T
Chasland 2017	BHS	CVD mortality	≥13.1	nmol/L	0.8	(0.48-1.35)	
Pye 2014	EMAS	CVD mortality	<11.65	nmol/L	1	(0.4-2.2)	
Pye 2014	EMAS	CVD mortality	11.65-14.61	nmol/L	0.5	(0.2-1.4)	
Pye 2014	EMAS	CVD mortality	14.61-17.28	nmol/L	0.4	(0.2-1.2)	
Pye 2014	EMAS	CVD mortality	17.28-21.20	nmol/L	1.1	(0.5-2.4)	
Pye 2014	EMAS	CVD mortality	>21.20	nmol/L	ref.		
Srinath 2015	ARIC	CVD mortality	≤288.4	ng/dL	1.36	(0.45-4.08)	
Srinath 2015	ARIC	CVD mortality	288.5-377.6	ng/dL	1.26	(0.73-3.7)	
Srinath 2015	ARIC	CVD mortality	377.7-480.1	ng/dL	0.57	(0.16-1.99)	
Srinath 2015	ARIC	CVD mortality	≥480.2	ng/dL	ref.		

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

Article	Study	Outcome	Testosterone	Units	HR	95% CI	Notes
Shores 2014b	CHS	CVD mortality	<278	ng/dL	1.28	(0.94-1.75)	
Shores 2014b	CHS	CVD mortality	≥278	ng/dL	ref.		
Yeap 2014b	HIMS	CVD mortality	0.25-9.82	nmol/L	ref.		
Yeap 2014b	HIMS	CVD mortality	9.82-12.53	nmol/L	0.82	(0.61-1.11)	
Yeap 2014b	HIMS	CVD mortality	12.56-15.75	nmol/L	0.79	(0.58-1.09)	
Yeap 2014b	HIMS	CVD mortality	15.79-46.50	nmol/L	0.79	(0.56-1.11)	

For peer review only

## Supplementary Material: Figures.

Figure S1. Meta-regression diagnostics.

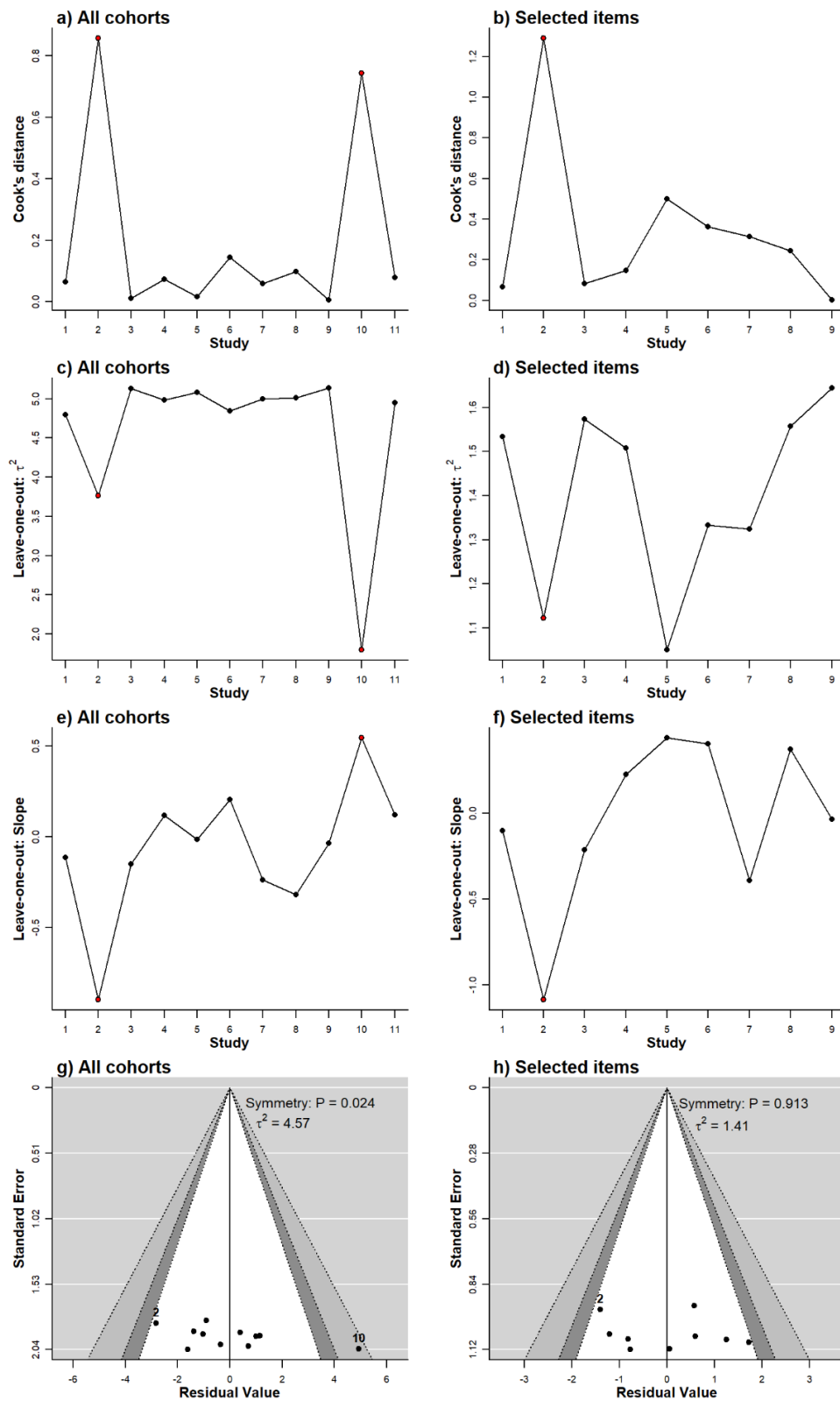


Figure S1: Meta-regression diagnostics showing the influence of studies on model fit (a,b),  $\tau^2$  (estimated amount of total heterogeneity: c,d), estimated slope (e,f), and distribution of residuals with funnel plots (g,h). Analysis repeated for all 11 cohort studies (a,c,e,g) and for 9

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

1  
2  
3 studies with selected articles (b,d,f,h). In cases where more than one article was available per  
4 cohort study, the article with the largest sample size was used. Highlighted estimates for  
5 cohort study 2 (BHS) were those from Chan et al.[18] (N=1,804) and for study 10 (FHS)  
6 were from Pencina et al.[12] (N=720). In funnel plots: light grey + dark grey + white shading  
7 = 99% pseudo confidence interval (CI); dark grey + white shading = 95% CI; white shading  
8 = 90% CI.  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

Figure S2. Funnel plots for dose-response meta-analyses.

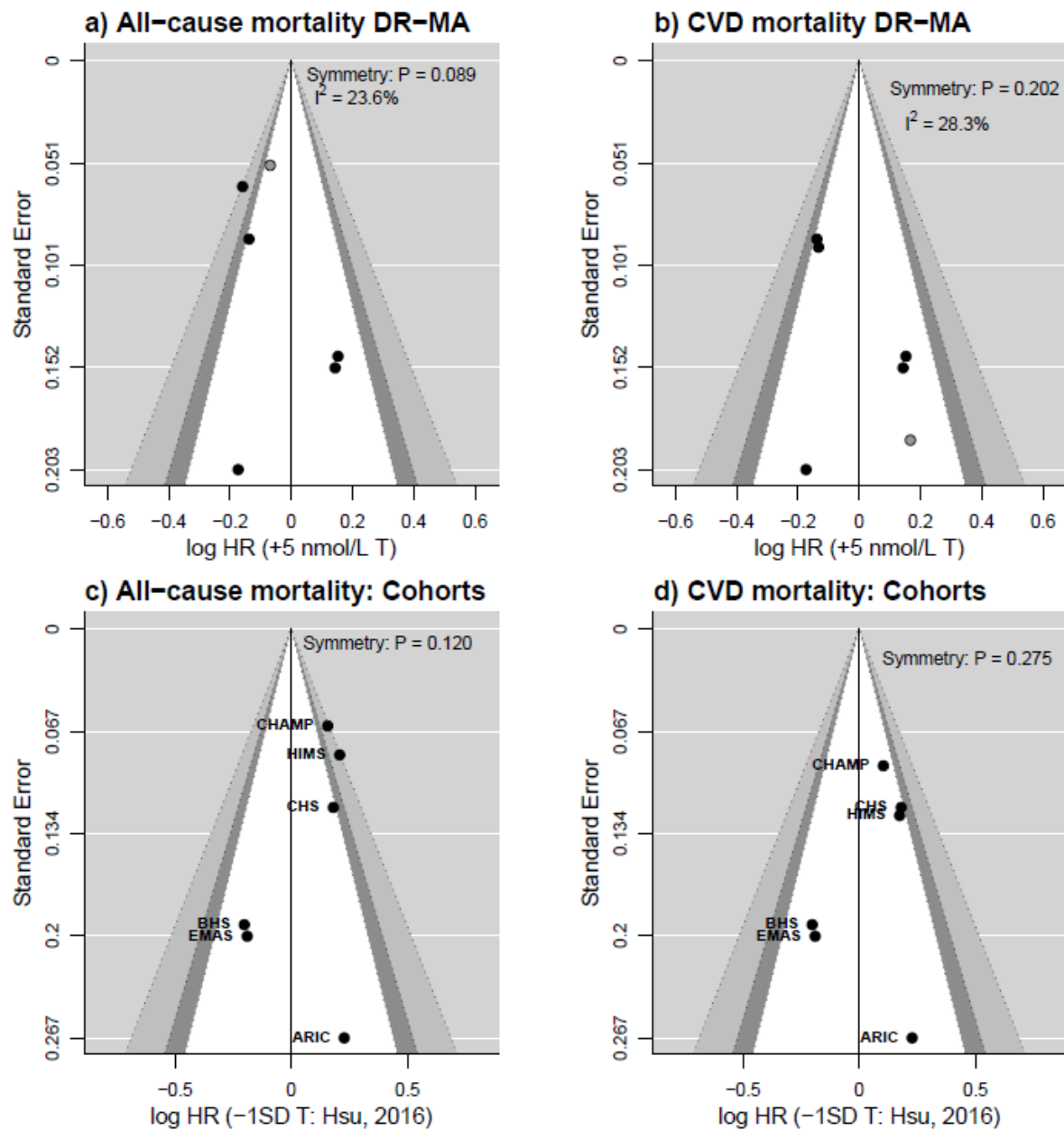


Figure S2: Contour-enhanced funnel plots showing the distribution of log hazard ratio (HR) estimates for all-cause mortality (a, c) and mortality caused by cardiovascular disease (CVD) (b, d) attributed to a 5 nmol/L increase (a, b), or a 1.9 nmol/L (1SD in Hsu et al. 2016[22]) decrease, in endogenous testosterone concentration. Log HR values and standard errors were calculated using generalised least squares regression of published estimates.[42, 43] In cases where more than one article was available per cohort study, the article with the largest sample size was used. Estimates represented by black dots in (a) and (b) were analysed in respective dose-response meta-analyses (DR-MA; results presented in Figs. 3, 4). The grey dot in (a) is the estimate for Yeap et al. 2014b[13] and in (b) is the estimate for Chasland et al. (2017)[19]; these estimates were substituted for others for the HIMS and BHS studies respectively for alternative summary estimates (i.e., the grey summary estimates presented in Figs. 3, 4). Estimates presented in (c) and (d) are shown for a more complete assessment of funnel plot symmetry: estimates are plotted for all studies with estimates, including those that

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

did not have sufficient information for including in the DR-MA. In funnel plots: light grey + dark grey + white shading = 99% pseudo confidence interval (CI); dark grey + white shading = 95% CI; white shading = 90% CI.

**References cited.**

1. MedNar: Deep Web Technologies.
2. OpenGrey: GreyNet International.
3. Yeap BB, Marriott RJ, Adams RJ, et al. Androgens In Men Study (AIMS): protocol for meta-analyses of individual participant data investigating associations of androgens with health outcomes in men. *BMJ Open* 2020; 10. <https://bmjopen.bmj.com/content/10/5/e034777> (accessed 19 Nov 2020).
4. Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the Journal of Clinical Endocrinology and Metabolism. *J Clin Endocrinol Metab* 2013;98(10):3971-73.
5. Endnote X8: Clarivate Analytics, 2019
6. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan - a web and mobile app for systematic reviews. *Syst Rev* 2016; 5.
7. Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to Meta-Analysis. West Sussex: John Wiley & Sons Ltd 2009.
8. Grant JF, Martin SA, Taylor AW, et al. Cohort profile: The men androgen inflammation lifestyle environment and stress (MAILES) study. *Int J Epidemiol* 2014;43(4):1040-53. doi: <https://dx.doi.org/10.1093/ije/dyt064>
9. Arnlov J, Pencina MJ, Amin S, et al. Endogenous sex hormones and cardiovascular disease incidence in men. *Ann Intern Med* 2006;145(3):176-84.
10. Haring R, Teng Z, Xanthakis V, et al. Association of sex steroids, gonadotrophins, and their trajectories with clinical cardiovascular disease and all-cause mortality in elderly men from the Framingham Heart Study. *Clin Endocrinol (Oxf)* 2013;78(4):629-34. doi: <https://dx.doi.org/10.1111/cen.12013>
11. Bhasin S, Pencina MJ, Kaur Jasuja G, et al. Reference ranges for testosterone in men generated using liquid chromatography tandem mass spectrometry in a community-base sample of healthy nonobese young men in the Framingham Heart Study and applied to three geographically distinct cohorts. *J Clin Endocrinol Metab* 2011;96(8):2430-39.
12. Pencina KM, Travison TG, Bhasin S, et al. Endogenous circulating testosterone and sex hormone-binding globulin levels and measures of myocardial structure and function: the Framingham Heart Study. *Andrology* 2019;7:307-14.
13. Yeap BB, Alfonso H, Chubb SAP, et al. In older men an optimal plasma testosterone is associated with reduced all-cause mortality and higher dihydrotestosterone with reduced ischemic heart disease mortality, while estradiol levels do not predict mortality. *J Clin Endocrinol Metab* 2014;99(1):E9-E18.
14. Sueoka KT, Ewing MS, Ensrud KE, et al. Higher endogenous testosterone levels associated with increased risk of coronary heart disease in elderly men: a prospective study. *Endocr Rev* 2010;31(3):S858.
15. Holmegard HN, Nordestgaard BG, Jensen GB, et al. Sex hormones and ischemic stroke: a prospective cohort study and meta-analyses. *J Clin Endocrinol Metab* 2016;101:69-78.
16. Srinath R, Hill Golden S, Carson KA. Endogenous Testosterone and its Relationship to Preclinical and Clinical Measures of Cardiovascular Disease in the Atherosclerosis Risk in Communities Study. *J Clin Endocrinol Metab* 2015;100(4):1602-02. doi: 10.1210/jc.2014-3934



*Supplementary Material: Systematic Review for the Androgens In Men Study.*

17. Srinath R, Gottesman RF, Hill Golden S, et al. Association Between Endogenous Testosterone and Cerebrovascular Disease in the ARIC Study (Atherosclerosis Risk in Communities). *Stroke* 2016;47(11):2682-88.
18. Chan YX, Knuiman MW, Hung J, et al. Neutral associations of testosterone, dihydrotestosterone and estradiol with fatal and non-fatal cardiovascular events, and mortality in men aged 17-97 years. *Clin Endocrinol (Oxf)* 2016;85(4):575-82. doi: <https://dx.doi.org/10.1111/cen.13089>
19. Chasland LC, Knuiman MW, Divitini ML, et al. Greater physical activity and higher androgen concentrations are independently associated with lower cardiometabolic risk in men. *Clin Endocrinol (Oxf)* 2017;87(5):466-74. doi: <https://dx.doi.org/10.1111/cen.13407>
20. Chan YX, Knuiman MW, Divitini ML, et al. Lower Circulating Androgens Are Associated with Overall Cancer Risk and Prostate Cancer Risk in Men Aged 25-84 Years from the Busselton Health Study. *Horm Cancer* 2018;9(6):391-98. doi: <https://dx.doi.org/10.1007/s12672-018-0346-5>
21. Hsu B, Cumming RG, Waite LM, et al. Longitudinal Relationships between Reproductive Hormones and Cognitive Decline in Older Men: The Concord Health and Ageing in Men Project. *J Clin Endocrinol Metab* 2015;100(6):2223-30. doi: <https://dx.doi.org/10.1210/jc.2015-1016>
22. Hsu B, Cumming RG, Naganathan V, et al. Temporal Changes in Androgens and Estrogens Are Associated With All-Cause and Cause-Specific Mortality in Older Men. *J Clin Endocrinol Metab* 2016;101(5):2201-10. doi: <https://dx.doi.org/10.1210/jc.2016-1025>
23. Hsu B, Cumming RG, Blyth FM, et al. Evaluating Calculated Free Testosterone as a Predictor of Morbidity and Mortality Independent of Testosterone for Cross-sectional and 5-Year Longitudinal Health Outcomes in Older Men: The Concord Health and Ageing in Men Project. *J Gerontol A Biol Sci Med Sci* 2018;73(6):729-36. doi: <https://dx.doi.org/10.1093/gerona/glx170>
24. Rosenberg MA, Shores MM, Matsumoto AM, et al. Serum androgens and risk of atrial fibrillation in older men: The Cardiovascular Health Study. *Clin Cardiol* 2018;41(6):830-36. doi: <https://dx.doi.org/10.1002/clc.22965>
25. Shores MM, Arnold AM, Biggs ML, et al. Testosterone and dihydrotestosterone and incident ischaemic stroke in men in the Cardiovascular Health Study. *Clin Endocrinol (Oxf)* 2014;81(5):746-53. doi: <https://dx.doi.org/10.1111/cen.12452>
26. Shores MM, Biggs ML, Arnold AM, et al. Testosterone, dihydrotestosterone, and incident cardiovascular disease and mortality in the cardiovascular health study. *J Clin Endocrinol Metab* 2014;99(6):2061-8. doi: <https://dx.doi.org/10.1210/jc.2013-3576>
27. Lee DM, Pye SR, Tajar A, et al. Cohort profile: the European Male Ageing Study. *Int J Epidemiol* 2013;42(2):391-401. doi: <https://dx.doi.org/10.1093/ije/dyr234>
28. Pye SR, Huhtaniemi IT, Finn JD, et al. Late-onset hypogonadism and mortality in aging men. *J Clin Endocrinol Metab* 2014;99(4):1357-66. doi: <https://dx.doi.org/10.1210/jc.2013-2052>
29. Chan YX, Alfonso H, Chubb SA, et al. Higher Dihydrotestosterone Is Associated with the Incidence of Lung Cancer in Older Men. *Horm Cancer* 2017;8(2):119-26. doi: <https://dx.doi.org/10.1007/s12672-017-0287-4>
30. Ford AH, Yeap BB, Flicker L, et al. Sex hormones and incident dementia in older men: The health in men study. *Psychoneuroendocrinology* 2018;98:139-47. doi: <https://dx.doi.org/10.1016/j.psyneuen.2018.08.013>
31. Yeap BB, Alfonso H, Chubb SA, et al. In older men, higher plasma testosterone or dihydrotestosterone is an independent predictor for reduced incidence of stroke but not myocardial infarction. *J Clin Endocrinol Metab* 2014;99(12):4565-73. doi: <https://dx.doi.org/10.1210/jc.2014-2664>

*Supplementary Material: Systematic Review for the Androgens In Men Study.*

32. Ohlsson C, Labrie F, Barrett-Connor E, et al. Low serum levels of dehydroepiandrosterone sulfate predict all-cause and cardiovascular mortality in elderly Swedish men. *J Clin Endocrinol Metab* 2010;95(9):4406-14. doi: <https://dx.doi.org/10.1210/jc.2010-0760>
33. Ohlsson C, Barrett-Connor E, Bhasin S, et al. High serum testosterone is associated with reduced risk of cardiovascular events in elderly men. The MrOS (Osteoporotic Fractures in Men) study in Sweden. *J Am Coll Cardiol* 2011;58(16):1674-81. doi: <https://dx.doi.org/10.1016/j.jacc.2011.07.019>
34. Tivesten Å, Vandenput L, Carlzon D, et al. Dehydroepiandrosterone and its sulfate predict the 5-year risk of coronary heart disease events in elderly men. *J Am Coll Cardiol* 2014;64(17):1801-10. doi: <https://dx.doi.org/10.1016/j.jacc.2014.05.076>
35. Kische H, Gross S, Wallaschofski H, et al. Associations of androgens with depressive symptoms and cognitive status in the general population. *PLoS ONE* 2017;12(5):e0177272. doi: <https://dx.doi.org/10.1371/journal.pone.0177272>
36. LeBlanc ES, Wang PY, Janowsky JS, et al. Association between sex steroids and cognition in elderly men. *Clin Endocrinol (Oxf)* 2010;72(3):393-403. doi: <https://dx.doi.org/10.1111/j.1365-2265.2009.03692.x>
37. Magnani JW, Moser CB, Murabito JM, et al. Association of sex hormones, aging, and atrial fibrillation in men: The Framingham Heart Study. *Circ Arrhythm Electrophysiol* 2014;7:307-12.
38. D'Agostino RB, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: The Framingham Heart Study. *Circulation* 2008;117:743-53.
39. Murabito JM, Rosenberg CL, Finger D, et al. A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham Heart Study. *BMC Med Genet* 2007; 8(Suppl I).
40. Li JJ, Wittert GA, Vincent A, et al. Muscle grip strength predicts incident type 2 diabetes: population-based cohort study. *Metab Clin Exp* 2016;65:883-92.
41. McGrath S, Zhao X, Steele R, et al. Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Stat Methods Med Res* 2020;29(9):2520-37. doi: 10.1177/0962280219889080
42. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135(11):1301-09.
43. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose-response data. *Stata J* 2006;6(1):40-57.



PRISMA Checklist for Systematic Review: the Androgens In Men Study.

Section/topic	#	Checklist item	Reported on page #
<b>TITLE</b>			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
<b>ABSTRACT</b>			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5. For this type of review it is PEO instead.
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5, Table S2
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Table S1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6, Tables S2-S3.

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6, Suppl. Data.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6, Suppl. Data.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6, Table S7.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$ ) for each meta-analysis.	7

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7
<b>RESULTS</b>			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Tables S5-6, S9
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9, Table S7
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9-11, Figs. 2-4
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-11, Figs. 3-4

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10-11, Fig. S2
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9-10, Fig. S1
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	11-14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	11-14
<b>FUNDING</b>			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	15

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

# BMJ Open

## Systematic review and meta-analyses on associations of endogenous testosterone concentration with health outcomes in community-dwelling men.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-048013.R1
Article Type:	Original research
Date Submitted by the Author:	11-Aug-2021
Complete List of Authors:	Marriott, Ross; The University of Western Australia, School of Population and Global Health Harse, Janis; The University of Western Australia, School of Population and Global Health Murray, Kevin; The University of Western Australia, School of Population and Global Health Yeap, Bu; The University of Western Australia, Medical School; Harry Perkins Institute of Medical Research
<b>Primary Subject Heading</b>:	Epidemiology
Secondary Subject Heading:	Research methods
Keywords:	EPIDEMIOLOGY, Sex steroids & HRT < DIABETES & ENDOCRINOLOGY, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3 1 Systematic review and meta-analyses on associations of endogenous testosterone  
4  
5 2 concentration with health outcomes in community-dwelling men.  
6  
7  
8 3

9  
10 4 Corresponding author:

11  
12 5 Ross J. Marriott PhD, School of Population and Global Health, Faculty of Health and  
13  
14 6 Medical Sciences, The University of Western Australia, Clifton Street Building, Clifton  
15  
16  
17 7 Street, Nedlands 6009, Western Australia, Australia.

18  
19 8 Email: [ross.marriott@uwa.edu.au](mailto:ross.marriott@uwa.edu.au); Telephone: +61 8 6488 1299  
20  
21  
22 9

23  
24 10 Authors:

25  
26 11 Ross J. Marriott, School of Population and Global Health, The University of Western  
27  
28 12 Australia, Nedlands 6009, Australia.

29  
30 13 Janis Harse, School of Population and Global Health, The University of Western Australia,  
31  
32 14 Nedlands 6009, Australia.

33  
34 15 Kevin Murray, School of Population and Global Health, The University of Western Australia,  
35  
36 16 Nedlands 6009, Australia.

37  
38 17 Bu B Yeap, Medical School, The University of Western Australia, Perth 6009, Australia.  
39  
40  
41  
42 18

43  
44 19 Short title:

45  
46 20 Systematic review: associations of testosterone with men's health  
47  
48  
49 21

50  
51 22 Keywords:

52  
53 23 Testosterone, Individual Participant Data, Cardiovascular Disease, Cancer, Mortality,  
54  
55 24 Dementia, Meta-analysis.  
56  
57  
58 25

1  
2  
3 26 Word count, excluding title page, abstract, strengths and limitations, references,  
4  
5 27 acknowledgements, contributions, figures and tables: 3,639 words.  
6  
7  
8 28

9  
10 29 **ABSTRACT**

11  
12 30 Objectives

13  
14 31 The overall study aim is to clarify the relation of endogenous sex hormones with major health  
15  
16 32 outcomes in men. This paper reports a systematic review focussing on published estimates for  
17  
18 33 testosterone associations.  
19

20  
21 34 Setting

22  
23 35 Community-dwelling men.

24  
25 36 Participants

26  
27 37 20,180 adult males participated in the final set of studies identified and selected from a  
28  
29 38 systematic review. Eligible studies included prospective cohort studies with plasma or serum  
30  
31 39 testosterone concentrations measured for adult males using mass spectrometry with at least 5  
32  
33 40 years of follow-up data and one of the specified outcome measures recorded. Only published  
34  
35 41 or grey literature items written in English were considered.  
36  
37

38  
39 42 Primary and secondary outcome measures

40  
41 43 Planned prospective outcome measures: cardiovascular disease (CVD) events, CVD deaths,  
42  
43 44 all-cause mortality, cancer deaths, cancer diagnoses, cognitive decline, dementia. Meta-  
44  
45 45 analyses were of the most frequently reported outcomes in selected studies: CVD deaths and  
46  
47 46 all-cause mortality. Succinct characterisations of testosterone associations with other  
48  
49 47 outcomes are also presented.  
50

51  
52 48 Results

53  
54 49 Screening of 1,994 de-duplicated items identified 9 suitable studies, with an additional two  
55  
56 50 identified by colleagues (11 in total). Summary estimates of mean testosterone concentration  
57  
58  
59  
60

1  
2  
3 51 and age at recruitment for 20,180 adult males were  $15.4 \pm 0.7$  nmol/L and  $64.9 \pm 3.3$  yr. Despite  
4  
5 52 considerable variation in mean testosterone, a meta-regression estimated no significant  
6  
7 53 dependence on mean age at recruitment among studies (Slope = -0.03, 95% CI -0.11 – 0.06).  
8  
9  
10 54 Meta-analyses demonstrated negligible heterogeneity and no significant effect of a 5 nmol/L  
11  
12 55 increase in testosterone on the risk of all-cause mortality (hazard ratio, HR = 0.96, 95% CI  
13  
14 56 0.89 – 1.03) or death from CVD (HR = 0.95, 95% CI 0.83 – 1.08).

## 17 57 Conclusions

18  
19 58 Analyses of published estimates did not demonstrate associations of endogenous testosterone  
20  
21 59 with CVD deaths or with all-cause mortality. Suggested further research includes the planned  
22  
23 60 individual participant data meta-analyses for selected studies, enabling the investigation of  
24  
25 61 non-linear summary effects.

## 28 62 Registration

29  
30 63 PROSPERO: CRD42019139668.  
31  
32 64

## 34 65 **STRENGTHS AND LIMITATIONS OF THIS STUDY**

- 36  
37 66 • This is the first systematic review on this topic to restrict selections to prospective cohort  
38  
39 67 studies of community-dwelling men with testosterone measured using mass spectrometry:  
40  
41 68 the “gold standard” method.  
42  
43  
44 69 • Systematic searches were made of both the published and grey literature using online  
45  
46 70 search tools.  
47  
48  
49 71 • Meta-analyses used estimates obtained from studies with at least five years of follow-up  
50  
51 72 data and from fitted models which controlled for (at least) the age, smoking status, and  
52  
53 73 body mass index or waist circumference of participants.  
54  
55  
56 74 • Meta-analyses of published estimates were limited to assuming linear relationships,  
57  
58 75 however subsequent IPD meta-analyses planned to arise from this work will look to  
59  
60 76 explore non-linear associations.



- 1  
2  
3 77 • Analyses are of observational data, and so summary estimates will not fully eliminate the  
4  
5 78 possibility of confounding arising from unadjusted effects.  
6  
7  
8 79

9  
10 80 **1. INTRODUCTION**

11  
12 81 What does a low testosterone level mean for a man's health? In men, levels of testosterone,  
13  
14 82 the key male sex hormone (androgen), decline with increasing age, yet the basis for and  
15  
16 83 health consequences of this phenomenon remain unclear.[1-5] Endogenous testosterone  
17  
18 84 concentrations reflect the function of the hypothalamic-pituitary-testicular (HPT) axis, and  
19  
20 85 are relatively lower in men who are obese, or with metabolic syndrome or diabetes.[6-8]  
21  
22 86 Others have reported associations of lower endogenous testosterone concentrations with  
23  
24 87 higher risk of incident diseases, such as cardiovascular disease (CVD), and death.[9-16]  
25  
26 88 Whether low testosterone concentrations might contribute directly to the risk of CVD or  
27  
28 89 death or whether it may be associated indirectly through its relationship with aging and  
29  
30 90 obesity is unknown.[17] And whether or not it is directly related, it is possible that  
31  
32 91 endogenous testosterone could be useful as a biomarker for diagnostic and/or prognostic  
33  
34 92 health care applications in men.[18-20] An improved understanding of the associations of  
35  
36 93 testosterone to health outcomes could inform further exploration and development of this  
37  
38 94 concept.  
39  
40  
41  
42  
43  
44  
45  
46

47 96 The Androgens In Men Study (AIMS) seeks to clarify the associations of androgens  
48  
49 97 (primarily testosterone) with key health outcomes in men (mortality, cardiovascular disease,  
50  
51 98 cancer, cognitive decline and dementia) by conducting a systematic review and a series of  
52  
53 99 individual participant data meta-analyses.[21] In this paper we present the systematic review  
54  
55  
56 100 and meta-analyses using published estimates from prospective cohort studies with at least 5  
57  
58  
59  
60

1  
2  
3 101 years of follow-up data and testosterone measured using only mass spectrometry, the most  
4  
5 102 reliable method.[22]  
6  
7  
8 103

## 104 **2. METHODS**

11  
12 105 This systematic review, conducted 14 June—31 December 2019, was of “etiology and/or risk  
13  
14 106 type” studies.[23-24] The pre-specified purpose of the systematic review was to identify  
15  
16 107 studies with suitable individual participant-level data (IPD) for collaborating with on a series  
17  
18 108 of IPD meta-analyses. The PEO (Population, Exposure, Outcomes) characteristics included:  
19  
20 109 adult men in the general community; endogenous circulating sex hormone concentration  
21  
22 110 (primarily testosterone); incident cardiovascular disease (CVD), mortality, cancers, cognitive  
23  
24 111 decline, dementia. Subgroup IPD meta-analyses are also planned for heart failure, myocardial  
25  
26 112 infarction, stroke; colorectal cancer, lung cancer, prostate cancer. A protocol was submitted  
27  
28 113 to PROSPERO on 23 July 2019 and registered on 20 November 2019 (registration number  
29  
30 114 CRD42019139668) and a protocol article has been published.[21]  
31  
32  
33  
34  
35  
36

### 37 116 **2.1. Literature search and screening**

38  
39 117 Four online search tools were used to identify available published (MEDLINE, EMBASE)  
40  
41 118 and grey literature (OpenGrey, Mednar) *items* (journal articles, reports, theses, webpage  
42  
43 119 articles) reporting on suitable prospective cohort *studies* (the underlying unique sources of  
44  
45 120 data). Two reviewers (RJM, JH) independently screened the de-duplicated items against pre-  
46  
47 121 specified criteria using Rayyan.[25] To optimise efficiency, title and abstract screenings were  
48  
49 122 initially conducted (Step 1), followed by full text screenings of the selected abstracts (Step 2).  
50  
51 123 Disagreements were resolved through subsequent discussions between reviewers and  
52  
53 124 agreement quantified using Cohen’s Kappa and percent agreement. Only items reporting on  
54  
55 125 prospective population-based cohort studies of adults (combined sexes or of men alone) with  
56  
57  
58  
59  
60

1  
2  
3 126 mass spectrometry measurements of testosterone and at least five years of subsequent follow-  
4  
5 127 up data on incident CVD events, cancer or dementia diagnoses, cognition assessments, or on  
6  
7  
8 128 all-cause, CVD, or cancer deaths were selected. The Newcastle-Ottawa Quality Assessment  
9  
10 129 Scale for Cohort Studies (NOS) was used to assess quality of the selected items.[26]

11  
12 130

13  
14 131 Additional details on the methods and results are provided in Supplementary Material  
15  
16 132 (Supplementary Section 1). Specifically, additional details on systematic searches and  
17  
18 133 screening (Supplementary Section 2; Supplementary tables 1-4), supplemental tables  
19  
20 134 (Supplementary Section 3), including the PRISMA checklist (Supplementary table 5),  
21  
22 135 supplemental figures (Supplementary Section 4), and references cited (Supplementary  
23  
24 136 Section 5) have been included.  
25  
26  
27  
28  
29

30 137

## 31 138 **2.2. Meta-analyses of published estimates**

32  
33 139 Published estimates (author names, publication year, cohort study name, number of  
34  
35 140 participants analysed, model covariates, testosterone statistics (overall and for individual  
36  
37 141 exposure levels), participant age statistics, numbers of outcome events, follow-up time,  
38  
39 142 hazard ratios (HRs) and 95% CIs of the most fully-adjusted model) were extracted from  
40  
41 143 selected articles by the first author (RJM). For the purpose of these analyses, we present  
42  
43 144 associations for endogenous total testosterone concentrations, comprising the sum of  
44  
45 145 testosterone in the circulation, whether bound to sex hormone-binding globulin or albumin, or  
46  
47 146 unbound. Testosterone statistics were converted into standard units (nmol/L) and values  
48  
49 147 representing categorical ranges were determined following Wang et al.[27] If not reported,  
50  
51 148 the numbers of participants and events within categories of testosterone, and the means of  
52  
53 149 participant ages and testosterone concentrations at baseline, were calculated. The numbers of  
54  
55 150 participants within quartile or quintile categories were calculated by dividing the total sample  
56  
57  
58  
59  
60

1  
2  
3 151 size by four or five. The numbers of events within categories were solved using Newton's  
4  
5 152 method by applying the algorithm of Greenland and Longnecker.[28] Means and standard  
6  
7  
8 153 deviations for testosterone and age were calculated from presented quartile estimates using  
9  
10 154 the Box-Cox method of McGrath et al., which does not make distributional assumptions.[29]

11  
12 155  
13  
14 156 A random effects meta-regression of mean baseline testosterone concentration on the mean  
15  
16 157 participant age at baseline was conducted using published estimates from: (i) only those items  
17  
18 158 identified in systematic searches; and (ii) all suitable articles, including those found outside of  
19  
20 159 systematic searches. A t-test of the meta-regression slope coefficient's departure from zero  
21  
22 160 was done after applying the Knapp and Hartung adjustment.  
23  
24  
25

26 161  
27  
28 162 Dose-response random effects meta-analyses (DR-MAs) were conducted to summarise  
29  
30 163 published HR estimates for the associations of baseline testosterone concentrations with  
31  
32 164 incident all-cause deaths and with CVD cause-specific deaths, as these were the most  
33  
34 165 frequently reported outcomes in selected articles. Estimates from an additional article that  
35  
36 166 had not been selected from systematic searches (Yeap et al[30]) were also used because it  
37  
38 167 reported suitable estimates from one of the selected studies, and had been published within  
39  
40 168 the literature search period. Contour-enhanced funnel plots were inspected for publication  
41  
42 169 bias and patterns in heterogeneity and Cochran Q tests for heterogeneity ( $I^2$ ), as well as  
43  
44 170 regression tests for funnel plot asymmetry,[31] were done. Forest plots were constructed to  
45  
46 171 represent single HR estimates for each study, per 5nmol/L increase in testosterone. For  
47  
48 172 completeness, HR estimates for the other outcomes are represented in a grouped forest plot,  
49  
50 173 and other effect sizes in tables.  
51  
52  
53  
54  
55

56 174  
57  
58 175 The "metafor" package was used for meta-regressions, forest plots and funnel plots, the  
59  
60

1  
2  
3 176 “dosermeta” package for DR-MAs, and the “estmeansd” package for calculating study  
4  
5 177 means and standard deviations from published quartile statistics in R version 4.0.2.[32-35]  
6  
7  
8 178

### 10 179 **2.3. Patient and public involvement**

12 180 This work uses existing published data. Patients and public were not involved in the design,  
13  
14 181 conduct, reporting, or dissemination plans of the systematic review or meta-analyses.  
15  
16  
17 182

## 19 183 **3. RESULTS**

### 24 185 **3.1. Literature search and study selection**

26 186 The literature search returned 2,177 items (1,738 published and 439 from grey literature),  
27  
28 187 with 1,994 items remaining after duplicates had been removed, and after excluding two  
29  
30 188 Mednar items that had insufficient information available to review (Fig. 1). These included  
31  
32 189 1,764 journal articles, 111 webpage articles, 81 theses, and 38 unpublished reports/other  
33  
34 190 documents. Systematic screening of the returned, deduplicated items excluded 1,968,  
35  
36 191 classified five as “Maybe”, and selected 20 as suitable. Most (92.1%) items were excluded  
37  
38 192 from screening titles and abstracts at Step 1, with a much smaller percentage (6.6%) excluded  
39  
40 193 from screening the 157 full text items in Step 2. One item could not be screened in Step 2  
41  
42 194 because the full text was not available. Inter-reviewer agreement was a Cohen’s Kappa  
43  
44 195  $\kappa = 0.69$  (or 96.0 percent agreement) for Step 1 and  $\kappa = 0.82$  (or 98.1 percent agreement) for  
45  
46 196 Step 2.  
47  
48  
49 197

51  
52 198 The 20 selected items collectively reported on eight prospective cohort studies: three from  
53  
54 199 Australia (Busselton Health Study *BHS*,[36-38] The Concord Health and Ageing in Men  
55  
56 200 Project *CHAMP*,[9, 39-40] The Health In Men Study *HIMS*);[14, 41-42] three from Europe  
57  
58  
59  
60

1  
2  
3 201 (European Male Ageing Study *EMAS*,[11, 43] The *MrOS* Osteoporotic Fractures in Men  
4  
5 202 study in Sweden,[10, 44-45] Study of Health in Pomerania *SHIP*);[46] and two from the USA  
6  
7 203 (Atherosclerosis Risk in Communities *ARIC*,[47-48] Cardiovascular Health Study *CHS*).[49-  
8  
9 204 51] Two of the five items classified as “Maybe” reported on the *MrOS USA* study, which  
10  
11 205 were found, after further investigation, to be suitable for selection.[52-53] Two additional  
12  
13 206 studies were identified as suitable based on information external to the systematic searches  
14  
15 207 and screenings: one from Australia (The Men Androgen Inflammation Lifestyle Environment  
16  
17 208 and Stress study *MAILES*);[54] and one from the USA (the Framingham Heart Study  
18  
19 209 *FHS*).[55] This is 11 cohort studies identified, in total. Additional details on returned and  
20  
21 210 screened items, and selected article attributes are provided in Supplementary Material  
22  
23 211 (Supplementary Section 2, Supplementary tables 4, 6-7).  
24  
25  
26  
27  
28  
29  
30

### 31 213 **3.2. Meta-analysis and summary of selected articles.**

32  
33 214 The quality of selected articles ranged from six to nine (out of nine) stars on the Newcastle-  
34  
35 215 Ottawa Scale. Relatively high scores reflected that all articles: were of population-based  
36  
37 216 studies; accurately measured the exposure (baseline testosterone concentration); included  
38  
39 217 multivariable models adjusting for participant age and other risk factors; had outcomes  
40  
41 218 measured or collected from record linkage, with or without expert adjudication; and had  
42  
43 219 sufficient follow-up, ranging from 5-20 years (Supplementary tables 6-8). Relevant outcomes  
44  
45 220 included: all-cause deaths (n=8 articles); CVD deaths (n=7); strokes or cerebrovascular  
46  
47 221 disease (n=6); cognitive function or cognitive decline (n=5); coronary heart disease (n=4);  
48  
49 222 CVD events (n=4); cancer deaths (n=4); cancer diagnoses (n=3); myocardial infarction (MI;  
50  
51 223 n=2); heart failure (HF; n=2); and dementia (n=1). However, one of these articles was a  
52  
53 224 cohort profile description that did not report effect size estimates but the availability of all-  
54  
55 225 cause deaths, cause-specific deaths, stroke, cognitive function, CVD, cancer, MI, and HF  
56  
57  
58  
59  
60

1  
2  
3 226 outcome data.[43]. Two articles reported testosterone not as the exposure but as a covariate in  
4  
5 227 analyses investigating associations with cerebrovascular events and with all-cause, cancer,  
6  
7 228 and CVD deaths.[44-45] The supplementary material for one article[11] was sought to obtain  
8  
9 229 effect size estimates for cancer deaths but these were not obtained as at the time of writing.  
10  
11  
12 230 All were published between 2010 and 2018, reflecting the relatively recent adoption of mass  
13  
14 231 spectrometry as the “gold standard” for measuring endogenous testosterone levels.[22]  
15  
16  
17 232  
18  
19 233 The mean age of men at baseline ranged from middle-aged (49-54yr: BHS, FHS, MAILES,  
20  
21 234 SHIP)[18, 36-38, 46, 56] to elderly (72-77yr: CHAMP, CHS, HIMS, MrOS Sweden, MrOS  
22  
23 235 USA).[9-10, 39, 41, 44-45, 49-50, 53] Across the 11 studies, summary estimates for 20,180  
24  
25 236 adult males at baseline were  $64.9 \pm 3.3$ yr for mean age and  $15.4 \pm 0.7$ nmol/L for mean  
26  
27 237 testosterone. Although there appeared to be a slight declining trend in mean testosterone with  
28  
29 238 mean age among studies (Meta-regression Slope=  $-0.07$ , 95% CI  $-0.21 - 0.07$ ), this estimate  
30  
31 239 was not significantly different from zero ( $P=0.27$ ; Fig. 2a). However, the distribution of  
32  
33 240 model residuals demonstrated significant heterogeneity ( $P<0.001$ ) and funnel plot asymmetry  
34  
35 241 ( $P=0.02$ ). Additional diagnostics highlighted a relatively high mean testosterone estimate  
36  
37 242 from Pencina et al.[57] (FHS) and a low mean testosterone estimate (relative to mean age)  
38  
39 243 from Chan et al.[37] (BHS), as compared to the other studies (Supplementary figure 1).  
40  
41  
42 244 When restricted to systematically selected items (reporting on ARIC, BHS, CHAMP, CHS,  
43  
44 245 EMAS, HIMS, MAILES, MrOS Sweden, SHIP studies), tests of residual heterogeneity were  
45  
46 246 significant ( $P<0.001$ ), funnel plot asymmetry ( $P=0.91$ ) was non-significant, and the slope  
47  
48 247 estimate (Meta-regression Slope=  $-0.03$ , 95% CI  $-0.11 - 0.06$ ) was not significantly different  
49  
50 248 from zero ( $P=0.50$ ; Fig. 2b). These results demonstrate that varying distributions of  
51  
52 249 participant age (likely reflecting differences in study-specific objectives and recruitment  
53  
54 250 methods) did not explain the observed heterogeneity in published estimates of testosterone  
55  
56  
57  
58  
59  
60



1  
2  
3 251 among the studies.  
4

5 252

7 253 Hazard ratios (HRs) for all-cause mortality were calculated from values in four of the

9 254 selected articles (ARIC,[48] BHS,[37] CHS,[51] EMAS[11]) and from one that was not

11 255 selected, but had reported on the HIMS study during the literature search period.[30] All HRs

13 256 were adjusted for the age, smoking status, and body mass index (BMI) or waist

15 257 circumference of participants. A DR-MA estimated a summary HR of 0.96 (95% CI 0.89-

17 258 1.03) per 5nmol/L increase in testosterone (Fig. 3). The summary estimate was similar when

19 259 calculated using an alternative estimate from Yeap et al[30] (HR=0.97, 95% CI 0.92-1.03).

21 260 For both analyses, tests for residual heterogeneity ( $I^2=23.6%$ ,  $P=0.26$ ;  $I^2=0.0%$ ,  $P=0.76$ ) and

23 261 funnel plot asymmetry ( $P=0.09$ ;  $P=0.39$ ) were non-significant (Supplementary figure 2a). A

25 262 comparable HR was calculated from a CHAMP study article[9] for inclusion in the forest plot

27 263 but not in the DR-MA, because a corresponding estimate of variance per 5nmol/L increase in

29 264 testosterone could not be calculated. An additional funnel plot, which included the HR

31 265 estimate from this CHAMP article[9] (per 1 standard deviation decrease in testosterone, as

33 266 reported in that article), also demonstrated no significant asymmetry (Supplementary figure

35 267 2b). These results demonstrate no overall effect of baseline testosterone concentration on the

37 268 relative hazard of death from any cause after adjusting for factors including age, smoking

39 269 status, and BMI or waist circumference.

41 270

43 271 HRs for death caused by CVD demonstrated similar findings. A DR-MA using estimates

45 272 from the same five articles estimated a summary HR of 0.95 (95% CI 0.83-1.08) per 5nmol/L

47 273 increase in testosterone, with no significant residual heterogeneity ( $I^2=28.3%$ ,  $P=0.23$ ) or

49 274 funnel plot asymmetry ( $P=0.20$ ; Fig. 4; Supplementary figure 2c). Again, all HRs were

51 275 adjusted for the age, smoking status, and BMI or waist circumference. The DR-MA repeated

53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 276 using an alternative estimate from Chasland et al.[38] for the BHS gave similar results  
4  
5 277 (summary HR=0.93, 95% CI 0.83-1.03; heterogeneity  $I^2=17.5\%$ ,  $P=0.30$ ; funnel plot  
6  
7 278 asymmetry  $P=0.17$ ; Supplementary figure 2d). These results demonstrate no overall effect of  
8  
9 279 baseline testosterone concentration on the relative hazard of death from CVD after adjusting  
10  
11 280 for factors including age, smoking status, and BMI or waist circumference.  
12  
13  
14

15 281  
16  
17 282 Summary estimates calculated for the combined outcome of incident stroke and  
18  
19 283 cerebrovascular disease (summary HR=0.93, 95% CI 0.83-1.03; heterogeneity  $I^2=43.3\%$ ,  
20  
21 284  $P=0.15$ ) and incident CVD diagnosis (summary HR=0.93, 95% CI 0.84-1.03; heterogeneity  
22  
23 285  $I^2=34.7\%$ ,  $P=0.22$ ) demonstrated no overall effect of testosterone (Supplementary figure 3).  
24  
25 286 Funnel plot asymmetry was not assessed due to the low number of studies ( $n \leq 4$ ),[58] and  
26  
27 287 95% confidence intervals could not be calculated for several studies[36-37, 41] using the  
28  
29 288 published information. Although a summary estimate could not be calculated, the study-  
30  
31 289 specific estimates demonstrated some significant associations with cancer outcomes  
32  
33 290 (Supplementary figure 3, Supplementary table 9). Estimates showed an increased risk of lung  
34  
35 291 cancer for men with higher concentrations,[41] an increased risk of death from cancer for  
36  
37 292 men with lower[9] or the lowest ( $<8\text{nmol/L}$ )[11] concentrations, and an increased risk of  
38  
39 293 diagnosis for any cancer or for prostate cancer for men with the lowest ( $<10.17\text{nmol/L}$ )  
40  
41 294 concentrations of testosterone.[36] However, results were varied and not all articles reported  
42  
43 295 these associations as being significant.[39] Furthermore, aside from an average increase in  
44  
45 296 MMSE of 0.067 per ng/mL decrease in testosterone concentration during follow-up,[40]  
46  
47 297 there were no significant associations of baseline testosterone with cognitive function, or with  
48  
49 298 change in cognitive function reported in the selected articles (Supplementary table 10).  
50  
51  
52  
53  
54  
55

56 299

#### 57 300 **4. DISCUSSION**

1  
2  
3 301 The systematic review identified nine studies, and when combined with an additional two  
4  
5 302 identified by colleagues, comprises 11 in total, with data for over 20,000 men from Australia,  
6  
7 303 Europe, USA and the United Kingdom. Meta-regressions revealed significant heterogeneity  
8  
9 304 in testosterone measurements at baseline, which was not explained by the mean age of  
10  
11 305 participants among studies. However, DR-MA summary estimates demonstrated no  
12  
13 306 significant effects of baseline testosterone on the relative hazard of death from any cause or  
14  
15 307 from CVD, with negligible heterogeneity present. The DR-MAs, which suitably accounted  
16  
17 308 for correlations between estimates for different exposure categories within studies, were of  
18  
19 309 published estimates that had been adjusted for age, smoking status, and BMI or waist  
20  
21 310 circumference. Furthermore, only published estimates from prospective cohort studies of  
22  
23 311 community-dwelling men that had measured testosterone accurately using mass spectrometry  
24  
25 312 and had observed at least five years of follow-up data were used. Despite some of these  
26  
27 313 studies having reported an association between testosterone and mortality,[9, 30] the  
28  
29 314 collective body of evidence demonstrated no overall associations of endogenous testosterone  
30  
31 315 concentration with mortality or CVD mortality.  
32  
33  
34  
35  
36  
37  
38  
39

40 317 Previous meta-analyses investigating associations of endogenous testosterone with the health  
41  
42 318 outcomes of interest looked at CVD outcomes,[59-61] all-cause mortality,[59] and prostate  
43  
44 319 cancer.[62] Boyle et al.[62] and Holmegard et al.[60] both reported negligible heterogeneity  
45  
46 320 in their estimates. Boyle et al. found no significant association of a 5nmol/L increase in  
47  
48 321 testosterone with prostate cancer and Holmegard et al. estimated a 43% increase in risk of  
49  
50 322 ischemic stroke for men with testosterone levels below the 10<sup>th</sup> percentile, as compared to  
51  
52 323 men in the 11<sup>th</sup>-90<sup>th</sup> percentile range, from a meta-analysis of four articles.[60, 62] Ruige et  
53  
54 324 al. estimated an 11% decrease in risk of a CVD event from a standard deviation increase in  
55  
56 325 testosterone, and reported that significant heterogeneity was explained by larger effect sizes  
57  
58  
59  
60

1  
2  
3 326 estimated for studies that recruited older men and for more recent articles.[61] Araujo et al  
4  
5 327 estimated a 35% increase in risk of all-cause mortality and a non-significant effect on CVD  
6  
7 328 mortality from a 2.18 standard deviation decrease in testosterone, although reported  
8  
9  
10 329 significant heterogeneity, and suggested that effects were driven by differences between the  
11  
12 330 cohorts, such as underlying health status.[59] Two of these meta-analyses did not restrict  
13  
14 331 selections to prospective cohort studies[59, 62] and none restricted selections based on  
15  
16 332 testosterone assay method, although Ruige et al.[61] did find that assay method did not  
17  
18 333 explain heterogeneity in that study.  
19  
20  
21  
22  
23  
24

25 335 The presented meta-analyses are the first to restrict selections to items of prospective cohort  
26  
27 336 studies of community-dwelling men with testosterone measured using mass spectrometry,  
28  
29 337 which is widely regarded as the reference method,[22] and with at least five years of follow-  
30  
31 338 up data. Accordingly, the presented summary estimates could arguably be viewed as the most  
32  
33 339 reliable to date. These restrictions also resulted in the selection of a relatively small number  
34  
35 340 of publications with estimates suitable for use in DR-MAs. Follow-up times for all-cause and  
36  
37 341 CVD mortality ranged from a median of 4.3 years (total = 5 years; EMAS)[11] to a mean of  
38  
39 342 14.9 years (total = 16 years; BHS).[37] The number of incident deaths ranged from 147  
40  
41 343 (EMAS)[40] to 777 (CHS),[51] or to 974 with the additional HIMS article[30] included. The  
42  
43 344 number of CVD deaths ranged from 29 (ARIC)[48] to 264 (CHS)[51], or to 325 with the  
44  
45 345 additional HIMS article.[30] However, despite these differences, there was negligible  
46  
47 346 heterogeneity in estimates and no significant funnel plot asymmetry detected.  
48  
49  
50  
51  
52  
53  
54

55  
56 348 Linear models were fitted because the HR estimates were reported for insufficient numbers  
57  
58 349 of testosterone categories to have fitted non-linear DR-MA models. This was a key limitation  
59  
60

1  
2  
3 350 of the analyses and likely to have resulted in an oversimplification of true effects. For  
4  
5 351 instance, although the 95% CI for the Pye et al[11] study (calculated from HR estimates for  
6  
7 352 quintile categories of testosterone) overlapped one, an alternative set of estimates in that  
8  
9  
10 353 article (which could not be included in the DR-MAs) reported a two-fold increase in the risk  
11  
12 354 of all-cause mortality for men with very low testosterone (<8nmol/L), as compared to  
13  
14 355 “eugonadal” men (>11nmol/L). Pye et al[11] postulated that their reported differences in  
15  
16 356 estimates might be reflective of a nonlinear association that emerges only when endogenous  
17  
18 357 testosterone declines into the lower part of the range (<8nmol/L). Furthermore, Yeap et  
19  
20 358 al.[30] estimated an “U”-shaped association between endogenous testosterone and all-cause  
21  
22 359 mortality, as consistent with a lower relative risk of health impacts for adult males with mid-  
23  
24 360 range levels of testosterone. However, Shores et al.[51] also used non-linear modelling but  
25  
26 361 did not find any significant associations of testosterone with all-cause or CVD mortality.  
27  
28 362 Clearly, the investigation of non-linear associations is required to more comprehensively  
29  
30 363 investigate the associations of testosterone concentrations with health outcomes in men.  
31  
32  
33  
34  
35  
36 364

37  
38  
39 365 In addition to the linearity assumption, there were other methodological limitations. Several  
40  
41 366 articles reported estimated HRs per increase or decrease in standard deviation (SD) and it was  
42  
43 367 not possible to use these estimates in DR-MAs. Although it was possible to convert the per  
44  
45 368 SD estimates to a standardised scale (i.e., per 5nmol/L increase), there was no information to  
46  
47 369 determine adjustments to respective estimates of precision. Estimates for those studies could  
48  
49 370 therefore not be included in the calculation of summary estimates and 95% confidence  
50  
51 371 intervals could not be calculated in forest plots. Summary estimates were calculated from a  
52  
53 372 relatively low number (n = 3-5) of articles and for most outcomes a summary estimate could  
54  
55 373 not be calculated, which impacts upon the generalisability of findings. Furthermore, these  
56  
57 374 analyses were of observational data so summary estimates will not fully eliminate the  
58  
59  
60

1  
2  
3 375 possibility of confounding arising from unadjusted effects.  
4  
5

6 376  
7  
8

9 377 The implications of these findings are that associations of endogenous testosterone  
10  
11 378 concentrations with key health outcomes should not be overstated, as they are not readily  
12  
13 379 portrayed by meta-analyses of summary estimates. A more nuanced approach may be  
14  
15  
16 380 required, to capture non-linear or U-shaped associations.[11, 30] Also, while testosterone  
17  
18 381 concentrations across ages were relatively stable when considering estimates from different  
19  
20 382 cohorts, associations of testosterone with health outcomes may differ with age, for example  
21  
22 383 with all-cause mortality in middle-aged men[37] compared to older men.[9, 30] A deeper  
23  
24 384 understanding of associations of endogenous testosterone concentrations with key health  
25  
26 385 outcomes, would provide a foundation for analyses of the effects of exogenous testosterone,  
27  
28 386 administered via therapeutic or pharmacologic interventions, on men's health.  
29

30  
31  
32 387  
33

34 388 Individual participant data (IPD) meta-analyses that incorporate flexible non-linear modelling  
35  
36 389 techniques will provide improved scope to clarify the nature of such associations. The ability  
37  
38  
39 390 to apply a consistent statistical model to all studies, incorporate a more extended set of  
40  
41 391 covariates than may have been included at the individual study level, and to estimate effects  
42  
43 392 with increased statistical power, should result in more reliable summary estimates with  
44  
45 393 reduced bias. Furthermore, other hitherto unpublished variables may be available for sharing  
46  
47  
48 394 by the collaborating studies to use in IPD meta-analyses, which could be useful for  
49  
50 395 constructing analysis covariates or outcome variables. For instance, articles from the ARIC  
51  
52 396 study that were identified from the systematic review reported on incident CVD event and  
53  
54 397 death outcomes, but documentation on the ARIC study website shows that data on other  
55  
56 398 prospective health outcomes, including cause-specific deaths and dementia diagnoses, are  
57  
58 399 also available upon request.[63] Although there have been recent advances with non-linear

1  
2  
3 400 modelling methods for the meta-analyses of published estimates,[32, 64] sufficient  
4  
5 401 information in the published articles, as is required for implementing these methods, was not  
6  
7 402 available. In future work, estimates from analyses of the IPD-level data will be used to  
8  
9 403 estimate and plot non-linear summary effects, and so will provide further improvements to  
10  
11 404 estimates of associations between androgen levels and health outcomes in men.  
12  
13  
14  
15 405

## 16 406 **ACKNOWLEDGEMENTS**

17  
18  
19 407 We thank Terena Solomons for valuable advice and guidance with conducting the literature  
20  
21 408 search and screening steps of the systematic review.  
22  
23  
24  
25 409

## 26 27 410 **AUTHORS' CONTRIBUTIONS**

28  
29 411 BBY, KM, RJM, JH contributed to the design of the systematic review. RJM conducted the  
30  
31 412 literature search and RJM, JH independently screened the returned items. RJM, KM  
32  
33 413 conducted the statistical analyses. All authors were involved in manuscript preparation and  
34  
35 414 subsequent revisions, and approved this submission.  
36  
37  
38  
39 415

## 40 41 416 **COMPETING INTERESTS**

42  
43 417 None declared.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 418  
4

5 419 **FUNDING**  
6

7  
8 420 This work was supported by: (i) Western Australian Health Translation Network Medical  
9  
10 421 Research Future Fund Rapid Applied Translation Grant (2018), Grant number N/A; (ii)  
11  
12 422 Lawley Pharmaceuticals, Western Australia, Grant number N/A.  
13

14  
15 423  
16

17 424 **DATA SHARING STATEMENT**  
18

19 425 No additional data available.  
20

21  
22 426  
23

24 427 **PATIENT CONSENT**  
25

26 428 This manuscript does not contain patient personal data.  
27

28  
29 429  
30

31 430 **PROVENANCE AND PEER REVIEW**  
32

33 431 Not commissioned; externally peer reviewed.  
34

35  
36 432  
37

38 433 **ETHICS STATEMENT**  
39

40 434 Not applicable. No human participants or animal subjects included.  
41

42  
43 435  
44

45 436  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



437 **REFERENCES CITED**

- 438  
439 1. Ahern T, Swiecicka A, Eendebak RJA, *et al.* Natural history, risk factors and clinical features of  
440 primary hypogonadism in ageing men: Longitudinal Data from the European Male Ageing Study.  
441 *Clin Endocrinol* 2016;85:891-901.
- 442 2. Feldman HA, Longcope C, Derby CA, *et al.* Age Trends in the Level of Serum Testosterone and  
443 Other Hormones in Middle-Aged Men: Longitudinal Results from the Massachusetts Male Aging  
444 Study. *J Clin Endocrinol Metab* 2002;87:589-98.
- 445 3. Handelsman DJ, Yeap BB, Flicker L, *et al.* Age-specific population centiles for androgen status in  
446 men. *Eur J Endocrinol* 2015;173:809-17.
- 447 4. Hsu B, Cumming RG, Hirani V, *et al.* Temporal Trend in Androgen Status and Androgen-Sensitive  
448 Outcomes in Older Men. *J Clin Endocrinol Metab* 2016;101:1836-46.
- 449 5. Yeap BB, Manning L, Chubb SAP, *et al.* Progressive impairment of testicular endocrine function in  
450 ageing men: Testosterone and dihydrotestosterone decrease, and luteinizing hormone increases,  
451 in men transitioning from the 8th to 9th decades of life. *Clin Endocrinol* 2018;88:88-95.
- 452 6. Brand JS, Rovers MM, Yeap BB, *et al.* Testosterone, sex hormone-binding globulin and the  
453 metabolic syndrome in men: An individual participant data meta-analysis of observational  
454 studies. *PLoS One* 2014;9:e100409.
- 455 7. Ding EL, Song Y, Malik VS, *et al.* Sex differences of endogenous sex hormones and risk of type 2  
456 diabetes: a systematic review and meta-analysis. *JAMA* 2006;295:1288-99.
- 457 8. Yeap BB, Marriott RJ, Antonio L, *et al.* Sociodemographic, lifestyle and medical influences on serum  
458 testosterone and sex hormone-binding globulin in men from UK Biobank. *Clin Endocrinol*  
459 2021;94:290-302.
- 460 9. Hsu B, Cumming RG, Naganathan V, *et al.* Temporal Changes in Androgens and Estrogens Are  
461 Associated With All-Cause and Cause-Specific Mortality in Older Men. *J Clin Endocrinol Metab*  
462 2016;101:2201-10.
- 463 10. Ohlsson C, Barrett-Connor E, Bhasin S, *et al.* High serum testosterone is associated with reduced  
464 risk of cardiovascular events in elderly men. The MrOS (Osteoporotic Fractures in Men) study in  
465 Sweden. *J Am Coll Cardiol* 2011;58:1674-81.
- 466 11. Pye SR, Huhtaniemi IT, Finn JD, *et al.* Late-onset hypogonadism and mortality in aging men. *J Clin*  
467 *Endocrinol Metab* 2014;99:1357-66.
- 468 12. Tivesten Å, Mellström D, Jutberger H, *et al.* Low serum testosterone and high serum estradiol  
469 associate with lower extremity peripheral arterial disease in elderly men. *J Am Coll Cardiol*  
470 2007;50:1070-6.
- 471 13. Tivesten Å, Vandenput L, Labrie F, *et al.* Low serum testosterone and estradiol predict mortality  
472 in elderly men. *J Clin Endocrinol Metab* 2009;94:2482-8.
- 473 14. Yeap BB, Alfonso H, Chubb SA, *et al.* In older men, higher plasma testosterone or  
474 dihydrotestosterone is an independent predictor for reduced incidence of stroke but not  
475 myocardial infarction. *J Clin Endocrinol Metab* 2014;99:4565-73.
- 476 15. Yeap BB, Hyde Zoë, Almeida OP, *et al.* Lower testosterone levels predict incident stroke and  
477 transient ischemic attack in older men. *J Clin Endocrinol Metab* 2009;94:2353-9.
- 478 16. Yeap BB, Knuiman MW, Divitini ML, *et al.* Differential associations of testosterone,  
479 dihydrotestosterone and oestradiol with physical, metabolic and health-related factors in  
480 community-dwelling men aged 17-97 years from the Busselton Health Survey. *Clin Endocrinol*  
481 2014;81:100-8.
- 482 17. Yeap BB, Araujo AB, Wittert GA. Do low testosterone levels contribute to ill-health during male  
483 ageing? *Crit Rev Clin Lab Sci* 2012;49:168-82.
- 484 18. Bhasin S, Pencina MJ, Kaur Jasuja G, *et al.* Reference ranges for testosterone in men generated  
485 using liquid chromatography tandem mass spectrometry in a community-base sample of healthy



- 1  
2  
3 486 nonobese young men in the Framingham Heart Study and applied to three geographically distinct  
4 487 cohorts. *J Clin Endocrinol Metab* 2011;96:2430-9.
- 5 488 19. Sikaris K, McLachlan RI, Kazlauskas R, *et al.* Reproductive hormone reference intervals for healthy  
6 489 fertile young men: evaluation of automated platform assays. *J Clin Endocrinol Metab*  
7 490 2005;90:5928-36.
- 8 491 20. Yeap BB, Alfonso H, Chubb SAP, *et al.* Reference ranges and determinants of testosterone,  
9 492 dihydrotestosterone, and estradiol levels measured using liquid chromatography-tandem mass  
10 493 spectrometry in a population-based cohort of older men. *J Clin Endocrinol Metab* 2012;97:4030-9.
- 11 494 21. Yeap BB, Marriott RJ, Adams RJ, *et al.* Androgens In Men Study (AIMS): protocol for meta-  
12 495 analyses of individual participant data investigating associations of androgens with health  
13 496 outcomes in men. *BMJ Open* 2020;10:e034777.
- 14 497 22. Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the  
15 498 Journal of Clinical Endocrinology and Metabolism. *J Clin Endocrinol Metab* 2013;98:3971-73.
- 16 499 23. Munn Z, Stern C, Aromataris E, *et al.* What kind of systematic review should I conduct? A  
17 500 proposed typology and guidance for systematic reviewers in the medical and health sciences.  
18 501 *BMC Med Res Methodol* 2018;18:5.
- 19 502 24. Stroup DF, Berlin JA, Morton SC, *et al.* Meta-analysis of observational studies in epidemiology: A  
20 503 proposal for reporting. *JAMA* 2000;283:2008-12.
- 21 504 25. Ouzzani M, Hammady H, Fedorowicz Z, *et al.* Rayyan - a web and mobile app for systematic  
22 505 reviews. *Syst Rev* 2016;5:210.
- 23 506 26. Wells GA, Shea B, O'Connell DO, *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the  
24 507 quality of nonrandomised studies in meta-analyses 2020.  
25 508 [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp) (accessed 10 Nov 2020).
- 26 509 27. Wang X, Ouyang Y, Liu J, *et al.* Fruit and vegetable consumption and mortality from all causes,  
27 510 cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of  
28 511 prospective cohort studies. *BMJ* 2014;349:g4490.
- 29 512 28. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response  
30 513 data, with applications to meta-analysis. *Am J Epidemiol* 1992;135:1301-9.
- 31 514 29. McGrath S, Zhao X, Steele R, *et al.* Estimating the sample mean and standard deviation from  
32 515 commonly reported quantiles in meta-analysis. *Stat Methods Med Res* 2020;29:2520-37.
- 33 516 30. Yeap BB, Alfonso H, Chubb SAP, *et al.* In older men an optimal plasma testosterone is associated  
34 517 with reduced all-cause mortality and higher dihydrotestosterone with reduced ischemic heart  
35 518 disease mortality, while estradiol levels do not predict mortality. *J Clin Endocrinol Metab*  
36 519 2014;99:E9-18.
- 37 520 31. Sterne JAC, Egger M. Regression methods to detect publication and other bias in meta-analysis.  
38 521 In: Rothstein HR, Sutton AJ, Borenstein M, eds. *Publication bias in meta-analysis: Prevention,*  
39 522 *assessment, and adjustments.* Chichester, England: Wiley 2005:99-110.
- 40 523 32. Crippa A, Orsini N. Multivariate dose-response meta-analysis: The dosresmeta R package. *J Stat*  
41 524 *Softw* 2016;72:1.
- 42 525 33. estmeansd: Estimating the Sample Mean and Standard Deviation from Commonly Reported  
43 526 Quantiles in Meta-Analysis. R package version 0.2.1, 2020.
- 44 527 34. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for  
45 528 Statistical Computing, 2020. <https://www.R-project.org/>
- 46 529 35. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:3.
- 47 530 36. Chan YX, Knuiman MW, Divitini ML, *et al.* Lower Circulating Androgens Are Associated with  
48 531 Overall Cancer Risk and Prostate Cancer Risk in Men Aged 25-84 Years from the Busselton Health  
49 532 Study. *Horm Cancer* 2018;9:391-8.
- 50 533 37. Chan YX, Knuiman MW, Hung J, *et al.* Neutral associations of testosterone, dihydrotestosterone  
51 534 and estradiol with fatal and non-fatal cardiovascular events, and mortality in men aged 17-97  
52 535 years. *Clin Endocrinol* 2016;85:575-82.
- 53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 536 38. Chasland LC, Knuiiman MW, Divitini ML, *et al.* Greater physical activity and higher androgen  
4 537 concentrations are independently associated with lower cardiometabolic risk in men. *Clin*  
5 538 *Endocrinol* 2017;87:466-74.  
6 539 39. Hsu B, Cumming RG, Blyth FM, *et al.* Evaluating Calculated Free Testosterone as a Predictor of  
7 540 Morbidity and Mortality Independent of Testosterone for Cross-sectional and 5-Year Longitudinal  
8 541 Health Outcomes in Older Men: The Concord Health and Ageing in Men Project. *J Gerontol A Biol*  
9 542 *Sci Med Sci* 2018;73:729-36.  
10 543 40. Hsu B, Cumming RG, Waite LM, *et al.* Longitudinal relationships between reproductive hormones  
11 544 and cognitive decline in older men: The Concord Health and Ageing in Men Project. *J Clin*  
12 545 *Endocrinol Metab* 2015;100:2223-30.  
13 546 41. Chan YX, Alfonso H, Chubb SA, *et al.* Higher Dihydrotestosterone Is Associated with the Incidence  
14 547 of Lung Cancer in Older Men. *Horm Cancer* 2017;8:119-26.  
15 548 42. Ford AH, Yeap BB, Flicker L, *et al.* Sex hormones and incident dementia in older men: The health  
16 549 in men study. *Psychoneuroendocrinology* 2018;98:139-47.  
17 550 43. Lee DM, Pye SR, Tajar A, *et al.* Cohort profile: the European Male Ageing Study. *Int J Epidemiol*  
18 551 2013;42:391-401.  
19 552 44. Ohlsson C, Labrie F, Barrett-Connor E, *et al.* Low serum levels of dehydroepiandrosterone sulfate  
20 553 predict all-cause and cardiovascular mortality in elderly Swedish men. *J Clin Endocrinol Metab*  
21 554 2010;95:4406-14.  
22 555 45. Tivesten Å, Vandenput L, Carlzon D, *et al.* Dehydroepiandrosterone and its sulfate predict the 5-  
23 556 year risk of coronary heart disease events in elderly men. *J Am Coll Cardiol* 2014;64:1801-10.  
24 557 46. Kische H, Gross S, Wallaschofski H, *et al.* Associations of androgens with depressive symptoms  
25 558 and cognitive status in the general population. *PLoS One* 2017;12:e0177272.  
26 559 47. Srinath R, Gottesman RF, Hill Golden S, *et al.* Association Between Endogenous Testosterone and  
27 560 Cerebrovascular Disease in the ARIC Study (Atherosclerosis Risk in Communities). *Stroke*  
28 561 2016;47:2682-8.  
29 562 48. Srinath R, Hill Golden S, Carson KA, *et al.* Endogenous Testosterone and its Relationship to  
30 563 Preclinical and Clinical Measures of Cardiovascular Disease in the Atherosclerosis Risk in  
31 564 Communities Study. *J Clin Endocrinol Metab* 2015;100:1602-8.  
32 565 49. Rosenberg MA, Shores MM, Matsumoto AM, *et al.* Serum androgens and risk of atrial fibrillation  
33 566 in older men: The Cardiovascular Health Study. *Clin Cardiol* 2018;41:830-6.  
34 567 50. Shores MM, Arnold AM, Biggs ML, *et al.* Testosterone and dihydrotestosterone and incident  
35 568 ischaemic stroke in men in the Cardiovascular Health Study. *Clin Endocrinol* 2014;81:746-53.  
36 569 51. Shores MM, Biggs ML, Arnold AM, *et al.* Testosterone, dihydrotestosterone, and incident  
37 570 cardiovascular disease and mortality in the cardiovascular health study. *J Clin Endocrinol Metab*  
38 571 2014;99:2061-8.  
39 572 52. LeBlanc ES, Wang PY, Janowsky JS, *et al.* Association between sex steroids and cognition in  
40 573 elderly men. *Clin Endocrinol* 2010;72:393-403.  
41 574 53. Sueoka KT, Ewing MS, Ensrud KE, *et al.* Higher endogenous testosterone levels associated with  
42 575 increased risk of coronary heart disease in elderly men: a prospective study. *Endocr Rev*  
43 576 2010;31:S858.  
44 577 54. Grant JF, Martin SA, Taylor AW, *et al.* Cohort profile: The men androgen inflammation lifestyle  
45 578 environment and stress (MAILES) study. *Int J Epidemiol* 2014;43:1040-53.  
46 579 55. Kannel WB, Feinleib M, McNamara PM, *et al.* An investigation of coronary heart disease in  
47 580 families. The Framingham offspring study. *Am J Epidemiol* 1979;110:281-90.  
48 581 56. Li JJ, Wittert GA, Vincent A, *et al.* Muscle grip strength predicts incident type 2 diabetes:  
49 582 population-based cohort study. *Metab Clin Exp* 2016;65:883-92.  
50 583 57. Pencina KM, Travison TG, Bhasin S, *et al.* Endogenous circulating testosterone and sex hormone-  
51 584 binding globulin levels and measures of myocardial structure and function: the Framingham  
52 585 Heart Study. *Andrology* 2019;7:307-14.

- 1  
2  
3 586 58. Sterne JAC, Sutton AJ, Ioannidis JPA, *et al.* Recommendations for examining and interpreting  
4 587 funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;342:d4002.  
5 588 59. Araujo AB, Dixon JM, Suarez EA, *et al.* Endogenous testosterone and mortality in men: a  
6 589 systematic review and meta-analysis. *J Clin Endocrinol Metab* 2011;96:3007-19.  
7 590 60. Holmegard HN, Nordestgaard BG, Jensen GB, *et al.* Sex hormones and ischemic stroke: a  
8 591 prospective cohort study and meta-analyses. *J Clin Endocrinol Metab* 2016;101:69-78.  
9 592 61. Ruige JB, Mahmoud AM, De Bacquer D, *et al.* Endogenous testosterone and cardiovascular  
10 593 disease in healthy men: a meta-analysis. *Heart* 2011;97:870-5.  
11 594 62. Boyle P, Koechlin A, Bota M, *et al.* Endogenous and exogenous testosterone and the risk of  
12 595 prostate cancer and increased prostate-specific antigen (PSA) level: a meta-analysis. *BJU Int*  
13 596 2016;118:731-41.  
14 597 63. Surveillance Dictionaries | Atherosclerosis Risk In Communities 2020.  
15 598 <https://sites.csc.unc.edu/atic/surveillance-dictionaries> (accessed 6 July 2020).  
16 599 64. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized  
17 600 dose-response data. *Stata J* 2006;6:40-57.  
18 601  
19  
20  
21

## 602 **FIGURE LEGENDS**

603  
604 Figure 1. Studies returned from systematic review of the published and grey literature. Step 1  
605 involved screening of titles and abstracts only and Step 2 the screening of full text items not  
606 excluded at Step 1 (see Supplementary tables 2, 3). “Items” are individual articles or reports,  
607 with multiple items returned for some studies (the purpose was to identify studies with  
608 suitable IPD-level data). \* = Mednar items with insufficient information available to review;  
609 \*\* = Additional studies identified through known contacts; \*\*\* = Screening criteria for five  
610 items selected as “Maybe” in Step 2 were further investigated using information external to  
611 systematic searches and screenings, resulting in the identification of one additional study with  
612 suitable IPD-level data.

613  
614 Figure 2. Meta-regression of mean testosterone on mean age for (a) all 11 cohort studies and  
615 (b) 9 studies with articles that were selected by systematic literature searches and screening.  
616 The size of plotted points refers are proportional to the inverse of the corresponding standard  
617 errors (indicative of relative weightings), with lines demonstrating the fitted model and 95%  
618 CIs. Plotted estimates are numbered as from the following articles (cohort studies):

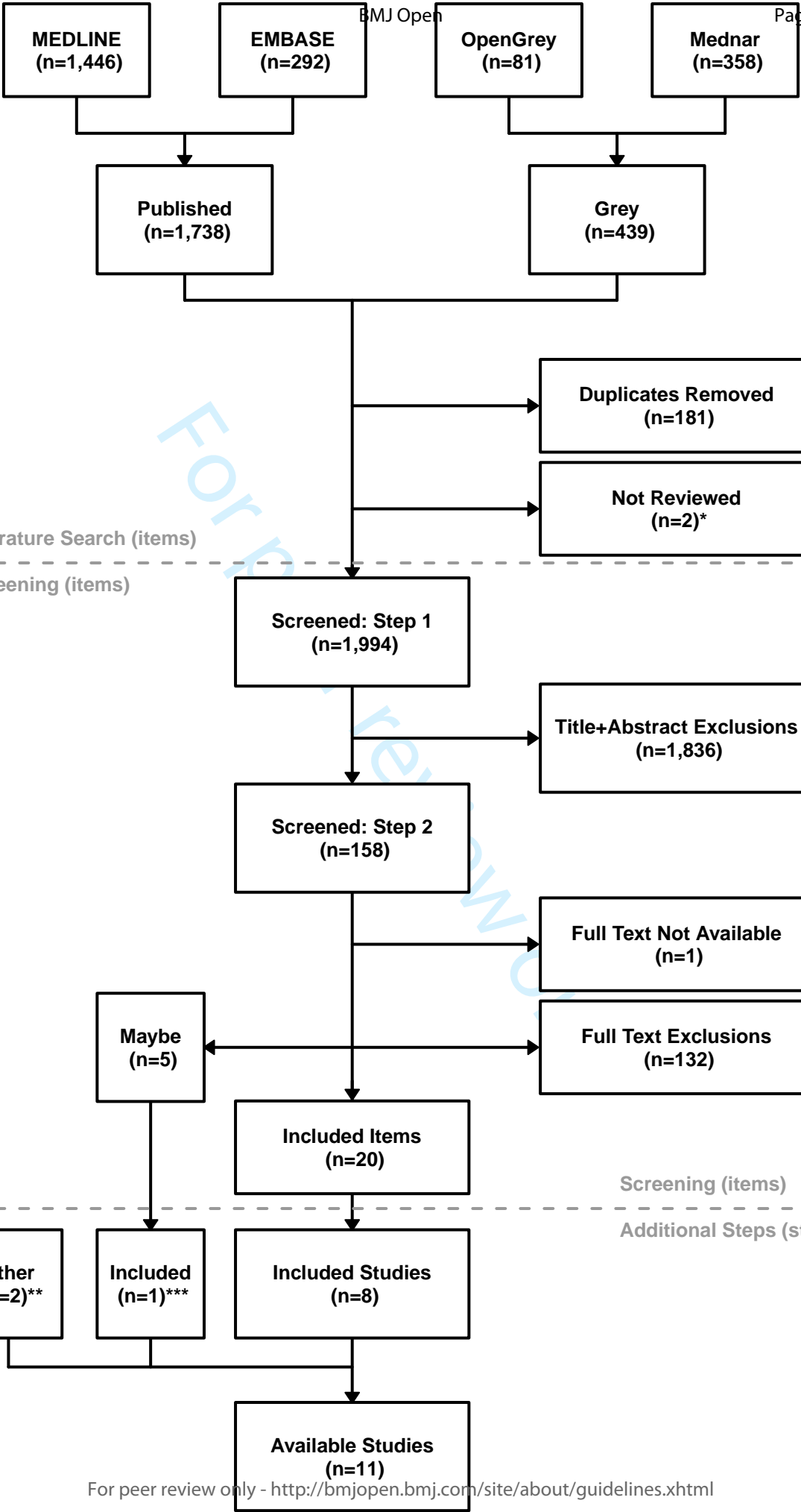
1  
2  
3 619 1= Srinath et al.[48] (ARIC); 2= Chan et al.[37] (BHS); 3= Hsu et al.[9] (CHAMP); 4=  
4  
5 620 Shores et al.[50] (CHS); 5= Lee et al.[43] (EMAS); 6= Chan et al.[41] (HIMS); 7= Ohlsson  
6  
7 621 et al.[44] (MrOS Sweden); 8= Kische et al.[46] (SHIP); 9= Sueoka et al.[53] (MrOS USA);  
8  
9 622 10= Pencina et al.[57] (FHS); 11= Li et al.[56] (MAILES). \* = includes articles from two  
10  
11 623 additional studies (FHS, MAILES) that were not identified from systematic searches but by  
12  
13 624 colleagues.  
14  
15  
16  
17  
18

19 626 Figure 3. Forest plot of a meta-analysis of published estimates: association of testosterone  
20  
21 627 with all-cause mortality. Plotted values are the estimated hazard ratios (HR) for death from  
22  
23 628 any cause, as attributed to an increase in endogenous testosterone concentration by 5 nmol/L.  
24  
25 629 The vertical reference line is HR=1. Study-specific estimates are presented for six of the  
26  
27 630 selected studies: BHS (Chan, 2016)[37]; EMAS (Pye, 2014)[11]; ARIC (Srinath, 2015)[48];  
28  
29 631 CHS (Shores, 2014b)[51]; HIMS (Yeap, 2014b)[30]; CHAMP (Hsu, 2016).[9] Summary  
30  
31 632 estimates are colour-coded as calculated using either the estimates from Yeap et al.[30]  
32  
33 633 calculated from the model including SHBG (black) or from the model including LH (grey). \*  
34  
35 634 This estimate from Hsu et al.[9] could not be used to calculate the summary estimate because  
36  
37 635 a variance estimate was not calculable for a 5nmol/L change in testosterone using the  
38  
39 636 published information.  
40  
41  
42  
43  
44  
45  
46

47 638 Figure 4. Forest plot of a meta-analysis of published estimates: association of testosterone  
48  
49 639 with mortality caused by cardiovascular disease. Plotted values are the estimated hazard  
50  
51 640 ratios (HR) for death from any cause, as attributed to an increase in endogenous testosterone  
52  
53 641 concentration by 5 nmol/L. The vertical reference line is HR=1. Study-specific estimates are  
54  
55 642 presented for six of the selected studies: BHS (Chan, 2016; Chasland, 2017)[37-38]; EMAS  
56  
57 643 (Pye, 2014)[11]; ARIC (Srinath, 2015)[48]; CHS (Shores, 2014b)[51]; HIMS (Yeap,  
58  
59  
60

1  
2  
3 644 2014b)[30]; CHAMP (Hsu, 2016).[9] Summary estimates are colour-coded as calculated  
4  
5 645 using either the estimates from Chan et al.[37] (black) or Chasland et al.[38] (grey) for the  
6  
7 646 BHS. \* This estimate from Hsu et al.[9] could not be used to calculate the summary estimate  
8  
9 647 because a variance estimate was not calculable for a 5nmol/L change in testosterone using the  
10  
11 648 published information.  
12  
13  
14  
15 649  
16  
17 650  
18 651  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

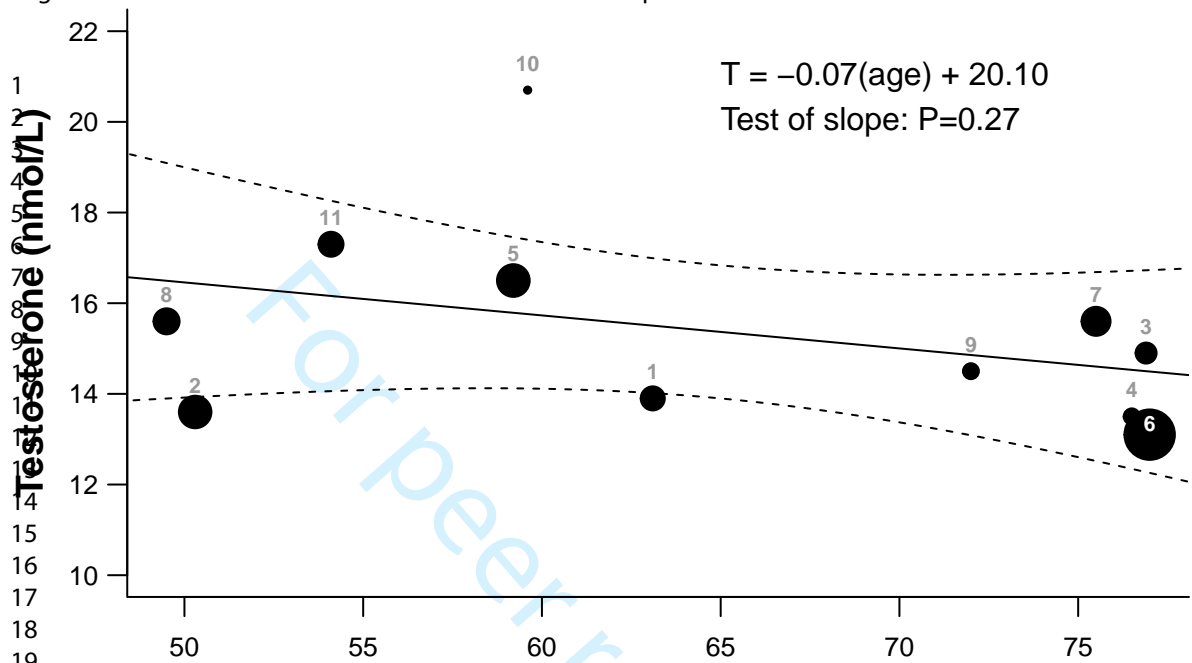
Literature Search (items)

Screening (items)

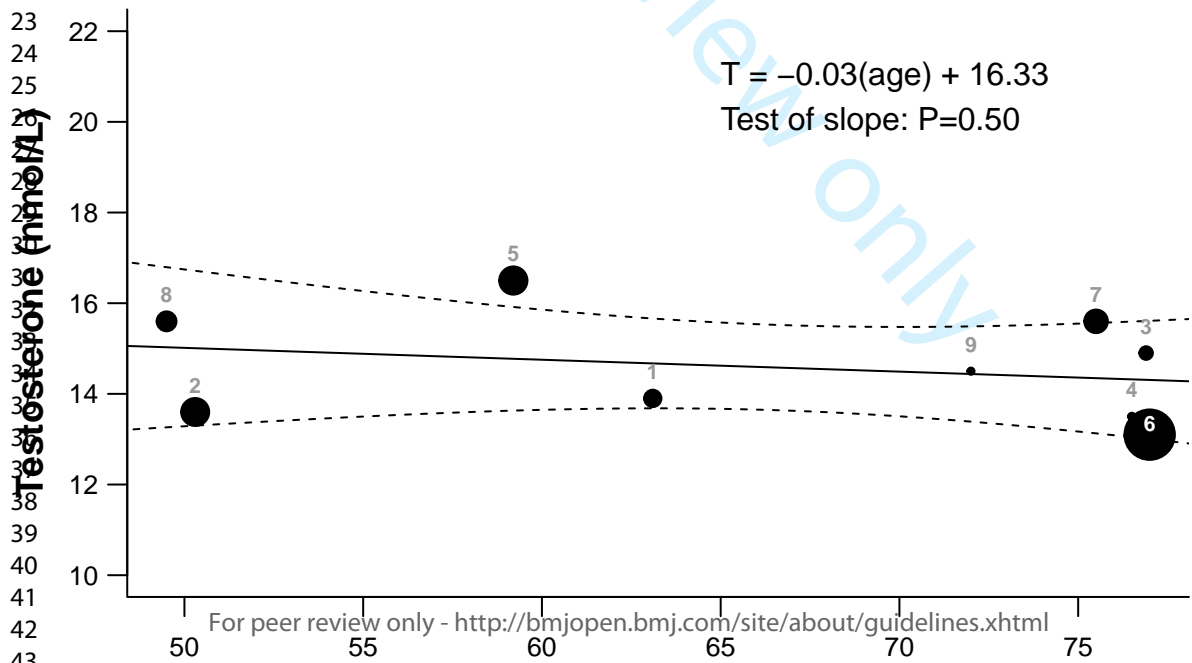
Screening (items)

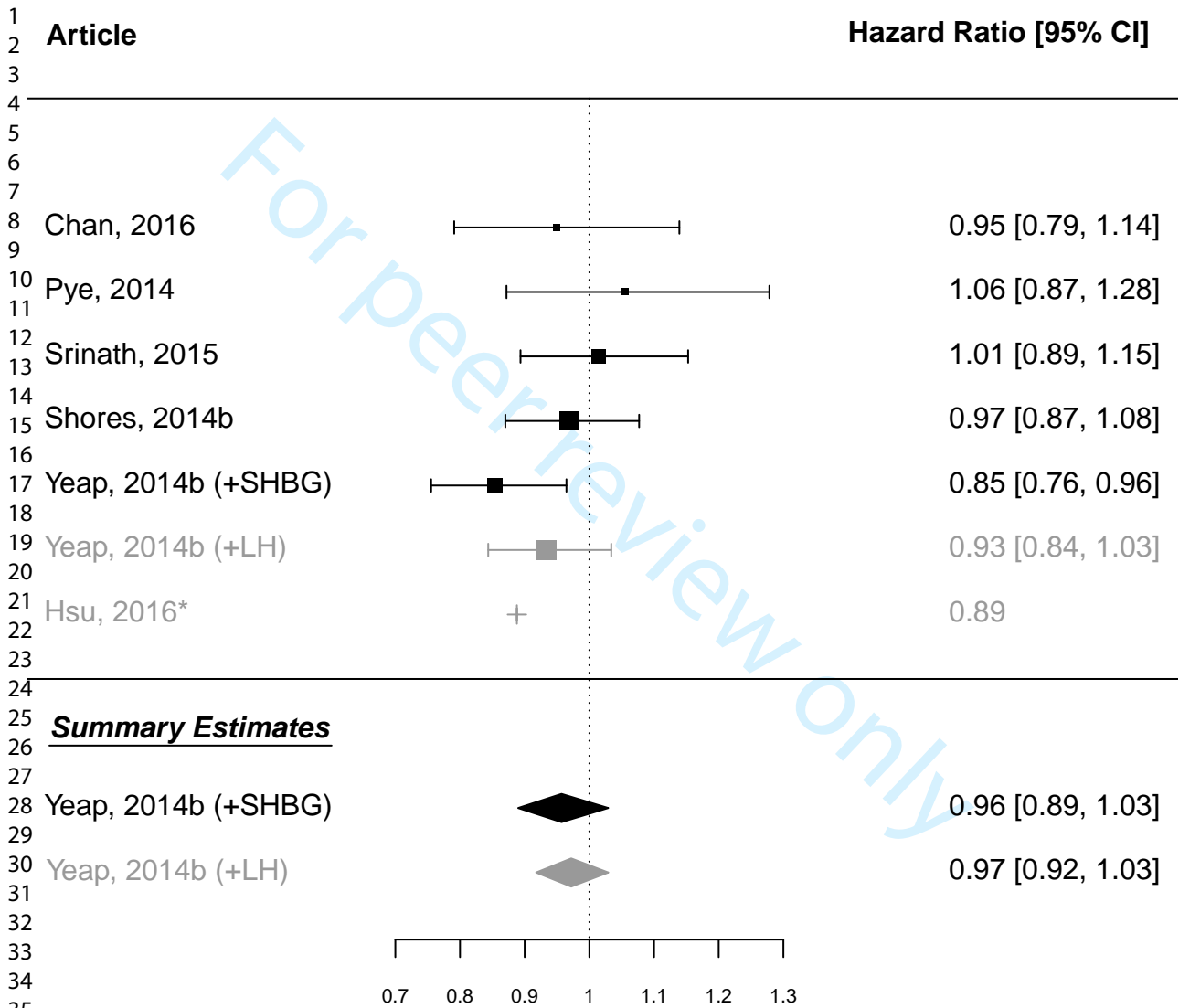
Additional Steps (studies)

### a) All cohorts\*



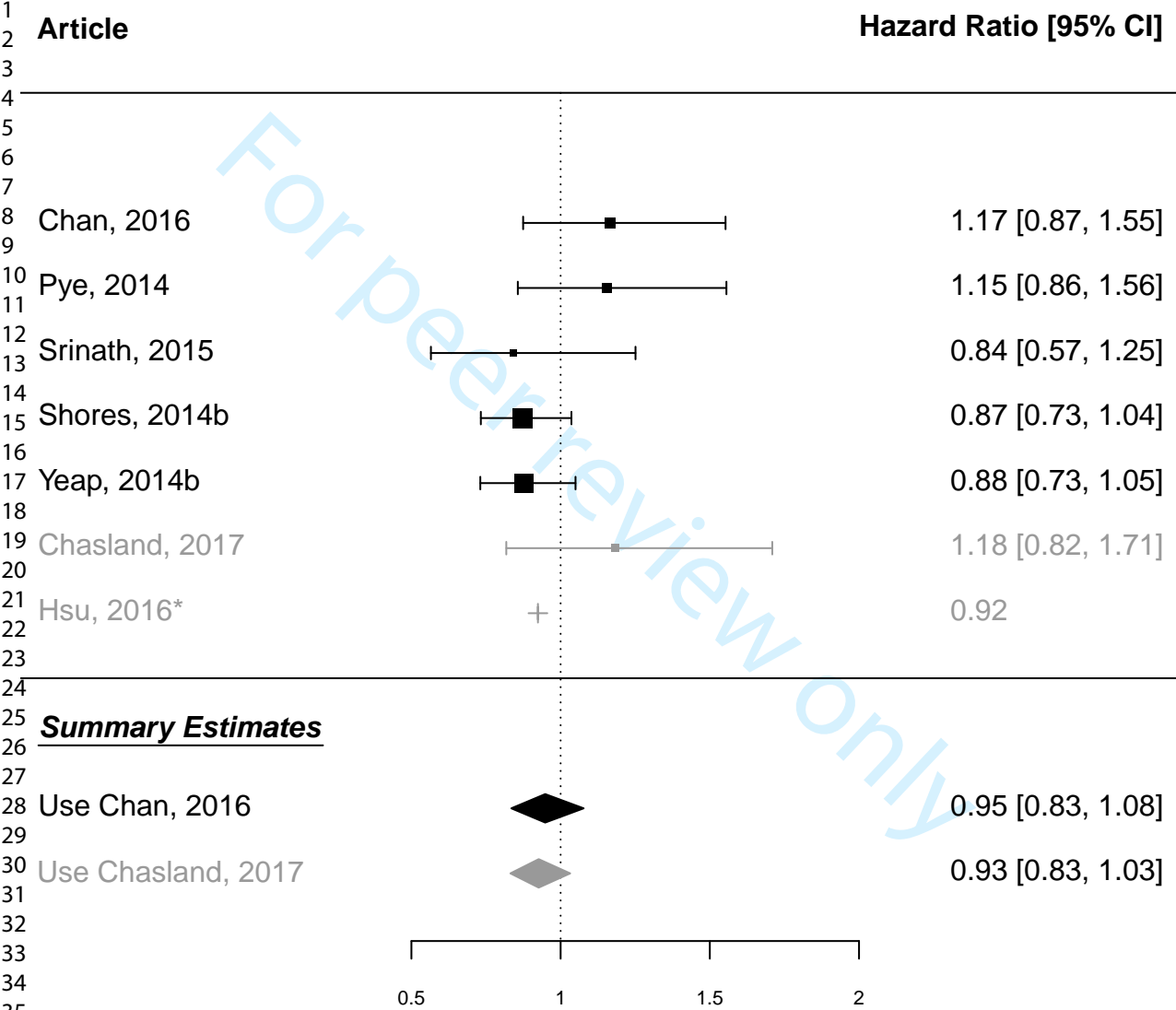
### b) Selected items





For peer review only





*Supplementary Material. Systematic review: associations of testosterone with men's health.***1. Table of Contents**

1	1	1. Table of Contents.....	1
2	2	2. Additional details on systematic searches and screening .....	2
3	3	3. Tables.....	7
4	4	Supplementary table 1. Full electronic search strategy used for MEDLINE database.....	7
5	5	Supplementary table 2: Selection criteria for screening items returned from the literature	
6	6	search. ....	8
7	7	Supplementary table 3: Adaptation of screening rules for different types of published and	
8	8	unpublished items. ....	9
9	9	Supplementary table 4. Words mentioned in the titles or abstracts of reviewed items. ....	10
10	10	Supplementary table 5. PRISMA Checklist.....	11
11	11	Supplementary table 6. Attributes of selected items.....	14
12	12	Supplementary table 7. Exposure levels, outcome assessment, covariates .....	18
13	13	Supplementary table 8. Newcastle-Ottawa Quality Assessment Scale for Cohort Studies:	
14	14	selected articles. ....	21
15	15	Supplementary table 9. Extracted hazard ratio data for dose-response meta-analyses (DR-	
16	16	MAAs). ....	22
17	17	Supplementary table 10. Published estimates for selected studies investigating associations	
18	18	of total testosterone with cognitive status or decline .....	26
19	19	4. Figures. ....	27
20	20	Supplementary figure 1. Meta-regression diagnostics.....	27
21	21	Supplementary figure 2. Funnel plots for dose-response meta-analyses. ....	29
22	22	Supplementary figure 3: Forest plot of published hazard ratio (HR) estimates: association	
23	23	of testosterone with other AIMS outcomes. ....	30
24	24	5. References cited.....	31

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

## 2. Additional details on systematic searches and screening

We used online search tools to identify available published (MEDLINE, EMBASE) and grey literature (OpenGrey, Mednar) *items* (journal articles, reports, theses, webpage articles) reporting on suitable prospective cohort *studies* (the underlying unique sources of data). We used OpenGrey and Mednar because both were free search tools that we considered likely to identify additional grey literature items and studies in an expanded search beyond the mainstream publications. Mednar is a medically-focussed search engine of public and deep web resources, excluding subscription services.[1] OpenGrey is a searchable database containing citations for items including technical or research reports, theses, conference papers, and other types of grey literature.[2] Literature searches were conducted on 18-22 July 2019, with no date restrictions set.

Where possible (as functionality varied among the different tools), we placed the following restrictions on the search: items reporting on the results of a research study, longitudinal or prospective cohort studies, not of hormone therapy or deprivation treatments. Due to study timeframe and language translation limitations, we opted to search for only those items that were reported in the English language. The terms and full criteria used for the MEDLINE search are provided in Supplementary table 1, and the PRISMA checklist as Supplementary table 5.

Selection criteria were set as applicable to the planned sets of IPD meta-analyses (Supplementary table 2).[3] Only items reporting on prospective population-based cohort studies, adults of combined sexes or of men alone, with individuals free of the disease at baseline, were sought. Items reporting a different design for the analysis of longitudinal data, such as nested case-control or case-cohort design, were also considered acceptable. A

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

1  
2  
3 55 minimum of five years follow-up was selected, to ensure a sufficient number of incident  
4  
5 56 events for statistical modelling. We excluded items that did not measure testosterone using  
6  
7 57 mass spectrometry, which is regarded to be the 'gold standard' method,[4] although  
8  
9 58 testosterone was not required to be mentioned in the title or abstract, nor modelled as the  
10  
11 59 primary exposure variable. Selected items were to be studies of humans, reported in English,  
12  
13 60 and reporting on analyses of at least one of the AIMS outcomes.  
14  
15  
16

17 61  
18  
19 62 Two reviewers (RJM, JH) independently screened the de-duplicated items against these pre-  
20  
21 63 specified criteria. To optimise efficiency, the selection of items proceeded in two steps. Title  
22  
23 64 and abstract screenings (Step 1) were followed by full text screening of items selected in Step  
24  
25 65 1 (Step 2). If an item was selected for exclusion, then the main reason for that decision was  
26  
27 66 recorded. If there was uncertainty in the decision to exclude, in Step 1 the reviewer selected  
28  
29 67 "include" (in Step 1) or "maybe" (in Step 2). At the end of each step, the two reviewers  
30  
31 68 sought to achieve consensus, through discussion, for each item that did not achieve  
32  
33 69 agreement. Exclusion reasons were used to inform discussions for achieving consensus. Items  
34  
35 70 with a consensus decision of "maybe" were further investigated by Reviewer 1 (RJM) using  
36  
37 71 information external to the systematic searches and screenings (reading further details of  
38  
39 72 methods used in cited articles, and from correspondence with authors or other researchers  
40  
41 73 currently working on the research study).  
42  
43  
44  
45  
46  
47  
48

49 75 This screening procedure was adjusted to accommodate the different types of items reviewed  
50  
51 76 (published articles, theses, webpage articles, unpublished reports; Supplementary table 3). A  
52  
53 77 pilot set of title-only screenings for 30 randomly chosen articles suggested that sufficient  
54  
55  
56  
57  
58  
59  
60

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

1  
2  
3 78 information was contained within the titles alone for the purpose of Step 1 screenings.<sup>a</sup>  
4  
5 79 Therefore, in cases when an abstract was not available, only the titles were screened. Website  
6  
7  
8 80 items identified by the Mednar search tool were the type of item that most often did not have  
9  
10 81 abstract or summary text, and in these cases the webpage text was reviewed in place of an  
11  
12 82 abstract (Supplementary table 3).  
13

14 83  
15  
16  
17 84 Endnote X8[5] was used for collating and storing the citations returned from literature  
18  
19 85 searches, and for de-duplicating and storing the selected references. The full citations,  
20  
21 86 including abstracts, were exported from Endnote for uploading into Rayyan[6], which is a  
22  
23 87 free web tool that was used for screening, recording exclusion decisions, and downloading  
24  
25 88 selection results.  
26  
27  
28 89

29  
30  
31 90 The literature search identified 2,177 items (1,738 published and 439 from grey literature),  
32  
33 91 with 1,994 items remaining after duplicates had been removed, and after excluding two  
34  
35 92 Mednar items that had insufficient information available to review (Fig. 1). Supplementary  
36  
37 93 table 4 shows the frequencies of returned items by search terms present in the titles and  
38  
39 94 abstracts. Most (72.7%) had the word “cancer”, and 1,107 (55.5%) of these had the word  
40  
41 95 “prostate cancer”, in the title or abstract. This, combined with frequent mentions of  
42  
43 96 “androgen deprivation” (29.2%), “radiotherapy” (18.6%), and “brachytherapy” (8.3%), show  
44  
45 97 that items reporting aspects of testosterone deprivation or suppression for treating prostate  
46  
47 98 cancer were a predominant feature of the returned items. Different types of returned items  
48  
49 99 included 1,764 published articles, 111 webpage articles, 81 theses, and 38 unpublished  
50  
51 100 reports/other documents, and the percentages without abstract or webpage text screened in  
52  
53  
54

---

55  
56  
57 <sup>a</sup> 30 titles were initially screened at random. 18 were flagged as not suitable, leaving 12 as potentially suitable.  
58 Subsequent Step 1 screening of titles with abstracts selected 25 of these articles for exclusion, with 5 retained  
59 for Step 2 (full text screening). All 5 were flagged as being potentially suitable in the pilot set of title-only  
60 screenings.

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

1  
2  
3 101 Step 1 were 2.6%, 1.8%, 24.7%, 65.8%, respectively (i.e., 4.7% overall).  
4  
5

6 102

7  
8 103 One thousand nine hundred sixty-eight items were excluded, five items were classified as  
9  
10 104 “Maybe”, and one item could not be screened because the full text version was not available,  
11  
12 105 leaving n = 20 suitable items selected (Fig. 1). Most (92.1%) of the exclusions were made  
13  
14 106 from reviewing titles and abstracts at Step 1, with a further 6.6% excluded from screening of  
15  
16 107 the 157 full text items in Step 2. Inter-reader agreement was a Cohen's Kappa  $\kappa = 0.69$   
17  
18 108 (or 96.0 percent agreement) for Step 1 and  $\kappa = 0.82$  (or 98.1 percent agreement) for Step 2.  
19  
20 109 Percentages of items with search terms (AIMS outcomes) in the title or abstract increased  
21  
22 110 after Step 1 in most cases except for “cancer” and “prostate cancer” (Supplementary table 4).  
23  
24 111 This reflects many exclusions in Step 1 that were of items reporting research on testosterone  
25  
26 112 deprivation or suppression treatments for prostate cancer.  
27  
28  
29  
30

31 113

32  
33 114 The systematic approach to literature searching and screening is widely held to be beneficial  
34  
35 115 to identifying studies that otherwise may not have been considered for inclusion, and thus to  
36  
37 116 minimise the prospect for reviewer biases affecting study selections and summary results.[7]  
38  
39 117 This process is not perfect though, and in our case it did not identify two prospective cohort  
40  
41 118 studies that were known to be suitable, prior to commencing this review (FHS, MAILES).[3]  
42  
43 119 In the case of MAILES, this was one of the more recently commenced of the selected studies,  
44  
45 120 with its cohort profile article published in 2014,[8] and accordingly has had a comparatively  
46  
47 121 short timeframe within which to analyse and publish suitable findings. In the case of FHS,  
48  
49 122 associations of endogenous testosterone with male health outcomes had previously been  
50  
51 123 investigated and published, but not using mass spectrometry for measuring testosterone.[9,  
52  
53 124 10] Those articles were identified in the literature search but had been excluded on account of  
54  
55 125 assay method. Only relatively recently have testosterone measures been re-assayed for FHS  
56  
57  
58  
59  
60

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

1  
2  
3 126 participants using mass spectrometry methods.[11] One article by Pencina et al[12] was  
4  
5 127 possibly within scope but not identified because it had not been entered into the MEDLINE  
6  
7 128 database prior to the literature search (article entry date = 14 May 2020). Furthermore, an  
8  
9  
10 129 article that presented suitable estimates from one of the selected studies by Yeap et al[13]  
11  
12 130 was not identified from the literature search because it did not have “prospective”, “follow-  
13  
14 131 up”, “cohort study” or “longitudinal study” terms in its title or abstract, nor any of the  
15  
16 132 corresponding MeSH terms listed (refer to Supplementary table 1 for search terms used).  
17  
18  
19 133  
20  
21 134 In expanding our literature search to unpublished grey literature, it successfully located one  
22  
23 135 suitable item, which was a link to a Web MD webpage article, with further details published  
24  
25 136 in a conference abstract by Sueoka et al[14] that would otherwise have not been returned  
26  
27 137 from searching only the MEDLINE and EMBASE databases.  
28  
29  
30  
31 138

1  
2  
3 139 **3. Tables**  
4

5 140 Supplementary table 1. Full electronic search strategy used for MEDLINE database.  
6

7 141  
8 142 The following is the search that was conducted on 18 July 2019 using MEDLINE.  
9 143

- 10 144 1. Testosterone/ or Androgens/  
11 145 2. (testosterone or androgen\* or sex hormone\* or sex steroid\*).ti.  
12 146 3. (testosterone or androgen\*).ab.  
13 147 4. cardiovascular diseases/ or heart diseases/ or heart failure/ or vascular diseases/ or stroke/  
14 148 or myocardial infarction/ or coronary disease/ or cerebrovascular disorders/  
15 149 5. (cardiovascular or stroke or myocardial infarction or heart failure).ti.  
16 150 6. neoplasms/ or colorectal neoplasms/ or lung neoplasms/ or prostatic neoplasms/  
17 151 7. cancer.ti.  
18 152 8. mortality/ or mortality.ti.  
19 153 9. dementia/ or cognition/ or dementia.ti. or cognit\*.ti.  
20 154 10. Aging/psychology or Neuropsychological Tests/  
21 155 11. 1 or 2 or 3  
22 156 12. 4 or 5 or 6 or 7 or 8 or 9 or 10  
23 157 13. 11 and 12  
24 158 14. longitudinal studies/ or prospective studies/ or follow-up studies/ or cohort studies/  
25 159 15. (prospective or follow-up or cohort study or longitudinal study).ti,ab.  
26 160 16. 14 or 15  
27 161 17. 13 and 16  
28 162 18. (exogenous or replacement or therapy or hormone treatment).ti.  
29 163 19. Hormone Replacement Therapy/  
30 164 20. 18 or 19  
31 165 21. 17 not 20  
32 166 22. limit 21 to humans  
33 167 23. limit 22 to english language  
34 168 24. limit 23 to (adaptive clinical trial or address or autobiography or bibliography or  
35 169 biography or case reports or clinical trial, all or clinical trial, phase i or clinical trial, phase ii  
36 170 or clinical trial, phase iii or clinical trial, phase iv or clinical trial, veterinary or clinical trials,  
37 171 veterinary as topic or clinical trial protocol or clinical trial or comment or controlled clinical  
38 172 trial or dictionary or editorial or lecture or legislation or meta analysis or practice guideline or  
39 173 pragmatic clinical trial or published erratum or randomized controlled trial or retracted  
40 174 publication or "retraction of publication" or "review" or "scientific integrity review" or  
41 175 "systematic review")  
42 176 25. Retrospective Studies/ or Case-Control Studies/ or (retrospective analysis or case-  
43 177 control).ti.  
44 178 26. 24 or 25  
45 179 27. 23 not 26  
46 180

47 181 Notes:  
48 182

49 183 Terms with a trailing "/" are MeSH terms and those with a trailing "\*" are truncated search  
50 184 strings. Beforehand, a search of PROSPERO was conducted for another suitable strategy but  
51 185 none were found. However, the above strategy is based upon one that has been used for a  
52 186 similar study.[15] This search strategy is also published in the protocol article for the  
53 187 Androgens In Men Study.[3]



188 **Supplementary table 2: Selection criteria for screening items returned from the literature search.** If neither Include nor Exclude could be selected for Step  
 189 1, then reviewer selected "Include".

	Exclude	Include	Rationale	Used in Step 1		Used in Step 2
				Title & Abstract		Full-text
				Title only (no abstract)	Title & Abstract	
<b>Article type:</b>	Reviews, comments/opinion pieces, systematic reviews, dictionary, fact sheet, website information about diseases, fact sheets, etc.	Research study article / report, or an article that specifically refers to the results of one (e.g., a webpage referring to unpublished data).	These searches were of both published and unpublished scientific literature for the purpose of identifying prospective cohort studies that are likely to have the relevant data for planned IPD meta-analyses	Yes	Yes	Yes
<b>Study type:</b>	Retrospective or cross-sectional designs, case studies, case-control, surveys, RCTs or other trials, experiments, evaluation of androgen / testosterone therapy / deprivation / HRT or the effectiveness of any other type of intervention / surgery / treatment, genetics, etc.	Prospective cohort study.	A prospective cohort study design is of incident health outcomes for investigating etiology or disease risk for a cohort free of disease at baseline, and ideally should be representative of the local population, but may or may not be some demographic subset: e.g., age range, sex, ethnicity type.	Yes	Yes	Yes
<b>Population</b> (at baseline/date of recruitment to study)	Studies of juveniles only Studies of females only Individuals with some specific health condition/characteristic or following surgery / other medical treatment for specific illness	Adults (18 yr or older) Not females only Community-dwelling men	The study is of community-dwelling men.	Yes	Yes	Yes
<b>Exposure</b> (at baseline)	Do not exclude studies that do not model testosterone as the exposure: although it should be shown that it was measured for participants. If not mentioned in Step 2 then Exclude.	Endogenous testosterone	This will be the focal exposure for all IPD meta-analyses. However, as we are focussing on the identification of only those studies who have suitable androgen measurements available in IPD data, then testosterone does not necessarily need to be modelled as the focal exposure in included items. It is likely that details on the methods will be available only from full-text review.	Only if available	Only if available	Yes
	Testosterone not measured using mass spectrometry	Testosterone assay of serum or plasma sample using mass spectrometry (lc-ms or gc-ms)		Only if available	Only if available	Yes
<b>Outcome</b> (at follow-up)	Incident outcome not one of those type of events specified for inclusion.	Diagnosis/event of: cardiovascular disease (any); cancer (any); dementia. Deaths (any cause); deaths due to any type of cardiovascular disease; deaths due to any type of cancer. Cognition change / outcome	These are the outcomes for the planned IPD meta-analyses so it is important to seek IPD datasets from those studies who have already modelled these outcomes.  We refer to these as the "AIMS outcomes".	Yes	Yes	Yes
	Less than 5 years of follow-up data	Five or more years of follow-up data, with outcomes identified using systematic follow-up or data linkage.	As consistent across all included studies for IPD meta-analyses and set <i>a priori</i> . Likely that this will be available only from full-text review so not included Step 1.	No	No	Yes
<b>Language</b>	Title and/or abstract of article not in English	Title and/or abstract of article in English	As limited by the timeframe of this study and the native language of reviewers (a practicality).	Yes	Yes	Yes
<b>Species</b>	Studies not of humans	Studies of humans	We are studying humans.	Yes	Yes	Yes

190

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

191 Supplementary table 3: Adaptation of screening rules for different types of published and  
 192 unpublished items.

193

<b>Item Type</b>	<b>Step 1</b>	<b>Step 2</b>
Published article	Screen title (and abstract <sup>a</sup> )	Screen full text article
Thesis	Screen title (and abstract <sup>a</sup> )	Screen full thesis
Unpublished report / other document	Screen title (and abstract <sup>a,b</sup> )	Screen full document
Webpage	Screen title and webpage <sup>c</sup>	Screen full text article/document as identified from the webpage, or from a google search of information provided about the article, from the webpage.

194

195 <sup>a</sup> = when an abstract was available, otherwise title-only decisions were made (see

196 Supplementary table 2).

197 <sup>b</sup> = or, if not an abstract, other suitable document summary, as returned by the search tool.

198 <sup>c</sup> = for webpage articles, the webpage text served as the proxy for an abstract, with the

199 proviso that the reviewer did not navigate to additional webpages during Step 1.

200

*Supplementary Material. Systematic review: associations of testosterone with men's health.*201 Supplementary table 4. Words mentioned in the titles or abstracts of reviewed items.<sup>a</sup>

Word(s)	Step 1 items (n=1,994)	Step 2 items (n=158)	Selected items (n=20)
<i>Search terms (AIMS outcomes)</i>			
cancer	1,449 (72.7)	72 (45.6)	6 (30.0)
colorectal cancer	9 (0.5)	4 (2.5)	2 (10.0)
lung cancer	10 (0.5)	6 (3.8)	2 (10.0)
prostate cancer	1,107 (55.5)	40 (25.3)	2 (10.0)
cardiovascular	219 (11.0)	49 (31)	15 (75.0)
heart failure	29 (1.5)	2 (1.3)	1 (5.0)
stroke	31 (1.6)	12 (7.6)	4 (20.0)
myocardial infarction	33 (1.7)	7 (4.4)	1 (5.0)
mortality	232 (11.6)	45 (28.5)	9 (45.0)
dementia	22 (1.1)	8 (5.1)	2 (10.0)
cognit*	87 (4.4)	20 (12.7)	4 (20.0)
<i>Other frequently observed (not search terms)</i>			
androgen deprivation	583 (29.2)	2 (1.3)	0 (0.0)
androgen receptor	235 (11.8)	10 (6.3)	0 (0.0)
brachytherapy	165 (8.3)	0 (0.0)	0 (0.0)
breast cancer	153 (7.7)	9 (5.7)	0 (0.0)
radiotherapy	371 (18.6)	0 (0.0)	0 (0.0)

202 <sup>a</sup> = Items summarised as numbers (percentages); \*= wildcard character designating truncation

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

Supplementary table 5. PRISMA Checklist.

Section/topic	#	Checklist item	Reported on page #
<b>TITLE</b>			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
<b>ABSTRACT</b>			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5. For this type of review it is PEO instead.
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5, Supplementary table 2
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary table 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-6, Supplementary tables 2-3.

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

Section/topic	#	Checklist item	Reported on page #
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6, Suppl. Data.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6, Suppl. Data.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6, Supplementary table 8.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$ ) for each meta-analysis.	7
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7
<b>RESULTS</b>			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Supplementary tables 6-7, 9-10
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9, Supplementary table 8
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	10-12, Figs 3-4; Supplementary figure 3

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

Section/topic	#	Checklist item	Reported on page #
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-12, Figs 3-4; Supplementary figure 3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10-11, Supplementary figure 2
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	10-12, Fig 2; Supplementary figures 1-3
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	12-16
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14-15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	13-16
<b>FUNDING</b>			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	17

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Supplementary Material. Systematic review: associations of testosterone with men's health.

Supplementary table 6. Attributes of selected items.

Item	Article	Country	Study name <sup>§</sup>	Baseline**				Follow-up (relevant outcomes)	
				No. adult males	Baseline period	Age (yr) Mean (sd)	T (nmol/L) Mean (sd)	Length of follow-up (yr) (person-years) <sup>¶</sup>	AIMS Longitudinal Outcomes (no. of events analysed)
<i>Selected from systematic review</i>									
1	Srinath, 2015[16]	USA	ARIC	1,558	1996-98	63.1 (5.6)	13.9 (5.7)	Md=12.8 (CHD); Md=13.1 (HF) (25,374; HF)	Coronary Heart Disease (CHD; 287) Heart Failure (HF; 104) CHD deaths (29) All-cause deaths (347)
2	Srinath, 2016[17]	USA	ARIC	1,558	1996-98	63.1 (5.6)	13.9 (5.7)	Md=14.1 (27,311)	Ischemic Stroke (79)
3	Chan, 2016[18]	Australia	BHS	1,804	1994-95	50.3 (16.8)	13.6 (4.9)	Mn=14.9 (31,930)	CVD events (234; 399)*** CVD deaths (71; 141)*** All-cause deaths (191; 319)***
4	Chasland, 2017[19]	Australia	BHS	1,649	1994-95	49.8 (15.3)	13.7 (4.9)	Tot=20	CVD events (415) CVD deaths (127)
5	Chan, 2018[20]	Australia	BHS	1,574	1994-95	51.1 (14.7)	13.5 (4.8)	Tot=20	Prostate cancer (116) Lung cancer (22) Colorectal cancer (48) Cancer (any; 289)
6	Hsu, 2015[21]	Australia	CHAMP	853	2005-07	76.9 (5.5)	14.6 (6.2)	Tot=5	Cognitive decline (95)
7	Hsu, 2016[22]	Australia	CHAMP	1,705	2005-07	76.9 (5.5)	14.9 (6.6)	Md=6.9; Tot=10 (11,764)	Cancer deaths (151) CVD deaths (185) Other deaths (174) All-cause deaths (510)
8	Hsu, 2018[23]	Australia	CHAMP	1,651	2005-07	76.9 (5.5)	14.7 (6.4)	Tot=5	All-cause deaths (382) CVD deaths (cases not reported) Cancer deaths (cases not reported) Other deaths (cases not reported) Change in: MMSE, SF-12 (Mental).

## Supplementary Material. Systematic review: associations of testosterone with men's health.

Item	Article	Country	Study name <sup>s</sup>	Baseline**				Follow-up (relevant outcomes)	
				No. adult males	Baseline period	Age (yr) Mean (sd)	T (nmol/L) Mean (sd)	Length of follow-up (yr) (person-years) <sup>¶</sup>	AIMS Longitudinal Outcomes (no. of events analysed)
9	Rosenberg, 2018[24]	USA	CHS	1,019	1994	76.3 (4.9)	13.2 (6.2)	Md=9.5 (10,716)	Atrial Fibrillation (304)
10	Shores, 2014a[25]	USA	CHS	1,032	1994	76.5 (5.2)	13.5 (6.1)	Md=10; Tot=16 (19,220)	Ischemic stroke (114)
11	Shores, 2014b[26]	USA	CHS	NR	1994	NR	NR	Md=8.9 (CVD events) Md=10.8 yr (All-cause deaths). (9,184; CVD events)	CVD events (436) CVD deaths (157) All-cause deaths (777)
12	Lee, 2013[27]	Europe <sup>§§</sup>	EMAS	2,736	2003-05	59.2 (10.7)	16.5 (6)	Md=4.3; Tot=5 (14,486)	Cancer (any) Myocardial Infarction (MI) Heart Failure, Other heart conditions Stroke Cognitive function All-cause deaths (193)
13	Pye, 2014[28]	Europe <sup>§§</sup>	EMAS	2,599	NR	60 (11)	NR	Md=4.3; Tot=5 (11,140)	Cancer deaths (60) CVD deaths (56) All-cause deaths (147)
14	Chan, 2017[29]	Australia	HIMS	3,690	2001-04	77 (3.6)	13.1 (4.9)	Md=9.1, 9.2; Tot=11 (38,665)	Prostate cancer (348) Lung cancer (107) Colorectal cancer (137)
15	Ford, 2018[30]	Australia	HIMS	4,069	2001-04	NR	NR	Md=10.5; Tot=12 (44,404)	Dementia (499)
16	Yeap, 2014[31]	Australia	HIMS	3,690	2001-04	NR	NR	Mn=6.6 (2.3 sd) (28,036)	MI (344) Stroke (300)



Supplementary Material. Systematic review: associations of testosterone with men’s health.

Item	Article	Country	Study name <sup>§</sup>	Baseline**				Follow-up (relevant outcomes)	
				No. adult males	Baseline period	Age (yr) Mean (sd)	T (nmol/L) Mean (sd)	Length of follow-up (yr) (person-years) <sup>¶</sup>	AIMS Longitudinal Outcomes (no. of events analysed)
17	Ohlsson, 2010[32]	Sweden	MrOS	2,644	2001-04	75.5 (3.2)	15.6 (6.5)	Mn=4.5† (11,880)	CVD deaths (123) Cancer deaths (127) All-cause deaths (328)
18	Ohlsson, 2011[33]	Sweden	MrOS	2,416	2001-04	75.4 (3.2)	15.7 (6)	Md=5.1 (11,605)	CVD events (485) Chronic Heart Disease events (302) Cerebrovascular events (225)
19	Tivesten, 2014[34]	Sweden	MrOS	2,416	2001-04	75.4 (3.2)	15.7 (6)	Md=5.2 (12,070; CHD) (12,137; CBD)	Chronic Heart Disease (302; CHD) Cerebrovascular Disease (225; CBD)
20	Kische, 2017[35]	Germany	SHIP	1,962	1997-01	49.5 (16.3)	15.6 (6.1)	Tot=10	Change in cognitive status
<b>Decision = “Maybe”. Item selected based on additional information</b>									
21	LeBlanc, 2010[36]	USA	MrOS	1,602	NR	NR	NR	Mn=4.5†† (26,977)	Cognitive function (and change in) Cognitive decline
22*	Sueoka, 2010[14]	USA	MrOS	697	2000-05	72 (5.5)	14.5 (5.1)	Av=3.9†† (6,247)	Coronary Heart Disease events (100)
<b>Other. Additional studies selected based on information external to the systematic review</b>									
	No articles were selected.	USA	FHS	3,352[12]	1998-05	59.6 (9.1)[12] 49.4 (13.8)[11]	20.7 (8.0)[12]	Tot=10 (for Atrial Fibrillation)[37]	Cardiovascular outcomes[37, 38] Deaths[37] Cause-specific deaths[38] Cancer[39]
	No articles were selected.	Australia	MAILES	1,632[40]	2002-06[8]	54.1 (11.4)[40]	17.3 (5.7)[40]	Md=4.95; IQR=4.35-5.00[40] (12,686)	CVD events Deaths (99)[8] Cause-specific deaths[8]

§ Study name abbreviations: ‘ARIC’= Atherosclerosis Risk in Communities; ‘BHS’=Busselton Health Study; ‘CHAMP’=The Concord Health and Ageing in Men Project; ‘CHS’= Cardiovascular Health Study; ‘FHS’= the Framingham Heart Study; ‘HIMS’=The Health In Men Study; ‘EMAS’=European Male Ageing Study; ‘MAILES’= The Men Androgen Inflammation Lifestyle Environment and Stress study; ‘MrOS Sweden’=The MrOS Osteoporotic Fractures in Men study in Sweden; ‘MrOS USA’ = The MrOS Osteoporotic Fractures in Men study USA; ‘SHIP’=Study of Health in Pomerania SHIP.

§§ = UK, Italy, Belgium, Poland, Sweden, Spain, Hungary, Estonia

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

¶ 'Md'=median; 'Mn'=mean; 'Av'=average; 'Tot'=total follow-up for the cohort (i.e., maximum, rounded down to nearest whole year); 'IQR'=interquartile range. Unless provided in text, person-years was calculated by multiplying the median, mean, or average length of follow-up by the total number of adult male participants.

\* = Note that this is a published conference abstract so is not technically a "Full Text" item.

\*\* = Baseline statistics reported for whole cohort; 'NR' = statistics not reported for whole cohort; Means and standard deviations calculated by firstly transforming into standard units (for T: nmol/L) and then, where required, transforming from quartile statistics using the Box-Cox method of McGrath et al.[41]

\*\*\* = First number is for individuals without CVD at baseline.

† = Total follow-up exceeded 5 years, from baseline visit (2001-04) to end of mortality data collection (March 1, 2008).

†† = Note that since there was no published follow-up estimate exceeding 5 years (a requirement for selection) and it was not clear, based on the article information alone, whether the total follow-up was at least 5 years, these items were initially classified as "Maybe". The length of follow-up for collection of AIMS outcome data was determined to be satisfactory from subsequent correspondence with MrOS USA researchers.

## Supplementary Material. Systematic review: associations of testosterone with men's health.

Supplementary table 7. Exposure levels, outcome assessment, covariates.\*

Study	Article	Longitudinal measure of association	Exposure (testosterone)	Outcome ascertainment	Covariates
ARIC	Srinath, 2015[16]	HR	T quartiles	CVD events and deaths identified by annual questionnaires and continuous surveillance, independent from hospital admissions data (ICD codes). Cause of death from death certificates.	Age, race/centre, BMI, waist circumference, smoking status, diabetes mellitus, hypertension, LDL, HDL.
	Srinath, 2016[17]	HR	T tertiles	Definite or probable stroke events identified from hospital admissions, annual phone calls, study examinations adjudicated by a physician, with secondary physician adjudication if it disagreed with a computer algorithm.	Age, race, centre, BMI, waist circumference, smoking status, diabetes mellitus, hypertension, LDL, HDL.
BHS	Chan, 2016[18]	HR	T quartiles (results not shown), Continuous T.	Linked hospital admissions and deaths records (ICD codes)	Age, smoking, vigorous exercise, alcohol, BMI, diabetes, CVD, COPD, non-skin cancer, systolic blood pressure, hypertension, lipid lowering therapy, cholesterol, HDL, triglycerides, C-reactive protein, creatinine
	Chasland, 2017[19]	HR	Categories: Low (L) v High (H) T, physical activity(PA) LT+LPA, LT+HPA, HT+LPA, HT+HPA	Linked hospital admissions and deaths records (ICD codes)	Age, prevalent CVD, smoking, waist circumference, cholesterol, HDL, lipids medication, diabetes, systolic blood pressure, hypertension medication
	Chan, 2018[20]	HR	T quartiles, Continuous T.	Linked cancer and death registry records (ICD codes)	Age, marital status, occupation, smoking, alcohol consumption, leisure time physical activity, BMI, diabetes
CHAMP	Hsu, 2015[21]	Slope estimate (change in MMSE on baseline hormone level or longitudinal change in hormone level)	Continuous T, cFT	Clinic assessment: MMSE, Informant Questionnaire on Cognitive Decline as initial screen, followed by clinical assessment to diagnosis categories: normal cognition, MCI, dementia. During follow-up: A decline in MMSE by $\geq 3$ points	Age, BMI, smoking status, years of education, depression score (GDS)
	Hsu, 2016[22]	RR	Continuous T, cFT	Deaths identified from 4-monthly phone calls or deaths registry. Cause of death identified on death certificates independently by 2 medical practitioners.	Age, BMI, smoking status, comorbidity score
	Hsu, 2018[23]	HR, RR (Death outcomes); Slope estimates (MMSE, SF-12 Mental)	Categories: Low (<20 <sup>th</sup> centile) v Normal T combinations with Low (<20 <sup>th</sup> centile) v Normal cFT	Cause of death identified on death certificates independently by 2 medical practitioners.	Age, BMI, smoking status, comorbidity score
CHS	Rosenberg, 2018[24]	HR	Continuous T and cFT, T and cFT quintiles	Independently verified from ECGs taken annually for participants and from hospital discharge diagnoses	Age (stratified), race, education, income, clinic, smoking status, diabetes mellitus, BMI, loop diuretics, height, hypertension, depressed left ventricular ejection fraction, kidney function, systolic blood pressure, SHBG

## Supplementary Material. Systematic review: associations of testosterone with men's health.

Study	Article	Longitudinal measure of association	Exposure (testosterone)	Outcome ascertainment	Covariates
	Shores, 2014a[25]	HR	Continuous T, cFT (linear & non-linear), T categories	Medicare data, hospital records, imaging studies, autopsy results, death certificates, physician interviews data used for adjudications by committee, which included a neurologist.	Age, systolic blood pressure, anti-hypertensive medications, atrial fibrillation, diabetes, smoking, lipid-lowering drugs, HDL, cholesterol, creatinine, fasting glucose, diabetes medications.
	Shores, 2014b[26]	HR	Continuous T, cFT (linear or non-linear) categories: Q1, Q2-4	Medicare data, hospital records, imaging studies, autopsy results, death certificates, physician interviews data used for adjudications by committee, which included a neurologist.	Age, race, site, smoking status, alcohol consumption, hypertensive use, HDL, BMI, waist circumference, diabetes, SHBG.
EMAS	Lee, 2013[27]	N/A	No modelling of longitudinal outcomes reported	MI, heart failure, other heart conditions, cancers, stroke identified from postal questionnaire, MMSE for participants $\geq 65$ yr old from clinic assessments Variable methods for data capture + validation among centres.	No modelling of longitudinal outcomes reported
	Pye, 2014[28]	HR	T, cFT categories: quintiles, low v eugonadal T, LOH status.	Deaths identified from follow-up postal questionnaire or enquiry if no reply received, with 89% of deaths verified from death certificates, death registers, or medical/hospital records.	Age, site, BMI, smoking status, general health.
HIMS	Chan, 2017[29]	SHR	Continuous T, cFT.	Linked hospital admissions, death and cancer registry records (ICD, ICD-O-3 codes).	Age, BMI, smoking status, physical activity, alcohol consumption, diabetes mellitus, HDL, triglycerides, prior cancer diagnosis.
	Ford, 2018[30]	HR	Continuous T, cFT Quartile categories of T, cFT	Linked data (ICD codes) from inpatient and outpatient mental health services, hospital admissions, community aged care services, cancer and death registries.	Age, baseline cognitive function, depression, BMI, hypertension, CVD, plasma homocysteine.
	Yeap, 2014[31]	HR	T, cFT as quartile categories	Linked hospital admissions, death and cancer registry records (ICD codes).	Age, education, smoking status, BMI, waist to hip ratio, hypertension, dyslipidemia, diabetes, creatinine, prior cancer or existing CVD. Also SHBG for models with T.
MrOS Europe	Ohlsson, 2010[32]	HR (in relation to DHEA, DHEA-S)	No: T modelled as a covariate only	Linked data (death and hospital discharge registries), death certificates.	Age, site, BMI, C-reactive protein, ApoB/A1, smoking status, diabetes, hypertension, prior CVD, prior cancer, low testosterone (in lowest quartile), low estradiol
	Ohlsson, 2011[33]	HR	T, cFT as quartile categories, T as binary categories.	Linked data (death and hospital discharge registries), death certificates.	Age, morning sample, site, BMI, ApoB/A1, physical activity, smoking status, diabetes, hypertension
	Tivesten, 2014[34]	HR (in relation to DHEA, DHEA-S)	No: T modelled as a covariate only	Linked data (death and hospital discharge registries), death certificates.	Age, morning sample, site, BMI, ApoB/A1, C-reactive protein, estradiol, testosterone (i.e., continuous T), SHBG, eGFR, smoking status, diabetes, hypertension.
SHIP	Kische, 2017[35]	Slope estimate (change in MMSE on baseline hormone)	T, cFT as continuous and as 10-year age group quartile categories.	MMSE score.	Age, BMI, smoking status, alcohol consumption, physical activity, hypertension, occupational status, education level, civil status, baseline MMSE.

Supplementary Material. Systematic review: associations of testosterone with men's health.

Study	Article	Longitudinal measure of association	Exposure (testosterone)	Outcome ascertainment	Covariates
MrOS USA	LeBlanc, 2010[36]	Change in mean score RR of clinically important decline	cFT quartiles and continuous cFT and T (data not shown)	Cognitive tests at the baseline and follow-up visit from Part B of the Trail Making Test (Trails B) and the Modified Mental State Examination (3MS). Calculated from pre-defined drop in scores.	Age group, education level, race, general health, alcohol consumption, clinic, physical and mental health, physical activity, medications used at baseline, other sex steroids, SHBG.
	Sueoka, 2010[14]	HR	T quartiles	CHD events identified from 3-monthly contacts with participants. Incident events were reviewed and adjudicated by cardiologist using clinical records.	Age, clinic, BMI, blood pressure, lipid levels, smoking, hypertension, diabetes, use of lipid-lowering agents
FHS	N/A – no items were selected.	N/A	N/A	AF measured and adjudicated by cardiologists. Mortality data from death certificates, hospital or institutional records, obituaries, or direct notification[37] Medical records of CVD events reviewed by panel of experienced investigators. A heart study neurologist examined most participants with suspected stroke[38] Medical records of cancer diagnoses reviewed by two independent reviewers, with majority confirmed by pathology reports.[39]	N/A
MAILES	N/A – no items selected.	N/A	N/A	Self-reported and clinical follow-up data, death registry (linked data)[8]	N/A

\* ApoB/A1 = apolipoprotein-B to apolipoprotein-A1 ratio; BMI = body mass index; cFT = calculated free testosterone; CHD = coronary heart disease; COPD = chronic obstructive pulmonary disease; CVD = cardiovascular disease; DHEA = dehydroepiandrosterone; DHEA-S = DHEA sulfate; ECG = electrocardiogram; eGFR = estimated glomerular filtration rate; GDS = Geriatric Depression Scale; HDL = high-density lipoprotein; HR = hazard ratio; ICD = International Classification of Diseases; ICD-O-3 = International Classification of Diseases for Oncology; LDL = low-density lipoprotein; LOH = late-onset hypogonadism; MCI = mild cognitive impairment; MMSE = mini-mental state examination; N/A = not applicable; Q1=quartile 1; Q2-4=quartiles 2 to 4 combined; RR = relative risk; SF-12 = The Short Form (12) Health Survey; SHBG = sex hormone binding globulin; SHR = subhazard ratio, as estimated from competing-risks regression; T = total endogenous testosterone.

## Supplementary Material. Systematic review: associations of testosterone with men's health.

Supplementary table 8. Newcastle-Ottawa Quality Assessment Scale for Cohort Studies: selected articles.

Article	Study	Selection (4 stars)	Comparability (2 stars)	Outcome (3 stars)	Notes on Selection	Notes on Outcome
Srinath 2015[16]	ARIC	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Srinath 2016[17]	ARIC	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Chan 2016[18]	BHS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Chasland 2017[19]	BHS	***	**	**	Prevalent cases not excluded <sup>a</sup>	Losses to f/u not mentioned; linked data <sup>b</sup>
Chan 2018[20]	BHS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Hsu 2015[21]	CHAMP	****	**	***		
Hsu 2016[22]	CHAMP	***	**	***	Prevalent cases not excluded <sup>c</sup>	
Hsu 2018[23]	CHAMP	***	**	***	Prevalent cases not excluded <sup>c</sup>	
Rosenberg 2018[24]	CHS	****	**	**		Losses to f/u not mentioned; linked data <sup>d</sup>
Shores 2014a[25]	CHS	****	**	**		Losses to f/u not mentioned; linked data <sup>d</sup>
Shores 2014b[26]	CHS	****	**	**		Losses to f/u not mentioned; linked data <sup>d</sup>
Lee 2013[27]	EMAS	NA	NA	NA	No modelling of longitudinal outcomes reported	
Pye 2014[28]	EMAS	***	**	***	Prevalent cases not excluded <sup>a</sup>	
Chan 2017[29]	HIMS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Ford 2018[30]	HIMS	****	**	**		Losses to f/u not mentioned; linked data <sup>b</sup>
Yeap 2014[31]	HIMS	***	**	**	Prevalent cases not excluded <sup>a</sup>	Losses to f/u not mentioned; linked data <sup>b</sup>
Ohlsson 2010[32]	MrOS Sw.	NA	NA	NA	Testosterone was not the exposure variable in this article	
Ohlsson 2011[33]	MrOS Sw.	****	**	***		
Tivesten 2014[34]	MrOS Sw.	NA	NA	NA	Testosterone was not the exposure variable in this article	
Kische 2017[35]	SHIP	***	**	***	Prevalent cases not excluded <sup>e</sup>	
LeBlanc 2010[36]	MrOS USA	****	**	*		Bias from loss to f/u; F/u OK: additional steps <sup>f</sup>
Sueoka 2010[14]	MrOS USA	***	**	*	Prevalent cases not excluded <sup>a</sup>	F/u OK: additional steps <sup>f</sup>
<b>Additional item not selected but included in DR-MA:</b>						
Yeap 2014b[13]	HIMS	***	**	**	Prevalent cases not excluded <sup>a</sup>	Losses to f/u not mentioned; linked data <sup>b</sup>

'NA' = Not applicable (see Notes); 'f/u' = follow-up (of incident events); 'DR-MA' = dose-response meta-analyses of published estimates.

<sup>a</sup> = The influence of prevalent cases was statistically adjusted by including prevalent status as a model predictor.

<sup>b</sup> = Follow-up of cases was assumed to be almost complete because analyses were of linked administrative data.

<sup>c</sup> = The influence of prevalent cases was statistically adjusted by incorporating into a comorbidity status model predictor.

<sup>d</sup> = Follow-up of cases was assumed to be almost complete because analyses were of linked administrative data (with expert adjudications).

<sup>e</sup> = Outcome was change in cognition score, with baseline score (prevalent status) included as a model predictor.

<sup>f</sup> = Total length of follow-up period was not reported but determined to be satisfactory from correspondence with MrOS USA researchers.

## Supplementary Material. Systematic review: associations of testosterone with men's health.

Supplementary table 9. Extracted hazard ratio data for dose-response meta-analyses (DR-MAs).\*

Article	Study	Outcome	Testosterone	Units	HR	95% CI	Notes
Chan 2016[18]	BHS	All-cause mortality	<10.20	nmol/L	ref.		
Chan 2016	BHS	All-cause mortality	10.20 - <13.04	nmol/L	0.84	(0.62-1.14)	
Chan 2016	BHS	All-cause mortality	13.04 - <16.58	nmol/L	0.86	(0.62-1.19)	
Chan 2016	BHS	All-cause mortality	≥16.58	nmol/L	0.9	(0.62-1.3)	
Pye 2014[28]	EMAS	All-cause mortality	<11.65	nmol/L	1.1	(0.6-1.8)	
Pye 2014	EMAS	All-cause mortality	11.65-14.61	nmol/L	0.7	(0.4-1.3)	
Pye 2014	EMAS	All-cause mortality	14.61-17.28	nmol/L	0.7	(0.4-1.3)	
Pye 2014	EMAS	All-cause mortality	17.28-21.20	nmol/L	1.2	(0.7-2)	
Pye 2014	EMAS	All-cause mortality	>21.20	nmol/L	ref.		
Srinath 2015[16]	ARIC	All-cause mortality	≤288.4	ng/dL	0.96	(0.7-1.34)	
Srinath 2015	ARIC	All-cause mortality	288.5-377.6	ng/dL	0.99	(0.72-1.35)	
Srinath 2015	ARIC	All-cause mortality	377.7-480.1	ng/dL	1	(0.74-1.35)	
Srinath 2015	ARIC	All-cause mortality	≥480.2	ng/dL	ref.		
Shores 2014b[26]	CHS	All-cause mortality	<278	ng/dL	1.06	(0.88-1.29)	
Shores 2014b	CHS	All-cause mortality	≥278	ng/dL	ref.		
Yeap 2014b[13]	HIMS	All-cause mortality	0.25-9.82	nmol/L	ref.		Fully-adjusted model + SHBG
Yeap 2014b	HIMS	All-cause mortality	9.82-12.53	nmol/L	0.81	(0.68-0.98)	
Yeap 2014b	HIMS	All-cause mortality	12.56-15.75	nmol/L	0.75	(0.61-0.92)	
Yeap 2014b	HIMS	All-cause mortality	15.79-46.50	nmol/L	0.77	(0.61-0.97)	
Yeap 2014b	HIMS	All-cause mortality	0.25-9.82	nmol/L	ref.		Fully-adjusted model + LH
Yeap 2014b	HIMS	All-cause mortality	9.82-12.53	nmol/L	0.84	(0.7-1.01)	
Yeap 2014b	HIMS	All-cause mortality	12.56-15.75	nmol/L	0.81	(0.67-0.97)	
Yeap 2014b	HIMS	All-cause mortality	15.79-46.50	nmol/L	0.89	(0.73-1.07)	
Hsu 2016[22]	CHAMP	All-cause mortality		ng/mL	1.17	(1.03-1.32)	Per SD decrease in T. RR estimate used.
Chan 2016	BHS	CVD mortality	<10.20	nmol/L	ref.		



## Supplementary Material. Systematic review: associations of testosterone with men's health.

Article	Study	Outcome	Testosterone	Units	HR	95% CI	Notes
Chan 2016	BHS	CVD mortality	10.20 - <13.04	nmol/L	1.12	(0.7-1.78)	
Chan 2016	BHS	CVD mortality	13.04 - <16.58	nmol/L	1.39	(0.86-2.25)	
Chan 2016	BHS	CVD mortality	≥16.58	nmol/L	1.25	(0.69-2.25)	
Chasland 2017[19]	BHS	CVD mortality	<13.1	nmol/L	ref.		Total PA, "Low" PA, NS PA x T: these estimates were used
	BHS	CVD mortality	≥13.1	nmol/L	1.25	(0.77-2.03)	
Chasland 2017	BHS	CVD mortality	<13.1	nmol/L	0.69	(0.4-1.2)	Total PA, "High" PA, NS PA x T
Chasland 2017	BHS	CVD mortality	≥13.1	nmol/L	0.8	(0.48-1.35)	
Pye 2014	EMAS	CVD mortality	<11.65	nmol/L	1	(0.4-2.2)	
Pye 2014	EMAS	CVD mortality	11.65-14.61	nmol/L	0.5	(0.2-1.4)	
Pye 2014	EMAS	CVD mortality	14.61-17.28	nmol/L	0.4	(0.2-1.2)	
Pye 2014	EMAS	CVD mortality	17.28-21.20	nmol/L	1.1	(0.5-2.4)	
Pye 2014	EMAS	CVD mortality	>21.20	nmol/L	ref.		
Srinath 2015	ARIC	CVD mortality	≤288.4	ng/dL	1.36	(0.45-4.08)	
Srinath 2015	ARIC	CVD mortality	288.5-377.6	ng/dL	1.26	(0.73-3.7)	
Srinath 2015	ARIC	CVD mortality	377.7-480.1	ng/dL	0.57	(0.16-1.99)	
Srinath 2015	ARIC	CVD mortality	≥480.2	ng/dL	ref.		
Shores 2014b	CHS	CVD mortality	<278	ng/dL	1.28	(0.94-1.75)	
Shores 2014b	CHS	CVD mortality	≥278	ng/dL	ref.		
Yeap 2014b	HIMS	CVD mortality	0.25-9.82	nmol/L	ref.		
Yeap 2014b	HIMS	CVD mortality	9.82-12.53	nmol/L	0.82	(0.61-1.11)	
Yeap 2014b	HIMS	CVD mortality	12.56-15.75	nmol/L	0.79	(0.58-1.09)	
Yeap 2014b	HIMS	CVD mortality	15.79-46.50	nmol/L	0.79	(0.56-1.11)	
Hsu 2016	CHAMP	CVD mortality		ng/mL	1.11	(0.93-1.32)	Per SD decrease in T. RR estimate used
Srinath 2016[17]	ARIC	Stroke / CBD	≤317.7	ng/dL	1.47	(0.83-2.61)	
Srinath 2016	ARIC	Stroke / CBD	317.8-441.2	ng/dL	ref.		
Srinath 2016	ARIC	Stroke / CBD	≥441.3	ng/dL	1.15	(0.62-2.14)	
Shores 2014a[25]	CHS	Stroke / CBD	<200	ng/dL	1.46	(0.77-2.75)	
Shores 2014a	CHS	Stroke / CBD	200-400	ng/dL	0.9	(0.56-1.45)	



*Supplementary Material. Systematic review: associations of testosterone with men's health.*

Article	Study	Outcome	Testosterone	Units	HR	95% CI	Notes
Shores 2014a	CHS	Stroke / CBD	400-600	ng/dL	ref.		
Shores 2014a	CHS	Stroke / CBD	600-800	ng/dL	1.73	(0.88-3.39)	
Shores 2014a	CHS	Stroke / CBD	>800	ng/dL	1.69	(0.51-5.60)	
Yeap 2014[31]	HIMS	Stroke / CBD	0.25-9.82	nmol/L	ref.		
Yeap 2014	HIMS	Stroke / CBD	9.82-12.53	nmol/L	0.8	(0.59-1.09)	
Yeap 2014	HIMS	Stroke / CBD	12.56-15.75	nmol/L	0.72	(0.52-0.99)	
Yeap 2014	HIMS	Stroke / CBD	15.79-46.5	nmol/L	0.56	(0.39-0.81)	
Ohlsson	MrOS(Sw)	Stroke / CBD	≤340	ng/dL	ref.		
2011[33]	MrOS(Sw)	Stroke / CBD	341-438	ng/dL	ref.		
Ohlsson 2011	MrOS(Sw)	Stroke / CBD	439-549	ng/dL	ref.		Quartile 4 vs. quartiles 1 to 3 of testosterone
Ohlsson 2011	MrOS(Sw)	Stroke / CBD	≥550	ng/dL	0.76	(0.55-1.05)	
Ohlsson 2011	MrOS(Sw)	CVD	≤340	ng/dL	ref.		
Ohlsson 2011	MrOS(Sw)	CVD	341-438	ng/dL	1.02	(0.80-1.30)	
Ohlsson 2011	MrOS(Sw)	CVD	439-549	ng/dL	0.96	(0.75-1.23)	
Ohlsson 2011	MrOS(Sw)	CVD	≥550	ng/dL	0.71	(0.54-0.93)	
Chan 2016	BHS	CVD		nmol/L	1.03	(0.92-1.15)	Per SD increase in T
Chasland 2017	BHS	CVD	<13.1	nmol/L	ref.		Total PA, "Low" PA, NS PA x T: these estimates were used
Chasland 2017	BHS	CVD	≥13.1	nmol/L	1.09	(0.83-1.44)	
Chasland 2017	BHS	CVD	<13.1	nmol/L	0.93	(0.70-1.23)	Total PA, "High" PA, NS PA x T
Chasland 2017	BHS	CVD	≥13.1	nmol/L	1.04	(0.78-1.38)	
Shores 2014b	CHS	CVD	<278	ng/dL	1.11	(0.87-1.43)	
Shores 2014b	CHS	CVD	≥278	ng/dL	ref.		
Yeap 2014	HIMS	CVD: MI	0.25-9.82	nmol/L	ref.		
Yeap 2014	HIMS	CVD: MI	9.82-12.53	nmol/L	1.07	(0.79-1.44)	
Yeap 2014	HIMS	CVD: MI	12.56-15.75	nmol/L	1.03	(0.76-1.41)	
Yeap 2014	HIMS	CVD: MI	15.79-46.5	nmol/L	0.92	(0.66-1.28)	
Srinath 2015	ARIC	CVD: HF	≤288.4	ng/dL	0.77	(0.46-1.29)	
Srinath 2015	ARIC	CVD: HF	288.5-377.6	ng/dL	0.72	(0.43-1.21)	

## Supplementary Material. Systematic review: associations of testosterone with men's health.

Article	Study	Outcome	Testosterone	Units	HR	95% CI	Notes
Srinath 2015	ARIC	CVD: HF	377.7-480.1	ng/dL	0.87	(0.53-1.42)	
Srinath 2015	ARIC	CVD: HF	≥480.2	ng/dL	ref.		
Chan 2018[20]	BHS	Cancer	<10.17	nmol/L	ref.		
Chan 2018	BHS	Cancer	10.17-<12.95	nmol/L	0.72	(0.53-0.99)	
Chan 2018	BHS	Cancer	12.95-<16.49	nmol/L	0.71	(0.51-0.98)	
Chan 2018	BHS	Cancer	≥16.49	nmol/L	0.81	(0.57-1.14)	
Chan 2018	BHS	Cancer: Prostate	<10.17	nmol/L	ref.		
Chan 2018	BHS	Cancer: Prostate	10.17-<12.95	nmol/L	0.62	(0.37-1.03)	
Chan 2018	BHS	Cancer: Prostate	12.95-<16.49	nmol/L	0.75	(0.46-1.23)	
Chan 2018	BHS	Cancer: Prostate	≥16.49	nmol/L	0.58	(0.33-1.01)	
Chan 2018	BHS	Cancer: Colorectal		nmol/L	1.04	(0.76-1.42)	Per SD increase in T
Chan 2018	BHS	Cancer: Lung		nmol/L	0.65	(0.39-1.09)	Per SD increase in T
Chan 2017[29]	HIMS	Cancer: Prostate		nmol/L	1.00	(0.90-1.12)	Per SD increase in T.
Chan 2017	HIMS	Cancer: Colorectal		nmol/L	0.96	(0.80-1.15)	Per SD increase in T
Chan 2017	HIMS	Cancer: Lung		nmol/L	1.30	(1.06-1.60)	Per SD increase in T
Hsu 2016	CHAMP	Cancer mortality		ng/mL	1.30	(1.02-1.65)	Per SD decrease in T. RR estimate used.
Ford 2018[30]	HIMS	Dementia	Not reported	nmol/L	1.39	(1.04-1.85)	
Ford 2018	HIMS	Dementia			1.31	(1.00-1.73)	
Ford 2018	HIMS	Dementia			1.23	(0.93-1.61)	
Ford 2018	HIMS	Dementia			ref.		
Ford 2018	HIMS	Dementia	SD for cohort not reported	nmol/L	1.11	(1.01-1.21)	Per SD decrease in T

\* = Estimates were also reported for all-cause and CVD mortality mortality for the CHAMP study in another of the selected articles,[23] but were not used because they were reported for combinations of free testosterone and total testosterone, and so were not comparable to the above published estimates.

CBD = cerebrovascular disease; CVD = cardiovascular disease; HF = heart failure; LH = luteinising hormone; MI = myocardial infarction; NS = non-significant result at *a priori* selected threshold for test; PA = physical activity level; ref. = referent level; RR = relative risk; SD = standard deviation; SHBG = sex hormone-binding globulin; T = endogenous total testosterone concentration.

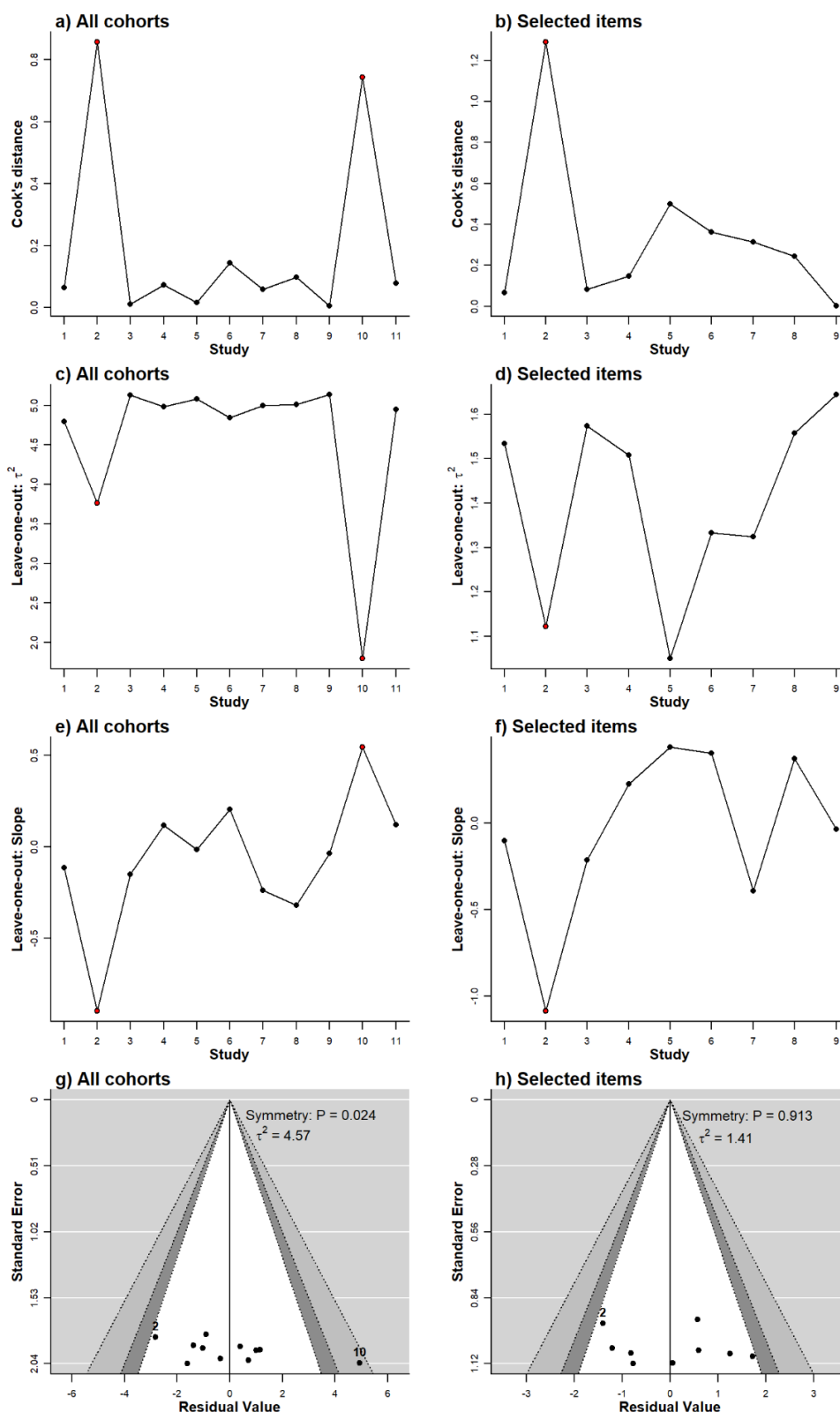
## Supplementary Material. Systematic review: associations of testosterone with men's health.

Supplementary table 10. Published estimates for selected studies investigating associations of total testosterone with cognitive status or decline.\*

Article	Study	Outcome	Testosterone	Units	Effect size parameter	Estimate	95% CI	p value	Notes
Hsu 2015[21]	CHAMP	Cognitive change	Baseline	ng/mL	Slope: Per	0.012	NR	0.7	Longitudinal change in MMSE score
Hsu 2015	CHAMP	Cognitive change	Longitudinal	ng/mL	unit decline	0.067	NR	0.03	Longitudinal change in MMSE score
Hsu 2015	CHAMP	Cognitive decline	Baseline	ng/mL	Odds Ratio	NR	NR	NR	Longitudinal decline in MMSE $\geq$ 3 points. Non-significant association (data not shown)
Hsu 2018[23]	CHAMP	Baseline cognition	Baseline NN	nmol/L	Slope	0	0	NR	MMSE** at baseline
Hsu 2018	CHAMP	Baseline cognition	Baseline NL	nmol/L	Slope	0.1	-0.8-1.1	NR	MMSE at baseline
Hsu 2018	CHAMP	Baseline cognition	Baseline LN	nmol/L	Slope	0.02	-1.0-1.02	NR	MMSE at baseline
Hsu 2018	CHAMP	Baseline cognition	Baseline LL	nmol/L	Slope	0.8	-0.5-0.4	NR	MMSE at baseline
Hsu 2018	CHAMP	Cognitive change	Baseline NN	nmol/L	Slope	0	0	NR	Longitudinal change in MMSE score
Hsu 2018	CHAMP	Cognitive change	Baseline NL	nmol/L	Slope	0.007	-0.5-0.5	NR	Longitudinal change in MMSE score
Hsu 2018	CHAMP	Cognitive change	Baseline LN	nmol/L	Slope	-0.2	-0.7-0.4	NR	Longitudinal change in MMSE score
Hsu 2018	CHAMP	Cognitive change	Baseline LL	nmol/L	Slope	-0.004	-0.3-0.3	NR	Longitudinal change in MMSE score
Kische 2017[35]	SHIP	Cognitive change	Baseline	nmol/L	Slope	0.02	-0.15-0.20	$\geq$ 0.05	Longitudinal change in MMSE score after 5 years
Kische 2017	SHIP	Cognitive change	Baseline	nmol/L	Slope	0.01	-0.22-0.24	$\geq$ 0.05	Longitudinal change in MMSE score after 10 years
LeBlanc 2010[36]	MrOS US	Baseline cognition	Baseline	nmol/L	-	NR	NR	$\geq$ 0.29	3MS score (includes MMSE)
LeBlanc 2010	MrOS US	Follow-up cognition	Baseline	nmol/L	-	NR	NR	$\geq$ 0.29	3MS score (includes MMSE)
LeBlanc 2010	MrOS US	Cognitive change	Baseline	nmol/L	-	NR	NR	$\geq$ 0.29	3MS score (includes MMSE)
LeBlanc 2010	MrOS US	Baseline cognition	Baseline	nmol/L	-	NR	NR	$\geq$ 0.63	Trails B: test of executive function and motor speed
LeBlanc 2010	MrOS US	Follow-up cognition	Baseline	nmol/L	-	NR	NR	$\geq$ 0.63	Trails B: test of executive function and motor speed
LeBlanc 2010	MrOS US	Cognitive change	Baseline	nmol/L	-	NR	NR	$\geq$ 0.63	Trails B: test of executive function and motor speed

\* NR = Not reported; NN = Normal ( $>10.2$  nmol/L) total testosterone (T), normal ( $>156$  pmol/L) calculated free testosterone (cFT); NL = normal T, low ( $<156$  pmol/L) cFT; LN = low ( $<10.2$  nmol/L) T, normal cFT; LL = low T, low cFT. MMSE = mini-mental state examination; 3MS = modified mini-mental state examination.

4. Figures.



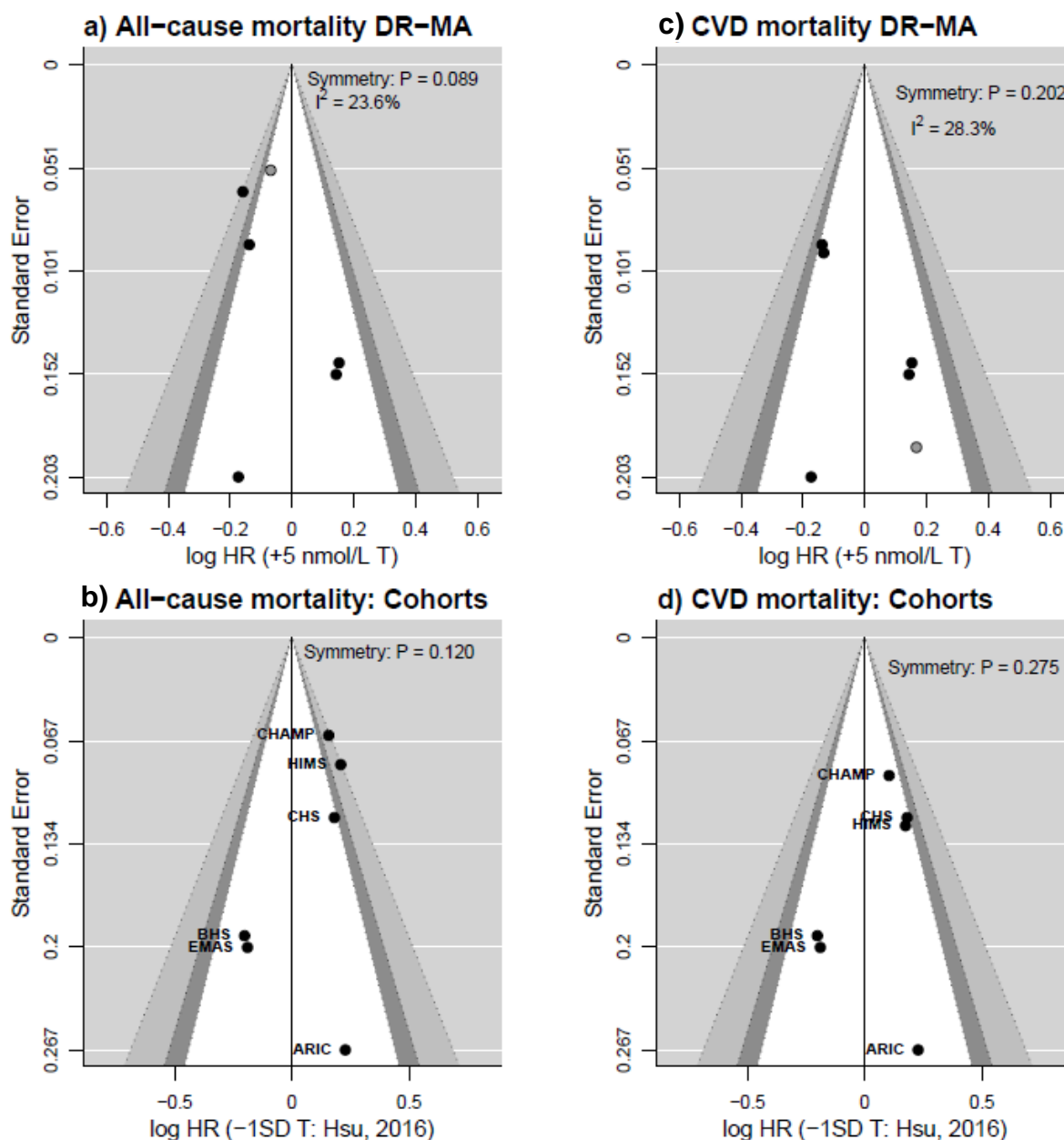
Supplementary figure 1. Meta-regression diagnostics. Meta-regression diagnostics showing the influence of studies on model fit (a,b),  $\tau^2$  (estimated amount of total heterogeneity: c,d), estimated slope (e,f), and distribution of residuals with funnel plots (g,h). Analysis repeated for all 11 cohort studies (a,c,e,g) and for 9 studies with selected articles (b,d,f,h). In cases where more than one article was available per cohort study, the article with the largest sample

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

size was used. Highlighted estimates for cohort study 2 (BHS) were those from Chan et al.[18] (N=1,804) and for study 10 (FHS) were from Pencina et al.[12] (N=720). In funnel plots: light grey + dark grey + white shading = 99% pseudo confidence interval (CI); dark grey + white shading = 95% CI; white shading = 90% CI.

For peer review only

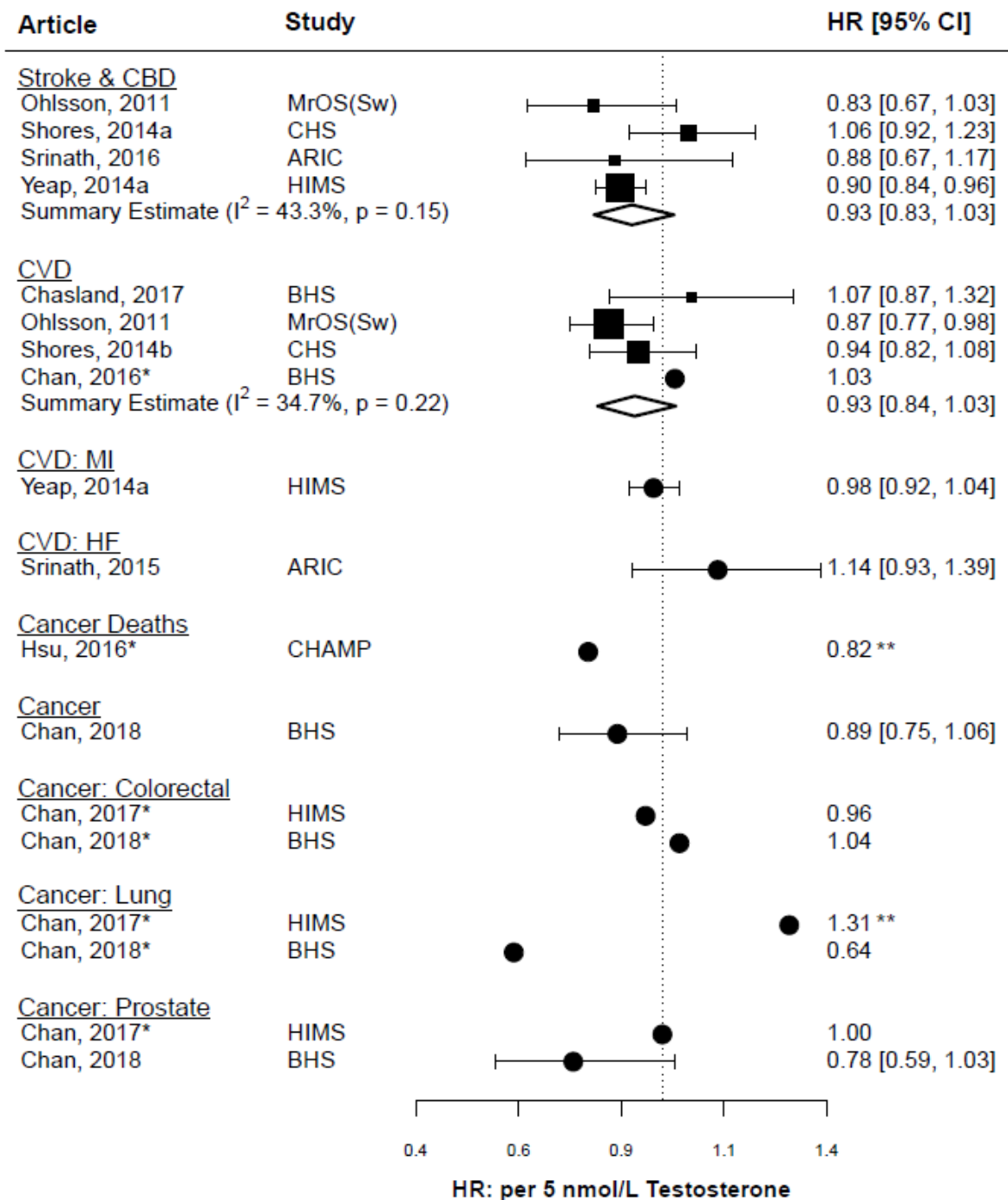
Supplementary Material. Systematic review: associations of testosterone with men's health.



Supplementary figure 2. Funnel plots for dose-response meta-analyses. Contour-enhanced funnel plots showing the distribution of log hazard ratio (HR) estimates for all-cause mortality (a, b) and mortality caused by cardiovascular disease (CVD) (c, d) attributed to a 5 nmol/L increase (a, c), or a 1.9 nmol/L (1SD in Hsu et al. 2016[22]) decrease (b, d), in endogenous testosterone concentration. Log HR values and standard errors were calculated using generalised least squares regression of published estimates.[42, 43] In cases where more than one article was available per cohort study, the article with the largest sample size was used. Estimates represented by black dots in (a) and (c) were analysed in respective dose-response meta-analyses (DR-MA; results presented in Figs. 3, 4). The grey dot in (a) is the estimate for Yeap et al. 2014b[13] and in (c) is the estimate for Chasland et al. (2017)[19]; these estimates were substituted for others for the HIMS and BHS studies respectively for alternative summary estimates (i.e., the grey summary estimates presented in Figs. 3, 4). Estimates presented in (b) and (d) are shown for a more complete assessment of funnel plot symmetry: estimates are plotted for all studies with estimates, including those that did not have sufficient information for including in the DR-MA. In funnel plots: light grey + dark

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

grey + white shading = 99% pseudo confidence interval (CI); dark grey + white shading = 95% CI; white shading = 90% CI.



Supplementary figure 3: Forest plot of published hazard ratio (HR) estimates: association of testosterone with other AIMS outcomes. Plotted estimates for other outcomes, as listed in the AIMS protocol article,[3] have been standardised to the HR for a 5nmol/L increase in testosterone. The size of squares are scaled to the precision of estimates, as used for obtaining the corresponding summary estimate for that outcome (diamonds). Estimates presented as circles were not used to obtain a summary estimate and so the size of circles is not scaled to estimated precision. \* = 95 % confidence intervals (CIs) were not calculable for these



*Supplementary Material. Systematic review: associations of testosterone with men's health.*

estimates, which were reported as the HR per standard deviation change (see Supplementary table 9).

\*\* = for 'per SD' estimates: to show that the published HR 95% CIs did not overlap 1. 'CBD' = Cerebrovascular disease; 'CVD' = cardiovascular disease; 'MI' = myocardial infarction; 'HF' = heart failure. Study-specific estimates presented for MrOS Sweden (Ohlsson, 2011)[33]; CHS (Shores, 2014a, b)[25-26]; ARIC (Srinath, 2015; 2016)[16-17]; HIMS (Yeap, 2014a, b)[13,31]; BHS (Chan, 2016; Chasland, 2017; Chan, 2018)[18-20].

## 5. References cited.

1. MedNar: Deep Web Technologies.
2. OpenGrey: GreyNet International.
3. Yeap BB, Marriott RJ, Adams RJ, *et al.* Androgens In Men Study (AIMS): protocol for meta-analyses of individual participant data investigating associations of androgens with health outcomes in men. *BMJ Open* 2020;10:e034777.
4. Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the Journal of Clinical Endocrinology and Metabolism. *J Clin Endocrinol Metab* 2013;98:3971-73.
5. Endnote X8: Clarivate Analytics, 2019
6. Ouzzani M, Hammady H, Fedorowicz Z, *et al.* Rayyan - a web and mobile app for systematic reviews. *Syst Rev* 2016;5:210.
7. Borenstein M, Hedges LV, Higgins JPT, *et al.* Introduction to Meta-Analysis. West Sussex: John Wiley & Sons Ltd 2009.
8. Grant JF, Martin SA, Taylor AW, *et al.* Cohort profile: The men androgen inflammation lifestyle environment and stress (MAILES) study. *Int J Epidemiol* 2014;43:1040-53.
9. Arnlov J, Pencina MJ, Amin S, *et al.* Endogenous sex hormones and cardiovascular disease incidence in men. *Ann Intern Med* 2006;145:176-84.
10. Haring R, Teng Z, Xanthakis V, *et al.* Association of sex steroids, gonadotrophins, and their trajectories with clinical cardiovascular disease and all-cause mortality in elderly men from the Framingham Heart Study. *Clin Endocrinol* 2013;78:629-34.
11. Bhasin S, Pencina MJ, Kaur Jasuja G, *et al.* Reference ranges for testosterone in men generated using liquid chromatography tandem mass spectrometry in a community-base sample of healthy nonobese young men in the Framingham Heart Study and applied to three geographically distinct cohorts. *J Clin Endocrinol Metab* 2011;96:2430-39.
12. Pencina KM, Travison TG, Bhasin S, *et al.* Endogenous circulating testosterone and sex hormone-binding globulin levels and measures of myocardial structure and function: the Framingham Heart Study. *Andrology* 2019;7:307-14.
13. Yeap BB, Alfonso H, Chubb SAP, *et al.* In older men an optimal plasma testosterone is associated with reduced all-cause mortality and higher dihydrotestosterone with reduced ischemic heart disease mortality, while estradiol levels do not predict mortality. *J Clin Endocrinol Metab* 2014;99:E9-E18.
14. Sueoka KT, Ewing MS, Ensrud KE, *et al.* Higher endogenous testosterone levels associated with increased risk of coronary heart disease in elderly men: a prospective study. *Endocr Rev* 2010;31:S858.
15. Holmegard HN, Nordestgaard BG, Jensen GB, *et al.* Sex hormones and ischemic stroke: a prospective cohort study and meta-analyses. *J Clin Endocrinol Metab* 2016;101:69-78.
16. Srinath R, Hill Golden S, Carson KA. Endogenous Testosterone and its Relationship to Preclinical and Clinical Measures of Cardiovascular Disease in the Atherosclerosis Risk in Communities Study. *J Clin Endocrinol Metab* 2015;100:1602-02.



*Supplementary Material. Systematic review: associations of testosterone with men's health.*

17. Srinath R, Gottesman RF, Hill Golden S, *et al.* Association Between Endogenous Testosterone and Cerebrovascular Disease in the ARIC Study (Atherosclerosis Risk in Communities). *Stroke* 2016;47:2682-88.
18. Chan YX, Knuiman MW, Hung J, *et al.* Neutral associations of testosterone, dihydrotestosterone and estradiol with fatal and non-fatal cardiovascular events, and mortality in men aged 17-97 years. *Clin Endocrinol* 2016;85:575-82.
19. Chasland LC, Knuiman MW, Divitini ML, *et al.* Greater physical activity and higher androgen concentrations are independently associated with lower cardiometabolic risk in men. *Clin Endocrinol* 2017;87:466-74.
20. Chan YX, Knuiman MW, Divitini ML, *et al.* Lower Circulating Androgens Are Associated with Overall Cancer Risk and Prostate Cancer Risk in Men Aged 25-84 Years from the Busselton Health Study. *Horm Cancer* 2018;9:391-98.
21. Hsu B, Cumming RG, Waite LM, *et al.* Longitudinal Relationships between Reproductive Hormones and Cognitive Decline in Older Men: The Concord Health and Ageing in Men Project. *J Clin Endocrinol Metab* 2015;100:2223-30.
22. Hsu B, Cumming RG, Naganathan V, *et al.* Temporal Changes in Androgens and Estrogens Are Associated With All-Cause and Cause-Specific Mortality in Older Men. *J Clin Endocrinol Metab* 2016;101:2201-10.
23. Hsu B, Cumming RG, Blyth FM, *et al.* Evaluating Calculated Free Testosterone as a Predictor of Morbidity and Mortality Independent of Testosterone for Cross-sectional and 5-Year Longitudinal Health Outcomes in Older Men: The Concord Health and Ageing in Men Project. *J Gerontol A Biol Sci Med Sci* 2018;73:729-36.
24. Rosenberg MA, Shores MM, Matsumoto AM, *et al.* Serum androgens and risk of atrial fibrillation in older men: The Cardiovascular Health Study. *Clin Cardiol* 2018;41:830-36.
25. Shores MM, Arnold AM, Biggs ML, *et al.* Testosterone and dihydrotestosterone and incident ischaemic stroke in men in the Cardiovascular Health Study. *Clin Endocrinol* 2014;81:746-53.
26. Shores MM, Biggs ML, Arnold AM, *et al.* Testosterone, dihydrotestosterone, and incident cardiovascular disease and mortality in the cardiovascular health study. *J Clin Endocrinol Metab* 2014;99:2061-8.
27. Lee DM, Pye SR, Tajar A, *et al.* Cohort profile: the European Male Ageing Study. *Int J Epidemiol* 2013;42(2):391-401.
28. Pye SR, Huhtaniemi IT, Finn JD, *et al.* Late-onset hypogonadism and mortality in aging men. *J Clin Endocrinol Metab* 2014;99:1357-66.
29. Chan YX, Alfonso H, Chubb SA, *et al.* Higher Dihydrotestosterone Is Associated with the Incidence of Lung Cancer in Older Men. *Horm Cancer* 2017;8:119-26.
30. Ford AH, Yeap BB, Flicker L, *et al.* Sex hormones and incident dementia in older men: The health in men study. *Psychoneuroendocrinology* 2018;98:139-47.
31. Yeap BB, Alfonso H, Chubb SA, *et al.* In older men, higher plasma testosterone or dihydrotestosterone is an independent predictor for reduced incidence of stroke but not myocardial infarction. *J Clin Endocrinol Metab* 2014;99:4565-73.
32. Ohlsson C, Labrie F, Barrett-Connor E, *et al.* Low serum levels of dehydroepiandrosterone sulfate predict all-cause and cardiovascular mortality in elderly Swedish men. *J Clin Endocrinol Metab* 2010;95:4406-14.
33. Ohlsson C, Barrett-Connor E, Bhasin S, *et al.* High serum testosterone is associated with reduced risk of cardiovascular events in elderly men. The MrOS (Osteoporotic Fractures in Men) study in Sweden. *J Am Coll Cardiol* 2011;58:1674-81.
34. Tivesten Å, Vandenput L, Carlzon D, *et al.* Dehydroepiandrosterone and its sulfate predict the 5-year risk of coronary heart disease events in elderly men. *J Am Coll Cardiol* 2014;64(17):1801-10.

*Supplementary Material. Systematic review: associations of testosterone with men's health.*

35. Kische H, Gross S, Wallaschofski H, *et al.* Associations of androgens with depressive symptoms and cognitive status in the general population. *PLoS ONE* 2017;12:e0177272.
36. LeBlanc ES, Wang PY, Janowsky JS, *et al.* Association between sex steroids and cognition in elderly men. *Clin Endocrinol* 2010;72:393-403.
37. Magnani JW, Moser CB, Murabito JM, *et al.* Association of sex hormones, aging, and atrial fibrillation in men: The Framingham Heart Study. *Circ Arrhythm Electrophysiol* 2014;7:307-12.
38. D'Agostino RB, Vasan RS, Pencina MJ, *et al.* General cardiovascular risk profile for use in primary care: The Framingham Heart Study. *Circulation* 2008;117:743-53.
39. Murabito JM, Rosenberg CL, Finger D, *et al.* A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham Heart Study. *BMC Med Genet* 2007; 8(Suppl I).
40. Li JJ, Wittert GA, Vincent A, *et al.* Muscle grip strength predicts incident type 2 diabetes: population-based cohort study. *Metab Clin Exp* 2016;65:883-92.
41. McGrath S, Zhao X, Steele R, *et al.* Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Stat Methods Med Res* 2020;29:2520-37.
42. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135:1301-09.
43. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose-response data. *Stata J* 2006;6:40-57.

PRISMA Checklist: Systematic review and meta-analyses on associations of endogenous testosterone concentration with health outcomes in community-dwelling men.

Section/topic	#	Checklist item	Reported on page #
<b>TITLE</b>			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
<b>ABSTRACT</b>			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5. For this type of review it is PEO instead.
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5, Supplementary table 2
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary table 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-6, Supplementary tables 2-3.

Section/topic	#	Checklist item	Reported on page #
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6, Suppl. Data.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6, Suppl. Data.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6, Supplementary table 8.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$ ) for each meta-analysis.	7
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7
<b>RESULTS</b>			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Supplementary tables 7-8, 9-10
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9, Supplementary table 8
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	10-12, Figs 3-4; Supplementary figure 3

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Section/topic	#	Checklist item	Reported on page #
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-12, Figs. 3-4, Supplementary figure 3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10-11, Supplementary figure 2
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	10-12, Fig 2; Supplementary figures 1-3
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	12-16
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14-15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	13-16
<b>FUNDING</b>			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	17

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097