



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

Ross James Marriott ¹, Janis Harse,¹ Kevin Murray ¹, Bu Beng Yeap^{2,3}

To cite: Marriott RJ, Harse J, Murray K, *et al.* Systematic review and meta-analyses on associations of endogenous testosterone concentration with health outcomes in community-dwelling men. *BMJ Open* 2021;**11**:e048013. doi:10.1136/bmjopen-2020-048013

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-048013>).

Received 11 January 2021
Accepted 14 October 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹School of Population and Global Health, The University of Western Australia, Perth, Western Australia, Australia
²Medical School, The University of Western Australia, Perth, Western Australia, Australia
³Department of Endocrinology and Diabetes, Fiona Stanley Hospital, Perth, Western Australia, Australia

Correspondence to

Dr Ross James Marriott;
ross.marriott@uwa.edu.au

ABSTRACT

Objectives The overall study aim is to clarify the relation of endogenous sex hormones with major health outcomes in men. This paper reports a systematic review focusing on published estimates for testosterone associations.

Setting Community-dwelling men.

Participants 20 180 adult men participated in the final set of studies identified and selected from a systematic review. Eligible studies included prospective cohort studies with plasma or serum testosterone concentrations measured for adult men using mass spectrometry with at least 5 years of follow-up data and one of the specified outcome measures recorded. Only published or grey literature items written in English were considered.

Primary and secondary outcome measures Planned prospective outcome measures: cardiovascular disease (CVD) events, CVD deaths, all-cause mortality, cancer deaths, cancer diagnoses, cognitive decline, dementia. Meta-analyses were of the most frequently reported outcomes in selected studies: CVD deaths and all-cause mortality. Succinct characterisations of testosterone associations with other outcomes are also presented.

Results Screening of 1994 deduplicated items identified 9 suitable studies, with an additional 2 identified by colleagues (11 in total). Summary estimates of mean testosterone concentration and age at recruitment for 20 180 adult men were 15.4±0.7 nmol/L and 64.9±3.3 year. Despite considerable variation in mean testosterone, a metaregression estimated no significant dependence on mean age at recruitment among studies (slope=−0.03, 95% CI −0.11 to 0.06). Meta-analyses demonstrated negligible heterogeneity and no significant effect of a 5 nmol/L increase in testosterone on the risk of all-cause mortality (HR=0.96, 95% CI 0.89 to 1.03) or death from CVD (HR=0.95, 95% CI 0.83 to 1.08).

Conclusions Analyses of published estimates did not demonstrate associations of endogenous testosterone with CVD deaths or with all-cause mortality. Suggested further research includes the planned individual participant data meta-analyses for selected studies, enabling the investigation of non-linear summary effects.

PROSPERO registration number PROSPERO: CRD42019139668.

Strengths and limitations of this study

- This is the first systematic review on this topic to restrict selections to prospective cohort studies of community-dwelling men with testosterone measured using mass spectrometry: the ‘gold standard’ method.
- Systematic searches were made of both the published and grey literature using online search tools.
- Meta-analyses used estimates obtained from studies with at least 5 years of follow-up data and from fitted models which controlled for (at least) the age, smoking status and body mass index or waist circumference of participants.
- Meta-analyses of published estimates were limited to assuming linear relationships; however, subsequent individual participant-level data meta-analyses planned to arise from this work will look to explore non-linear associations.
- Analyses are of observational data, and so summary estimates will not fully eliminate the possibility of confounding arising from unadjusted effects.

INTRODUCTION

What does a low testosterone level mean for a man’s health? In men, levels of testosterone, the key male sex hormone (androgen), decline with increasing age, yet the basis for and health consequences of this phenomenon remain unclear.^{1–5} Endogenous testosterone concentrations reflect the function of the hypothalamic-pituitary-testicular axis, and are relatively lower in men who are obese, or with metabolic syndrome or diabetes.^{6–8} Others have reported associations of lower endogenous testosterone concentrations with higher risk of incident diseases, such as cardiovascular disease (CVD), and death.^{9–16} Whether low testosterone concentrations might contribute directly to the risk of CVD or death or whether it may be associated indirectly through its relationship with ageing and obesity is unknown.¹⁷ And whether or not it is directly related, it is possible that

endogenous testosterone could be useful as a biomarker for diagnostic and/or prognostic healthcare applications in men.^{18–20} An improved understanding of the associations of testosterone to health outcomes could inform further exploration and development of this concept.

The Androgens In Men Study seeks to clarify the associations of androgens (primarily testosterone) with key health outcomes in men (mortality, CVD, cancer, cognitive decline and dementia) by conducting a systematic review and a series of individual participant data meta-analyses.²¹ In this paper, we present the systematic review and meta-analyses using published estimates from prospective cohort studies with at least 5 years of follow-up data and testosterone measured using only mass spectrometry, the most reliable method.²²

METHODS

This systematic review, conducted 14 June—31 December 2019, was of ‘etiology and/or risk type’ studies.^{23 24} The prespecified purpose of the systematic review was to identify studies with suitable individual participant-level data (IPD) for collaborating with on a series of IPD meta-analyses. The population, exposure, outcomes characteristics included: adult men in the general community; endogenous circulating sex hormone concentration (primarily testosterone); incident CVD, mortality, cancers, cognitive decline, dementia. Subgroup IPD meta-analyses are also planned for heart failure (HF), myocardial infarction (MI), stroke; colorectal cancer, lung cancer, prostate cancer. A protocol was submitted to PROSPERO on 23 July 2019 and registered on 20 November 2019 (registration number CRD42019139668) and a protocol article has been published.²¹

Literature search and screening

Four online search tools were used to identify available published (Medline, Embase) and grey literature (Open-Grey, Mednar) *items* (journal articles, reports, theses, webpage articles) reporting on suitable prospective cohort *studies* (the underlying unique sources of data). Two reviewers (RJM, JH) independently screened the deduplicated items against prespecified criteria using Rayyan.²⁵ To optimise efficiency, title and abstract screenings were initially conducted (step 1), followed by full text screenings of the selected abstracts (step 2). Disagreements were resolved through subsequent discussions between reviewers and agreement quantified using Cohen’s Kappa and percent agreement. Only items reporting on prospective population-based cohort studies of adults (combined sexes or of men alone) with mass spectrometry measurements of testosterone and at least 5 years of subsequent follow-up data on incident CVD events, cancer or dementia diagnoses, cognition assessments, or on all-cause, CVD, or cancer deaths were selected. The Newcastle-Ottawa Quality Assessment Scale for Cohort Studies (NOS) was used to assess quality of the selected items.²⁶

Additional details on the methods and results are provided in online supplemental material, section 1.

Specifically, additional details on systematic searches and screening (online supplemental tables 1–4, section 2), supplemental tables (online supplemental material, section 3), including the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist (online supplemental table 5), supplemental figures (online supplemental material, section 4), and references cited (online supplemental material, section 5) have been included.

Meta-analyses of published estimates

Published estimates (author names, publication year, cohort study name, number of participants analysed, model covariates, testosterone statistics (overall and for individual exposure levels), participant age statistics, numbers of outcome events, follow-up time, HRs and 95% CIs of the most fully-adjusted model) were extracted from selected articles by the first author (RJM). For the purpose of these analyses, we present associations for endogenous total testosterone concentrations, comprising the sum of testosterone in the circulation, whether bound to sex hormone-binding globulin or albumin, or unbound. Testosterone statistics were converted into standard units (nmol/L) and values representing categorical ranges were determined following Wang *et al.*²⁷ If not reported, the numbers of participants and events within categories of testosterone, and the means of participant ages and testosterone concentrations at baseline, were calculated. The numbers of participants within quartile or quintile categories were calculated by dividing the total sample size by four or five. The numbers of events within categories were solved using Newton’s method by applying the algorithm of Greenland and Longnecker.²⁸ Means and SD for testosterone and age were calculated from presented quartile estimates using the Box-Cox method of McGrath *et al.*, which does not make distributional assumptions.²⁹

A random effects metaregression of mean baseline testosterone concentration on the mean participant age at baseline was conducted using published estimates from: (1) only those items identified in systematic searches; and (2) all suitable articles, including those found outside of systematic searches. A t-test of the metaregression slope coefficient’s departure from zero was done after applying the Knapp and Hartung adjustment.

Dose–response random effects meta-analyses (DR-MAs) were conducted to summarise published HR estimates for the associations of baseline testosterone concentrations with incident all-cause deaths and with CVD cause-specific deaths, as these were the most frequently reported outcomes in selected articles. Estimates from an additional article that had not been selected from systematic searches (Yeap *et al.*³⁰) were also used because it reported suitable estimates from one of the selected studies, and had been published within the literature search period. Contour-enhanced funnel plots were inspected for publication bias and patterns in heterogeneity and Cochran Q tests for heterogeneity (I^2), as well as regression tests for funnel plot asymmetry,³¹ were done. Forest plots were constructed to represent single HR estimates for each study, per 5 nmol/L increase in testosterone. For completeness, HR estimates for the other outcomes are

represented in a grouped forest plot, and other effect sizes in tables.

The ‘metafor’ package was used for metaregressions, forest plots and funnel plots, the ‘doseremeta’ package for DR-MAs, and the ‘estmeansd’ package for calculating study means and SD from published quartile statistics in R V.4.0.2.^{32–35}

Patient and public involvement

This work uses existing published data. Patients and public were not involved in the design, conduct, reporting or dissemination plans of the systematic review or meta-analyses.

RESULTS

Literature search and study selection

The literature search returned 2177 items (1738 published and 439 from grey literature), with 1994 items remaining after duplicates had been removed, and after excluding two Mednar items that had insufficient information available to review (figure 1). These included 1764 journal articles, 111 webpage articles, 81 theses and 38 unpublished reports/other documents. Systematic screening of the returned, deduplicated items excluded 1968, classified 5 as ‘Maybe’, and selected 20 as suitable. Most (92.1%) items were excluded from screening titles and abstracts at step 1, with a much smaller percentage (6.6%) excluded from screening the 157 full-text items in step 2. One item could not be screened in step 2 because the full text was not available. Inter-reviewer agreement was a Cohen’s Kappa $\kappa = 0.69$ (or 96.0% agreement) for step 1 and $\kappa = 0.82$ (or 98.1% agreement) for step 2.

The 20 selected items collectively reported on eight prospective cohort studies: three from Australia (Busselton Health Study *BHS*,^{36–38} The Concord Health and Ageing in Men Project *CHAMP*,^{9 39 40} The Health In Men Study *HIMS*)^{14 41 42}; three from Europe (European Male Ageing Study *EMAS*,^{11 43} The *MrOS* Osteoporotic Fractures in Men study in Sweden,^{10 44 45} Study of Health in Pomerania *SHIP*)⁴⁶; and two from the USA (Atherosclerosis Risk in Communities *ARIC*,^{47 48} Cardiovascular Health Study *CHS*).^{49–51} Two of the five items classified as Maybe reported on the *MrOS USA* study, which were found, after further investigation, to be suitable for selection.^{52 53} Two additional studies were identified as suitable based on information external to the systematic searches and screenings: one from Australia (The Men Androgen Inflammation Lifestyle Environment and Stress study *MAILES*)⁵⁴; and one from the USA (the Framingham Heart Study *FHS*).⁵⁵ This is 11 cohort studies identified, in total. Additional details on returned and screened items, and selected article attributes are provided in Supplementary Material (Supplementary Section 2, online supplementary tables 4, 6 and 7).

Meta-analysis and summary of selected articles

The quality of selected articles ranged from six to nine (out of nine) stars on the NOS. Relatively high scores reflected

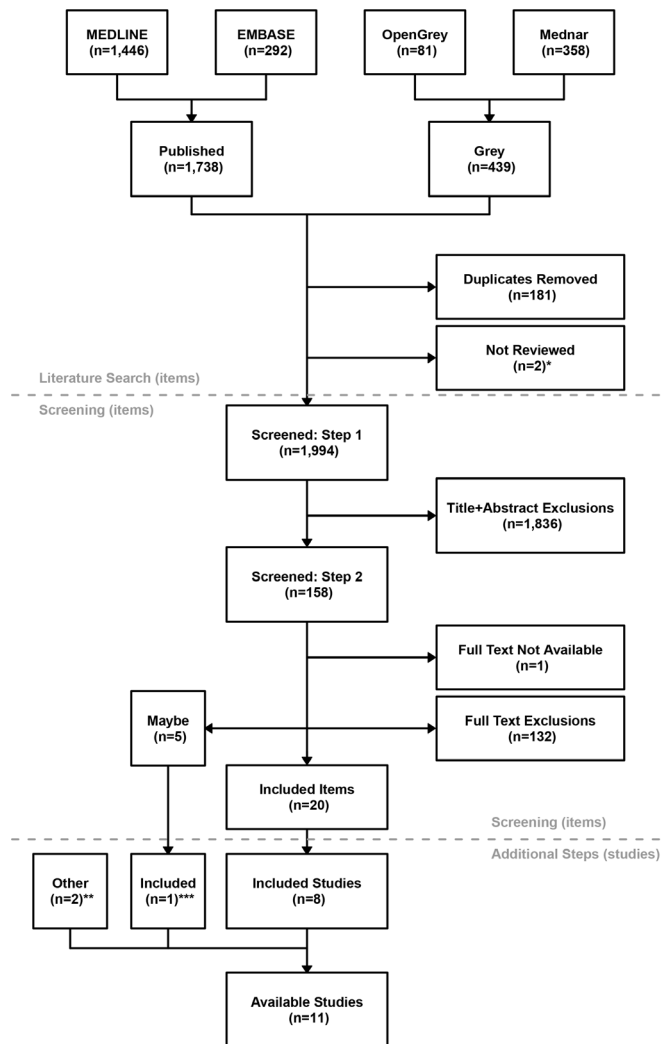


Figure 1 Studies returned from systematic review of the published and grey literature. Step 1 involved screening of titles and abstracts only and step 2 the screening of full-text items not excluded at step 1 (see online supplemental tables 2 and 3). ‘Items’ are individual articles or reports, with multiple items returned for some studies (the purpose was to identify studies with suitable individual participant (IPD)-level data). *Mednar items with insufficient information available to review; **additional studies identified through known contacts; ***screening criteria for five items selected as ‘Maybe’ in step 2 were further investigated using information external to systematic searches and screenings, resulting in the identification of one additional study with suitable IPD-level data.

that all articles: were of population-based studies; accurately measured the exposure (baseline testosterone concentration); included multivariable models adjusting for participant age and other risk factors; had outcomes measured or collected from record linkage, with or without expert adjudication; and had sufficient follow-up, ranging from 5 to 20 years (online supplemental tables 6–8). Relevant outcomes included: all-cause deaths (n=8 articles); CVD deaths (n=7); strokes or cerebrovascular disease (n=6); cognitive function or cognitive decline (n=5); coronary heart disease (n=4); CVD events (n=4); cancer deaths (n=4); cancer diagnoses

(n=3); MI (n=2); HF (n=2); and dementia (n=1). However, one of these articles was a cohort profile description that did not report effect size estimates but the availability of all-cause deaths, cause-specific deaths, stroke, cognitive function, CVD, cancer, MI and HF outcome data.⁴³ Two articles reported testosterone not as the exposure but as a covariate in analyses investigating associations with cerebrovascular events and with all-cause, cancer, and CVD deaths.^{44 45} The supplementary material for one article¹¹ was sought to obtain effect size estimates for cancer deaths but these were not obtained as at the time of writing. All were published between 2010 and 2018, reflecting the relatively recent adoption of mass spectrometry as the 'gold standard' for measuring endogenous testosterone levels.²²

The mean age of men at baseline ranged from middle-aged (49–54 year: BHS, FHS, MAILES, SHIP)^{18 36–38 46 56} to elderly (72–77 year: CHAMP, CHS, HIMS, MrOS Sweden, MrOS USA).^{9 10 39 41 44 45 49 50 53} Across the 11 studies, summary estimates for 20180 adult males at baseline were 64.9±3.3 year for mean age and 15.4±0.7 nmol/L for mean testosterone. Although there appeared to be a slight declining trend in mean testosterone with mean age among studies (metaregression slope=−0.07, 95% CI −0.21 to 0.07), this estimate was not significantly different from zero (p=0.27; figure 2a). However, the distribution of model residuals demonstrated significant heterogeneity (p<0.001) and funnel plot asymmetry (p=0.02). Additional diagnostics highlighted a relatively high mean testosterone estimate from Pencina *et al*⁵⁷ (FHS) and a low mean testosterone estimate (relative to mean age) from Chan *et al*³⁷ (BHS), as compared with the other studies (online supplemental figure 1). When restricted to systematically selected items (reporting on ARIC, BHS, CHAMP, CHS, EMAS, HIMS, MAILES, MrOS Sweden, SHIP studies), tests of residual heterogeneity were significant (p<0.001), funnel plot asymmetry (p=0.91) was non-significant, and the slope estimate (metaregression slope=−0.03, 95% CI −0.11–0.06) was not significantly different from zero (p=0.50; figure 2b). These results demonstrate that varying distributions of participant age (likely reflecting differences in study-specific objectives and recruitment methods) did not explain the observed heterogeneity in published estimates of testosterone among the studies.

HRs for all-cause mortality were calculated from values in four of the selected articles (ARIC,⁴⁸ BHS,³⁷ CHS,⁵¹ EMAS¹¹) and from one that was not selected, but had reported on the HIMS study during the literature search period.³⁰ All HRs were adjusted for the age, smoking status, and body mass index (BMI) or waist circumference of participants. A DR-MA estimated a summary HR of 0.96 (95% CI 0.89 to 1.03) per 5 nmol/L increase in testosterone (figure 3). The summary estimate was similar when calculated using an alternative estimate from Yeap *et al*³⁰ (HR=0.97, 95% CI 0.92 to 1.03). For both analyses, tests for residual heterogeneity ($I^2=23.6%$, p=0.26; $I^2=0.0%$, p=0.76) and funnel plot asymmetry (p=0.09; p=0.39) were non-significant (online supplemental figure 2a). A comparable HR was calculated from a CHAMP

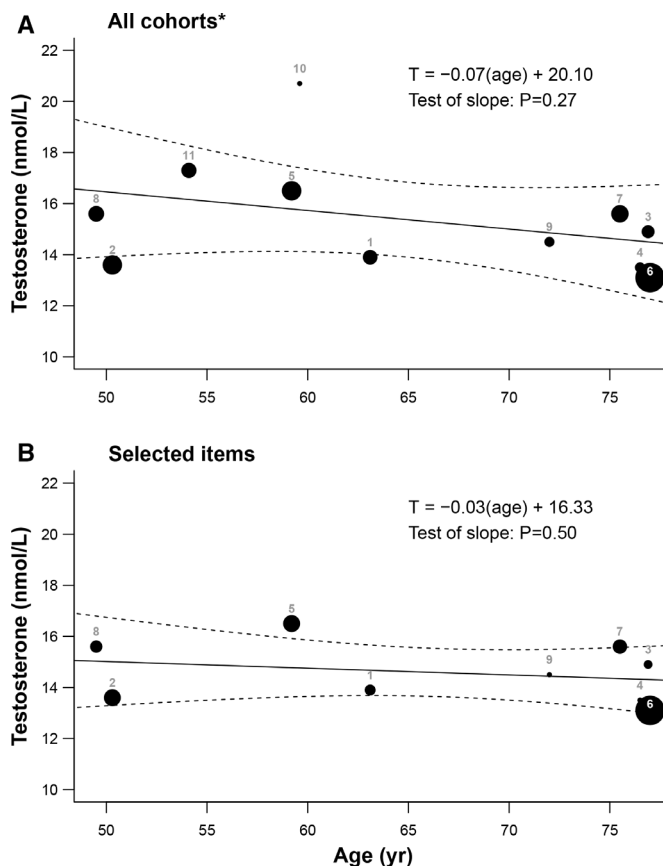


Figure 2 Metaregression of mean testosterone on mean age for (A) all 11 cohort studies and (B) 9 studies with articles that were selected by systematic literature searches and screening. The size of plotted points refers are proportional to the inverse of the corresponding SEs (indicative of relative weightings), with lines demonstrating the fitted model and 95% CIs. Plotted estimates are numbered as from the following articles (cohort studies): 1=Srinath *et al*⁴⁸ (ARIC); 2=Chan *et al*³⁷ (BHS); 3=Hsu *et al*⁹ (CHAMP); 4=Shores *et al*⁵⁰ (CHS); 5=Lee *et al*⁴³ (EMAS); 6=Chan *et al*⁴¹ (HIMS); 7=Ohlsson *et al*⁴⁴ (MrOS Sweden); 8=Kische *et al*⁴⁶ (SHIP); 9=Sueoka *et al*⁵³ (MrOS USA); 10=Pencina *et al*⁵⁷ (FHS); 11=Li *et al*⁵⁶ (MAILES). * = includes articles from two additional studies (FHS, MAILES) that were not identified from systematic searches but by colleagues.

study article⁹ for inclusion in the forest plot but not in the DR-MA, because a corresponding estimate of variance per 5 nmol/L increase in testosterone could not be calculated. An additional funnel plot, which included the HR estimate from this CHAMP article⁹ (per 1 SD decrease in testosterone, as reported in that article, also demonstrated no significant asymmetry (online supplemental figure 2b). These results demonstrate no overall effect of baseline testosterone concentration on the relative hazard of death from any cause after adjusting for factors including age, smoking status and BMI or waist circumference.

HRs for death caused by CVD demonstrated similar findings. A DR-MA using estimates from the same five articles estimated a summary HR of 0.95 (95% CI 0.83 to 1.08) per 5 nmol/L increase in testosterone, with no significant

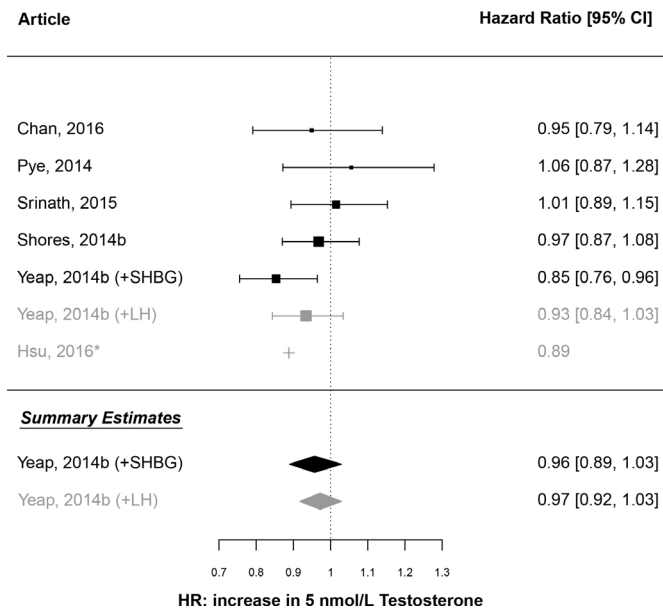


Figure 3 Forest plot of a meta-analysis of published estimates: association of testosterone with all-cause mortality. Plotted values are the estimated HR for death from any cause, as attributed to an increase in endogenous testosterone concentration by 5 nmol/L. The vertical reference line is HR=1. Study-specific estimates are presented for six of the selected studies: BHS (Chan, 2016)³⁷; EMAS (Pye, 2014)¹¹; ARIC (Srinath, 2015)⁴⁸; CHS (Shores, 2014b)⁵¹; HIMS (Yeap, 2014b)³⁰; CHAMP (Hsu, 2016).⁹ Summary estimates are colour coded as calculated using either the estimates from Yeap *et al*³⁰ calculated from the model including SHBG (black) or from the model including LH (grey). *This estimate from Hsu *et al*⁹ could not be used to calculate the summary estimate because a variance estimate was not calculable for a 5 nmol/L change in testosterone using the published information.

residual heterogeneity ($I^2=28.3\%$, $p=0.23$) or funnel plot asymmetry ($p=0.20$; figure 4; online supplemental figure 2c). Again, all HRs were adjusted for the age, smoking status, and BMI or waist circumference. The DR-MA repeated using an alternative estimate from Chasland *et al*³⁸ for the BHS gave similar results (summary HR=0.93, 95% CI 0.83 to 1.03; heterogeneity $I^2=17.5\%$, $p=0.30$; funnel plot asymmetry $p=0.17$; online supplemental figure 2d). These results demonstrate no overall effect of baseline testosterone concentration on the relative hazard of death from CVD after adjusting for factors including age, smoking status, and BMI or waist circumference.

Summary estimates calculated for the combined outcome of incident stroke and cerebrovascular disease (summary HR=0.93, 95% CI 0.83 to 1.03; heterogeneity $I^2=43.3\%$, $p=0.15$) and incident CVD diagnosis (summary HR=0.93, 95% CI 0.84 to 1.03; heterogeneity $I^2=34.7\%$, $p=0.22$) demonstrated no overall effect of testosterone (online supplemental figure 3). Funnel plot asymmetry was not assessed due to the low number of studies ($n\leq 4$),⁵⁸ and 95% CIs could not be calculated for several studies^{36 37 41} using the published information. Although a summary estimate could not be calculated, the study-specific estimates

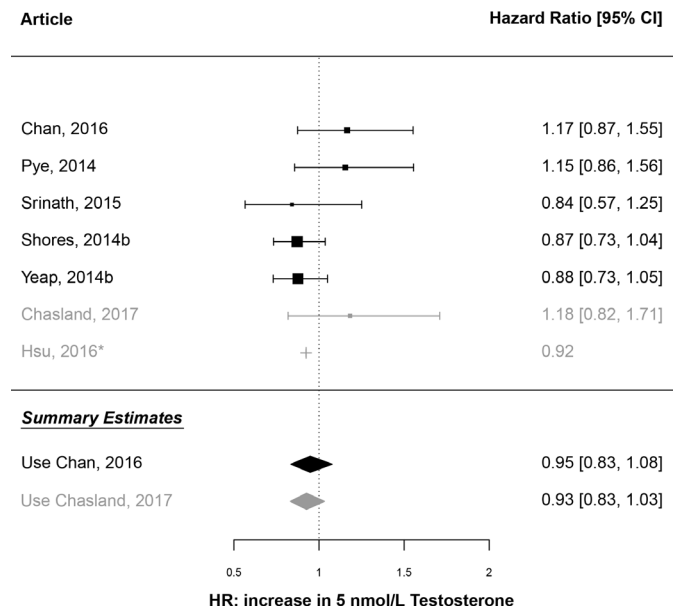


Figure 4 Forest plot of a meta-analysis of published estimates: association of testosterone with mortality caused by cardiovascular disease (CVD). Plotted values are the estimated HR for death from CVD, as attributed to an increase in endogenous testosterone concentration by 5 nmol/L. The vertical reference line is HR=1. Study-specific estimates are presented for six of the selected studies: BHS (Chan, 2016; Chasland, 2017)^{37 38}; EMAS (Pye, 2014)¹¹; ARIC (Srinath, 2015)⁴⁸; CHS (Shores, 2014b)⁵¹; HIMS (Yeap, 2014b)³⁰; CHAMP (Hsu, 2016).⁹ Summary estimates are colour coded as calculated using either the estimates from Chan *et al*³⁷ (black) or Chasland *et al*³⁸ (grey) for the BHS. *This estimate from Hsu *et al*⁹ could not be used to calculate the summary estimate because a variance estimate was not calculable for a 5 nmol/L change in testosterone using the published information.

demonstrated some significant associations with cancer outcomes (online supplemental figure 3, online supplemental table 9). Estimates showed an increased risk of lung cancer for men with higher concentrations,⁴¹ an increased risk of death from cancer for men with lower⁹ or the lowest (<8 nmol/L)¹¹ concentrations, and an increased risk of diagnosis for any cancer or for prostate cancer for men with the lowest (<10.17 nmol/L) concentrations of testosterone.³⁶ However, results were varied and not all articles reported these associations as being significant.³⁹ Furthermore, aside from an average increase in mini-mental state examination score (MMSE) of 0.067 per ng/mL decrease in testosterone concentration during follow-up,⁴⁰ there were no significant associations of baseline testosterone with cognitive function, or with change in cognitive function reported in the selected articles (online supplemental table 10).

DISCUSSION

The systematic review identified nine studies, and when combined with an additional two identified by colleagues, comprises 11 in total, with data for over 20 000 men from Australia, Europe, USA and the UK. Metaregressions

revealed significant heterogeneity in testosterone measurements at baseline, which was not explained by the mean age of participants among studies. However, DR-MA summary estimates demonstrated no significant effects of baseline testosterone on the relative hazard of death from any cause or from CVD, with negligible heterogeneity present. The DR-MAs, which suitably accounted for correlations between estimates for different exposure categories within studies, were of published estimates that had been adjusted for age, smoking status, and BMI or waist circumference. Furthermore, only published estimates from prospective cohort studies of community-dwelling men that had measured testosterone accurately using mass spectrometry and had observed at least 5 years of follow-up data were used. Despite some of these studies having reported an association between testosterone and mortality,^{9 30} the collective body of evidence demonstrated no overall associations of endogenous testosterone concentration with mortality or CVD mortality.

Previous meta-analyses investigating associations of endogenous testosterone with the health outcomes of interest looked at CVD outcomes,^{59–61} all-cause mortality⁵⁹ and prostate cancer.⁶² Boyle *et al*⁶² and Holmegard *et al*⁶⁰ both reported negligible heterogeneity in their estimates. Boyle *et al* found no significant association of a 5 nmol/L increase in testosterone with prostate cancer and Holmegard *et al* estimated a 43% increase in risk of ischaemic stroke for men with testosterone levels below the 10th percentile, as compared with men in the 11th–90th percentile range, from a meta-analysis of four articles.^{60 62} Ruige *et al* estimated an 11% decrease in risk of a CVD event from a SD increase in testosterone, and reported that significant heterogeneity was explained by larger effect sizes estimated for studies that recruited older men and for more recent articles.⁶¹ Araujo *et al* estimated a 35% increase in risk of all-cause mortality and a non-significant effect on CVD mortality from a 2.18 SD decrease in testosterone, although reported significant heterogeneity, and suggested that effects were driven by differences between the cohorts, such as underlying health status.⁵⁹ Two of these meta-analyses did not restrict selections to prospective cohort studies^{59 62} and none restricted selections based on testosterone assay method, although Ruige *et al*⁶¹ did find that assay method did not explain heterogeneity in that study.

The presented meta-analyses are the first to restrict selections to items of prospective cohort studies of community-dwelling men with testosterone measured using mass spectrometry, which is widely regarded as the reference method,²² and with at least 5 years of follow-up data. Accordingly, the presented summary estimates could arguably be viewed as the most reliable to date. These restrictions also resulted in the selection of a relatively small number of publications with estimates suitable for use in DR-MAs. Follow-up times for all-cause and CVD mortality ranged from a median of 4.3 years (total=5 years; EMAS)¹¹ to a mean of 14.9 years (total=16 years; BHS).³⁷ The number of incident deaths ranged from 147

(EMAS)⁴⁰ to 777 (CHS),⁵¹ or to 974 with the additional HIMS article³⁰ included. The number of CVD deaths ranged from 29 (ARIC)⁴⁸ to 264 (CHS),⁵¹ or to 325 with the additional HIMS article.³⁰ However, despite these differences, there was negligible heterogeneity in estimates and no significant funnel plot asymmetry detected.

Linear models were fitted because the HR estimates were reported for insufficient numbers of testosterone categories to have fitted non-linear DR-MA models. This was a key limitation of the analyses and likely to have resulted in an oversimplification of true effects. For instance, although the 95% CI for the Pye *et al*¹¹ study (calculated from HR estimates for quintile categories of testosterone) overlapped one, an alternative set of estimates in that article (which could not be included in the DR-MAs) reported a two-fold increase in the risk of all-cause mortality for men with very low testosterone (<8 nmol/L), as compared with 'eugonadal' men (>11 nmol/L). Pye *et al*¹¹ postulated that their reported differences in estimates might be reflective of a nonlinear association that emerges only when endogenous testosterone declines into the lower part of the range (<8 nmol/L). Furthermore, Yeap *et al*³⁰ estimated an 'U'-shaped association between endogenous testosterone and all-cause mortality, as consistent with a lower relative risk of health impacts for adult males with mid-range levels of testosterone. However, Shores *et al*⁵¹ also used non-linear modelling but did not find any significant associations of testosterone with all-cause or CVD mortality. Clearly, the investigation of non-linear associations is required to more comprehensively investigate the associations of testosterone concentrations with health outcomes in men.

In addition to the linearity assumption, there were other methodological limitations. Several articles reported estimated HRs per increase or decrease in SD and it was not possible to use these estimates in DR-MAs. Although it was possible to convert the per SD estimates to a standardised scale (ie, per 5 nmol/L increase), there was no information to determine adjustments to respective estimates of precision. Estimates for those studies could therefore not be included in the calculation of summary estimates and 95% CIs could not be calculated in forest plots. Summary estimates were calculated from a relatively low number (n=3–5) of articles and for most outcomes a summary estimate could not be calculated, which impacts on the generalisability of findings. Furthermore, these analyses were of observational data so summary estimates will not fully eliminate the possibility of confounding arising from unadjusted effects.

The implications of these findings are that associations of endogenous testosterone concentrations with key health outcomes should not be overstated, as they are not readily portrayed by meta-analyses of summary estimates. A more nuanced approach may be required, to capture non-linear or U-shaped associations.^{11 30} Also, while testosterone concentrations across ages were relatively stable when considering estimates from different cohorts, associations of testosterone with health outcomes

may differ with age, for example with all-cause mortality in middle-aged men³⁷ compared with older men.^{9 30} A deeper understanding of associations of endogenous testosterone concentrations with key health outcomes, would provide a foundation for analyses of the effects of exogenous testosterone, administered via therapeutic or pharmacologic interventions, on men's health.

IPD meta-analyses that incorporate flexible non-linear modelling techniques will provide improved scope to clarify the nature of such associations. The ability to apply a consistent statistical model to all studies, incorporate a more extended set of covariates than may have been included at the individual study level, and to estimate effects with increased statistical power, should result in more reliable summary estimates with reduced bias. Furthermore, other hitherto unpublished variables may be available for sharing by the collaborating studies to use in IPD meta-analyses, which could be useful for constructing analysis covariates or outcome variables. For instance, articles from the ARIC study that were identified from the systematic review reported on incident CVD event and death outcomes, but documentation on the ARIC study website shows that data on other prospective health outcomes, including cause-specific deaths and dementia diagnoses, are also available on request.⁶³ Although there have been recent advances with non-linear modelling methods for the meta-analyses of published estimates,^{32 64} sufficient information in the published articles, as is required for implementing these methods, was not available. In future work, estimates from analyses of the IPD-level data will be used to estimate and plot non-linear summary effects, and so will provide further improvements to estimates of associations between androgen levels and health outcomes in men.

Acknowledgements We thank Terena Solomons for valuable advice and guidance with conducting the literature search and screening steps of the systematic review.

Contributors All authors contributed to the design of the systematic review. RJM conducted the literature search and RJM and JH independently screened the returned items. RJM and KM conducted the statistical analyses. All authors were involved in manuscript preparation and subsequent revisions, and approved this submission. RJM is the study guarantor.

Funding This work was supported by: (i) Western Australian Health Translation Network Medical Research Future Fund Rapid Applied Translation Grant (2018), Grant number N/A; (ii) Lawley Pharmaceuticals, Western Australia, Grant number N/A.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Ross James Marriott <http://orcid.org/0000-0002-8805-8498>

Kevin Murray <http://orcid.org/0000-0002-8856-6046>

REFERENCES

- 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* 2016;85:891–901.
- 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;87:589–98.
- Handelsman DJ, Yeap B, Flicker L, *et al*. Age-Specific population centiles for androgen status in men. *Eur J Endocrinol* 2015;173:809–17.
- 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;101:1836–46.
- 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* 2018;88:88–95.
- Brand JS, Rovers MM, Yeap BB, *et al*. Testosterone, sex hormone-binding globulin and the metabolic syndrome in men: an individual participant data meta-analysis of observational studies. *PLoS One* 2014;9:e100409.
- Ding EL, Song Y, Malik VS, *et al*. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA* 2006;295:1288–99.
- Yeap BB, Marriott RJ, Antonio L, *et al*. Sociodemographic, lifestyle and medical influences on serum testosterone and sex hormone-binding globulin in men from UK Biobank. *Clin Endocrinol* 2021;94:290–302.
- 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.
- 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.
- Pye SR, Huhtaniemi IT, Finn JD, *et al*. Late-Onset hypogonadism and mortality in aging men. *J Clin Endocrinol Metab* 2014;99:1357–66.
- Tivesten A, Mellström D, Jutberger H, *et al*. Low serum testosterone and high serum estradiol associate with lower extremity peripheral arterial disease in elderly men. The MROS study in Sweden. *J Am Coll Cardiol* 2007;50:1070–6.
- Tivesten A, Vandenput L, Labrie F, *et al*. Low serum testosterone and estradiol predict mortality in elderly men. *J Clin Endocrinol Metab* 2009;94:2482–8.
- Yeap BB, Alfonso H, Chubb SAP, *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.
- Yeap BB, Hyde Z, Almeida OP, *et al*. Lower testosterone levels predict incident stroke and transient ischemic attack in older men. *J Clin Endocrinol Metab* 2009;94:2353–9.
- Yeap BB, Knuijan MW, Divitini ML, *et al*. Differential associations of testosterone, dihydrotestosterone and oestradiol with physical, metabolic and health-related factors in community-dwelling men aged 17–97 years from the Busselton health survey. *Clin Endocrinol* 2014;81:100–8.
- Yeap BB, Araujo AB, Wittert GA. Do low testosterone levels contribute to ill-health during male ageing? *Crit Rev Clin Lab Sci* 2012;49:168–82.
- Bhasin S, Pencina M, Jasuja GK, *et al*. Reference ranges for testosterone in men generated using liquid chromatography tandem mass spectrometry in a community-based sample of healthy nonobese young men in the Framingham heart study and applied

- to three geographically distinct cohorts. *J Clin Endocrinol Metab* 2011;96:2430–9.
- 19 Sikaris K, McLachlan RI, Kazlauskas R, *et al.* Reproductive hormone reference intervals for healthy fertile young men: evaluation of automated platform assays. *J Clin Endocrinol Metab* 2005;90:5928–36.
 - 20 Yeap BB, Alfonso H, Chubb SAP, *et al.* Reference ranges and determinants of testosterone, dihydrotestosterone, and estradiol levels measured using liquid chromatography-tandem mass spectrometry in a population-based cohort of older men. *J Clin Endocrinol Metab* 2012;97:4030–9.
 - 21 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.
 - 22 Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the Journal of clinical endocrinology and metabolism. *J Clin Endocrinol Metab* 2013;98:3971–3.
 - 23 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;18:5.
 - 24 Stroup DF, Berlin JA, Morton SC, *et al.* Meta-Analysis of observational studies in epidemiology: a proposal for reporting. meta-analysis of observational studies in epidemiology (moose) group. *JAMA* 2000;283:2008–12.
 - 25 Uzzani M, Hammady H, Fedorowicz Z, *et al.* Rayyan—a web and mobile APP for systematic reviews. *Syst Rev* 2016;5:210.
 - 26 Wells GA, Shea B, O'Connell DO. The Newcastle–Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses, 2020. Available: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [Accessed 10 Nov 2020].
 - 27 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:g4490.
 - 28 Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135:1301–9.
 - 29 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.
 - 30 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:E9–18.
 - 31 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.
 - 32 Crippa A, Orsini N. Multivariate Dose-Response Meta-Analysis: The *dosresmeta R* Package. *J Stat Softw* 2016;72:1.
 - 33 *estmeansd*: estimating the sample mean and standard deviation from commonly reported Quantiles in meta-analysis. R package version 0.2.1 2020.
 - 34 R Foundation for Statistical Computing. *R: a language and environment for statistical computing*. Vienna, Austria, 2020. <https://www.R-project.org/>
 - 35 Viechtbauer W. Conducting Meta-Analyses in R with the *metafor* Package. *J Stat Softw* 2010;36:3.
 - 36 Chan YX, Knuiam 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;9:391–8.
 - 37 Chan YX, Knuiam 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.
 - 38 Chasland LC, Knuiam 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.
 - 39 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.
 - 40 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.
 - 41 Chan YX, Alfonso H, Chubb SAP, *et al.* Higher dihydrotestosterone is associated with the incidence of lung cancer in older men. *Horm Cancer* 2017;8:119–26.
 - 42 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.
 - 43 Lee DM, Pye SR, Tajar A, *et al.* Cohort profile: the European male ageing study. *Int J Epidemiol* 2013;42:391–401.
 - 44 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.
 - 45 Tivesten Åsa, 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;64:1801–10.
 - 46 Kische H, Gross S, Wallaschowski H, *et al.* Associations of androgens with depressive symptoms and cognitive status in the general population. *PLoS One* 2017;12:e0177272.
 - 47 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;47:2682–8.
 - 48 Srinath R, Hill Golden S, Carson KA, *et al.* 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;100:1602–8.
 - 49 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;41:830–6.
 - 50 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.
 - 51 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.
 - 52 LeBlanc ES, Wang PY, Janowsky JS, *et al.* Association between sex steroids and cognition in elderly men. *Clin Endocrinol* 2010;72:393–403.
 - 53 Sueoka KT, Ewing MS, Ensrud KE. Higher endogenous testosterone levels associated with increased risk of coronary heart disease in elderly men: a prospective study. *Endocr Rev* 2010;31:S858.
 - 54 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.
 - 55 Kannel WB, Feinleib M, McNamara PM, *et al.* An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 1979;110:281–90.
 - 56 Li JJ, Wittert GA, Vincent A, *et al.* Muscle grip strength predicts incident type 2 diabetes: population-based cohort study. *Metabolism* 2016;65:883–92.
 - 57 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.
 - 58 Sterne JAC, Sutton AJ, Ioannidis JPA, *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
 - 59 Araujo AB, Dixon JM, Suarez EA, *et al.* Clinical review: endogenous testosterone and mortality in men: a systematic review and meta-analysis. *J Clin Endocrinol Metab* 2011;96:3007–19.
 - 60 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.
 - 61 Ruige JB, Mahmoud AM, De Bacquer D, *et al.* Endogenous testosterone and cardiovascular disease in healthy men: a meta-analysis. *Heart* 2011;97:870–5.
 - 62 Boyle P, Koehlin A, Bota M, *et al.* Endogenous and exogenous testosterone and the risk of prostate cancer and increased prostate-specific antigen (PSA) level: a meta-analysis. *BJU Int* 2016;118:731–41.
 - 63 Surveillance Dictionaries | atherosclerosis risk in communities, 2020. Available: <https://sites.csc.unc.edu/aric/surveillance-dictionaries> [Accessed 6 July 2020].
 - 64 Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized Dose–response data. *Stata J* 2006;6:40–57.