Testosterone Effects and Safety in Men with Low Testosterone levels (TESTES): An evidence synthesis and economic evaluation

Research protocol prepared by:

Miriam Brazzelli[1]

Katie Gillies[1]

Rodolfo Hernández[2]

Lorna Aucott[1]

Cynthia Fraser[1]

Siladitya Bhattacharya[3]

Channa Jayasena[4]

1 Health Services Research Unit, University of Aberdeen

2 Health Economics Research Unit, University of Aberdeen

3 School of Medicine, Cardiff University

4 Imperial College, London

NIHR HTA Programme (project no. 17/68/01)

Date of revision: 28/01/2020

Version: 2

Project Investigators

Channa Jayasena (Clinical Senior Lecturer, Imperial College, London)

Waljit Dhillo (NIHR Research Professor, Imperial College, London)

Nicholas Oliver (Consultant Diabetologist; Imperial College Healthcare NHS Trust, London)

Frederick Wu (Division of Diabetes, Endocrinology & Gastroenterology, University of Manchester)

Richard Quinton (Consultant Endocrinologist, Newcastle upon Tyne University & Hospitals NHS Foundation Trust).

Miriam Brazzelli (Senior Research Fellow, Health Services Research Unit, University of Aberdeen)

Katie Gillies (Senior Research Fellow & Director of Health Care Assessment Programme, Health Services Research Unit, University of Aberdeen)

Rodolfo Hernández (Research Fellow, Health Economics Research Unit, University of Aberdeen)

Lorna Aucott (Senior Statistician, Health Services Research Unit, University of Aberdeen)

Siladitya Bhattacharya (Head of School of Medicine, Cardiff University)

Patient representative: Justin Gold (London).

Background

Approximately 30% of men aged 40-79 years have low levels of circulating testosterone, the major male sex hormone produced in the testes. Testosterone is essential for sexual function, muscle growth and bone mineralisation, and has important behavioural effects in men. Low testosterone is associated with sexual dysfunction, hot flushes, reduced physical energy, cognitive and mood disturbance, reduced muscle strength, osteoporosis and increased body fat content. However, these symptoms are non-specific and may be caused by co-morbidities such as obesity and depression rather than low testosterone itself. Androgen replacement therapy (ART) has been used for decades to treat men with symptomatic low testosterone. However, serious concerns have been raised whether ART increases the risk of cardiovascular disease and prostate cancer in men with symptomatic low testosterone.

Symptomatic low testosterone may impair quality of life, cognition, mental health and daily function in affected patients. In men aged <40 years with low a priori risk of cardiovascular or prostate disease, ART can be given with low risks to health in most patients. However, an increasing number of middle aged and elderly men are being prescribed ART, and many of these men have co-morbidities, which might make therapy more risky. Middle-aged and elderly men and their clinicians therefore face uncertainty between symptomatic benefits and safety risks when taking ART.

Several RCTs have investigated the effects of ART in men with symptomatic low testosterone; however, they studied a variety of patient groups (e.g., baseline testosterone, patient age, comorbidities), and have used a range of validated symptom score questionnaires. A number of meta-analyses been performed for ART in men with symptomatic low testosterone. An important limitation of conventional meta-analyses of published results is the inability to estimate effects of treatment in different subgroups (e.g., older men, men with varying degrees of testosterone deprivation), as these are often not analysed and reported in sufficient detail in individual trials. A key advantage of our proposed IPD MA is the ability to determine which patient groups (stratified by patient age, symptoms, co-morbidities, baseline

testosterone) have the highest probability of experiencing benefits and adverse effects during ART.

A limited number of qualitative studies have also explored the perceptions of men who receive ART, but these data have not been systematically summarised previously. Understanding men's expectations and experiences of ART and the influence the therapy has on their quality of life would contribute further evidence to determining for whom this intervention may be most relevant and have the most potential for benefit. It is likely that the symptomatic effects of ART are dependent on factors including patient age, the severity of their low testosterone pre-treatment, and pre-existing comorbidities (e.g., depression, type 2 diabetes, poor mobility).

NHS prescriptions of androgen replacement therapy (ART, commonly known as 'testosterone therapy') for men have doubled since 2001, at an increased annual cost of £8M; however, the incidence of low testosterone remains unchanged.[1]

Published RCTs and systematic reviews of summary RCT data have yielded conflicting results regarding the safety and effectiveness of ART for men with low testosterone. As a consequence, deep divisions in clinical practice between practitioners are obvious, and the annual NHS costs of ART for men with low testosterone are escalating rapidly; these expose men with low testosterone to inconsistent treatment and potential harm.

In 2008, the US National Institute of Health (NIH) funded a series of interlinked, multicentre RCTs, investigating the safety and efficacy of ART in men with symptomatic lows testosterone (T Trials). Data from the T Trials have recently published[2-6] and profoundly change the balance of evidence in the field of ART. Results from the recently published T Trials have not been included in any of the published meta-analyses.

In conventional analyses based on aggregated data from published reports it is usually very challenging to get sufficient data to be able to undertake meaningful subgroup analyses (e.g., older men, men with varying degrees of testosterone deprivation). When subgroup analyses are presented, the definition of subgroups may vary across

individual trials and results may be reported inconsistently. The methodology of individual participant data (IPD) meta-analysis allows more robust evaluation of treatment effects in patient subgroups. Moreover, where trials have used different scales to measure outcomes, IPD permits meaningful translation between scales and useful combination of data. The IPD approach is also known to bring substantial improvements to the quality and quantity of data (e.g., by including more trials, participants or outcome measures).[7, 8] It also increases consistency across trials and enables detailed data checking.[8] At present, no IPD MA has investigated the clinical effectiveness of ART in symptomatic men with testosterone deficiency. The key advantage of IPD MA is the ability to determine which patient groups (stratified by patient age, symptoms, co-morbidities, baseline testosterone) have the highest probability of experiencing benefits and adverse effects during ART.

Inclusion of IPD from the T Trials, as well as cost effectiveness and qualitative data offer a unique opportunity to enhance the quality and quantity of the current evidence base and improve clinical decision making around the treatment of men with testosterone deficiency.

We propose to conduct an IPD meta-analysis to identify for the first time which specific patient groups will most benefit from ART, and which have the highest risk of harmful effects. The IPD meta-analysis will be complemented by i) a qualitative evidence synthesis to ascertain the experience and motivations of men using ART, and ii) the development of an economic model to inform decision making regarding use of ART in men with low testosterone.

Aims and objectives

To determine the clinical effectiveness, safety, cost-effectiveness and acceptability of androgen replacement therapy (ART) in men with testosterone deficiency.

Specific objectives are as follows:

i) To conduct a comprehensive systematic review and Individual Participant Data (IPD) meta-analysis to estimate the clinical effectiveness and safety of ART for men with testosterone deficiency syndrome and to provide the key parameters for the development of a decision model;
ii) To conduct a systematic review of existing qualitative evidence, which reports men's experience and acceptability of ART, and an analysis of patient reported outcome measures (PROMs);
iii) To develop a decision model to estimate the cost-effectiveness of ART for the treatment of symptomatic men with low testosterone.

Methods

i) Systematic review and IPD meta-analysis

Criteria for considering studies for this evidence synthesis

Types of studies

Evidence will be considered from randomised placebo-controlled clinical trials evaluating the effects of ART in men with testosterone deficiency. Only trials with a duration of at least 3 months for all intervention groups will be considered suitable for inclusion. This is in line with the current recommendation of the Endocrine Society Clinical Practice Guideline, which recommends evaluating men 3 to 6 months after ART initiation and then annually thereafter.[9]

Target population

Men presenting with a proven low level of serum testosterone.

Participant characteristics: Studies must include men with all of the following characteristics:

- Aged 18 years or over with no upper age limit
- Low levels of testosterone. There has never been a consensus definition of a low testosterone, which is reflected by the participant characteristics of trials in this field. However, all current clinical guidelines are in broad agreement that men with a serum level of total testosterone >12nmol/L are unlikely to have clinical features of low testosterone.[10] This criterion will be adopted in the proposed project. The following information should be available for included studies must be available for trials to be considered for inclusion:
  - When samples were collected and assayed (since dates may differ)
  - Details of any extraction method used prior to testosterone assay
  - Details of the assay method and manufacturer
  - Details of any local correction made to adjust assay measurements
  - Relevant local validation data for the assay e.g. external quality assurance

Intervention

Androgen Replacement Therapy (ART) with any testosterone formulation, dose, frequency and route of administration (e.g., intramuscular, subdermal, transdermal, oral and buccal preparations of testosterone). Studies that use other androgens apart from testosterone and studies allowing concurrent treatment with other hormones will not be deemed suitable for inclusion.

Setting

Any relevant clinical setting (e.g., primary care, secondary care).

Outcome measures

We anticipate studies will provide data on any of the following outcome measures:

- **Sexual function** e.g., self-reported early morning erections, ability to maintain erection during intercourse, frequency of intercourse. Where possible, these will

be quantified by validated scores such as, but not limited to, the International Index of Erectile Function (IIEF).

- **Physical parameters** e.g., muscle mass and strength, exercise tolerance, body weight, body mass index, total lean body mass, fat mass.
- **Functional activities** e.g., running, walking, kneeling; quantified where possible by validated scores such as the SF-36.
- **Psychological symptoms**: e.g., cognition, mood and behaviour assessed by validated scores.
- **Cardiovascular and cerebrovascular events** e.g., fatal and non-fatal myocardial infarction, acute coronary syndrome, fatal and non-fatal stroke, transient ischaemic attack
- **Other co-morbidities** e.g., diabetes mellitus, psychiatric disease, hypertension, dyslipidaemia, erectile dysfunction, obstructive sleep apnoea, reduced bone mineral density (osteoporosis or osteopenia).
- **Prostate-related outcomes** e.g., prostate-specific antigen levels, prostate volume, increase in the International Prostate Symptoms Score
- **Physiological markers** e.g., blood pressure, haemoglobin concentration, haematocrit; total serum lipid profile, plasma glucose, bone mineral density.
- **Quality of life** measured through validated scores, whether generic and/or disease-specific.
- **Mortality** from any cause during the study period.

As many outcomes will be assessed by a variety of tools, we will restrict inclusion to validated scales or measurement tools only. Outcomes, which will be included in the IPD meta-analysis, will be limited to those identified as most pertinent by the Advisory Group for this project. In particular, primary outcomes for the IPD meta-analysis will be:

- **Sexual function**
- **Adverse events** e.g., Major Adverse Cardiac Events (MACE), type 2 diabetes, fractures, prostate cancer
- **Quality of Life**

Search methods for identification of relevant RCTs

Comprehensive literature searches, using an appropriate combination of controlled vocabulary and text terms, will be conducted to identify reports of published, ongoing and unpublished studies reporting the clinical effectiveness of ART in men with testosterone deficiency. Highly sensitive search strategies will be designed, using appropriate subject headings and text word terms, the clinical intervention under consideration and relevant study designs. The searches will be conducted from 1992 (year of the first published randomised placebo controlled study of testosterone administration) to the present, in order to reflect the introduction of ART in clinical practice, and restricted to reports published in English. In particular, the Cochrane Highly Sensitive Search Strategy for identifying randomised controlled trials will be used in MEDLINE and adapted for other electronic databases. The following databases will be searched to identify relevant clinical trials: MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, MEDLINE Epub Ahead of Print, EMBASE, Science Citation Index, and the Cochrane Controlled Trials Register (CENTRAL. Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Review of Effects (DARE) and the HTA databases will be searched for evidence syntheses. Recent conference proceedings of key professional organisations in the fields of endocrinology (e.g., American Endocrine Society), cardiology (e.g., American College of Cardiology), and men's health (e.g., European Menopause and Andropause Society, International Society of Men's Health).

Reference lists of all included studies will be perused in order to identify additional potentially relevant reports. We will also contact our panel of experts for details of any additional potentially relevant reports. A preliminary MEDLINE search strategy is detailed in Appendix 1.

Ongoing studies will be identified through searching Current Controlled Trials, Clinical Trials and WHO International Clinical Trials Registry. Websites of professional organisations, regulatory bodies and HTA organisations will also be searched to identify additional relevant reports.

Inclusion of studies

Study selection

Two reviewers will independently screen the titles and abstracts of all citations identified by the search strategies. Full text copies of all potentially relevant studies will be retrieved, and assessed independently by the same two reviewers for eligibility. Any disagreements will be resolved by discussion or arbitration by a third reviewer. References will be stored using the Endnote software. Studies that do not meet the inclusion criteria will be excluded. Their bibliographic details will be listed in an appendix and main reasons for exclusion will be provided (e.g., 'not a RCT', 'not appropriate intervention', 'inadequate duration of treatment').

Risk of bias assessment

We will assess the risk of bias of included RCTs by means of the Cochrane risk of bias tool.[11] Two reviewers will independently assess the risk of bias of each included trial according to the following domains: sequence generation (selection bias), allocation concealment (selection bias), blinding of participants, personnel and outcome assessors (performance and detection bias), incomplete outcome data (attrition bias), selective outcome reporting (reporting bias), and other potential thread to validity. Judgements about risk of bias for each of the domains in the tool will be based on the criteria detailed in the Cochrane Handbook for Systematic Reviews of Interventions.[11] In all cases an answer of YES will indicate a low risk of bias, an answer of NO a high risk of bias. If insufficient information are available the judgement will be UNCLEAR risk of bias. To establish an overall risk of bias we will consider the assessment of each individual domain as well as the relative importance of different domains for the current evidence synthesis. Any difference in the assessment of the risk of bias of included RCTs or any issue of uncertainty will be resolved by discussion and consensus between reviewers. We will also seek additional information from the authors of selected trials through a collaborative approach in order to improve the risk of bias assessment of each individual trial. In particular, we will gather further information on randomisation procedures and blinding as well as on completeness of outcome data.

Data collection and data checking

We aim to establish a collaborative group of all trials investigators. Authors of eligible studies will be invited to join the collaboration by providing the individual participant data to be included in the IPD meta-analysis. For each relevant RCT, we will identify contact information from the published report of the trial and through electronic searches. We will initially contact the principal investigators (corresponding authors) of eligible trials by email and provide them with a brief summary of the project and a cover letter explaining the rationale, objectives and general plan of the project. Reminders will be sent to non-responders after one week. If reminders do not prove to be successful we will attempt other communication channels (e.g. letter, phone) or to contact other investigators. After obtaining a memorandum of understanding, preferably electronically, trials investigators will be asked to agree to transfer and share their anonymised data by signing a Data Transfer and Sharing Agreement form, which specifies that the data will be anonymous, stored securely and used exclusively for the purpose of the project, with access restricted to the members of the project team. At present, we have contacted the authors of the relevant RCTs identified by our scoping search and published within the last 10 years. An informal agreement to collaborate have been received from all authors of 35 trials, and no authors have declined to collaborate (see Appendix 2).

The procedures for collection, organisation and checking of data will be coordinated by the project team based at the University of Aberdeen. Trials investigators will be asked to provide anonymised data (without information such as name or date of birth) for all randomised patients. We will seek outcome data for all participants at all relevant time points (six months and over) together with information on baseline patients' demographic and clinical characteristics. Variables not reported in the published trial reports will be requested as they may be useful for the conduct of subgroup analyses. We will provide authors with a list of data items we definitely require. To ensure maximum participation, trials investigators will be allowed to supply data in whatever format convenient to them. The methodological team at the University of Aberdeen will take responsibility for converting the data to the required format. All data supplied will be subjected to a range of consistency checks. Any missing data, errors and inconsistencies between variables or outlying values will be queried and rectified, where necessary, through direct communication with the trial

investigators. The investigators will be asked to confirm accuracy of the individual trial data. Any received data set will be also compared with the existing published reports. These consistency checks will ensure that we have the most up-to-date, unbiased data on the effects of ART for the treatment of men with symptomatic testosterone deficiency. After completion of the consistency checks, we will combine individual trial datasets into a master dataset; we will add a variable/code to allow identification of the original trial.

Secure methods will be put in place to ensure secure transfer and storage of data (encrypted/password protected files; secure computer server). Access to the data will be limited to the members of the research team based at the University of Aberdeen who work directly on the project. Study data will not be used for any other purpose than that of the project. Copying data on memory sticks, tablets, or smart phones will be prohibited.

Data analysis

Aggregate data

The main source of data to assess the effects and safety of ART will be the IPD. However, if by month 14 of the project some authors have not agreed to share and transfer their data we will proceed with the analyses of the available data sets. We will combine available IPD with aggregate data when IPD are not available. We will compare the results obtained from these analyses with those obtained from analyses of IPD only and of aggregate data only by means of sensitivity analyses. For any eligible trial for which we will not able to collect IPD, aggregate data will be extracted from the trial published report. A data extraction form will be developed and piloted for this purpose. Information on study design, characteristics of participants, settings, characteristics of interventions and outcome measures will be recorded. One reviewer will complete the data extraction form for all relevant trials and a second reviewer will check the data extracted. Any disagreements will be resolved by discussion or arbitration by a third reviewer.

IPD meta-analysis

The following section provides information on the main IPD analyses we propose to conduct. However, in view of the complexity of the IPD analyses some adjustments or

additional analyses may be necessary during the course of the project. We will rely on the current recommendations for performing IPD meta-analyses.[7, 8, 12, 13]

We will perform IPD meta-analysis for the pre-specified outcome measures from all eligible trials. In particular, the IPD approach will allow us to investigate whether the observed effects of ART are consistent across participants with certain characteristics, for example men of specific age group or with concomitant disease (e.g., diabetes) or with a higher baseline risk (e.g., at risk of cardiovascular/cerebrovascular events).
A feature of the IPD approach is to preserve the clustering of participants within trials. Two methods are currently recommended to retain clusters during statistical analyses: the two step or one step approach. With regard to the two-step approach, estimates of effects are initially derived from IPD for each trial using statistical methods appropriate for the type of data being analysed. In the second stage aggregate data for each trial are synthesised using suitable model for meta-analysis of aggregate date (e.g., assuming fixed or random effects across trials). Meta-analyses results are displayed on forest plots. The one-step approach permits modelling of IPD from all trials simultaneously while stratifying for or taking into account the differences between trials; it is better for non-normal outcomes and is less affected by some or many studies being small, but can be computationally intensive and prone to convergence problems. The two approaches often produce similar meta-analyses results especially if the study estimates are approximately normally distributed with known variances but may differ for several reasons.[7, 8] However, a one-step approach, which allows for most sophisticated modelling of covariates and has the best performance in terms of power,[14] will be the preferred approach for this project. This may need revision after consideration of the data collected for the selected trials and will depend on the outcome of the collaborators workshop. All analyses will follow the principle of intention-to-treat as closely as possible. We will include all randomised participants with outcome data. Dependent on the specific outcome date type, the analyses will typically be regression models (such as linear, logistic, survival, Poisson or non-parametric equivalents should they be more appropriate) with either a separate term for each trial or one that varies across trial via a random effect.[8] Patient-level factors not used in the randomisation will be determined as being fixed or random prior to any analysis plan as appropriate. A random effects approach to the intervention effect will be the preferred over a fixed effect approach; however, if the

between studies standard deviation is very low, fixed effect one-stage models will be considered to reduce failure of model convergence. For any time-to-event outcome, appropriate models, which take into account censored data, will be used (such as Cox regression model).

The impact of participants level characteristics (such as age of participants e.g., <60, 60-75, >75 years ; level of total serum testosterone at baseline e.g.,<6nmol/L, 6-8nmol/L and 8-12nmol/L; presence of baseline co-morbidities e.g. diabetes mellitus, hypertension, cardiovascular disease, obstructive sleep apnoea, dyslipidaemia, prostate disease and / or urinary symptoms) and trial-level characteristics (such as duration of treatment e.g., <6 months, 6-12 months, > 12 months, route of administration: oral, transdermal, injection) will be assessed if possible, initially, by grouping trials and participants into subgroups for each covariate of interest and performing meta-analyses within each subgroup. Where possible a one-step approach will be used to assess the interactions between treatment covariates.

Every effort will be made to minimise the amount of missing data. We will request information on any enrolled participants who were subsequently excluded from the original trials. Where data are missing for some participants in the master dataset, a complete data analysis will be conducted in the first instance. If there are substantial missing data (10% for any relevant outcome or covariate), sensitivity analyses will be considered to assess the impact of missing data.

Publication policy

Any publication outputs from the IPD meta-analysis will be in the name of the collaborative group, with all contributors listed. All trial principal investigators will be invited to attend a collaborators workshop. The purpose of the workshop is to present and refine the research protocol, discuss which data are to be collected, reach an agreement on data checking procedures, statistical analyses to be performed, project timetable and publication policy. Throughout the project, communication (via email, phone and tele conference) will be maintained with all collaborators. A meeting will also be organised in the second year of the project to discuss preliminary results.

ii) Synthesis of qualitative evidence and analysis of patient reported outcomes

Understanding the experiences and expectations of men (their partners or their healthcare professionals) in relation to ART could give critical insight into how certain factors enable or disable effectiveness of interventions, especially for those outcomes which are patient reported. Whilst the key focus of this project is to assess the clinical benefits, risks and costs of ART, a synthesis of qualitative studies reporting men's (or relevant others) experiences of ART may further elucidate aspects of interventions not considered previously. In addition to the primary qualitative studies exploring men's ART experience, we will analyse existing patient reported outcome measures (PROMs) related to low testosterone, then will investigate and compare patient-relevance of each PROMs included items. Determining the relevance of items (to men on ART) included in PROMs will add further depth to the quantitative data by ensuring that existing PROMs capture relevant items important for men to live well with low testosterone.

Criteria for inclusion of eligible studies

The initial scoping search developed for this evidence synthesis will be further refined and run across several databases from 1992 (year of the first published randomised placebo controlled study of testosterone administration) till present. Main electronic databases will include: MEDLINE, EMBASE, CINAHL, ASSIA, and Psycinfo.
The parameters of the search and identification of eligible studies for the qualitative review will be defined using the SPIDER tool, a PICO alternative for application in qualitative or mixed-methods research syntheses to optimise identification of qualitative studies.[15] The SPIDER tool specifies the research question by identifying Sample, Phenomenon of Interest, Design, Evaluation, and Research type. For the purposes of this evidence synthesis these items will be defined as follows:

- *Sample:* Men, their partners or health professionals, who are eligible for ART to treat low testosterone.
- *Phenomenon of interest:* Androgen Replacement Therapy for men with low testosterone.
- *Design:* Any qualitative method for collecting data; interviews; focus groups; observations; case studies; surveys.

- *Evaluation:* Understanding experiences, opinions, views, beliefs, attitudes, or expectations in relation to ART
- *Research type:* Any primary studies that have explored any aspect of ART for low testosterone from the perspective of men, their partners, or their clinicians. Mixed methods studies will be included if the qualitative element's methods and results are reported separately.

Eligibility of studies

One reviewer will independently assess all the citations identified by the literature searches. Secondary screening will be carried out by another reviewer who will screen a random sample (20%) of the identified citations. Copies of all potentially relevant articles, which meet the pre-specified inclusion criteria or for which there is insufficient information in the title and abstract to make a decision, will be retrieved in full. Any disagreement between reviewers on the eligibility of included articles will be resolved through discussion or arbitration by a third reviewer (KG).

Data extraction

Data will be extracted independently by one reviewer using a data extraction from developed *ad hoc* for the purpose of this evidence synthesis. Double data extraction by another reviewer will be carried out on a random sample (20%) of studies. Reviewers will review extracted data together to assess consensus and ensure all relevant information has been collected. Disagreement will be resolved through discussion or consultation with a third reviewer (KG). Information on objectives, methods, characteristics of the patient population, characteristics of the intervention, and outcomes in terms of patient's perspective and experience will be recorded for each relevant study.

Assessment of quality

Qualitative data synthesis

Several well described approaches exist for synthesis of primary qualitative studies, with the rationale for choice of specific methods often driven by the overall objective of the synthesis. The purpose of the qualitative synthesis in this review is to identify data-driven themes within existing primary studies which have relevance in terms of

experiences and expectations of ART in men with low testosterone and may provide explanatory data to support intervention effectiveness. For that reason we are proposing to use thematic synthesis as an appropriate method using both inductive and deductive approaches to analysis.[12] The formal thematic analysis of the content of the included studies will start with close reading of the publications to identify main recurring 'descriptive' themes, followed by the generation of higher level 'analytical' themes capturing the phenomena described across the identified literature. Finally, we will aim to map the relationships between the analytical themes to summarise the existing qualitative evidence 'landscape'. This process will involve constant comparison of the emerging constructs within the data from the analysed publications. The analysis will be conducted by two reviewers, with the initial reading and coding undertaken independently with any disagreements discussed until consensus is reached. If studies are identified that report the perspectives of multiple stakeholders (i.e., men, their partners and their healthcare providers) individual analyses will be conducted by considering each group in isolation to generate both descriptive and analytical themes. The data would then be considered in juxtaposition to allow comparisons and contrasts to be developed. However, the perspectives of men would always be considered as superior to other groups and would be brought to the fore in the synthesis.

Analysis of Patient Reported Outcome Measures (PROMs)

A published review of the measurement properties of health related (generic and disease specific) quality of life instruments previously identified four disease specific measures of testosterone deficiency.[16] However, whist this review examined the clinical face validity of these instruments it did not compare the content of items across instruments. Such an investigation would examine the homogeneity of these disease-specific instruments and provide evidence as to whether such instruments measure outcomes that can be meaningfully combined in a meta-analysis and their relevance to men with low testosterone.

Identification of eligible studies

Disease specific patient reported outcome measures will be identified from the previous review of measurement properties and from trials included in the IPD meta-analysis and will be collated together.

Data extraction

Data will be extracted independently by one reviewer. Double data extraction by another reviewer will be carried out on a random sample (20%) of studies. Reviewers will then review extracted data together to assess consensus and ensure all relevant information has been collected. Disagreement will be resolved through discussion or consultation with a third reviewer (KG).

Data will be extracted on the name of the PROM(s), the reported PRO scales and individual verbatim items.

Data Analysis

Analysis will be informed by previous studies that have analysed PROMs into individual outcome domains.[17, 18] The individual verbatim items from each PROM will be analysed by using an inductive content analysis approach. All PROM items will be examined and systematically categorised into conceptual health domains according to the aspect which they aim to capture. Health domains will be generated inductively from the identified individual items, but likely focus on aspects of sexual health, mental health, and physical and social functioning. Domains will be further defined until all individual items are mapped onto a domain. Domain mapping will be conducted by two reviewers independently with any conflicts resolved through discussion or inclusion of a third reviewer as appropriate. Synthesising the content of individual items from PROMs in this way can provide a framework for future PROM development or, in this case, ensures domains contained within the tool (and therefore measured in trials) will be used to present the outcomes and investigate their relevance for men by mapping back onto the qualitative synthesis.

iii) Economic evaluation of ART in men with low testosterone

In order to make informed decisions regarding the optimal management for men with symptomatic testosterone deficiency, information is required on the cost-effectiveness of ART. The economic evidence on ART will be assessed through a systematic review of economic evaluations as well as a new model based economic evaluation comparing ART with standard care (e.g., no treatment). A cost-utility analysis will be conducted following best practice in decision modelling.[19-22] The analysis will adopt a

NHS and personal and social services perspective on costs, and consider health consequences for patients over a lifetime horizon.[21]

Systematic literature review of economic evaluations

Inclusion criteria

Full economic evaluation will be included. These are studies reporting cost and consequences of at least two alternative care pathways (i.e., ART compared to 'standard care' – no treatment) Cost-consequences, cost-effectiveness, cost-utility and cost-benefit analysis will be included.

Search strategy

Sensitive electronic searches using an appropriate combination of controlled vocabulary and text terms will be developed, which assess ART in the treatment of men with symptomatic low testosterone. MEDLINE, Embase, NHS Economic Evaluations Database (NEED), the HTA Database, Cost-effectiveness Analysis Registry, and Research Papers in Economics (RePEc) will be searched from 1992 onwards. A draft MEDLINE strategy is reproduced in Appendix 1 and will be adapted for other databases. Recent conference proceedings of key professional organisations in the fields of endocrinology (e.g., American Endocrine Society), cardiology (e.g., American College of Cardiology), and men's health (e.g., European Menopause and Andropause Society, International Society of Men's Health) for the last three years (2016-2018) will also be scrutinised as well as the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Scientific Presentations Database. Reference lists of all included studies will be scrutinised and experts in the field contacted for details of additional reports.

Study selection and data extraction

After electronic deduplication, two reviewers will independently screen titles and abstracts. All studies identified for full text will be retrieved. Two reviewers will independently select studies for final inclusion. Discussion will be used to resolve any disagreement. If disagreement persists, a third member of the project team will be addressed to make a final decision on the studies to include.

Data from the included studies will be extracted by one reviewer following the Drummond economic evaluation checklist.[23] This checklist will be complemented by the Philips checklist for decision model based studies.[24] A second reviewer will check the data extracted for 20% of the studies for mistakes and consistency.

Study quality will be assessed using the Drummond and Philips critical appraisal questions.[23, 24] Results will be reported in a narrative manner with no attempt to synthesise quantitatively the extracted data. Applicability and generalisability of the results to the UK setting will be considered when reporting the findings of this evidence synthesis. The focus will be on cost-effectiveness of ART but also on the identification of those care pathway characteristics that could improve the cost-effectiveness of ART.

Model based economic evaluation

The model structure will incorporate relevant care and event pathways for individuals with low testosterone, informed by existing guidelines,[10] the IPD meta-analysis, and discussions with experts in the project Advisory Group. A final agreement on the model structure and strategies will be sought from this group. It is anticipated that the model will need to consider a number of important event pathways such as cardiovascular and cerebrovascular disease, diabetes, and sexual function. We therefore expect that an individual sampling model will be required in this instance. While these models are often more resource and data intensive to develop than Markov models, they are better suited to modelling and tracking multiple event histories and comorbidities.[25] This type of model can also make the best use of the individual participant data to reflect heterogeneity in risks and outcomes.

The economic model will include costs, clinical outcomes and health state utility values associated with modelled outcomes and events. Model input data on clinical outcomes will be primarily obtained from the IPD meta-analysis. We will coordinate the economic and statistical analyses to ensure that the probabilities of clinically and statistically significant events are incorporated in the model appropriately. Primary care resource use (e.g., general practitioner visits, diagnostic tests requested by a GP) and secondary care resources use (e.g., outpatient visits) associated with administration and monitoring of the alternative treatment strategies will be informed

by guidelines and expert opinion. Resource use associated with the clinical events and care pathways included in the model will be sourced from focussed reviews of costing and studies and existing economic models. Standard sources, such as NHS reference costs,[26] the British National Formulary[27] and the Unit costs of Health and Social Care[28]will be used to value resource use. We will follow a micro-costing approach to estimate cost of ART according to the mode of administration, dose and frequency. Health state utility values for the cost-utility analysis will be obtained from structured reviews of the literature if suitable data are not available from the IPD analysis. We will use these to model quality adjusted life years (QALYs) accruing to patients under the alternative treatment strategies. Expected costs and QALYs will be estimated for each model strategy. Cost-effectiveness will be expressed in terms of the incremental cost per QALY gained with ART versus no ART. Uncertainty will be addressed using deterministic and probabilistic sensitivity analysis. The latest methodological guidelines will be followed to define probability distributions for input parameters in the model.[21] The access to IPD presents an opportunity to characterise uncertainty in the decision model more accurately compared to models based on standard aggregate data meta-analysis[29, 30] i.e., the IPD will be used to estimate correlations between key model input parameters. Results for the probabilistic analysis will be reported as point estimates but also using cost-effectiveness scatterplots and cost-effectiveness acceptability curves.[23]

Integration of findings from the evidence syntheses and economic model

Individual outputs from the data analysis in each work package described above will be synthesised together (through juxtaposing in a matrix) to provide an overall summative conclusion considering the integrated evidence. The benefits of this integration could have significant potential in providing a more in-depth understanding of the individual components. We are unaware of any other evidence syntheses that have combined individual patient data evidence with other forms of review evidence. Several methods exist for synthesising quantitative and qualitative research evidence in a systematic review but as yet no one single method has been has been agreed.[31] The datasets, which will be both integrative and interpretive, will be combined and juxtaposed (i.e., discussed side- by side) as appropriate to produce a detailed narrative summary that contributes to a more nuanced understanding of differences in outcomes across sub-groups through a focus on patient-centred

outcomes and experiences relevant to men on ART. Narrative summaries can vary in their methodology.[31] We are proposing to utilise a process that follows an interpretive approach that includes explicit reflexive accounts, which can provide opportunities for higher levels of data construction and unpack complex dynamic processes or experiences. The narrative synthesis conducted will be informed by existing guidance in the area.[31] Specifically, integration will be developed using a framework covering the following key steps: developing a theory of how the intervention works, why and for whom; developing a preliminary synthesis; exploring relationships within and between studies (or in this case, data types); assessing robustness of the synthesis product.[31] Ultimately, this integration will allow the intervention effects (both clinical, patient experience and economic) to be viewed collectively which can better inform a holistic approach to the use of ART in men with low testosterone.

Project management

The project will be managed through a multidisciplinary Steering Group, which includes members from the Health Services Research Unit and the Health Economic Research Unit at the University of Aberdeen, clinicians from the Imperial College, London and the School of Medicine, University of Cardiff, and through a Project Team, which comprises senior and junior staff at both the University of Aberdeen and the Imperial College. The Project Team will be responsible for contacting and liaising with trials investigators; collecting and checking data; conducting statistical, qualitative and cost-effectiveness analyses; interpreting and disseminating results.

The academic staff at both institutions, the University of Aberdeen and the Imperial College, London, will adhere to standard University procedures (e.g., registration of the research protocol on PROSPERO database, confidentiality of data collection and secure storage of data and final report production).

Advisory Group

In addition to the members of the Steering Group and the Project Team, an Advisory Group comprising of methodologists, health economists, endocrinologists, and lay members have been set up to provide guidance on the care pathways, advise on important outcomes, and assist in the interpretation of the clinical effectiveness

findings. The Advisory Group will be convened at least three times during the duration of the project.

Ethical approval

As secondary data will be used for the IPD meta-analysis, no formal ethical approval will be required as informed consent has already been obtained by the investigators of the original trials, and our meta-analysis will address very similar research questions to those for which the data were originally collected and to which patients gave consent.

Patient and Public Involvement

Roles of patient representatives: To ensure the perspective of patients is central to the project, two patient representatives will be actively involved in it. Their expertise will be required for the following:

- Interpreting study findings (e.g., qualitative data, cost-effectiveness model).
- Co-leading a focus group to explore the implications of the study results for men with symptomatic low testosterone.
- Advising on the content of scientific publications and presentations.
- Contributing to the integration of results in the final study report.
- Contributing to the writing of the Plain English Summary.

A member of the Project Team will hold regular meetings with the patient representatives to review study progress and address concerns.

Dissemination plans

Our strategy involves disseminating the results of our research to researchers (through sharing of data and results with the aim to improve quality of research); to health care professionals (through provision of evidence which may improve clinical practice and the management of men with testosterone deficiency); to policy makers (through provision of evidence which may contribute to clinical guidelines for the management of men with testosterone deficiency); to patients, carers and members of the public. This project will build an international consortium of clinical investigators led by experts on testosterone replacement from multiple centres in the UK. By definition,

the outputs of this project would constitute a consensus statement likely to attract publication in a major academic journal, and change clinical guidelines of professional bodies such as the Society for Endocrinology, and subsequently policy bodies such as the NICE Technology Assessment Committees (TAC).

We will seek advice and recommendations from our Collaborators Group, which comprises authors of existing RCTs from different countries, as well as members from our Advisory Group, which includes also lay members, on the best and most effective way to disseminate our findings in order to ensure they are accessible to patients and carers as well as to policy makers, health professionals and their professional bodies.

The planned output of this research will be a detailed final report, which will be published in the NIHR Journals Library and in peer reviewed scientific journals, including open access publications. This will ensure that the research is reported fully and is publicly available. Members of the project team will present study results at relevant national and international meetings (e.g., annual NICE conference; annual NHS Scotland conference; Society for Endocrinology annual meeting; British Menopause Society annual meeting; European Menopause and Andropause Society annual meeting; Royal College of GPs annual conference). In collaboration with members of the research team, the patient representatives will be invited to co-lead a clinical symposium at the Society for Endocrinology (SFE) BES 2020 annual meeting, which will involve endocrinologists, nurses and patients.

Timetable

Project duration is 24 months with a starting date of 1st September 2018. Key project milestones are as follows:

Pre Grant (6 months): Protocol Development, Scoping searches, Initial contact with trials' collaborators.

TESTES Project Management Plan

	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
Recruitment of RFs, RA & health economist																									
Registration of protocol in PROSPERO																									
Consolidation of Advisory Group																									
Establish formal collaboration with trials' investigators																									
Commence literature searching																									

	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
Design data extraction forms																									
Discuss economic model scope																									
Systematic reviews of quantitative data:																									
Complete literature searching																									
Data extraction																									
Qualitative assessment																									
IPD collection																									

	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
IPD data checking																									
Systematic reviews of qualitative data:																									
Complete literature searching																									
Data extraction & summary																									
Comparison of main themes across reports																									
Economic evaluation & modelling:																									
Complete literature searching																									

	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
SR of economic evaluations: selection & data extraction																									
Conceptual modelling																									
Modelling implementation starts																									
Complete data collection & data checking																									
Data analysis																									
Economic model verification & calibration																									
SR of economic evaluation complete																									

	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
Complete IPD meta-analysis																									
Commence mixed-methods synthesis																									
Comparison of emerging theoretical constructs																									
Independent identification and coding of themes																									
Economic modelling: run model & sensitivity analysis																									
Prepare final report																									
Prepare papers for journal publication																									

	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
Symposium at national meeting																									
Meetings during the course of the project																									
Regular team meetings (dates TBC)																									
Three Advisory Group meetings (dates TBC)																									
Two collaborators meetings/workshops (dates TBC)																									

References

1. Gan EH, Pattman S, S HSP, Quinton R. A UK epidemic of testosterone prescribing, 2001-2010. *Clin Endocrinol* 2013;**79:**564-70.

2. Budoff MJ, Ellenberg SS, Lewis CE, Mohler ER, Wenger NK, Bhasin S, et al. Testosterone treatment and coronary artery plaque volume in older men with low testosterone. *JAMA* 2017;**317:**708-16.

3. Resnick SM, Matsumoto AM, Stephens-Shields AJ, Ellenberg SS, Gill TM, Shumaker SA, et al. Testosterone treatment and cognitive function in older men with low testosterone and age-associated memory impairment. *JAMA* 2017;**317:**717-27.

4. Roy CN, Snyder PJ, Stephens-Shields AJ, Artz AS, Bhasin S, Cohen HJ, et al. Association of testosterone levels with anemia in older men a controlled clinical trial. *JAMA Intern Med* 2017;**177:**480-90.

5. Snyder PJ, Bhasin S, Cunningham GR, Matsumoto AM, Stephens-Shields AJ, Cauley JA, et al. Effects of testosterone treatment in older men. *New Engl J Med* 2016;**374:**611-24.

6. Snyder PJ, Kopperdahl DL, Stephens-Shields AJ, Ellenberg SS, Cauley JA, Ensrud KE, et al. Effect of testosterone treatment on volumetric bone density and strength in older men with low testosterone a controlled clinical trial. *JAMA Intern Med* 2017;**177:**471-9.

7. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;**25:**76-97.

8. Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, et al. Individual participant data (IPD) metaanalyses of randomised controlled trials: Uidance on their use. *PLoS Med* 2015;**12:**e1001855.

9. Bhasin S, Cunningham GR, Hayes FJ, Matsumoto AM, Snyder PJ, Swerdloff RS, et al. Testosterone therapy in men with androgen deficiency syndromes: An endocrine society clinical practice guideline. *J Clin Endocrinol Metabol* 2010;**95:**2536-59.

10. Miner M, Barkin J, Rosenberg MT. Testosterone deficiency: myth, facts, and controversy. *Can J Urol* 2014;**21:**39-54.

11. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [document on the Internet]* 2011. Available from: URL:http://handbook-5-1.cochrane.org/ [Accessed August 2018]

12. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ* 2010;**340:**521-5.

13. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Stat Med* 1995;**14:**2057-79.

14. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* 2014;**36:**855-75.

15. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qual Health Res* 2012;**22:**1435-43.

16. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;**8:**45.

17. Langham S, Maggi M, Schulman C, Quinton R, Uhl-Hochgraeber K. Health-related quality of life instruments in studies of adult men with testosterone deficiency syndrome: A critical assessment. *J Sex Med* 2008;**5:**2842-52.

18. Macefield RC, Jacobs M, Korfage IJ, Nicklin J, Whistance RN, Brookes ST, et al. Developing core outcomes sets: Methods for identifying and including patient-reported outcomes (PROs). *Trials* 2014;**15** 49.

19. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Med Decis Making* 2012;**32:**722-32.

20. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices - Overview: A report of the ISPOR-SMDM modeling good research practices task force-1. *Value Health* 2012;**15:**796-803.

21. Kaltenthaler E, Tappenden P, Paisley S, Squires H. *NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models*. University of Sheffield: NICE Decision Support Unit; 2011. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD-13-model-parameters.pdf [Accessed August 2018]

22. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8:**36.

23. Drummond MF, O'Brien B, Sitoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes Oxford: Oxford University Press; 2015

24. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment. *PharmacoEconomics* 2006;**24:**355-71.

25. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006;**15:**1295-310.

26. *NHS reference costs* London: UK Department of Health; 2017. URL: https://www.gov.uk/government/collections/nhs-reference-costs#published-reference-costs [Accessed August 2018]

27. *British National Formulary* Medicines Complete; 2018. https://www.medicinescomplete.com/mc/bnf/current/index.htm [Accessed August 2018]

28. Curtis L, Burns A. *Unit Costs of Health and Social Care 2017*. Canterbury: Personal Social Services Research Unit.; 2016. URL:http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/ [Accessed August 2018]

29. Saramago P, Manca A, Sutton AJ. Deriving input parameters for cost-effectiveness modeling: taxonomy of data types and approaches to their statistical synthesis. *Value Health* 2012;**15:**639-49.

30. *Guide to the methods of technology appraisal*. London: National Institute for Health and Care Excellence; 2013. https://www.nice.org.uk/article/pmg9/resources/non-guidance-guide-to-the-methods-of-technology-appraisal-2013-pdf [Accessed August 2018]

31. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: A review of possible methods. *JHealth Serv Res Policy* 2005;**10:**45-53.

APPENDIX 1: MEDLINE SEARCH STRATEGY FOR IDENTIFICATION OF CLINICAL EFFECTIVENESS EVIDENCE (DRAFT)

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

1 exp androgens/tu (7740)
2 hormone replacement therapy/ (9222)
3 2 and (men or androgen? or testosterone).af. (2550)
4 androgen replacement therapy.tw,kw. (327)
5 testosterone.tw,kw. (75982)
6 or/1,3-5 (81511)
7 Erectile Dysfunction/ (16987)
8 testosterone/df (1214)
9 Libido/ (4656)
10 Hypogonadism/ (8218)
11 (erectile adj3 dysfunction).tw,kw. (13715)
12 (libido adj3 (low$ or decreas$ or reduc$ or loss)).tw,kw. (1836)
13 (impotence or impotent).tw,kw. (6483)
14 hypogonadism.tw,kw. (9730)
15 or/7-14 (42216)
16 6 and 15 (7264)
17 randomized controlled trial.pt. (477230)
18 controlled clinical trial.pt. (96153)
19 randomi?ed.ab. (499077)
20 placebo.ab. (195119)
21 drug therapy.fs. (2046487)
22 randomly.ab. (289527)
23 trial.ab. (439518)
24 groups.ab. (1784459)
25 or/17-24 (4238653)
26 exp animals/ not humans/ (4532769)
27 25 not 26 (3665980)
28 16 and 27 (3617)

MEDLINE SEARCH STRATEGY FOR IDENTIFICATION OF COST EFFECTIVENESS EVIDENCE (DRAFT)

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

1 Erectile Dysfunction/ (16938)
2 testosterone/df (1218)
3 Libido/ (4524)
4 Hypogonadism/ (8142)
5 (erectile adj3 dysfunction).tw,kw. (13967)
6 (libido adj3 (low$ or decreas$ or reduc$ or loss)).tw,kw. (1837)
7 (impotence or impotent).tw,kw. (6391)
8 hypogonadism.tw,kw. (9874)
9 or/1-8 (42251)
10 exp "costs and cost analysis"/ (216194)
11 economics/ (26926)
12 exp economics,hospital/ (22918)
13 exp economics,medical/ (14031)
14 economics,pharmaceutical/ (2773)
15 exp models, economic/ (13355)
16 exp decision theory/ (11103)
17 monte carlo method/ (25353)
18 markov chains/ (12804)
19 exp technology assessment, biomedical/ (10408)
20 (cost$ adj2 (effective$ or utilit$ or benefit$ or minimis$)).ab. (123462)
21 economics model$.tw. (46)
22 (economic$ or pharmacoeconomic$).tw. (225455)
23 (price or prices or pricing).tw. (33429)
24 budget$.tw. (25790)
25 (value adj1 money).tw. (32)
26 (expenditure$ not energy).tw. (25822)
27 markov$.tw. (20723)
28 monte carlo.tw. (42568)
29 (decision$ adj2 (tree? or analy$ or model$)).tw. (19025)

30 (metabolic adj cost).tw. (1225)

31 ((energy or oxygen) adj (cost or expenditure)).tw. (25357)

32 (letter or editorial or note or comment).pt. (1641416)

33 or/10-29 (651560)

34 33 not (30 or 31 or 32) (622759)

35 9 and 34 (549)

36 limit 35 to yr="1992 -Current" (475)

APPENDIX 2: RESULTS OF FEASIBILITY STUDY FOR INDIVIDUAL PARTICIPANT DATA (IPD) META-ANALYSIS.

Authors of 35 randomised placebo clinical trials (RCTs) published within the last 10 years were contacted by Dr. Jayasena from 5th September 2017 onwards, to participate in the proposed IPD MA. In the three weeks since requests were sent, agreements to collaborate have been received for 32 (91%) of the studies. No authors have declined to collaborate. Response is awaited for the authors of the remaining 3 papers.

Agreements to collaborate:

Professor Peter Snyder, University of Pennsylvania, USA & Chief investigator of NIH T Trials

JAMA. 2017 Feb 21;317(7):717-727. doi: 10.1001/jama.2016.21044.

JAMA. 2017 Feb 21;317(7):708-716. doi: 10.1001/jama.2016.21043.

JAMA Intern Med. 2017 Apr 1;177(4):480-490. doi: 10.1001/jamainternmed.2016.9540.
JAMA Intern Med. 2017 Apr 1;177(4):471-479. doi: 10.1001/jamainternmed.2016.9539.
J Clin Endocrinol Metab. 2016 Aug;101(8):3096-104. doi: 10.1210/jc.2016-1645. Epub 2016 Jun 29.
N Engl J Med. 2016 Feb 18;374(7):611-24. doi: 10.1056/NEJMoa1506119.

Professor Shalender Bhasin, University of Harvard, USA

J Clin Endocrinol Metab. 2017 Feb 1;102(2):583-593. doi: 10.1210/jc.2016-2771.

JAMA. 2015 Aug 11;314(6):570-81. doi: 10.1001/jama.2015.8881.

Andrology. 2013 May;1(3):475-82. doi: 10.1111/j.2047-2927.2013.00075.x. Epub 2013 Mar 15.

Ann Intern Med. 2012 Nov 20;157(10):681-91. doi: 10.7326/0003-4819-157-10-201211200-00004.
N Engl J Med. 2010 Jul 8;363(2):109-22. doi: 10.1056/NEJMoa1000485. Epub 2010 Jun 30.

Professor Geoffrey Hackett, Good Hope Hospital & University of Bedfordshire
BJU Int. 2016 Nov;118(5):804-813. doi: 10.1111/bju.13516. Epub 2016 May 27.

Int J Clin Pract. 2014 Feb;68(2):203-15. doi: 10.1111/ijcp.12235. Epub 2013 Dec 20.

J Sex Med. 2014 Mar;11(3):840-56. doi: 10.1111/jsm.12404. Epub 2013 Dec 6.

J Sex Med. 2013 Jun;10(6):1612-27. doi: 10.1111/jsm.12146. Epub 2013 Apr 3.

Professor Peter Liu, David Geffen School of Medicine at UCLA, USA
Andrology. 2016 Jan;4(1):55-61. doi: 10.1111/andr.12132. Epub 2015 Nov 26.

Clin Endocrinol (Oxf). 2012 Oct;77(4):599-607. doi: 10.1111/j.1365-2265.2012.04413.x.
Professor Pierre-Marc Bouloux, Royal Free Hospital & University College London
Aging Male. 2015;18(3):157-63. doi: 10.3109/13685538.2015.1032925. Epub 2015 Jun 1.
Aging Male. 2013 Jun;16(2):38-47. doi: 10.3109/13685538.2013.773420. Epub 2013 Apr 12.
Eur J Endocrinol. 2009 May;160(5):821-31. doi: 10.1530/EJE-08-0634. Epub 2009 Feb 11.

Professor Mathis Grossmann, University of Melbourne, Australia
J Clin Endocrinol Metab. 2014 Oct;99(10):3821-8. doi: 10.1210/jc.2014-1872. Epub 2014 Jun 30.
Diabetes Care. 2014 Aug;37(8):2098-107. doi: 10.2337/dc13-2845. Epub 2014 May 7.

Dr. Erik Giltay, University of Leiden, Netherlands.

J Sex Med. 2010 Jul;7(7):2572-82. doi: 10.1111/j.1743-6109.2010.01859.x. Epub 2010 May 26.

Prof Frederick Wu, University of Manchester

J Clin Endocrinol Metab. 2010 Feb;95(2):639-50. doi: 10.1210/jc.2009-1251. Epub 2010 Jan 8.

P**rofessor Hugh Jones, Barnsley Hospital & University of Sheffield**

Diabetes Care. 2011 Apr;34(4):828-37. doi: 10.2337/dc10-1233. Epub 2011 Mar 8.

Eur J Endocrinol. 2009 Sep;161(3):443-9. doi: 10.1530/EJE-09-0092. Epub 2009 Jun 19.

Professor Antonio Aversa, Sapienza University of Rome, Italy

J Sex Med. 2010 Oct;7(10):3495-503. doi: 10.1111/j.1743-6109.2010.01931.x.

Professor Dato Tan Hui Meng, Subang Jaya Medical Centre, Selangor, Malaysia

BJU Int. 2013 Jun;111(7):1130-40. doi: 10.1111/bju.12037. Epub 2013 Apr 12.

Asian J Androl. 2012 Jul;14(4):604-11. doi: 10.1038/aja.2011.178. Epub 2012 May 28.

BJU Int. 2012 Jul;110(2):260-5. doi: 10.1111/j.1464-410X.2011.10755.x. Epub 2011 Nov 17.

Dr. Marielle Emmelot-Vonk & Professor Ir. Y.T. van der Schouw, UMC Utrecht, Netherlands

Int J Impot Res. 2009 Mar-Apr;21(2):129-38. doi: 10.1038/ijir.2009.5. Epub 2009 Feb 19.

Professor Mario Maggi, University of Firenze, Florence, Italy (E-mailed Sept 11th 2017)

J Urol. 2016 Mar;195(3):699-705. doi: 10.1016/j.juro.2015.10.083. Epub 2015 Oct 20.
J Sex Med. 2016 Aug;13(8):1220-6.

Professor Robert McLachlan, Hudson Institute of Medical Research, Melbourne, Australia (E-mailed Sept 11th 2017)
Int J Impot Res. 2008 Jul-Aug;20(4):396-401. doi: 10.1038/ijir.2008.22. Epub 2008 Jun 5.

Declined invitation to collaborate:
None

APPENDIX 3: AMENDMENT HISTORY

Amendment number	Protocol version number	Date issued	Author(s) of changes	Section(s) amended	Details of change(s)