ODPRN Comprehensive Research Plan:

Testosterone Replacement Therapy

Systematic Review

April 29, 2014

Background

The androgen testosterone has been used clinically to treat patients for various medical disorders since the early 1960s (Handelsman, 2006). Currently, testosterone is indicated as a replacement therapy in adult men with diagnosed androgen deficiency. Androgen deficiency occurs when the testicles do not produce enough testosterone, and can be attributed to either primary testicular failure (primary hypergonadism) or a problem in the hypothalamus or pituitary gland (secondary hypergonadism). Both can be congenital or acquired later in life and the causes vary greatly within each type or related syndrome.

Hypogonadism in men may alter certain masculine physical characteristics, impair sexual function and impact mental and emotional well-being.(Reference: Up-to-Date 2014). The goal of testosterone replacement therapy (TRT) is to restore quality of life, sexual function, strength, prevent bone loss and to mimic the overall long-term health outcomes of eugonadal men. Although there is considerable evidence to support the benefit of TRT in men with diagnosed androgen deficiency syndrome (ADS), the risks of this type of therapy remain poorly understood.

Objective

To determine the comparative clinical benefit and harms of testosterone replacement therapy in adult men with androgen deficiency syndromes through a systematic review and Bayesian network meta-analysis.

Research Questions

RQ1. What is the current evidence for the efficacy and safety of testosterone replacement therapy in adult men with primary and secondary androgen deficiency syndromes?

RQ2. Does the efficacy or safety of testosterone replacement therapy vary in adult men based on:
   a. Age
   b. Duration of Therapy (>,<6 mo)
   c. Dose

PICO

The population, intervention, comparator, and outcome (PICO) statement, including the study designs of interest, is as follows:

Population	Adult males with primary or secondary diagnoses of an androgen deficiency syndrome[a]			
Interventions	Testosterone replacement therapies currently available in Canada			
	HC-Approved Drug	**Brand Name**	**Generic Available?**	**Dosage Form**
	Testosterone undecanoate	Andriol	Y	Oral
	Testosterone cypionate	Depo-Testosterone	Y	IM Depot
	Testosterone enanthate	Delatestryl	N	IM Depot
	Testosterone	Androderm	N	Patch
	Testosterone	Testim	N	Topical Gel
	Testosterone	Androgel	N	Topical Gel
	Testosterone	Axiron	N	Topical Solution
Comparator	Placebo All intervention comparisons			
Outcomes: Efficacy	• Proportion achieving normal serum testosterone level (i.e.,> 12 nMol/L)* • Quality of Life			
Outcomes: Safety	• Cardiovascular death • Myocardial Infarction • Stroke • Erythrocytosis • Serious Adverse Events • Newly diagnosed disease (diabetes/heart disease/prostate cancer) • Skin or site reactions			
Study Types: Efficacy	Randomized controlled trials (RCTs) will be included. No limits placed on sample size, study duration, patient follow-up			
Study Types: Safety	RCTs, high-quality non-randomized studies with a comparator. No limits placed on sample size, study duration, patient follow-up.			
Exclusions	• Studies not conducted in English • Testosterone products or delivery formulations (buccal cavity, implantable pellets) not currently approved by Health Canada. • Adult men with andropause, no formal diagnosis of ADS, off-label use of TRT in new patient populations (i.e. male contraception) or men who are asymptomatic. • Studies published in abstract format • Uncontrolled non-randomized studies			

DISCLAIMER	Efficacy and safety outcome lists may be truncated if we identify many studies for inclusion, as this is a rapid review. We will not perform a meta-analysis (or network meta-analysis) on all of these outcomes and will work with all stakeholders to select the two most important efficacy outcomes and safety outcomes with sufficient data to conduct network meta-analysis. Prior to conducting network meta-analysis, we will ensure that all factors are considered as this analysis only is valid when homogenous studies and patient populations are included.

Methods

De Novo Systematic Review – Efficacy (RCTs only)

The strategy for building and analyzing the evidence base for the efficacy of TRT in adult men with androgen deficiency syndrome consists of two fundamental steps:

1. A broad systematic review and meta-analysis of the available randomized evidence in the published and grey literature will be conducted, following the methods and procedures outlined in the Cochrane Handbook for Systematic Reviews for Interventions*.

2. A Bayesian network meta-analysis of randomized evidence will be conducted relating the various preparations and dosages of TRT in a network, for each of the efficacy outcomes specified *a priori*. The methods and procedures to be followed are those developed by the Canadian Collaboration for Drug Safety, Effectiveness and Network Meta-Analysis (ccNMA), funded by the Drug Safety and Effectiveness Network (DSEN) of the Canadian Institute of Health Research.

* Input from interested stakeholders made available through evidence submission packages will also be considered.

Safety – Stepped Approach (RCTs and NRS)

The strategy for building and analyzing the evidence base for the safety of TRT in adult males with androgen deficiency syndrome consists of two fundamental steps, the first of which will take a stepped approach:

1. First, we will look broadly for a well-conducted recent evidence synthesis that fits the PICO for safety. Next steps will be determined by the results of the search.

    a. If we are able to update an existing high-quality systematic review and meta-analysis of the available randomized and non-randomized evidence we will build onto the studies included in the existing review. A new literature search will capture studies published from the date of the last literature search to present.

    b. If we do not locate an existing evidence synthesis, we will need to conduct a de novo systematic review for safety.

All methods and procedures will follow those outlined in the Cochrane Handbook for Systematic Reviews for Interventions (Chapter 13/14 and related papers in Research Synthesis Methods 2013: [Link to series](#)).*

2. A Bayesian network meta-analysis of randomized and non-randomized evidence will be considered relating the various preparations and dosages of TRT for each of the safety outcomes specified *a priori*. The methods and procedures to be followed are those developed by the Canadian Collaboration for Drug Safety, Effectiveness and Network Meta-Analysis (CCNMA), funded by the Drug Safety and Effectiveness Network (DSEN) of the Canadian Institute of Health Research.

* Input from interested stakeholders made available through evidence submission packages will also be considered.

GENERAL STEPS IN THE SYSTEMATIC REVIEW:
This protocol was developed using guidance from the PRISMA Statement (1) and follows the methods and procedures outlined in the Cochrane Handbook for Systematic Reviews for Interventions (2). It has been peer-reviewed by experts in pharmacology, statistics, and systematic review methodology.

Our protocol, once approved, will be registered in the PROSPERO database and the protocol will be updated with the assigned registration number (3).

The specific steps for the systematic review are as follows.

Electronic Search Strategy:

The literature search will be conducted by a professional Information Scientist (MLIS). Literature search strategies for efficacy and safety will be developed using medical subject headings (MeSH) and text words related to the PICO statement. Searches will employ validated filters for RCTs and non-randomized studies (for safety only). All studies will be included regardless of publication status (i.e., unpublished studies) and year of publication. Included articles will be limited to English language studies.

Proposed electronic databases are listed below, and may be augmented based on feedback from experts and/or the Information Scientist.
- Cochrane CENTRAL
- Medline (indexed, in-process and other non-indexed)
- Embase (via OVID)

A limited grey literature search will be carried out by searching the websites of health technology assessment and related agencies, professional associations, and other specialized databases (following "CADTH Grey Matters: A practical search tool for evidence based medicine" as closely as time permits). Additionally, Google will be employed to search for

additional Internet-based materials and information.

Eligibility and Study Selection:

Studies will be included if the population, intervention, comparator, and outcome (PICO) criteria and type of study are appropriate.

Selection eligibility criteria will be applied to each title and abstract identified in the literature search by two independent review authors in a standardized manner. Any uncertainties will be resolved by discussion and consensus with a third review author. All studies passing the selection criteria will be obtained in full-text format. The eligibility criteria will then be applied and a final decision made for inclusion. The reviewers will not remain blinded to study authors or centre of publication prior to study selection as this can complicate the review process and only weak evidence suggests this would improve results.

Data Extraction and Management:

All information will be extracted using a standardized data extraction form, which will be developed, piloted and modified as necessary.

Abstraction will include:
1. Characteristics of trial participants;
2. Study characteristics;
3. Details on each study arm, including but not limited to: pharmacological intervention, dose, frequency, route of administration, duration and co-medication; and,
4. Results of the clinical safety and efficacy outcomes for the overall study population and the
   a priori subgroups identified.

All extracted data will be checked for accuracy by two independent review authors. The original, primary publication for each unique study included will be used for data extraction, except where multiple publications for a single RCT or NRS are found. Multiple publications for a unique RCT/NRS (e.g. supplemental online appendices, companion publications of specific outcomes or populations from the original study) will be handled by extracting the most recently adjudicated data for each outcome specified *a priori*.

Risk of Bias Assessment:

Cochrane Collaboration's tool for assessing risk of bias (ROB)(2) will be employed. Systematic reviews used to form the bases for the safety analyses will be evaluated by AMSTAR(4).

Assessment of Heterogeneity:

Results will be assessed for both clinical and methodological diversity. Clinical diversity will be assessed by checking that the participants, interventions, and comparators are not too different from each other such that combining them is not appropriate. Methodological diversity will be assessed by checking that the studies are similar in terms of study design and risk of bias.

Once satisfied that the studies are minimally diverse and that it makes sense to pool them together in a meta-analysis, an assessment of the statistical heterogeneity will be undertaken by examining the forest plot and result of the I^2 statistic; the forest plots providing a visual sense of heterogeneity and the I^2 statistic indicating the presence of statistical heterogeneity. If the effects observed across trials are inconsistent, and vary to a large extent (say $I^2 > 50\%$), the results will again be explored to assess whether the differences can be explained by some clinical or methodological feature.

Inconsistency that cannot be reduced by pre-specified subgroup or meta-regression analyses will lead to an overall estimate with less confidence when interpreting the inference from the meta-analysis. In this case, a more conservative random-effects model approach would be used so that the uncertainty of the single effect estimate is reflected in wider confidence intervals.

Assessment of Reporting Bias:

Reporting bias will be assessed by constructing funnel plots, as well as bias indicators (e.g. Egger, Harbold-Egger) for each outcome.

Data Synthesis:

Data will first be summarized descriptively. A meta-analysis will be undertaken using fixed or random-effects models when data are available, sufficiently similar and of sufficient quality. The effect sizes for the identified dichotomous outcomes will be expressed in terms of the risk ratio (RR) or odds ratio (OR). In cases when events are rare, the Peto odds ratio will be used. For continuous outcomes such as QoL, the effect size will be expressed in terms of the mean difference (MD) and standardized mean difference (SMD). Pair wise meta-analyses will be conducted using RevMan or R. Absolute differences in the important benefits and harms, absolute mean difference and relative percent change from baseline will be included in a summary of findings table.

Subgroup Analysis:

Major outcomes will be assessed in identified subgroups in the specific populations of adults with diagnosed androgen deficiency syndromes. Subgroups were selected to confirm clinically sound hypotheses and as few subgroups as possible were pre-specified and justified against the criteria proposed by Sun et al.; wherein the greater the number of criteria that are satisfied for each subgroup and outcome, the more plausible is the hypothesized subgroup effect (5, 6).

Planned subgroups include:
- Age (over/under 65)
- Duration of Therapy ($><6$ mo)
- Dose

Sensitivity Analysis:

Sensitivity analysis will be conducted based on aspects of the PICO statement and study methodology to examine the robustness of the results to the risk of bias and the influence of other variables. In particular, the results of the low-risk-of-bias-studies will be compared to

studies with a higher-risk-of-bias and if they differ substantively, the conclusions of the review will be based on analyses of low risk studies only.

*Published literature suggests that sensitivity analysis based on funding may be appropriate to consider (Xu, 2013).

Grading of Evidence:

To help in the understanding of the strength of the evidence included in the review, grading of the evidence for each major outcome will be considered using The 'Grading of Recommendation Assessment Development and Evaluation' (GRADE) approach (7).

Bayesian Network Meta-Analysis Methods

Bayesian network meta-analyses will be conducted using WinBUGS software (MRC Biostatistics Unit, Cambridge, UK) (8). The use of a Bayesian network meta-analysis offers several advantages, including:

1. TRTs have not been compared directly with each other in a large number of studies, and Bayesian network meta-analysis permits combination of all active and placebo-controlled evidence; and

2. The number of individual pair-wise comparisons between testosterone treatments for androgen deficiency syndromes is unwieldy given the large number of available treatment options.

As a result, summary effect estimates against a common comparator are likely to be of greater utility for clinical and policy decisions. Further, we will also construct graphical aids to assist in decision making.

We will conduct a Bayesian network meta-analysis using a model which accommodates complex interventions (9). The advantage of using the approach by Welton et al. (9) is that the proposed models (e.g., additive) may allow us to estimate treatment effects for comparisons that may not have been compared directly.

The essential methods for conducting the Bayesian mixed treatment comparison are summarized in Box 1. Both fixed and random-effects network meta-analyses will be conducted; model fit for Bayesian analyses will be based on the Deviance Information Criterion (DIC) and comparison of residual deviance to number of unconstrained data points (10-13) . Selection of model/measure will depend on the outcome of interest and availability of data. Heterogeneity across trials in terms of patient characteristics, trial methodologies, and treatment protocols will be carefully assessed. To further investigate heterogeneity, subgroup analyses and meta-regressions (12, 13) will be conducted exploring the effect of various characteristics including but not limited to the variables considered for the subgroup and sensitivity analyses. We will also perform analyses including removal of studies from the network of therapies that were not scored as being of high quality. We will formally (12) and informally assess consistency between direct and indirect evidence by comparing direct estimates obtained from pair wise meta-analysis with estimates from the Bayesian network meta-analysis (14). Model diagnostics including trace plots and the Brooks-Gelman-Rubin statistic will be assessed to ensure model

convergence. At least two chains will be fit in WinBUGS for each analysis, each employing at least 40,000 iterations, with a burn-in of at least 20,000 iterations (8, 11).(Exhibit 1)

Deliverables

We will provide a written report detailing methods adopted, results, discussion and key outcome highlights. The report will comprise of a two page executive summary followed by a detailed technical report.

Timelines

On acceptance of this proposal, work will commence. The systematic review, meta-analysis and Bayesian network meta-analysis will be completed in approximately 12 weeks to 16 weeks. The preliminary report with efficacy and safety results will also be delivered on or around August 15, 2014. Any reanalyses and a revised final report will be available 4 weeks after receipt of stakeholder reviews.

Exhibit 1: Methods for Bayesian mixed treatment comparison

- Bayesian NMAs will be conducted for outcomes pre-specified in the DSEN request, following careful assessment of heterogeneity across trials in terms of subject characteristics, trial methodologies, and treatment protocols.
- The effect estimate chosen (e.g., relative risk) will depend on the outcome of interest and availability of data.
- For reference case network meta-analyses, appropriate comparators will be considered and some comparators may be stratified by dose.
- Both fixed and random-effects models will be conducted; model selection will be based on the Deviance Information Criterion (DIC) and residual deviance.
- R (R Foundation for Statistical Computing, Vienna, Austria) and WinBUGS (MRC Biostatistics Unit, Cambridge, UK) will be used for Bayesian network meta-analyses according to the routine which accommodates evidence structures which may consist of multi-arm trials as developed at the Universities of Bristol and Leicester (www.bris.ac.uk/cobm/research/mpes/).
- Specific therapy(ies) will be identified as the reference group for all Bayesian network meta-analyses.
- Posterior densities for unknown parameters will be estimated using Markov Chain Monte Carlo (MCMC) methods.
- Basic parameters will be assigned non-informative or vague prior distributions; more informative priors will be considered after evaluation of the information base and clinical expert advice.
- Point estimates and 95% credible intervals will be used to summarize findings.
- The probability of a comparator being optimal will be estimated for each outcome based on the proportion of MCMC simulations in which its relative measure of effect was best.
- The mean rank for each comparator will also be calculated.
- Consistency between direct and indirect evidence will be formally assessed using back-calculation and node splitting techniques [14].
- Graphical methods and numerical summaries will be developed for presenting results from network meta-analysis [15].
- Model diagnostics will also include trace plots and the Brooks-Gelman-Rubin statistic (reference) to assess and ensure model convergence.
- Two chains will be fit in WinBUGS for each analysis, each usually employing ≥20,000 iterations, with a burn-in of ≥20,000 iterations.
- Provided sufficient data is available to inform the evidence network, meta-regression and/or sub-groups analyses will be conducted to adjust for key demographic, medical, and study design characteristics to test the robustness of reference case analyses.
- In other sensitivity analyses, studies will be removed from the network that are of poor methodological quality, study design, etc.
- Examine whether novel agent effects are present and estimate their magnitude of effect [16].

References

1.   Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009 Oct;62(10):e1-34.

2.   Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0. Higgins JPT, Green S, editors. West Sussex: Wiley-Blackwell, A John Wiley & Sons, Ltd.

3.   Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev.1(1):2.

4.   Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009 Oct;62(10):1013-20.

5.   Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. BMJ.344:e1553.

6.   Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ.340:c117.

7.   Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924-6.

8.   Spiegelhalter D, Thomas A, Best N, Lunn D. WinBUGS User Manual. Version 1.42003 [cited 2011 June 1]: Available from: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual114.pdf.

9.   Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. Am J Epidemiol. 2009 May 1;169(9):1158-65.

10.   Dias S, Sutton A, Welton N, Ades A, Golfinopoulos V, Kyrgiou M, et al. NICE DSU Technical Support Document 1: Introduction to evidence synthesis for decision making.May 2011 [cited 2012 June 13]: Available from: http://www.nicedsu.org.uk/TSD6%20Software.final.08.05.12.pdf.

11.   Dias S, Sutton A, Welton N, Ades A, Golfinopoulos V, Kyrgiou M, et al. NICE DSU Technical Support Document 2: Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials.May 2011 [cited 2012 June 13]: Available from: http://www.nicedsu.org.uk/TSD6%20Software.final.08.05.12.pdf.

12.   Dias S, Sutton A, Welton N, Ades A, Golfinopoulos V, Kyrgiou M, et al. NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials.May 2011 [cited 2012 June 13]: Available from:

http://www.nicedsu.org.uk/TSD6%20Software.final.08.05.12.pdf.

13. Dias S, Sutton A, Welton N, Ades A, Golfinopoulos V, Kyrgiou M, et al. NICE DSU Technical Support Document 3: Heterogeneity: Subgroups, meta-regression, bias and bias-adjustmentMay 2011 [cited 2012 June 13]: Available from: http://www.nicedsu.org.uk/TSD6%20Software.final.08.05.12.pdf.

14. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. Mar 30;29(7-8):932-44.