UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF ILLINOIS

IN RE: TESTOSTERONE REPLACEMENT THERAPY PRODUCTS LIABILITY LITIGATION	**CASE NO. 1:14-CV-01748** **MDL 2545** **JUDGE MATTHEW F. KENNELLY**

ABBVIE INC. AND ABBOTT LABORATORIES' PROPOSAL FOR SELECTION OF BELLWETHER CASES FOR TRIAL

I. INTRODUCTION

Since the early stages of the bellwether selection process, the Court has been explicit that an essential goal is to select plaintiffs who are representative of the bellwether pool as a whole, and productive for both pretrial and trial purposes. To be sure, the Court also has preserved a place for attorney preference. But the Court has consistently carved out a role for itself as the arbiter who can assure that judicial interests in fair and effective management of the MDL claims as a whole are served as well. Those interests militate in favor of selecting bellwethers whose claims reflect key cross-cutting issues and whose case-specific facts and personal demographic characteristics are not so distinctive that they don't represent the pool of claimants. This balanced approach was articulated by the Court last August and has been followed since. In particular, the Court emphasized the importance of identifying a "representative" pool of bellwether cases reflecting "issues that are going to affect lots of cases" that would, in turn, "get litigated pretrial." (Aug. 18, 2015 Hr'g Tr. at 30:23-31:8.)

As the number of discovery bellwethers has now decreased from 32 to 24, and the number of trial bellwethers from up to 12 to 8, this makes representativeness of the selected cases all the more important. At the same time, the discovery process has produced substantial data, which can be used to evaluate the candidates carefully to that end. AbbVie therefore proposes that the eight cases be selected to ensure representativeness.

Mirroring the selection of the discovery bellwethers last year, AbbVie Inc. and Abbott Laboratories (together "AbbVie") below propose 3 cardiovascular injury ("CV") and 3 thromboembolic clotting injury ("Clot") bellwether cases to be used in

filling 2 of 4 CV and 2 of 4 Clot trial slots. Also consistent with last year's process and the Court's July 21, 2016 Minute Order (Docket No. 1398) ("July 21 Minute Order"), AbbVie identifies 4 non-representative "outlier" CV cases and 5 outlier Clot cases. The balance of 9 cases that are neither proposed nor identified as outliers comprise cases that are, by AbbVie's assessment, to varying degrees favorable as trial candidates and representative of the pool. AbbVie cannot fairly say they are outliers and should not be available for selection by the Court or by Plaintiffs to fill 2 of 4 CV and 2 of 4 Clot cases. They too are described below. AbbVie also has furnished a summary table, attached as Exhibit A, which might be useful in keeping track of AbbVie's assessment.

II. PROCEDURAL BACKGROUND

The request for proposals regarding discovery bellwether case selection. Over a year ago, Amended Case Management Order ("CMO") No. 14 required the parties to propose a process for selecting AbbVie-only plaintiffs in two types of cases that would proceed to bellwether discovery: (1) those involving thromboembolic clotting injuries ("Clot cases"); and (2) those involving cardiovascular injuries ("CV cases"). (Docket. No. 793.) The stated goal was to "ensure fairness to all parties" and to "maximize the likelihood that the bellwether selection and trial process will be both representative and productive." (*Id.* at 1.) The process was to result in the selection of 32 discovery bellwethers divided evenly among CV and Clot cases. At that time, six of the discovery bellwether cases were contemplated to be selected for initial trials. (*Id.* at 4-5.)

The parties' proposals. The parties' proposals for the selection process reflected two substantially different approaches. (*See* Docket Nos. 932, 933.) Plaintiffs proposed that the attorneys should have unfettered discretion to choose all 32 discovery

bellwether cases from the set of cases filed on or before June 15, 2015, so long as a Plaintiff Fact Sheets had been timely completed. Plaintiffs' proposal thus contemplated no mechanism whatsoever for ensuring that the bellwether process would be representative or productive.

By contrast, AbbVie proposed that the discovery bellwether selection process should be "guided by data, use available statistical methods, and apply objective criteria." (Docket No. 932 at 2.) AbbVie also urged that the Court should be involved in the selection of the discovery bellwethers—not merely advocates. (*Id.*) Specifically, AbbVie recommended a random selection technique be adopted by the Court to identify 32 discovery bellwether cases reflecting key cross-cutting issues and demographic characteristics representative of the larger claimant pool, issues, and demographics. (*Id.* at 932-1 at 2-3.) The trial bellwethers would then later be selected by counsel.

The Court's decision on process. After considering these proposals, the Court directed the parties to first randomly select 100 cases from the pool of more than 470 "AbbVie-only" cases, and then select from that 100 a subset of 32 bellwether discovery cases based on case categories identified by the parties. (Aug. 18, 2015 Hr'g Tr. at 86:5-17, 87:1-24, 89:20-90:21.) To the extent that the parties could not agree on the pool of 32 cases, the Court would "arbitrate those disputes and make any final decision." (*Id.* at 69:21-70:9.) The Court again emphasized the importance of identifying a "representative" pool of cases and reiterated its expectation that as part of the bellwether process, "issues that are going to affect lots of cases are going to come up and are going to get litigated pretrial" so that "[we] will have that information." (*Id.* at

30:23-31:8, 84:14-15 ("I agree that it's important to do it right, and 'right' means getting something that's representative.").)

The parties select the discovery bellwether pool. The parties went on to select 100 cases at random, and each side submitted their proposed slates of sixteen bellwether discovery cases on November 2, 2015. (Docket No. 1038 at 6; Docket No. 1039 at 4-5.) Each proposed slate included eight CV cases and eight Clot cases.

On November 12, 2015, the Court heard the parties' proposals and objections to candidates on representativeness grounds. By the end of the process, 32 discovery bellwether cases had been selected and the parties were able to commence "core bellwether discovery." (Docket No. 793 at 5.)

The Court proposes increasing the number of trial bellwethers. During the November 12, 2015 hearing, the Court also took up AbbVie's proposal to amend CMO No. 9 increasing the number of trial bellwethers to 16. (Nov. 12, 2015 Hr'g Tr. at 49:8-51:6.) This would better enable expert discovery, dispositive motion practice, and ultimately trial, as needed, that adequately covered cross-cutting issues. (*Id.*) At the hearing, the Court expressed concern that six bellwether trials might not be sufficient for these purposes and thus agreed to double the number of trial bellwethers from six to up to twelve. (*Id.* at 69:7-70:16.) The Court's proposed change was intended to "increase the likelihood of getting decisions on more [cross-cutting] issues further down the road when we get to the *Daubert* [and] summary judgment stage." (*Id.*; *see also* Second Amended CMO 14 at 3 (Docket No. 1089).)

The pool of discovery bellwethers decreases. By the Spring, however, the pool of discovery bellwethers shrank and with this, the number of trial bellwethers

was decreased as well. In April 2016, AbbVie informed the Court that, upon reviewing medical records not previously provided, it discovered that a number of the 32 current bellwether plaintiffs had used testosterone therapy products other than AndroGel prior to their alleged injuries. (Docket No. 1242 at 11-12.) Accordingly, the Court excluded six of these cases from the bellwether trial pool and left two cases "subject to later exclusion." (Docket No. 1268.) The Court also made clear that mixed use cases should not be included in the bellwether pool. (Apr. 13, 2016 Hr'g Tr. 19:13-21 ("[W]hat I decided to do after hearing -- having briefs and after hearing argument was, no, we're just going to have this first round be people that just used AbbVie's product because I didn't think it would -- I thought it would complicate things and make it harder to get a representative sample if we introduced a complicating factor. . . . I mean, I don't know what a jury is going to think about that. Let's get all of that out. That was the decision.").) The parties subsequently agreed jointly to remove additional cases from the bellwether pool for similar reasons, leaving 24 cases. (*See* Joint Status Report, July 22, 2016 (Docket No. 1400); Joint Status Report, May 26, 2016 (Docket No. 1320).)

At the same time that the Court reduced the number of bellwether cases, it also reduced the number of bellwether trials from up to twelve to eight to ensure that the bellwether discovery, pre-trial briefing, and trial schedule would not be extended further. (*See* CMO No. 29 at 1-2 (Docket No. 1270); *see also* Third Amended CMO 14 at 3 (Docket No. 1287).)

III. REPRESENTATIVENESS SHOULD BE JUDGED BY REFERENCE TO THE POOL OF 100

The parties have now completed core discovery, including production of dozens of AbbVie custodial files, and depositions of 22 plaintiffs, 24 prescribers, 5

treaters, and 28 sales representatives (as well as 20 headquartered employees). As provided by CMO 14, this discovery process was "designed to provide information to enable the parties to assess the larger pool of cases" and "to provide information to the Court to enable the Court to select which cases shall serve as the first bellwether trials." (Third Amended CMO 14 at 3 (Docket No. 1287).) The best and perhaps only benchmark for that assessment is the randomly selected pool of 100 cases. (Aug. 18, 2015 Hr'g Tr. at 84:14-86:4, 89:20-90:21 (discussing random selection of 100 cases against the whole pool as method for ensuring representativeness); *see also* Nov. 12, 2015 Hr'g Tr. at 53:19-54:21.)

A. Cross-Cutting Issues from the Pool of 100 Should Be Reflected in the Trial Bellwether Cases to Ensure Productive Bellwether Trials

Using the pool of 100 as a benchmark requires characterizing the 100 claimants according to criteria that then can be applied to the proposed trial bellwethers. Again, this mirrors the selection of the 32 discovery bellwethers. *See supra* Section II. As then, the cross-cutting issues should be used. Since the discovery bellwether selection, AbbVie has used the allegations in the Complaint, the data reflected in Plaintiff Fact Sheets, and the medical records and testimony obtained during core discovery to reduce its cross-cutting issues to focus on nine key groups that not only drive the resolution of factual and legal issues in the bellwether cases, but also are broadly represented in the pool of 100. Based on the information available to AbbVie at this time, the key cross-cutting issues broadly reflected in the pool of 100 plaintiffs are:

1. CV Medical Causation Groups

- Group 1: Plaintiff under 65 with history of prior CV disease (17 plaintiffs, 26% of CV cases).

- Group 2: Plaintiff under 65 with no known history of prior CV disease (31 plaintiffs, 48% of CV cases).

2. CV Warning Groups

- Group 3: Plaintiff's alleged CV injury occurred prior to publication of the first study to suggest a possible risk on July 8, 2010 (33 plaintiffs, 50% of CV cases).
- Group 4: Plaintiff's alleged CV injury occurred after July 8, 2010 but before publication of the second study to suggest a possible risk on November 6, 2013 (29 plaintiffs, 45% of CV cases).

3. Clot Medical Causation Groups

- Group 5: Plaintiff had hematocrit level below 50% at the time of injury (22 plaintiffs, 61% of clot cases).

4. Clot Warning Groups

- Group 6: Plaintiff's alleged clot injury occurred between December 2007 labeling and April 2011 labeling (4 plaintiffs, 11% of clot cases).
- Group 7: Plaintiff's alleged clot injury occurred after April 2011 labeling (30 plaintiffs, 83% of clot cases).

5. The Special Population of Plaintiffs 65 or Older at Time of First Prescription

- Group 8: Plaintiff first prescribed AndroGel when 65 years old or older in December 2007 or thereafter when a special population language was added for such older men in labeling (16 plaintiffs, 16%).

6. Marketing Group

- Group 9: Plaintiff first prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009 (74 plaintiffs, 74%).

IV. ABBVIE PROPOSAL

As indicated at the outset, AbbVie proposes that the Court identify the 8 bellwether trials using an approach that closely follows the one it adopted in selecting the discovery bellwethers. Again, each side should propose half of the cases in each of the two disease categories, taking care to select cases that are representative of the underlying pool of 100 cases. Recognizing that the Court may disagree with the parties' selections

or that there could be overlap between AbbVie's selections and Plaintiffs', AbbVie has proposed 3 CV and 3 Clot cases. Based upon the Court's review, it will reduce the AbbVie proposal to 2 of each.

Consistent with the selection of discovery bellwethers (and as suggested in the July 21 Minute Order), the Court should again act as the arbiter of representativeness. AbbVie suggests that the Court consider the following in conducting its review.

Factual setting for the review. *First*, as this Court has recognized, a smaller number of trial bellwethers makes it inherently more difficult to test and generate pre-trial decisions on key cross-cutting issues. (*See, e.g.*, Nov. 12, 2015 Hr'g Tr. 70:5-16 (increasing the number of trial bellwether cases from 6 to 12 in order to "increase the likelihood of getting decisions on more [cross-cutting] issues further down the road when we get to the Daubert [and] summary judgment stage.").) Although the discovery process has been robust, it remains the case that a process that was previously designed to involve a pool of 32 leading to a set of up to 12 trials must now apply to a reduced pool of 24 leading to only 8 trials.

Second, Plaintiffs still necessarily have superior knowledge of their own cases. This has been manifest in proceedings to date. For example, as described further above, AbbVie learned only several months into the fact discovery process that a number of the bellwether plaintiffs had used testosterone therapy products other than AndroGel prior to their alleged injuries—a fact that should have disqualified these plaintiffs at the outset. Although the parties have completed substantial discovery to date, because of the limited nature of core discovery, including the number of depositions that were practically feasible during the allotted time period and the lack of expert discovery, the

playing field remains unbalanced. The imbalance also was compounded by Plaintiffs' access to the prescribers and treaters outside of depositions. (Docket No. 1209.)

Third, the developing science underscores the importance of the cross-cutting issues identified by the parties. For example, a recent retrospective cohort study of 83,010 male veterans reported *fewer* heart attacks in men with low testosterone levels who were treated with TRT. R. Sharma et al., *Normalization of Testosterone Level is Associated with Reduced Incidence of Myocardial Infarction and Mortality in Men*, 36 Eur. Heart J. 40, 2706-15, 2714 (Aug. 2015) (Exhibit B); *id.* at 2714 ("Results from our present study suggest that in men without a history of previous [myocardial infarction] or stroke who have low [testosterone] levels, TRT might be associated with decreased risks of [myocardial infarction], ischaemic stroke, and all-cause mortality in long-term follow-up.").

Another recent study published in the New England Journal of Medicine examining the effects of testosterone therapy in men 65 years or older based on results from the Testosterone Trial ("T-Trial") reported positive benefits with respect to sexual function and mood and depressive symptoms. P.J. Snyder et al., *Effects of Testosterone Treatment in Older Men*, 374 New Eng. J. Med. 7, 611-24, 611 (Feb. 2016) (Exhibit C). The study also reported that the number of major cardiovascular events in the treatment and placebo group was the same (seven) during the treatment period. *Id.* at 616 (further reporting that during the subsequent year, two men in the treatment group and nine men in the placebo group had major cardiovascular events). Based on these results, the study authors concluded that there was no pattern of increased cardiovascular risk associated

with testosterone treatment, though they acknowledged that the T-Trial "was too small to exclude other than a large increase." *Id.* at 622.

With respect to thromboembolic clotting injuries, a retrospective cohort study of more than 100,000 men treated with testosterone therapy published in October 2015 reported: "No significant association was found between [exogenous testosterone therapy] and incidents of idiopathic or overall [venous thrombotic events] in men with hypogonadism." H. Li, et al., *Association between Use of Exogenous Testosterone Therapy and Risk of Venous Thrombotic Events among Exogenous Testosterone Treated and Untreated Men with Hypogonadism*, 195 J. Urology 4, 1065-72 (Oct. 2015) (accepted manuscript) (Exhibit D); *see also* R. Sharma, et al., *Association between Testosterone Replacement Therapy and the Incidence of Deep Vein Thrombosis and Pulmonary Embolism: A Retrospective Cohort Study of the Veterans Administration Database*, CHEST (2016), doi: 10.1016/j.chest.2016.05.007 (accepted manuscript) (reporting that study "did not detect a significant association between testosterone replacement therapy and risk of DVT/PE in adult men with low [testosterone] who were at low-moderate baseline risk of DVT/PE") (Exhibit E).

These developments underscore the importance of using the pretrial litigation to present the full picture of the relevant, reliable science on these issues.

Methodology. The Court's July 21 Minute Order directed the parties to state the reasons why the bellwethers they do not propose are "not representative or otherwise should not be selected." (Docket No. 1398.) AbbVie understands this to be a continuation of the approach taken last year when "outliers" were identified, but AbbVie

is unclear as to whether something else is required. Accordingly, it has taken the following approach:

- Outliers: AbbVie identifies these and both states the cross-cutting issues they implicate but, importantly, the characteristics of their cases that are not only unique, but so substantial as to detract from their value in representing the pool—these cases are not just different but distractingly so. Compared to the larger pool, their results likely would not be meaningful for purposes of generalization. We have identified the key outlier cases in Section IV.B below.

- Other cases not proposed: There are a number of cases that AbbVie does not propose but it does not seek to exclude as an outlier. These include cases that AbbVie believes less clearly frame cross-cutting issues or implicate issues that are less broadly represented in the pool of 100 cases.[1] With respect to this group of cases that are not proposed, AbbVie identifies any cross-cutting issues they implicate and any distinguishing features they present.

To facilitate the Court's analysis and arbitration of selection disputes, AbbVie has prepared a chart, attached as Exhibit A, detailing the applicable cross-cutting issues for each of the remaining 24 plaintiffs. These cases are discussed in further detail below, including AbbVie's proposed bellwether cases.

[1] In this regard, it should be noted that because only 8 trial bellwether cases will be selected, the most representative cases should implicate issues that are reflected in more than 12.5 percent of the broader pool of 100 cases.

A. AbbVie's Proposed Bellwether Cases

AbbVie requests that the Court select 8 trial bellwether cases and proposes the following 3 CV cases and 3 Clot cases to be selected. As shown below, AbbVie's proposals cover almost all of the cross-cutting issues described above, and, notably, all of the cross-cutting issues covered by the 24 discovery bellwether cases are also covered by at least one of AbbVie's proposed trial bellwethers.

1. CV Cases

Edward Cribbs (No. 1:15-cv-01056): Mr. Cribbs allegedly used AndroGel 1.0% from March 2010, when he was 59, until February 2012. He then allegedly used AndroGel 1.62% from February 2012 to April 2014. Mr. Cribbs was diagnosed with a myocardial infarction ("MI") in May 2012. Mr. Cribbs's case reflects two key cross-cutting issues related to CV injuries, including (a) he was under 65 at the time of injury with no apparent history of heart disease, and (b) he experienced a CV injury after July 2010 but before November 2013 (significant dates in the timeline of TRT CV science). Because he was prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009, Mr. Cribbs's case reflects a third, non-CV-specific cross-cutting issue as well. Mr. Cribbs does not have any prominent individual or demographic characteristics that would make his case unrepresentative. **Groups represented: 2, 4, and 9**.

Cecile Frost (No. 15-cv-01484): Mr. Frost allegedly used AndroGel 1.62% from January 2012 to February 2013 when he experienced a stroke. Mr. Frost's case reflects two key cross-cutting issues related to CV injuries, including (a) he was under 65 at the time of injury with no apparent history of heart disease, and (b) he

experienced a CV injury after July 2010 but before November 2013. Because he was prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009, Mr. Frost's case reflects a third, non-CV-specific cross-cutting issue as well. Mr. Frost does not have any prominent individual or demographic characteristics that would make his case unrepresentative. **Groups represented: 2, 4, and 9**.

Jeffrey Konrad (No. 15-cv-00966): Mr. Konrad allegedly used AndroGel 1.0% from May 2010 to July 2010 to treat diagnosed hypogonadism. On July 9, 2010, he was diagnosed with a MI. Mr. Konrad's case reflects two key cross-cutting issues related to CV injuries, including (a) he was under 65 with a history of prior CV disease when he was first prescribed AndroGel, and (b) he experienced a CV injury after publication of the first study to suggest a possible CV risk (July 8, 2010) but before publication of the second study to suggest a possible risk. Because he was prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009, Mr. Konrad's case reflects a third, non-CV-specific cross-cutting issue as well. Mr. Konrad does not have any prominent individual or demographic characteristics that would make his case unrepresentative. **Groups represented: 1, 4, and 9**.

2. Clot Cases

Froylan Garcia (No. 1:15-cv-01086): Mr. Garcia allegedly used AndroGel 1.62% from May 2013, at age 66, until August 2013. In September 2013, he was diagnosed with deep vein thrombosis ("DVT") in his right and left legs. Mr. Garcia's case reflects two key cross-cutting issues related to Clot injuries, including (a) he was over 65 at the time of his AndroGel prescription and after the December 2007 label change for such a special population, and (b) he experienced a Clot injury after the

April 2011 label change related to hematocrit levels. Because he was prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009, Mr. Garcia's case reflects a third, non-Clot-specific cross-cutting issue as well. Mr. Garcia does not have any prominent individual or demographic characteristics that would make his case unrepresentative. **Groups represented: 7, 8, and 9**.

Robert Rowley (No. 1:15-cv-02760): Mr. Rowley allegedly used AndroGel 1.62% from April 2012 to April 2013 when he was diagnosed with DVTs in both legs. Mr. Rowley was selected by the Court to ensure representation of plaintiffs older than 65 in the bellwether pool. (Nov. 20, 2015 Hr'g Tr. at 5:24-6:16; *see also* Nov. 12, 2015 Hr'g Tr. 75:16-76:1.) Mr. Rowley's case implicates three key cross-cutting issues related to Clot injuries, including (a) he had hematocrit below 50 percent at the time of injury, (b) he was over 65 at the time of his AndroGel prescription and after the December 2007 label change for such a special population, and (c) he experienced a Clot injury after the April 2011 label change related to hematocrit levels. Because he was prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009, Mr. Rowley's case reflects a fourth, non-Clot-specific cross-cutting issue as well. Mr. Rowley does not have any prominent individual or demographic characteristics that would make his case unrepresentative. **Groups represented: 5, 7, 8, and 9**.

Dale Shepherd (No. 1:15-cv-00404): Mr. Shepherd allegedly used AndroGel 1.0% from February 2011 to September 2011. He was diagnosed with a left leg DVT in March 2011. Mr. Shepherd's case reflects two key cross-cutting issues related to Clot injuries, including (a) he had hematocrit below 50 percent at the time of injury, and (b) he experienced an injury after the December 2007 label change but before

the April 2011 label change (both label changes related to hematocrit levels). Because he was prescribed AndroGel after the first disease-state awareness television commercial aired in May 2009, Mr. Shepherd's case reflects a third, non-Clot-specific cross-cutting issue as well. Mr. Shepherd does not have any prominent individual or demographic characteristics that would make his case unrepresentative. **Groups represented: 5, 6, and 9**.

B. Outlier Cases

For the reasons discussed below, AbbVie submits that the following cases would not facilitate the Court's stated goal of trying representative and productive issues and should be excluded from the bellwether trial pool.

1. CV Cases

David Deel (No. 14-cv-10435): Mr. Deel allegedly used AndroGel 1.0% from August 2008 until October 2012 and AndroGel 1.62% from October 2012 through December 2013. He experienced a MI in January 2014, which he alleges was caused by AndroGel. Mr. Deel's case is an outlier that should be excluded from trial because medical records and testimony indicate that he was prescribed the *Androderm* patch in 2008 and subsequently reported symptom improvement. Mr. Deel's non-AbbVie TRT use is a "complicating factor" that led the Court to decide Mr. Deel's case was "subject to later exclusion" from the bellwether pool. (Apr. 13, 2016 Hr'g Tr. 19:6-21; Order Regarding Rule to Show Cause, Apr. 21, 2016 (Docket No. 1268).) Exclusion of this case is necessary to protect the MDL process and ensure representative trials. **Groups represented: 2**.

Anthony Long (No. 14-cv-06996): Mr. Long intermittently used AndroGel during the 2010 to 2013 time period. He experienced a stroke in December 2013, stopped using AndroGel, and then had another stroke in April 2014 (after which he has been unable to work). Mr. Long's case is an outlier because he has a complex medical history that will require numerous additional depositions to ready the case for trial, including his spouse, additional prescribers (he had 5 prescribers in Florida and Tennessee), a neurologist, and an orthopedist, to name a few. That Mr. Long purports to have suffered two distinct injuries also creates unique causation issues, since a jury will need to determine what injuries are the result of the first stroke (alleged to have been caused by AndroGel) and which are the result of the subsequent stroke (at which time he was no longer using AndroGel). Finally, Mr. Long's case presents another outlying issue that threatens to distort the representativeness of any verdict: one of his prescribers is now a speaker for AbbVie. This is the case for only 7% of the doctors who prescribed to the plaintiff pool of 100—far less than 1 in 8 cases and so not representative of the bellwether pool of 100. **Groups represented: 2, 9**.

Roccie Truax (No. 14-cv-02935): Mr. Truax claims that his use of AndroGel 1.0% for only two months in 2013 caused his MI during the relevant July 2010 to November 2013 period. Mr. Truax's case is an outlier because of his short-term AndroGel use—he only filled one prescription for AndroGel—and pertinent medical history—including a 2006 heart attack and triple bypass and another heart attack in May 2016—that would complicate a jury's causation decision. Additional facts that could complicate a potential trial include that Mr. Truax is illiterate and his prescriber has had his medical license suspended. Selection of Mr. Truax's case will also require litigation

of novel questions of West Virginia law. West Virginia adopted the learned intermediary doctrine by statute in February 2016. No court has yet applied or addressed the retroactivity of this statutory provision. **Groups represented: 1, 4, and 9**.

David White (No. 14-cv-03818): Mr. White died as a result of a heart attack that his estate alleges was caused by AndroGel use. AbbVie has reviewed the two death cases in the pool of 24—Mr. White and Gene Dial (No. 15-cv-02190)—to determine whether they are sufficiently representative of the pool of 100. They are not. Only 7 other CV plaintiffs in the pool of 100 died as a result of their alleged injury, making this issue alone a disqualifying outlier. However, even if the Court decides it is appropriate to select a death case for trial, other issues render Mr. White's case unsuitable for bellwether selection as well. There are more than five years of medical records missing from Mr. White's files, including those relevant to Mr. White's medical condition prior to beginning testosterone treatment. Causation and efficacy will be especially difficult to evaluate in Mr. White's case because there are no records of testosterone level testing while he was using AndroGel. **Groups represented: 2, 4, and 9**.

2. Clot Cases

Lance Blanck (No. 1:15-cv-01077): Mr. Blanck allegedly used AndroGel 1.0% from June 2012 to January 2014 and suffered a DVT and pulmonary embolism in December 2013. Mr. Blanck's case is an outlier because of the unusual circumstances of his first AndroGel prescription, which will complicate pre-trial litigation and trial on the issue of failure to warn. In particular, Mr. Blanck testified that he received the prescription for AndroGel from a nurse practitioner who was a friend of his wife and that

he received the prescription for this controlled substance without ever having met or spoken to the nurse practitioner who wrote it. **Groups represented: 5, 7, and 9**.

Richard Cannon (No. 15-cv-01853): Mr. Cannon developed a DVT in March 2014, after allegedly using AndroGel 1.62% from February 2012 until March 2014. After stopping AndroGel, Mr. Cannon developed another DVT in September of 2015 and was diagnosed with chronic DVT in March of 2016. Mr. Cannon's case involves an individual fact that is not representative, but that will greatly effect a potential trial: Mr. Cannon has been disabled since 1982, after he broke his neck when diving into a shallow pond. Furthermore, Mr. Cannon's AndroGel was prescribed by nurse practitioners, not licensed physicians. In addition, there will be significant further work required to ready Mr. Cannon's case for trial: the medical records documenting his injury, which occurred in Puerto Rico, are incomplete and will likely require translation, and not all of the doctors who diagnosed and treated him there have been deposed. **Groups represented: 5, 7, and 9**.

Robert Cripe (No. 1:14-cv-00843): As discussed in AbbVie's November 9, 2015 Response to Plaintiffs' Bellwether Selections, Mr. Cripe's case is inappropriate for bellwether selection for a number of reasons, including the *extremely* short duration of his alleged AndroGel use, and the type and severity of his alleged injury. (Docket No. 1055 at 5-6.) Mr. Cripe claims that he used AndroGel for just five days from February 18, 2011 until February 23, 2011. A few days after he began treatment with AndroGel, he was diagnosed with transverse myelitis, which resulted in paraplegia—two conditions that no other plaintiffs have claimed. Mr. Cripe's socioeconomic circumstances and alleged damages—an annual salary in the mid-six figures and alleged

out-of-pocket damages alone exceeding $200,000—are far out of line with those of the vast majority of the other plaintiffs.

In addition, Mr. Cripe's post-Complaint medical treatment presents unique challenges that would interfere with the bellwether trial process. One month after filing his Complaint, Mr. Cripe, a Kansas resident, was treated by Cincinnati-based Dr. Charles Glueck. Dr. Glueck authored a number of articles regarding the potential risk of clots with testosterone therapy and is a potential expert witness. The Court is aware of these and other outlying complications with Mr. Cripe's case, including that, before he began using AndroGel, Mr. Cripe told his doctor he was using an "over-the-counter preparation to boost his testosterone." For those reasons, in April, the Court stated that Mr. Cripe's case was "subject to later exclusion" from the bellwether pool. (Order Regarding Rule to Show Cause, Apr. 21, 2016 (Docket No. 1268).) AbbVie submits that exclusion is now necessary to ensure productive and representative trials. **Groups represented: 5, 6, and 9**.

Robert Nolte (No. 14-cv-08894): Mr. Nolte allegedly used AndroGel 1.0% from August 2012 to November 2012 and suffered a pulmonary embolism in November 2012 at the age of 72. Mr. Nolte has a rare genetic predisposition to the development of blood clots, and he suffered multiple DVTs and a PE before beginning his AndroGel use. According to the available medical records, no other plaintiff in the pool of 100 has that genetic predisposition. Mr. Nolte's predisposition to thrombophilia, considered with the relatively short term of AndroGel use, creates a significant causation question that is likely inapplicable to most cases. **Groups represented: 5, 7, 8, and 9**.

Jesse Patridge (No. 14-cv-07960): Mr. Patridge was prescribed and allegedly used AndroGel from April 2010 until May 2011. He claims two injuries on two separate dates: a deep vein thrombosis in January 2011 and a pulmonary embolism in January 2013. Mr. Patridge's case is not representative of other Clot cases because (a) he continued to use AndroGel after his first alleged injury, and also alleges a second injury years after he discontinued using AndroGel; (b) he is one of only three clot plaintiffs who had a hematocrit level above 50% at the time of his injury; and (c) his case also presents discrete statute of limitations risk since his prescriber testified that Mr. Patridge's wife called him in 2011 stating that she was concerned AndroGel was the cause of his injuries. **Groups represented: 6, 8, and 9**.

C. Other Cases Not Proposed

1. CV Cases

Gene Dial (No. 1:15-cv-02190): Mr. Dial allegedly used AndroGel 1.62% for 9 months from June 2012 until a MI caused his death in March 2013. He was under 65 years old with no history of heart disease. Mr. Dial suffered from total testicular failure, which his doctor believed would require lifelong administration of TRT. His case also presents novel issues of West Virginia law. **Groups represented: 2, 4, and 9**.

Randy Martina (No. 14-cv-08598): Mr. Martina was prescribed AndroGel 1.62% from March 2012 through January 2014. He suffered a MI while he allegedly was using AndroGel and a stroke about six weeks after discontinuing use. Mr. Martina was under 65 years old and had no history of heart disease. Mr. Martina suffered two different kinds of injuries occurring during two different timeframes of

significance in the litigation—his MI occurred between July 2010 and November 2013 and his stroke occurred in March of 2014. Despite counsel's repeated requests, AbbVie has not been able to secure the deposition of Mr. Martina's second prescriber. **Groups represented: 2, 4, and 9**.

Jesse Mitchell (No. 14-cv-09718): Mr. Mitchell allegedly used AndroGel 1.0% from December 2007 until November 2012. He alleges that, as a result of his AndroGel use, he experienced a severe MI and cardiac arrest. The seriousness and effect of Mr. Mitchell's cardiac event is extraordinary among the group of 100 plaintiffs—medical records reflect that Mr. Mitchell was clinically dead in the ER (his heart stopped beating) and that he has experienced significant psychiatric issues after the cardiac event. **Groups represented: 1, 4**.

Joe Trusty (No. 15-cv-01015): Mr. Trusty alleges a "chest pain" injury as a result of his use of AndroGel 1.0% from November 2008 through August 2013. As with one of AbbVie's proposed cases, Mr. Trusty was under 65 with a history of prior CV disease when he was first prescribed AndroGel. And like many plaintiffs in this litigation, Mr. Trusty suffered his injury after July 2010 but before November 2013. **Groups represented: 1, 4**.

2. Clot Cases

Theodore Diesslin (No. 15-cv-01853): Mr. Diesslin was prescribed AndroGel 1.0% from August 2011 to July 2012 and AndroGel 1.62% from July 2012 to September 2012. Mr. Diesslin suffered a pulmonary embolism on September 10, 2012, at the age of 52. At the time of his pulmonary embolism, Mr. Diesslin's hematocrit was below 50. Mr. Diesslin alleges that his pulmonary embolism led to the development of a

fainting disorder that was ultimately treated with placement of a pacemaker. **Groups represented: 5, 7, and 9**.

Michael Ennis (No. 15-cv-00624): Mr. Ennis allegedly used AndroGel for just 17 days in 2007. Only three other clot plaintiffs in the pool of 100 were prescribed AndroGel before the December 2007 label change. In addition, Mr. Ennis testified that he suspected AndroGel caused his injury as early as 2007. This testimony creates a statute of limitations issue under California law unique to Mr. Ennis's case and potentially dispositive at the pre-trial stage. **Groups represented: 6**.

Kevin Hession (No. 14-cv-08222): Mr. Hession allegedly used AndroGel 1.0% from February 2012 through May 2012 and 1.62% from May 2013 through July 2013. Mr. Hession alleges he suffered a DVT in October 2012 at the age of 44 as a consequence of his AndroGel use. His physician testified, however, that this was not a new DVT but a residual clot from a DVT he suffered in November 2011 before beginning AndroGel. After discontinuing AndroGel, Mr. Hession experienced another DVT in late 2013 and a DVT with pulmonary embolism in 2014. **Groups represented: 7, 9**.

Arthur Myers (No. 15-cv-01085): Mr. Myers was prescribed AndroGel 1.0% from June 2003 until September 2008. He developed a pulmonary embolism in February of 2008 at the age of 42. Mr. Myers is one of the youngest AndroGel users in the pool. **Groups represented: 5, 6**.

Michael Romanik (No. 1:14-cv-08202): Mr. Romanik allegedly used AndroGel 1.62% from July 2011 to April 2012 when he suffered a pulmonary embolism at the age of 46. At the time of his pulmonary embolism, Mr. Romanik's hematocrit was

below 50. Prior to his pulmonary embolism, Mr. Romanik suffered from stage 3 kidney disease, nephrotic syndrome, and vasculitis. Despite counsel's efforts, AbbVie has been unable to secure the depositions of Mr. Romanik's prescribing and treating physicians. **Groups represented: 5, 7, and 9**.

V. CONCLUSION

For the foregoing reasons, AbbVie respectfully requests that the Court select the following cases to serve as the first bellwether trials: Edward Cribbs (No. 1:15-cv-01056), Cecile Frost (No. 15-cv-01484), Froylan Garcia (No. 1:15-cv-01086), Jeffrey Konrad (No. 15-cv-00966), Robert Rowley (No. 1:15-cv-02760), and Dale Shepherd (No. 1:15-cv-00404).

AbbVie further requests that the court *exclude* the following cases from selection as bellwether trial cases: Lance Blanck (No. 15-cv-01077), Richard Cannon, Sr. (No. 15-cv-01835), Robert Cripe (No. 1:14-cv-00843), David Deel (No. 14-cv-10435), Robert Nolte (No. 14-cv-08894); Jesse Patridge (No. 14-cv-07960), Anthony Long (No. 14-cv-06996), Roccie Truax (No. 14-cv-02935), and David White (No. 14-cv-03818).

Dated: July 25, 2016

Respectfully submitted,

By: /s/ *David M. Bernick*

David M. Bernick
PAUL, WEISS, RIFKIND,
WHARTON & GARRISON LLP
1285 Avenue of the Americas
New York, NY 10019-6064
Tel: (212) 373-3000
Fax: (212) 757-3990
dbernick@paulweiss.com

Attorney for AbbVie Inc.

Hope S. Freiwald
Friedrich-Wilhelm W. Sachse
DECHERT LLP
2929 Arch St., Cira Centre
Philadelphia, PA 19104-2808
Tel: (215) 994-4000
Fax: (215) 665-4000
hope.freiwald@dechert.com
will.sachse@dechert.com

Michelle Hart Yeary
DECHERT LLP
902 Carnegie Center, Suite 500
Princeton, NJ 08540-6531
Tel: (609) 955-3200
Fax: (609) 955-3259
michelle.yeary@dechert.com

Attorneys for AbbVie Inc. and Abbott Laboratories

CERTIFICATE OF SERVICE

I, Christopher Burrichter, hereby certify that on July 25, 2016, the foregoing document was filed via the Court's CM/ECF system, which will automatically serve and send email notification of such filing to all registered attorneys of record.

/s/*Christopher Burrichter*
Christopher Burrichter

Exhibit A

PROPOSAL FOR SELECTION OF BELLWETHER CASES FOR TRIAL

CV Cases

Plaintiff	NDIL Case No.	Injury	Groups Represented	Plaintiffs' Proposal	AbbVie's Proposal
Cribbs, Edward	15-cv-01056	MI	2, 4, 9		X
Deel, David	14-cv-10435	MI	2		
Dial, Gene	15-cv-02190	MI, Death	2, 4, 9		
Frost, Cecile	15-cv-01484	Stroke	2, 4, 9		X
Konrad, Jeffrey	15-cv-00966	MI	1, 4, 9		X
Long, Anthony	14-cv-06996	Stroke	2, 9		
Martina, Randy	14-cv-08598	MI, Stroke	2, 4, 9		
Mitchell, Jesse	14-cv-09178	MI	1, 4		
Truax, Roccie	14-cv-02935	MI	1, 4, 9		
Trusty, Joe	15-cv-01015	Chest pain	1, 4		
White, Dave	14-cv-03818	MI, Death	2, 4, 9		

Clot Cases

Plaintiff	NDIL Case No.	Injury	Groups Represented	Plaintiffs' Proposal	AbbVie's Proposal
Blanck, Lance	15-cv-01077	DVT, PE	5, 7, 9		
Cannon, Sr., Richard	15-cv-01835	DVT	5, 7, 9		
Cripe, Robert	14-cv-00843	Blood Clot	5, 6, 9		
Diesslin, Theodor	14-cv-06770	PE	5, 7, 9		
Ennis, Michael	15-cv-00624	DVT, PE	6		
Garcia, Froylan	15-cv-01086	DVT	7, 8, 9		X
Hession, Kevin	14-cv-08222	DVT	7, 9		
Myers, Arthur	15-cv-01085	PE	5, 6		
Nolte, Robert	14-cv-08135	PE	5, 7, 8, 9		
Patridge, Jesse	14-cv-07960	DVT, PE	6, 8, 9		
Romanik, Michael	14-cv-08202	PE	5, 7, 9		
Rowley, Robert	15-cv-02760	DVT	5, 7, 8, 9		X
Shepherd, Dale	15-cv-00404	DVT	5, 6, 9		X

Exhibit B

European Heart Journal (2015) **36**, 2706–2715
doi:10.1093/eurheartj/ehv346

FASTTRACK CLINICAL RESEARCH
Coronary artery disease

Normalization of testosterone level is associated with reduced incidence of myocardial infarction and mortality in men

Rishi Sharma[1], Olurinde A. Oni[1], Kamal Gupta[2], Guoqing Chen[3], Mukut Sharma[1], Buddhadeb Dawn[2], Ram Sharma[1], Deepak Parashara[2,4], Virginia J. Savin[5], John A. Ambrose[6], and Rajat S. Barua[1,2,4]*

[1]Division of Cardiovascular Research, Kansas City VA Medical Center, Kansas City, MO, USA; [2]Division of Cardiovascular Diseases, University of Kansas Medical Center, Kansas City, KS, USA; [3]Division of Health Services Research, University of Kansas Medical Center, Kansas City, KS, USA; [4]Division of Cardiovascular Medicine, Kansas City VA Medical Center, 4801 E. Linwood Boulevard, Kansas City, MO 64128, USA; [5]Division of Nephrology, Kansas City VA Medical Center, Kansas City, MO, USA; and [6]Division of Cardiovascular Medicine, University of California San Francisco, Fresno, CA, USA

Received 2 June 2015; revised 1 July 2015; accepted 6 July 2015; online publish-ahead-of-print 6 August 2015

Aims There is a significant uncertainty regarding the effect of testosterone replacement therapy (TRT) on cardiovascular (CV) outcomes including myocardial infarction (MI) and stroke. The aim of this study was to examine the relationship between normalization of total testosterone (TT) after TRT and CV events as well as all-cause mortality in patients without previous history of MI and stroke.

Methods and results We retrospectively examined 83 010 male veterans with documented low TT levels. The subjects were categorized into (Gp1: TRT with resulting normalization of TT levels), (Gp2: TRT without normalization of TT levels) and (Gp3: Did not receive TRT). By utilizing propensity score-weighted Cox proportional hazard models, the association of TRT with all-cause mortality, MI, stroke, and a composite endpoint was compared between these groups. The all-cause mortality [hazard ratio (HR): 0.44, confidence interval (CI) 0.42–0.46], risk of MI (HR: 0.76, CI 0.63–0.93), and stroke (HR: 0.64, CI 0.43–0.96) were significantly lower in Gp1 ($n = 43\,931$, median age = 66 years, mean follow-up = 6.2 years) vs. Gp3 ($n = 13\,378$, median age = 66 years, mean follow-up = 4.7 years) in propensity-matched cohort. Similarly, the all-cause mortality (HR: 0.53, CI 0.50–0.55), risk of MI (HR: 0.82, CI 0.71–0.95), and stroke (HR: 0.70, CI 0.51–0.96) were significantly lower in Gp1 vs. Gp2 ($n = 25\,701$, median age = 66 years, mean follow-up = 4.6 years). There was no difference in MI or stroke risk between Gp2 and Gp3.

Conclusion In this large observational cohort with extended follow-up, normalization of TT levels after TRT was associated with a significant reduction in all-cause mortality, MI, and stroke.

Keywords Testosterone replacement therapy • Myocardial infarction • Stroke

Introduction

Professional guidelines recommend testosterone replacement therapy (TRT) in patients with signs and symptoms of hypogonadism and documented evidence of low testosterone (T) levels.[1] The diagnosis of late-onset hypogonadism is on the rise with estimates that nearly 2.4 million men aged 40–69 suffer from hypogonadism in the USA.[2] Even though late-onset hypogonadism is not a universally accepted concept, and FDA has advised against T supplementation in men on the basis of age alone. However, in the last decade there has been a nearly 400% increase in the number of TRT prescriptions creating a billion dollar market.[3] With such widespread and ever increasing use of TRT, there has been growing concern regarding its effect on mortality and cardiovascular (CV) outcomes.

Recent retrospective studies, multiple meta-analyses, and a few small prospective studies have presented conflicting results and

* Corresponding author. Tel: +1 816 922 2442, Email: rajat.barua@va.gov

contributed to this uncertainty.[4–14] Observational studies suggested that low serum T level is associated with increased CV events.[4,15,16] Clinical trials examining TRT have been relatively small, and these trials were underpowered to provide conclusive evidence related to CV events.[9] For instance, a small prospective study in frail elderly men showed an increased incidence of CV events with TRT and was stopped early.[10] Two separate retrospective studies of men in the Veterans Affairs (VA) Health System using two different databases reported opposite effects of TRT on all-cause mortality.[11,14] In two very recent studies, Vigen *et al.*[11] using a VA database and Finkle *et al.*[12] using a healthcare database reported that men receiving TRT had an increased risk of myocardial infarction (MI). It is important to note that in many of these studies repeat measurements to document normalization of T levels after TRT were lacking. On the heels of these recently published data, the FDA issued a drug safety alert related to TRT (http://www.fda.gov/Drugs/DrugSafety/ucm383904.htm).

In light of these conflicting results and uncertainty concerning the safety of TRT, we have conducted a large retrospective study with long-term follow-up to address this knowledge gap. The objective of our study was to examine the association between TRT with documented normalization of total testosterone (TT) levels and all-cause mortality and adverse CV events defined by MI and stroke.

Methods

This is a retrospective cohort study of male veterans who received their medical care at the Veterans Health Administration (VHA) between December 1999 and May 2014. The data of study patients were retrieved from VHA Veterans Administrations Corporate Data Warehouse (CDW) through the Veterans Administrations Informatics and Computing Infrastructure (VINCI) [http://www.hsrd.research.va.gov/for_researchers/vinci/default.cfm (cited 21 June 2014)]. The study complies with the Declaration of Helsinki, and the Institutional Review Board of Kansas City Veterans Affairs Medical Center, MO, USA, approved the study. Additional details are provided in the Supplementary material online, *Appendix*.

Study design

This study was designed to examine the effect of TRT on CV outcomes by comparing the incidences of MI, stroke, and all-cause mortality among different sub-populations of treated and untreated patients. All patients' CV events and co-existing conditions were based on the International Classification of Diseases 9th Revision (ICD-9) codes. All of the study patients had TT levels checked at least on two separate occasions as recommended by guideline.[1]

Ascertainment of testosterone replacement therapy exposure

Use of TRT was ascertained from the medication prescription of patient medical records. For this study, patients who received any form of TRT (injection, gel or patch) were considered as treated.

Determination of total testosterone level

Low TT was determined to be present when TT level was less than the lower limit of normal laboratory reference range (NLRR) reported for that particular test result. This method was adopted to include results from a large number of laboratories in the entire VA Health System over a period of 14 plus years that used different test assays and had different reference ranges and reporting units. Data from position statement of Endocrine Society and several other sources suggest that testosterone levels can vary significantly between different laboratories, even when they use same commercial kits. Moreover, because of assay ambiguities and biological variations, no single cut-off T value can clearly distinguish between hypogonadism and eugonadism[17,18] There is also a lack of standardization when it comes to T levels and other tests using the stoichiometric measurements.[19,20] Hence, we classified each test result as low or normal based on its respective laboratory reference range reported. This approach permitted inclusion of testosterone values obtained using different assay methods and minimized the investigator bias likely introduced by an arbitrary cut-off value.

Outcome measures

Primary outcome measures were (i) the incidence of MI (ICD-9 410.x0 and 410.x1), (ii) the incidence of ischaemic stroke [ICD-9 433.x1, 434 (excluding 434.x0), or 436], and (iii) the all-cause mortality determined using dates of death in CDW data augmented with vital status files.

Additional details are provided in the Supplementary material online, *Appendix*.

Study population

Figure 1 presents the patient selection process.

Inclusion criteria

We included patients whose first tested TT level was lower than the respective laboratory NLRR.

Exclusion criteria

We excluded (i) females, (ii) those who received TRT before the first available low TT, (iii) those who had MI or ischaemic stroke before the first day of study, and (iv) those who on repeat testing had normal TT level before any treatment was started.

Eligible study patients were classified into three groups: Gp1: TRT with resulting normalization of TT levels (normalized-TRT); Gp2: TRT without normalization of TT levels (non-normalized-TRT); and Gp3: Did not receive TRT (no-TRT). Additional details are provided in the Supplementary material online, *Appendix*.

Statistical analysis

Continuous variables were reported as means and standard deviation (SD), categorical variables as percentages. Chi square test and Student's *t*-test were used to compare normally distributed baseline characteristics of patients. Non-parametric tests were used for non-normally distributed variables. We performed univariate and multivariable Cox proportional hazard regression analyses to assess the differences between groups. Furthermore, propensity scores were used to correct for potential systematic differences between treated and untreated patients. Each study patient's propensity scores for receiving the TRT were computed and adjusted for the covariates in a logistic regression analysis. The covariates included were age, body mass index (BMI), hypertension (HTN), diabetes mellitus (DM), chronic obstructive pulmonary disease (COPD), obstructive sleep apnoea (OSA), congestive heart failure (CHF), peripheral vascular disease (PVD), coronary artery disease (CAD), low density lipoprotein (LDL), use of aspirin, beta-blockers, and statins. All individuals with missing data on these matching covariates were excluded from the analysis. For robust analysis of our data, we utilized propensity score-weighted, stabilized inverse probability of treatment weights (IPTW); this allowed us to keep all patients in the study while using the propensity scores to achieve balance between each pair of subgroups we studied.[21–23]

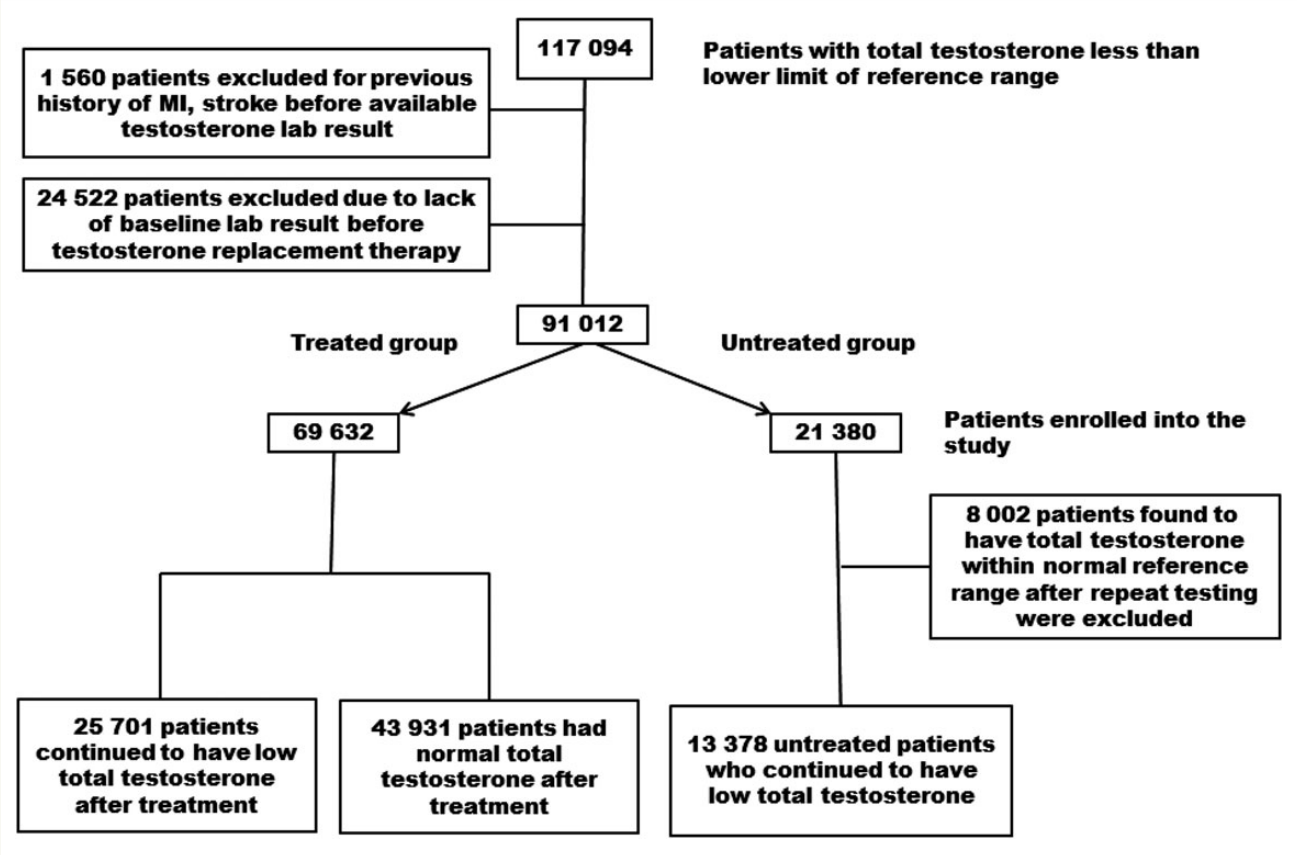


Figure 1 Methodology and patients selection process.

We also applied the stabilized IPTW to obtain Kaplan–Meier (KM) survival curves and to compare event-free survival time between the groups, along with log-rank *P*-value. SAS 9.4 was used for statistical analyses while Stata 12 was used to plot KM curves with TRT as a time-varying exposure variable. The study hypotheses were tested at two-sided level of significance with a *P*-value of <0.05. The use of IPTW effectively controlled for the imbalances in the groups as shown by the *P*-values (*Table 1*). Further details regarding how IPTW model was utilized in our study are described in the Supplementary material online, *Appendix*. Supplementary material online, *Figures S5–S7*, *Appendix* show how variations in low and high propensity scores in the unmatched pairs of cohorts were controlled for by IPTW.

Results

Cohort description

As shown in *Figure 1*, the initial cohort consisted of 117 094 patients with low TT. One thousand five hundred and sixty patients were excluded as they had a MI or stroke prior to the assessment of TT levels. These individuals were excluded because our study was focused on incident events. We then excluded 24 522 patients whose pretreatment baseline TT levels could not be ascertained. The remaining 91 012 patients were included in the study and categorized into those who received TRT at any time after they were determined to have low testosterone (81.5%) and those who did not receive TRT (18.5%). Testosterone replacement therapy achieved normalization of TT levels in 43 931 (63.1%) patients while the rest of this group continued to have low TT. Mean duration of treatment for normalized-TRT group was 3.0 ± 2.7 years and for non-normalized group was 1.5 ± 1.9 years.

In the untreated cohort, we identified certain individuals whose TT levels normalized at repeat testing ($n = 8002$). Though there was no record of treatment for these people, we could not rule out the possibility of non-VA prescriptions which could have been responsible for this finding. To prevent misclassification bias, individuals with these spuriously normalized TT levels were excluded leaving an *N* of 83 010. The percentage of people showing normal TT levels on repeat testing was around 30%; this number is consistent with the findings from population-based studies in which 1/3 of subjects showed normal TT levels on repeat testing.[24]

Baseline characteristics of the patients

Table 1 presents the baseline characteristics of the three groups. By means of stabilized IPTW, while performing Cox proportional hazard regression analyses, we controlled for discrepancies related to age, BMI, HTN, DM, COPD, OSA, CHF, PVD, CAD, LDL, use of aspirin, beta-blockers, and statins in the study groups by ensuring the cohorts were well matched ($P > 0.05$).

Table 1 **Baseline characteristics of all study subjects**

	Unmatched cohort			Propensity-matched cohort (stabilized IPTW)		
	Normalized treated vs. untreated					
	Normalized treated ***N*** **= 43 931**	**Untreated** ***N*** **= 13 378**	***P*-value**	**Normalized treated** ***N*** **= 40 852**	**Untreated** ***N*** **= 11 957**	***P*-value**
Age ≥ 50 years, *n* (%)	38 968 (89.4)	11 998 (90.3)	0.0055	36 641 (89.7)	10 716 (89.6)	0.8229
Age, median (Years)	66.0	67.0		66.0	67.0	
Body mass index ≥ 30 kg/m^2	28 670 (65.8)	8117 (63.7)	<0.0001	26 854 (65.7)	7871 (65.8)	0.8527
Body mass index, kg/m^2, mean (SD)	33.0 (6.6)	32.8 (6.9)		33.0 (6.6)	33.0 (6.8)	
Follow-up time (years), mean (SD)	6.2 (3.3)	4.7 (3.1)		6.0 (3.1)	4.6 (2.9)	
Hypertension, *n* (%)	7465 (17.0)	2342 (17.5)	0.1671	7251 (17.8)	2128 (17.8)	0.9118
Diabetes mellitus, *n* (%)	13 318 (30.3)	4228 (31.6)	0.0046	12 826 (31.4)	3762 (31.5)	0.8983
Chronic obstructive pulmonary disease, *n* (%)	528 (1.2)	215 (1.6)	0.0003	546 (1.3)	161 (1.3)	0.9676
Obstructive sleep apnoea, *n* (%)	801 (1.8)	279 (2.1)	0.0509	814 (2.0)	240 (2.0)	0.9428
Congestive heart failure, *n* (%)	713 (1.6)	353 (2.6)	<0.0001	779 (1.9)	228 (1.9)	0.9846
Peripheral vascular disease, *n* (%)	357 (0.8)	165 (1.2)	<0.0001	379 (0.9)	111 (0.9)	0.9759
Coronary artery disease	2141 (4.9)	738 (5.5)	0.0029	2146 (5.3)	629 (5.3)	0.9804
Depression, *n* (%)	3590 (8.2)	844 (6.3)	<0.0001	3284 (8.0)	957 (8.0)	0.8917
LDL > 100 mg/dL, *n* (%)	21 403 (51.6)	6085 (48.6)	<0.0001	20 779 (50.9)	6087 (50.9)	0.9297
Concomitant therapy with						
Antiplatelet agents (ASA), *n* (%)	12 410 (28.3)	3916 (29.3)	0.0217	11 904 (29.1)	3480 (29.1)	0.9451
B-blockers, *n* (%)	16 022 (36.5)	5041 (37.7)	0.0110	15 439 (37.8)	4515 (37.8)	0.9555
Statins, *n* (%)	25 260 (57.5)	7716 (57.7)	0.7161	24 334 (59.6)	7117 (59.5)	0.9237
	Normalized treated vs. non-normalized treated					
	Normalized treated ***N*** **= 43 931**	**Non-normalized treated** ***N*** **= 25 701**	***P*-value**	**Normalized treated** ***N*** **= 40 852**	**Non-normalized treated** ***N*** **= 23 953**	***P*-value**
Age ≥ 50 years, *n* (%)	38 968 (89.4)	22 692 (88.8)	0.0189	36 484 (89.3)	21 389 (89.3)	0.9945
Age, median (Years)	66.0	66.0		66.0	65.0	
Body mass index ≥ 30 kg/m^2	28 670 (65.8)	17 460 (69.0)	<0.0001	27 554 (67.4)	16 161 (67.5)	0.9327
Body mass index, kg/m^2, mean (SD)	33.0 (6.6)	33.6 (6.9)		33.2 (6.6)	33.4 (6.9)	
Follow-up time (years), mean (SD)	6.2 (3.3)	4.6 (3.1)		6.0 (3.1)	4.5 (3.0)	
Hypertension, *n* (%)	7465 (17.0)	5114 (19.9)	<0.0001	7655 (18.7)	4492 (18.8)	0.9502
Diabetes mellitus, *n* (%)	13 318 (30.3)	9233 (35.9)	<0.0001	13 512 (33.1)	7971 (33.1)	0.9967
Chronic obstructive pulmonary disease, *n* (%)	528 (1.2)	460 (1.8)	<0.0001	608 (1.5)	358 (1.5)	0.9509
Obstructive sleep apnoea, *n* (%)	801 (1.8)	712 (2.8)	<0.0001	936 (2.3)	549 (2.3)	0.9977
Congestive heart failure, *n* (%)	713 (1.6)	666 (2.6)	<0.0001	836 (2.1)	490 (2.0)	0.9892
Peripheral vascular disease, *n* (%)	357 (0.8)	291 (1.1)	<0.0001	386 (1.0)	227 (1.0)	0.9916

Continued

Table I *Continued*

	Unmatched cohort			Propensity-matched cohort (stabilized IPTW)		
	Normalized treated vs. untreated					
	Normalized treated ***N*** **= 43 931**	**Untreated** ***N*** **= 13 378**	***P*-value**	**Normalized treated** ***N*** **= 40 852**	**Untreated** ***N*** **= 11 957**	***P*-value**
Coronary artery disease	2141 (4.9)	1623 (6.3)	<0.0001	2304 (5.6)	1352 (5.6)	0.9742
Depression, *n* (%)	3590 (8.2)	2249 (8.8)	0.0078	3539 (8.7)	2079 (8.7)	0.9437
LDL > 100 mg/dL, *n* (%)	21 403 (51.6)	11 676 (47.8)	<0.0001	20 473 (50.1)	11 997 (50.1)	0.9621
Concomitant therapy with						
Antiplatelet agents (ASA), *n* (%)	12 410 (28.3)	7808 (30.4)	<0.0001	12 125 (29.7)	7111 (29.7)	0.9763
B-blockers, *n* (%)	16 022 (36.5)	10 532 (41.0)	<0.0001	15 947 (39.0)	9350 (39.0)	0.9884
Statins, *n* (%)	25 260 (57.5)	15 775 (61.4)	<0.0001	24 809 (60.7)	14 541 (60.7)	0.9675
	Non-normalized treated vs. untreated					
	Non-normalized treated ***N*** **= 25 701**	**Untreated** ***N*** **= 13 378**	***P*-value**	**Non-normalized treated** ***N*** **= 23 953**	**Untreated** ***N*** **= 11 957**	***P*-value**
Age ≥ 50 years, *n* (%)	22 692 (88.8)	11 998 (90.3)	<0.0001	21 391 (89.3)	10 677 (89.3)	0.9613
Age, median (Years)	66.0	67.0		66.0	67.0	
Body mass index ≥ 30 kg/m^2	17 460 (69.0)	8117 (63.7)	<0.0001	16 191 (67.6)	8086 (67.6)	0.9634
Body mass index, kg/m^2, mean (SD)	33.6 (6.9)	32.8 (6.9)		33.5 (6.9)	33.3 (6.9)	
Follow-up time (years), mean (SD)	4.6 (3.1)	4.7 (3.1)		4.5 (2.9)	4.5 (2.9)	
Hypertension, *n* (%)	5114 (19.9)	2342 (17.5)	<0.0001	4740 (19.8)	2370 (19.8)	0.9431
Diabetes mellitus, *n* (%)	9233 (35.9)	4228 (31.6)	<0.0001	8470 (35.4)	4231 (35.4)	0.9671
Chronic obstructive pulmonary disease, *n* (%)	460 (1.8)	215 (1.6)	0.1884	431 (1.8)	214 (1.8)	0.9718
Obstructive sleep apnoea, *n* (%)	712 (2.8)	279 (2.1)	<0.0001	645 (2.7)	323 (2.7)	0.9563
Congestive heart failure, *n* (%)	666 (2.6)	353 (2.6)	0.7806	644 (2.7)	324 (2.7)	0.9054
Peripheral vascular disease, *n* (%)	291 (1.1)	165 (1.2)	0.3771	288 (1.2)	145 (1.2)	0.9190
Coronary artery disease	1623 (6.3)	738 (5.5)	0.0017	1510 (6.3)	756 (6.3)	0.9504
Depression, *n* (%)	2249 (8.8)	844 (6.3)	<0.0001	1966 (8.2)	984 (8.2)	0.9342
LDL > 100 mg/dL, *n* (%)	11 676 (47.8)	6085 (48.6)	0.1484	11 489 (48.0)	5746 (48.1)	0.8731
Concomitant therapy with						
Antiplatelet agents (ASA), *n* (%)	7808 (30.4)	3916 (29.3)	0.0233	7359 (30.7)	3676 (30.8)	0.9649
B-blockers, *n* (%)	10 532 (41.0)	5041 (37.7)	<0.0001	9775 (40.8)	4875 (40.8)	0.9429
Statins, *n* (%)	15 775 (61.4)	7716 (57.7)	<0.0001	14 868 (62.1)	7419 (62.0)	0.9541

Relationship between testosterone replacement therapy and all-cause mortality

All-cause mortality in the three groups was as follows: normalized-TRT (Gp 1) (1654), non-normalized-TRT (Gp2) (3004), and no-TRT (Gp3) (3635) per 100 000 person-years. Normalized-TRT group had significantly fewer deaths than no-TRT (stabilized IPTW, hazard ratio, HR: 0.44, confidence interval, CI 0.42–0.46, $P < 0.0001$) and non-normalized-TRT (stabilized IPTW, HR: 0.53, CI 0.50–0.55, $P < 0.0001$) groups (*Table 2*). Mortality was also significantly lower in the non-normalized-TRT group compared with those in no-TRT group (stabilized IPTW, HR: 0.84, CI 0.80–0.89, $P < 0.0001$). The KM curves showed that the normalized-TRT group was associated with significantly increased all-cause mortality-free survival (log-rank, $P < 0.05$) compared with the non-normalized-TRT or no-TRT groups (*Figure 2*).

Relationship between testosterone replacement therapy and myocardial infarction

Table 2 presents result of the unadjusted and adjusted risk of MI in the study groups. Incidence of MI in the three groups was as follows: normalized-TRT group (189), non-normalized-TRT group (261), and no-TRT group (263) per 100 000 person-years. In the stabilized IPTW, normalized-TRT group showed lower risk of MI than non-normalized-TRT (HR: 0.82, CI 0.71–0.95, $P = 0.008$) and no-TRT (HR: 0.76, CI 0.63–0.93, $P = 0.005$) groups. However, non-normalized-TRT group was not different from no-TRT group (HR: 0.98, CI 0.80–1.19, $P = 0.811$). *Figure 3* shows a comparison of the probability of MI-free survival among the three groups. The KM curves show that normalized-TRT group was associated with significantly increased MI-free survival (log-rank, $P < 0.01$) compared with non-normalized-TRT and no-TRT groups. We performed additional analysis for MI-free survival after truncating the follow-up beyond 10 years. Although we lost a significant proportion of the study population, the findings remained fairly consistent after these analyses. See results in Supplementary material online, *Table S5* and *Figure S8*, Appendix.

Relationship between testosterone replacement therapy and ischaemic stroke

The incidence of ischaemic stroke was as follows: normalized-TRT group (43), non-normalized-TRT group (57), and no-TRT group (59) per 100 000 person-years. Stabilized IPTW showed that normalized-TRT group had significantly lower stroke events compared with non-normalized-TRT (HR: 0.70, CI 0.51–0.96, $P = 0.028$) and no-TRT (HR: 0.64, CI 0.43–0.96, $P = 0.031$) groups (*Table 2*). There was no difference in the risk of stroke between non-normalized-TRT group and no-TRT group. Overall, there was a protective effect against stroke in normalized-TRT group, as suggested by KM curves in Supplementary material online, *Figure S4, Appendix*.

Discussion

In this study of men with low TT levels and without prior MI or stroke, normalization of TT levels using TRT is associated with lower all-cause mortality, fewer MIs, and ischaemic strokes. This

Table 2 Unadjusted and adjusted hazard ratios for all-cause mortality, MI, and stroke

Model	All-cause mortality			Myocardial infarction			Stroke		
	Hazard ratio	95% CI	*P*	Hazard ratio	95% CI	*P*	Hazard ratio	95% CI	*P*
Comparing normalized treated vs. untreated (ref = untreated)									
Univariate *N* = 43 931 vs. 13 378	0.40	0.39–0.43	<0.001	0.70	0.59–0.83	<0.001	0.57	0.40–0.82	0.002
Propensity matched (stabilized inverse probability of treatment weights) *N* = 40 852 vs. 11 957	0.44	0.42–0.46	<0.001	0.76	0.63–0.93	0.005	0.64	0.43–0.96	0.031
Comparing normalized treated vs. non-normalized treated (ref = non-normalized treated)									
Univariate *N* = 43 931 vs. 25 701	0.49	0.47–0.51	<0.001	0.74	0.64–0.85	<0.001	0.64	0.48–0.87	0.004
Propensity matched (stabilized inverse probability of treatment weights) *N* = 40 852 vs. 23 953	0.53	0.50–0.55	<0.001	0.82	0.71–0.95	0.008	0.70	0.51–0.96	0.028
Comparing non-normalized treated vs. untreated (ref = untreated)									
Univariate *N* = 25 701 vs. 13 378	0.83	0.79–0.87	<0.001	0.95	0.79–1.15	0.599	0.90	0.61–1.34	0.610
Propensity matched (stabilized inverse probability of treatment weights) *N* = 23 953 vs. 11 957	0.84	0.80–0.89	<0.001	0.98	0.80–1.19	0.811	0.94	0.61–1.44	0.675

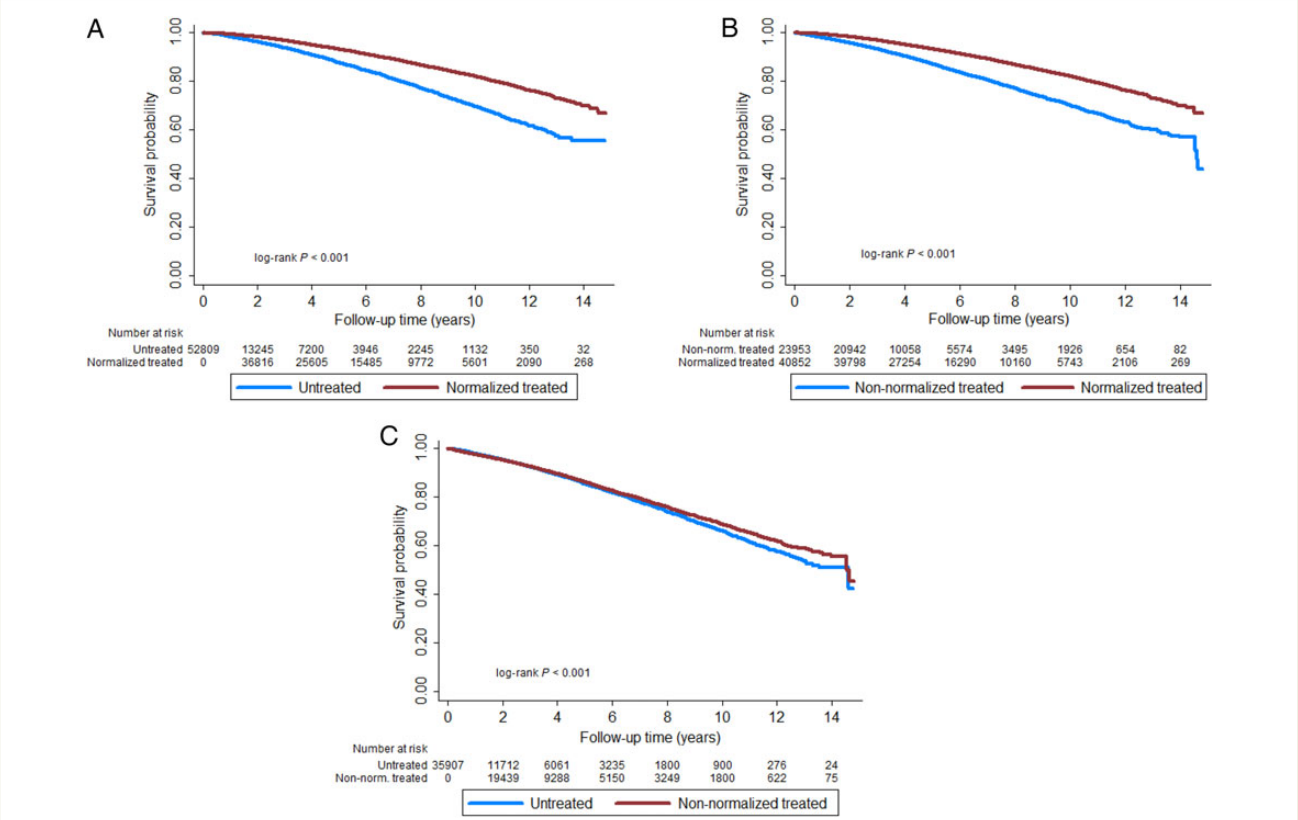


Figure 2 (*A*–*C*) Kaplan–Meier curve depicting the all-cause mortality among different propensity-matched study groups.

retrospective study describes the largest cohort of such patients and the longest follow-up for TRT to date. It is the first study to demonstrate that significant benefit is observed only if the dose is adequate to normalize the TT levels. Patients who failed to achieve the therapeutic range after TRT did not see a reduction in MI or stroke and had significantly less benefit on mortality. We selected patients without any previous history of MI or stroke prior to initiation of TRT to reduce bias related to CV outcomes. Further, we employed stabilized IPTW to decrease confounders by balancing measurable covariates between the groups. We modelled TRT as a time-varying covariate to account for the interval contributed by the treated individuals between enrollment (first low TT laboratory results) and the onset of TRT. We attempted to overcome the potential limitation of inadequate treatment by using follow-up TT levels as a marker of adequacy of dosing and compliance. We believe that our design criteria permit confidence in interpretation.

Several recent retrospective studies have investigated the association of TRT with CV outcomes. Vigen *et al.*[11] utilized the VA Clinical Assessment Reporting and Tracking (CART-CL) database that collects data from VA cardiac catheterization laboratories.[25] Their study enrolled patients who had cardiac catheterization done between 2005 and 2011 and also had low TT. The authors compared those who received TRT with those who did not. In that population, TRT was associated with significantly higher adverse events (MI, strokes, and death; HR: 1.29, 95% CI 1.05–1.58, $P = 0.02$). Our study differs from this study in several important ways. We included all patients who had their TT level checked, and we divided them into two groups: Gp1, patients who showed a documented appropriate rise in testosterone level post-TRT and Gp2, patients who did not achieve an appropriate rise. In comparison, Vigen *et al.* only included hypogonadal men who had undergone coronary angiography. This inclusion criterion may have introduced selection bias towards inclusion of a high CV-risk population. In this study, nearly 40% of the cohort had no repeat TT levels checked. Additionally, on the basis of the mean TT levels reported in the study by Vigen *et al.*, a number of patients likely did not achieve normalization of TT levels following TRT and, thus, may have reflected a subsequent risk of non-normalized hypogonadal men rather than a cohort with normalized TT level after TRT. Our study population was relatively healthier with lower average age (~64.2 years). Furthermore, we assessed hard end points (MI, stroke, and all-cause mortality).

The Testosterone in Older Men (TOM) trial[10] was a prospective, randomized placebo controlled study that was designed to determine the effects of TRT on lower extremity strength and physical function in older men with limitations in mobility and low serum levels of TT or free T. This trial was stopped prematurely at 6 months because of increased CV-related events in the TRT group. This trial had a small sample size (209 men), higher than average prevalence of chronic diseases (DM, HTN, and dyslipidaemia) in the cohort, and advanced age (mean age ~74 years); the adverse CV events were diverse and some were of variable clinical importance such as peripheral oedema, ectopy on ECG, and elevated BP. An additional

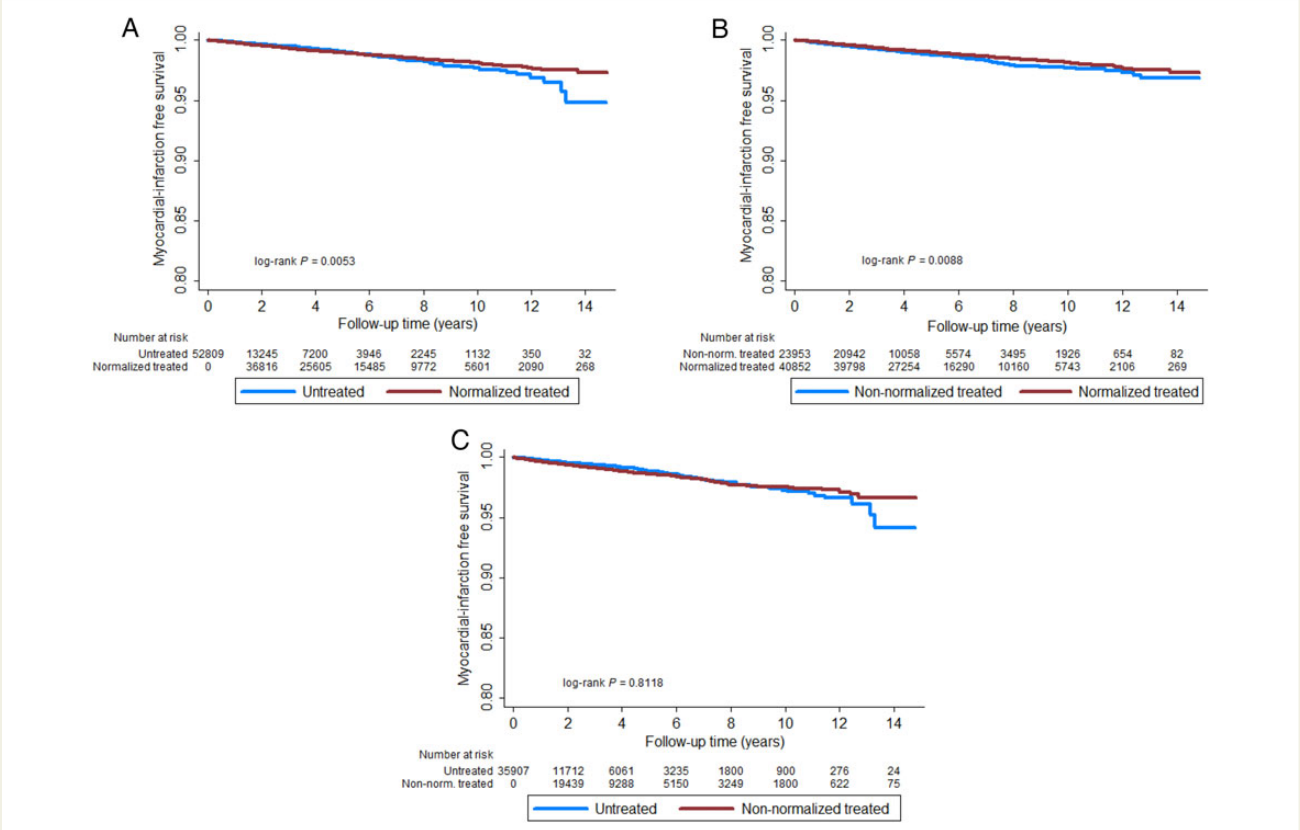


Figure 3 (*A*–*C*) Kaplan–Meier curve depicting the myocardial infarction-free survival among different propensity-matched study groups.

point worth noting was that early termination of the TOM trial may have contributed to an overestimation of the differences ascribed to treatment. In fact, some previous similar trials did not show an increased risk of adverse CV events with TRT therapy.[9,26]

A recent study by Finkle *et al.*[12] reported the risk of non-fatal MI in 90 days following a T prescription and compared it with the MI risk in the year preceding the prescription. They found that in older men (≥75 years) and in younger men with pre-existing heart disease, the risk of CV events was higher following a T prescription. However, this study did not take into account T levels. Thus, it is unclear how many people were adequately treated. Testosterone replacement therapy usually is a long-term therapy. These investigators limited the follow-up to 3 months of therapy. It is unknown whether this short duration of follow-up was sufficient to capture the outcomes of interest.

Our results do concur with a previous VA study. Shores *et al.*[14] analysed data from seven VA medical centers. They found that TRT was associated with a significant decrease in all-cause mortality (HR: 0.61, CI 95%, $P < 0.0001$). While supporting the results of Shores *et al.*, our study adds significantly to its conclusions both due to much larger sample size and also by more accurately identifying those who actually received and responded to the TRT. Shores *et al.* obtained data from the VA pharmacy records on T prescriptions, and those who received prescription were classified as treated. However, information regarding post-treatment TT level was not available in this study. Our study utilized post-TRT normalization in TT levels as a surrogate for administration of adequate therapy.

While our data found that normalization of TT levels after TRT was beneficial against CV risk and all-cause mortality, the mechanisms for these effects remain speculative. It can be postulated that the beneficial effect of normal T levels on adipose tissue, insulin sensitivity, and lipid profiles or by its anti-inflammatory and anticoagulant properties, as reported by other investigators, might have contributed to our findings.[27–29] However, there are other potential mechanisms such as sodium retention, CHF, increased platelet aggregation, or adverse changes in HDL through which T may increase the CV risks.[6] Therefore, additional studies will be needed to appropriately identify the mechanisms responsible for the outcomes noted in our study.

Finally, off-label use of TRT remains a concern. Recent FDA analyses suggest that currently only half of the men on TRT had been diagnosed with hypogonadism.[30] Furthermore, 25% of users did not have their T concentrations tested prior to initiating therapy, and 21% of those prescribed TRT did not have their levels tested at any time during treatment. Recently, a second advisory from the FDA posted caution about using testosterone products for low T due to ageing and requires labelling change to inform of possible increased risk of heart attack and stroke with use.[31] However, two very recent meta-analyses suggested a lack of convincing evidence posed by TRT.[32,33] Therefore, for now, to maximize the benefits of TRT and to mitigate potential risks, there is a need

for guideline-directed TRT with continuous active surveillance for potential risk in various cohorts of patients.

Study limitations

This was an observational study. Thus, unmeasured confounding or hidden bias might be present. A significant limitation of retrospective studies on TRT has been the inability to fully ascertain whether patients in the treatment arm actually took the medications in an adequate dose. This current study mostly overcomes this limitation by assessing follow-up TT levels. Normalization of follow-up TT levels is in our judgment a reliable surrogate for adequacy of dosing and compliance. Additionally, we could not ascertain the time of the day when the specimens for TT levels were drawn. Blood samples are usually collected during morning hours in the VA healthcare system. If some patients had their blood drawn after the morning hours, their levels would be underestimated. Furthermore, entry criteria and outcomes were determined using ICD-9 codes, and the VA cohort ICD-9 codes have been shown to be valid in determining outcomes.[11]

Another limitation of our study is that there was no randomization. Also our database does not have all the clinical data regarding indications for initiating TRT and not initiating TRT. Therefore, we cannot rule out the possibility that TRT may have been offered by a physician to healthier subjects and not to men who were less well. In our study, data regarding clinical response to TRT were also not available. Similarly, the available data do not permit us to ascertain the quality of care and/or poor compliance as reason(s) for persistent low testosterone levels observed in some individuals.

Despite the limitations associated with a retrospective study, our study has the advantages of having a large subject population with extensive follow-up. Our findings show that effective TRT is associated with lower rates of CV events in men without previous history of MI or stroke, in whom low TT levels are documented and effective TRT is provided. Safety and outcome of TRT in other populations remain to be determined.

Conclusion

Results from our present study suggest that in men without a history of previous MI or stroke who have low TT levels, TRT might be associated with decreased risks of MI, ischaemic stroke, and all-cause mortality in long-term follow-up. Our study also highlights that TRT should aim for doses resulting in normalization of TT level as this was shown to be associated with reduction in adverse CV events. In the future, adequately powered, prospective, well-designed trials with a long-term follow-up will be needed to reach a conclusive agreement regarding the effect of TRT on CV risk.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgment

Authors gratefully acknowledge the valuable discussions with Dr Sharad Mathur, M.D., Chief, Pathology and Laboratory Medicine, Kansas City VA Medical center, Kansas City, MO.

Funding

This study was supported by the Kansas City Veterans Administration Medical Center, Kansas City, MO and the Midwest Biomedical Research Foundation (MBRF), Kansas City, MO. M.S. received support from the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development (VA BX001037 awarded to V.J.S.).

Conflict of interest: The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States Government. None of the authors has any conflict of interest to declare regarding the contents of the paper.

References

1. Bhasin S, Cunningham GR, Hayes FJ, Matsumoto AM, Snyder PJ, Swerdloff RS, Montori VM, Task Force ES. Testosterone therapy in men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2010;**95**:2536–2559.
2. Araujo AB, O'Donnell AB, Brambilla DJ, Simpson WB, Longcope C, Matsumoto AM, McKinlay JB. Prevalence and incidence of androgen deficiency in middle-aged and older men: estimates from the Massachusetts Male Aging Study. *J Clin Endocrinol Metab* 2004;**89**:5920–5926.
3. Baillargeon J, Urban RJ, Ottenbacher KJ, Pierson KS, Goodwin JS. Trends in androgen prescribing in the United States, 2001 to 2011. *JAMA Intern Med* 2013;**173**:1465–1466.
4. Haddad RM, Kennedy CC, Caples SM, Tracz MJ, Bolona ER, Sideras K, Uraga MV, Erwin PJ, Montori VM. Testosterone and cardiovascular risk in men: a systematic review and meta-analysis of randomized placebo-controlled trials. *Mayo Clin Proc* 2007;**82**:29–39.
5. Calof OM, Singh AB, Lee ML, Kenny AM, Urban RJ, Tenover JL, Bhasin S. Adverse events associated with testosterone replacement in middle-aged and older men: a meta-analysis of randomized, placebo-controlled trials. *J Gerontol A Biol Sci Med Sci* 2005;**60**:1451–1457.
6. Xu L, Freeman G, Cowling BJ, Schooling CM. Testosterone therapy and cardiovascular events among men: a systematic review and meta-analysis of placebo-controlled randomized trials. *BMC Med* 2013;**11**:108.
7. Cawthon PM, Ensrud KE, Laughlin GA, Cauley JA, Dam TT, Barrett-Connor E, Fink HA, Hoffman AR, Lau E, Lane NE, Stefanick ML, Cummings SR, Orwoll ES. Sex hormones and frailty in older men: the osteoporotic fractures in men (MrOS) study. *J Clin Endocrinol Metab* 2009;**94**:3806–3815.
8. Ruige JB, Mahmoud AM, De Bacquer D, Kaufman JM. Endogenous testosterone and cardiovascular disease in healthy men: a meta-analysis. *Heart* 2011;**97**:870–875.
9. Fernandez-Balsells MM, Murad MH, Lane M, Lampropulos JF, Albuquerque F, Mullan RJ, Agrwal N, Elamin MB, Gallegos-Orozco JF, Wang AT, Erwin PJ, Bhasin S, Montori VM. Clinical review 1: adverse effects of testosterone therapy in adult men: a systematic review and meta-analysis. *J Clin Endocrinol Metab* 2010;**95**:2560–2575.
10. Basaria S, Coviello AD, Travison TG, Storer TW, Farwell WR, Jette AM, Eder R, Tennstedt S, Ulloor J, Zhang A, Choong K, Lakshman KM, Mazer NA, Miciek R, Krasnoff J, Elmi A, Knapp PE, Brooks B, Appleman E, Aggarwal S, Bhasin G, Hede-Brierley L, Bhatia A, Collins L, LeBrasseur N, Fiore LD, Bhasin S. Adverse events associated with testosterone administration. *N Engl J Med* 2010;**363**:109–122.
11. Vigen R, O'Donnell CI, Baron AE, Grunwald GK, Maddox TM, Bradley SM, Barqawi A, Woning G, Wierman ME, Plomondon ME, Rumsfeld JS, Ho PM. Association of testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. *JAMA* 2013;**310**:1829–1836.
12. Finkle WD, Greenland S, Ridgeway GK, Adams JL, Frasco MA, Cook MB, Fraumeni JF Jr, Hoover RN. Increased risk of non-fatal myocardial infarction following testosterone therapy prescription in men. *PLoS ONE* 2014;**9**:e85805.
13. Baillargeon J, Urban RJ, Kuo YF, Ottenbacher KJ, Raji MA, Du F, Lin YL, Goodwin JS. Risk of myocardial infarction in older men receiving testosterone therapy. *Ann Pharmacother* 2014;**48**:1138–1144.
14. Shores MM, Smith NL, Forsberg CW, Anawalt BD, Matsumoto AM. Testosterone treatment and mortality in men with low testosterone levels. *J Clin Endocrinol Metab* 2012;**97**:2050–2058.
15. Yeap BB, Alfonso H, Chubb SA, Hankey GJ, Handelsman DJ, Golledge J, Almeida OP, Flicker L, Norman PE. 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**:4565–4573.
16. Kaufman JM. Mortality associated to late-onset hypogonadism: reasons not to treat with testosterone? *J Clin Endocrinol Metab* 2014;**99**:1161–1163.

17. Lazarou S, Reyes-Vallejo L, Morgentaler A. Wide variability in laboratory reference values for serum testosterone. *J Sex Med* 2006;**3**:1085–1089.
18. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. *J Clin Endocrinol Metab* 2007;**92**:405–413.
19. Wang C, Catlin DH, Demers LM, Starcevic B, Swerdloff RS. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry. *J Clin Endocrinol Metab* 2004;**89**:534–543.
20. Vesper HW, Botelho JC. Standardization of testosterone measurements in humans. *J Clin Endocrinol Metab* 2010;**121**:513–519.
21. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;**46**:399–424.
22. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;**11**:550–560.
23. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods* 2010;**15**:234–249.
24. Brambilla DJ, O'Donnell AB, Matsumoto AM, McKinlay JB. Intraindividual variation in levels of serum testosterone and other reproductive and adrenal hormones in men. *Clin Endocrinol (Oxf)* 2007;**67**:853–862.
25. Byrd JB, Vigen R, Plomondon ME, Rumsfeld JS, Box TL, Fihn SD, Maddox TM. Data quality of an electronic health record tool to support VA cardiac catheterization laboratory quality improvement: the VA Clinical Assessment, Reporting, and Tracking System for Cath Labs (CART) program. *Am Heart J* 2013;**165**:434–440.
26. Srinivas-Shankar U, Roberts SA, Connolly MJ, O'Connell MD, Adams JE, Oldham JA, Wu FC. Effects of testosterone on muscle strength, physical function, body composition, and quality of life in intermediate-frail and frail elderly men: a randomized, double-blind, placebo-controlled study. *J Clin Endocrinol Metab* 2010;**95**:639–650.
27. Kapoor D, Goodwin E, Channer KS, Jones TH. Testosterone replacement therapy improves insulin resistance, glycaemic control, visceral adiposity and hypercholesterolaemia in hypogonadal men with type 2 diabetes. *Eur J Endocrinol* 2006;**154**:899–906.
28. Saad F, Gooren L, Haider A, Yassin A. An exploratory study of the effects of 12 month administration of the novel long-acting testosterone undecanoate on measures of sexual function and the metabolic syndrome. *Arch Androl* 2007;**53**:353–357.
29. Channer KS. Endogenous testosterone levels and cardiovascular disease in healthy men. *Heart* 2011;**97**:867–869.
30. FDA background documents for the discussion of two major issues in testosterone replacement therapy (TRT): 1. The appropriate indicated population for TRT, and 2. The potential for adverse cardiovascular outcomes associated with use of TRT. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ReproductiveHealthDrugsAdvisoryCommittee/UCM412536.pdf (12 December 2014).
31. FDA Drug Safety Communication: FDA cautions about using testosterone products for low testosterone due to aging; requires labeling change to inform of possible increased risk of heart attack and stroke with use. http://www.fda.gov/Drugs/DrugSafety/ucm436259.htm
32. Morgentaler A, Miner MM, Caliber M, Guay AT, Khera M, Traish AM. Testosterone therapy and cardiovascular risk: advances and controversies. *Mayo Clin Proc* 2015;**90**:224–251.
33. Corona G, Maseroli E, Rastrelli G, Isidori AM, Sforza A, Mannucci E, Maggi M. Cardiovascular risk associated with testosterone-boosting medications: a systematic review and meta-analysis. *Expert Opin Drug Saf* 2014;**13**:1327–1351.

Exhibit C

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 FEBRUARY 18, 2016 VOL. 374 NO. 7

Effects of Testosterone Treatment in Older Men

P.J. Snyder, S. Bhasin, G.R. Cunningham, A.M. Matsumoto, A.J. Stephens-Shields, J.A. Cauley, T.M. Gill, E. Barrett-Connor, R.S. Swerdloff, C. Wang, K.E. Ensrud, C.E. Lewis, J.T. Farrar, D. Cella, R.C. Rosen, M. Pahor, J.P. Crandall, M.E. Molitch, D. Cifelli, D. Dougar, L. Fluharty, S.M. Resnick, T.W. Storer, S. Anton, S. Basaria, S.J. Diem, X. Hou, E.R. Mohler III, J.K. Parsons, N.K. Wenger, B. Zeldow, J.R. Landis, and S.S. Ellenberg, for the Testosterone Trials Investigators*

ABSTRACT

BACKGROUND

Serum testosterone concentrations decrease as men age, but benefits of raising testosterone levels in older men have not been established.

METHODS

We assigned 790 men 65 years of age or older with a serum testosterone concentration of less than 275 ng per deciliter and symptoms suggesting hypoandrogenism to receive either testosterone gel or placebo gel for 1 year. Each man participated in one or more of three trials — the Sexual Function Trial, the Physical Function Trial, and the Vitality Trial. The primary outcome of each of the individual trials was also evaluated in all participants.

RESULTS

Testosterone treatment increased serum testosterone levels to the mid-normal range for men 19 to 40 years of age. The increase in testosterone levels was associated with significantly increased sexual activity, as assessed by the Psychosexual Daily Questionnaire ($P<0.001$), as well as significantly increased sexual desire and erectile function. The percentage of men who had an increase of at least 50 m in the 6-minute walking distance did not differ significantly between the two study groups in the Physical Function Trial but did differ significantly when men in all three trials were included (20.5% of men who received testosterone vs. 12.6% of men who received placebo, $P=0.003$). Testosterone had no significant benefit with respect to vitality, as assessed by the Functional Assessment of Chronic Illness Therapy–Fatigue scale, but men who received testosterone reported slightly better mood and lower severity of depressive symptoms than those who received placebo. The rates of adverse events were similar in the two groups.

CONCLUSIONS

In symptomatic men 65 years of age or older, raising testosterone concentrations for 1 year from moderately low to the mid-normal range for men 19 to 40 years of age had a moderate benefit with respect to sexual function and some benefit with respect to mood and depressive symptoms but no benefit with respect to vitality or walking distance. The number of participants was too few to draw conclusions about the risks of testosterone treatment. (Funded by the National Institutes of Health and others; ClinicalTrials.gov number, NCT00799617.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Snyder at pjs@mail.med.upenn.edu.

*A complete list of investigators in the Testosterone Trials is provided in the Supplementary Appendix, available at NEJM.org.

Drs. Bhasin, Cunningham, Matsumoto, Stephens-Shields, and Ellenberg contributed equally to this article.

N Engl J Med 2016;374:611-24.
DOI: 10.1056/NEJMoa1506119

Testosterone concentrations in men decrease with increasing age.[1,2] Many symptoms and conditions similar to those that are caused by low testosterone levels in men with pituitary or testicular disease become more common with increasing age. Such symptoms include decreases in mobility, sexual function, and energy. These parallels suggest that the lower testosterone levels in older men may contribute to these conditions.

A Quick Take is available at NEJM.org

Previous trials of testosterone treatment in men 65 years of age or older, however, have yielded equivocal results. Although testosterone treatment consistently increased muscle mass and decreased fat mass,[3,4] effects on physical performance,[3,5,6] sexual function,[3,6,7] and energy[3,6,8] have been inconsistent.

In 2003, an Institute of Medicine panel concluded that there was insufficient evidence that testosterone treatment was beneficial in older men[9] and recommended a coordinated set of clinical trials to determine whether testosterone would benefit older men who had low testosterone levels for no known reason other than age and who had clinical conditions to which low testosterone might contribute. The Testosterone Trials were designed to implement that recommendation.[10]

Methods

Study Design and Oversight

The Testosterone Trials are a coordinated set of seven double-blind, placebo-controlled trials that are being conducted at 12 sites.[10] To enroll in these trials overall, participants had to qualify for at least one of the three main trials (the Sexual Function Trial, the Physical Function Trial, or the Vitality Trial), but they could participate in more than one if they qualified. Participants were assigned to receive testosterone gel or placebo gel for 1 year. Efficacy was assessed at baseline and at 3, 6, 9, and 12 months. Data on adverse events were collected during the treatment period and for 12 months afterward. This report describes the efficacy results for the three main trials and adverse events in all the participants in these trials.

The protocol and consent forms were approved by the institutional review boards at the University of Pennsylvania and each participating trial site. All participants provided written informed consent. A data and safety monitoring board monitored data in an unblinded fashion every 3 months. The protocol, consent forms, and statistical analysis plan are available with the full text of this article at NEJM.org.

The investigators developed the protocol with assistance from the National Institutes of Health. AbbVie, one of the funders of the trial, donated the testosterone and placebo gels but did not participate in the design or conduct of the trials or in the analysis, review, or reporting of the data before the manuscript was submitted for publication. All the authors participated in the design and conduct of the trials. Trial statisticians performed all data analyses. The first author wrote the first draft of the manuscript, and all the authors contributed to subsequent drafts.

Participants

Participants were recruited principally through mass mailings.[11] Respondents were screened first by telephone interview and then during two clinic visits. Eligibility criteria included an age of 65 years or older and serum testosterone levels that averaged less than 275 ng per deciliter. Exclusion criteria were a history of prostate cancer, a risk of all prostate cancer of more than 35% or of high-grade prostate cancer of more than 7% as determined according to the Prostate Cancer Risk Calculator,[12] an International Prostate Symptom Score (IPSS; range, 0 to 35, with higher scores indicating more severe symptoms of benign prostatic hyperplasia) of more than 19, conditions known to cause hypogonadism, receipt of medications that alter the testosterone concentration, high cardiovascular risk (myocardial infarction or stroke within the previous 3 months, unstable angina, New York Heart Association class III or IV congestive heart failure, a systolic blood pressure >160 mm Hg, or a diastolic blood pressure >100 mm Hg), severe depression (defined by a score of ≥20 on the Patient Health Questionnaire 9 [PHQ-9; range, 0 to 27, with higher scores indicating greater severity of depressive symptoms]), and conditions that would affect the interpretation of the results.

Inclusion in the Sexual Function Trial required self-reported decreased libido, a score of 20 or less on the sexual-desire domain (range, 0 to 33, with higher scores indicating greater desire) of the Derogatis Interview for Sexual Functioning in Men–II (DISF-M-II),[13] and a part-

ner willing to have intercourse twice a month. Inclusion in the Physical Function Trial required self-reported difficulty walking or climbing stairs and a gait speed of less than 1.2 m per second on the 6-minute walk test.[14] Men who were not ambulatory or who had disabling neuromuscular or arthritic conditions were excluded. Inclusion in the Vitality Trial required self-reported low vitality and a score of less than 40 on the Functional Assessment of Chronic Illness Therapy (FACIT)–Fatigue scale (range, 0 to 52, with higher scores indicating less fatigue).[15]

STUDY TREATMENT

We assigned participants to testosterone or placebo by means of a minimization technique, with participants assigned to the study treatment that best balanced the balancing factors between groups with 80% probability.[16,17] Balancing variables included participation in the main trials, trial site, screening testosterone concentration (≤200 or >200 ng per deciliter), age (≤75 or >75 years), use or nonuse of antidepressants, and use or nonuse of phosphodiesterase type 5 inhibitors.

The testosterone preparation was AndroGel 1% in a pump bottle (AbbVie). The initial dose was 5 g daily. The placebo gel was formulated to have a similar application and appearance. Serum testosterone concentration was measured at months 1, 2, 3, 6, and 9 in a central laboratory (Quest Clinical Trials), and the dose of testosterone gel was adjusted after each measurement in an attempt to keep the concentration within the normal range for young men (19 to 40 years of age). To maintain blinding when the dose was adjusted in a participant receiving testosterone, the dose was changed simultaneously in a participant receiving placebo.

ASSESSMENTS

At the end of the trials, the serum concentrations of total testosterone, free testosterone, dihydrotestosterone, estradiol, and sex hormone–binding globulin were measured in serum samples frozen at −80°C (see the Supplementary Appendix, available at NEJM.org). Steroid assays were performed at the Brigham Research Assay Core Laboratory (Boston) by liquid chromatography with tandem mass spectroscopy, and free testosterone was measured by equilibrium dialysis. All samples from each participant were measured in the same assay run.

Serum prostate-specific antigen (PSA) was measured and a digital rectal examination was performed at months 3 and 12, and PSA was measured at month 18. Detection of a prostate nodule or a confirmed increase in the PSA level by at least 1.0 ng per milliliter above baseline led to referral to the site urologist for consideration of prostate biopsy. The IPSS was determined at months 3 and 12. At every visit, adverse events were recorded and a cardiovascular-event questionnaire (see the protocol) was administered. Cardiovascular events were adjudicated by two cardiologists and two neurologists (see the Supplementary Appendix).

OUTCOMES

Efficacy outcomes were assessed at baseline and after 3, 6, 9, and 12 months of treatment. Dichotomous outcomes were used when a clinically important difference had previously been established. The primary efficacy outcome of each trial and the secondary outcomes of the Physical Function Trial were assessed in all participants; secondary outcomes for the other trials were assessed only in participants in those trials.

The primary outcome of the Sexual Function Trial was the change from baseline in the score for sexual activity (question 4) on the Psychosexual Daily Questionnaire (PDQ-Q4; range, 0 to 12, with higher scores indicating a greater number of activities).[10,18] Secondary outcomes were changes in the score on the erectile-function domain (range, 0 to 30, with higher scores indicating better function) of the International Index of Erectile Function (IIEF)[19] and the sexual-desire domain of the DISF-M-II.[13] Details on the assessments in the Sexual Function Trial are provided in the protocol. The primary outcome of the Physical Function Trial was the percentage of men who increased the distance walked in the 6-minute walk test by at least 50 m.[10,14] Secondary outcomes were the percentage of men whose score on the physical-function domain (PF-10; range, 0 to 100, with higher scores indicating better function) of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) increased by at least 8 points[20] and changes from baseline in the 6-minute walking distance and PF-10 score. The primary outcome of the Vitality Trial was the percentage of men whose score on the FACIT–Fatigue scale increased by at least

4 points[10,15]; secondary outcomes were the change from baseline in the FACIT–Fatigue, the score on the vitality scale (range, 0 to 100, with higher scores indicating more vitality) of the SF-36,[21] scores on the Positive and Negative Affect Schedule (PANAS) scales (range, 5 to 50 for positive affect and for negative affect, with higher scores indicating a greater intensity of the affect),[22] and the PHQ-9 depression score.[23] Every 3 months, participants were asked about their general impression of the change in sexual desire, walking ability, or energy (depending on the trial) and in overall health.

STATISTICAL ANALYSIS

Participants were evaluated according to the intention-to-treat principle. Each outcome was prespecified. Primary analyses of outcomes at all time points were performed with random-effects models for longitudinal data. Models included visit time as a categorical variable and a single main effect for treatment. For linear models of continuous outcomes, the treatment effect denoted the average difference in response between study groups across all four visits. For logistic models of binary outcomes, the treatment effect was the log odds ratio of a positive versus negative outcome for participants who received testosterone versus those who received placebo, averaged over all visits. Additional fixed effects were the baseline value for each outcome and balancing variables. Random intercepts were included for participant.

We analyzed the three trials as independent studies, without adjusting analyses of the primary outcomes for multiple comparisons. We also did not adjust the analyses of the primary and secondary outcomes within each trial for multiple comparisons, because the correlations among outcomes within a trial were expected to be very high, making such adjustment excessively conservative. Analyses of the primary outcomes that included all participants, however, were adjusted for multiple comparisons; we report the nominal P value only when it was lower than the threshold specified by the multiple-comparisons procedure.[24] The sensitivity of results to missing data was assessed with the use of pattern-mixture models[25] and shared random-effects models.[26] The effect of change in total testosterone level on primary outcomes was assessed with the use of instrumental variables by two-stage residual inclusion,[27] with study-group assignment as the instrument and change in testosterone level from baseline as the exposure of interest.

Sample sizes were calculated such that the studies would have 90% power, with the use of a two-sided test at a type I error rate of 0.05,[10] to detect the following differences between the placebo group and the testosterone group: 15% versus 30% in the proportion of men with an increase of at least 50 m in the 6-minute walking distance, 20% versus 35% in the proportion of men with an increase of at least 4 points in the FACIT–Fatigue score, and a difference in change of 0.75 in the PDQ-Q4 score. These differences were conservatively based on comparisons between baseline and 12 months. Enrollment targets were 275 men for the Sexual Function Trial, 366 for the Physical Function Trial, and 420 for the Vitality Trial.

RESULTS

PARTICIPANTS AND STUDY TREATMENT

We screened 51,085 men and enrolled 790 who met all the criteria (Fig. S1 in the Supplementary Appendix).[11] Relatively few men had a sufficiently low testosterone level to qualify; only 4700 of 21,940 men (21.4%) who had blood sampled qualified by the first measurement and 1490 of 2163 men (68.9%) qualified by the second, for an overall inclusion rate by testosterone level of 14.7%.[11]

At baseline, the enrollees had unequivocally low serum testosterone concentrations according to criteria for healthy young men (Fig. S2 in the Supplementary Appendix). The participants had relatively high rates of coexisting conditions: 62.9% were obese, 71.6% had hypertension, and 14.7% had a history of myocardial infarction (Table S1 in the Supplementary Appendix). The two study groups, however, had similar rates of these and other coexisting conditions; other baseline characteristics were also similar in the two groups.

Of the 790 men who were enrolled, 705 completed 12 months of study treatment. The characteristics of men who completed 12 months and those who did not complete 12 months did not differ appreciably (Table S2 in the Supplementary Appendix).

Testosterone treatment increased the median

testosterone concentration to the mid-normal range for young men and maintained that range during the treatment period (Fig. S2 in the Supplementary Appendix). A total of 91% of men assigned to testosterone maintained a mean testosterone concentration above the lower limit of the normal range from month 3 through month 12. Testosterone treatment also increased levels of free testosterone, estradiol, and dihydrotestosterone but did not increase levels of sex hormone–binding globulin (Fig. S2 in the Supplementary Appendix).

EFFICACY

Sexual Function Trial

Averaged over all follow-up visits, sexual activity, as determined by the PDQ-Q4 score, increased more with testosterone treatment than with placebo, both among men enrolled in the Sexual Function Trial (treatment effect [the mean difference in the change from baseline between participants assigned to testosterone and those assigned to placebo], 0.58; P<0.001) (Fig. 1A) and among all Testosterone Trials participants (treatment effect, 0.62; P<0.001) (Table 1). A greater increase in testosterone level during treatment was associated with a greater increment in the PDQ-Q4 score (P<0.001 by instrumental variable analysis) (Fig. S3 in the Supplementary Appendix). The response was somewhat less at month 12 (P=0.08 for the interaction between time and treatment). Testosterone treatment was also associated with increased sexual desire according to the DISF-M-II (treatment effect, 2.93; P<0.001) and increased erectile function according to the IIEF (treatment effect, 2.64; P<0.001) (Table 1). Men in the testosterone group were more likely than those in the placebo group to report that their sexual desire had improved since the beginning of the trial (P<0.001) (Fig. S4 in the Supplementary Appendix).

Physical Function Trial

Among men enrolled in the Physical Function Trial, there were no significant differences between the testosterone group and the placebo group in the percentage of men whose 6-minute walking distance increased by at least 50 m (primary outcome) (odds ratio, 1.42; P=0.20) (Fig. 1B), the change from baseline in the 6-minute walking distance (mean difference, 4.09 m; P=0.28) (Table 2), or the percentage of men whose PF-10 score increased by at least 8 points (odds ratio, 1.34; P=0.15); there was a significant between-group difference in the change from baseline in the PF-10 score (mean difference, 2.75 points; P=0.03) (Table 2). Among all Testosterone Trials participants, there was a significant between-group difference in all four measures: the percentage of men whose 6-minute walking distance increased by at least 50 m (odds ratio, 1.76; P=0.003), the change from baseline in the 6-minute walking distance (mean difference, 6.69 m; P=0.007), the percentage of men whose PF-10 score increased by at least 8 points (odds ratio, 1.50; P=0.02), and the change from baseline in the PF-10 score (mean difference, 3.06 points; P=0.002). Men who received testosterone were more likely than those who received placebo to perceive that their walking ability had improved since the beginning of the trial (P=0.002) (Fig. S4 in the Supplementary Appendix).

Vitality Trial

Among men enrolled in the Vitality Trial, testosterone treatment showed no significant benefit over placebo with respect to vitality, as determined by an increase of at least 4 points in the FACIT–Fatigue score (primary outcome) (odds ratio, 1.23; P=0.30) (Fig. 1C). However, there appeared to be a small effect on the change from baseline in the FACIT–Fatigue score that did not reach significance (mean difference, 1.21 points; P=0.06) (Table 3). In addition, a greater increase in testosterone level was associated with a greater increment in the score (P=0.02 by instrumental variable analysis) (Fig. S3 in the Supplementary Appendix), and the effect of testosterone on the change from baseline in the score in the participants in the three trials combined was significant (P=0.006). Among participants in the Vitality Trial, there were significant differences between the testosterone group and the placebo group in the SF-36 vitality score (mean difference, 2.41 points; P=0.03), the PANAS positive affect score (mean difference, 0.47 points; P=0.04), the PANAS negative affect score (mean difference, −0.49 points; P<0.001), and the PHQ-9 depression score (mean difference, −0.72 points; P=0.004) (Table 3). The effect sizes (the mean between-group differences in outcome divided by the baseline standard deviations) were all below 0.20. The men who received testosterone were more likely than men who received placebo

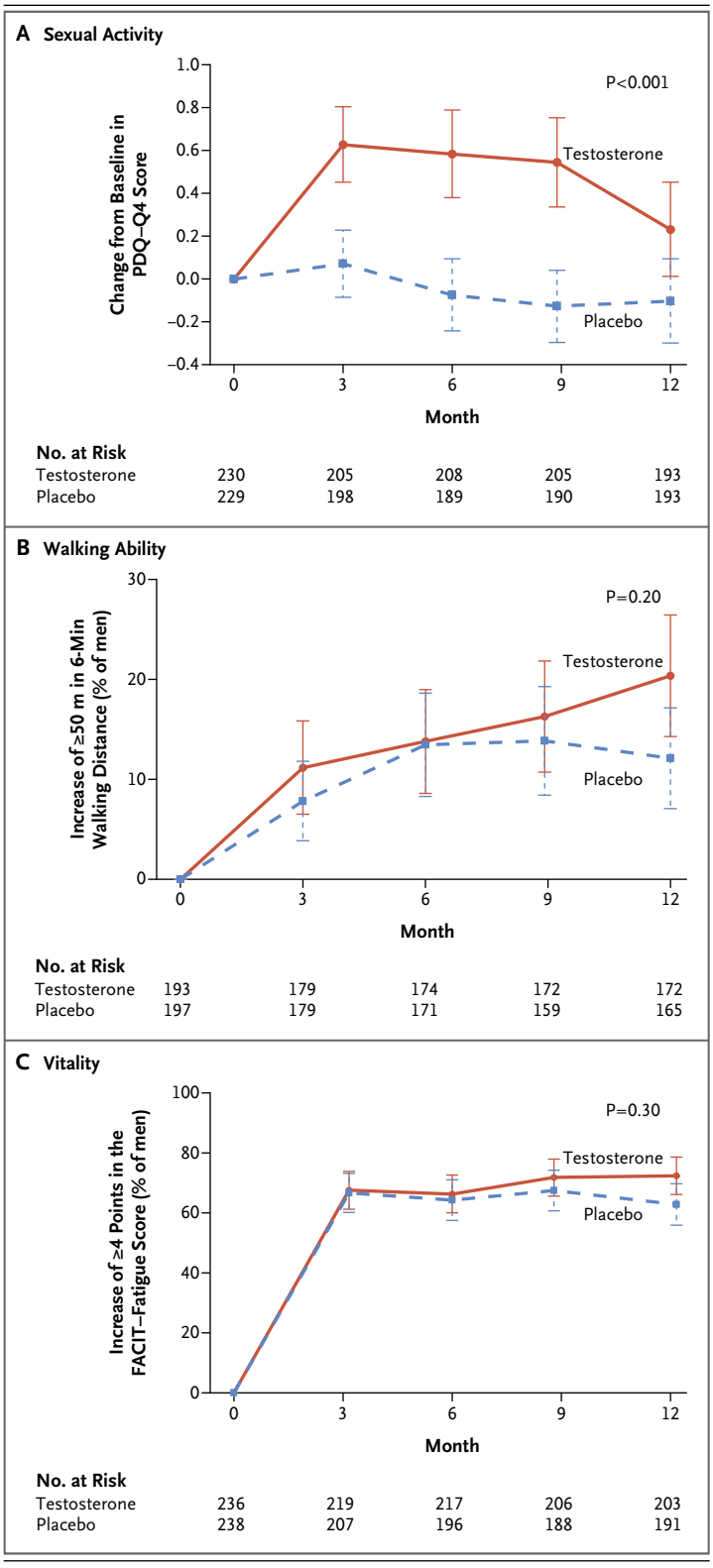


Figure 1. Primary Outcomes in the Three Main Trials of the Testosterone Trials.

The primary outcome of the Sexual Function Trial (Panel A) was the change from baseline in the score for sexual activity (question 4) on the Psychosexual Daily Questionnaire (PDQ-Q4; range, 0 to 12, with higher scores indicating more activity). The primary outcome of the Physical Function Trial (Panel B) was the percentage of men who had an increase of at least 50 m in the distance walked during the 6-minute walk test. The primary outcome of the Vitality Trial (Panel C) was the percentage of men who had an increase of at least 4 points in the score on the Functional Assessment of Chronic Illness Therapy (FACIT)–Fatigue scale (range, 0 to 52, with higher scores indicating less fatigue). P values were calculated with the use of a linear random-effects model for sexual activity and logistic random-effects models for walking ability and vitality. The I bars represent standard deviations.

to report that their energy was better at the end of the trial (P<0.001) (Fig. S4 in the Supplementary Appendix).

All Trials

Sensitivity analyses of the primary outcomes did not suggest that missing values affected any conclusions appreciably (Table S3 in the Supplementary Appendix). We found no significant interactions of treatment with age (P values ranged from 0.45 to 0.78 in the three trials), body-mass index (P values ranged from 0.35 to 0.85), or race (P values ranged from 0.49 to 0.72).

ADVERSE EVENTS

Although more men assigned to testosterone than those assigned to placebo had an increment in the PSA level of 1.0 ng per milliliter or more during the treatment period (23 vs. 8), only 1 man (in the testosterone group) received a diagnosis of prostate cancer during that time. Two men in the testosterone group and 1 in the placebo group received a diagnosis during the subsequent year (Table 4, and Table S4 in the Supplementary Appendix). The change in the IPSS did not differ significantly between the two groups. A hemoglobin level of 17.5 g per deciliter or more was observed in 7 men in the testosterone group and none in the placebo group.

Seven men in each study group were adjudicated to have had major cardiovascular events (myocardial infarction, stroke, or death from cardiovascular causes) during the treatment period and two men in the testosterone group and nine men in the placebo group were adjudicated to have had major cardiovascular events during the subsequent year (Table 4, and Table S4 in the Supplementary Appendix). There was no pattern

Table 1. Sexual Function Trial Outcomes.*

Cohort and Outcome	No. of Men	Baseline Value	Change from Baseline Value				Treatment Effect (95% CI)†	Effect Size (95% CI)‡	P Value§
			Month 3	Month 6	Month 9	Month 12			
Men enrolled in Sexual Function Trial									
Primary outcome: PDQ-Q4 score¶									
Testosterone	230	1.4±1.3	0.6±1.3	0.6±1.5	0.5±1.5	0.2±1.6	0.58 (0.38–0.78)	0.45 (0.30–0.60)	<0.001
Placebo	229	1.4±1.3	0.1±1.1	−0.1±1.2	−0.1±1.2	−0.1±1.4			
Secondary outcomes									
DISF-M-II sexual desire score‖									
Testosterone	234	11.9±6.7	3.5±6.3	3.5±6.0	4.0±7.4	2.6±6.5	2.93 (2.13–3.74)	0.44 (0.32–0.56)	<0.001
Placebo	236	11.6±6.6	0.7±5.8	0.8±5.6	0.9±5.5	0.0±5.0			
IIEF erectile function score**									
Testosterone	234	8.0±8.2	3.4±6.1	3.3±6.5	3.4±6.9	3.1±6.9	2.64 (1.68–3.61)	0.32 (0.20–0.44)	<0.001
Placebo	236	7.7±8.2	1.0± 5.3	0.5±6.1	0.5±7.1	1.0±6.0			
All Testosterone Trials participants††									
PDQ-Q4 score¶									
Testosterone	387	1.5±1.3	0.7±1.3	0.6±1.6	0.6±1.6	0.3±1.7	0.62 (0.45–0.79)	0.45 (0.33–0.58)	<0.001
Placebo	384	1.5±1.4	0.0±1.2	−0.1±1.3	−0.1±1.3	−0.1±1.4			

* Plus–minus values are means ±SD.
† The treatment effect is the mean difference in change from baseline for participants assigned to testosterone versus those assigned to placebo, with adjustment for balancing factors: baseline total testosterone level (≤200 or >200 ng per deciliter), age (≤75 or >75 years), trial site, participation in the main trials, use or nonuse of antidepressants, and use or nonuse of phosphodiesterase type 5 inhibitors.
‡ The effect size is the treatment effect divided by the baseline standard deviation.
§ The P value for the treatment effect was determined with the use of a linear mixed model with a random effect for participant.
¶ Scores for sexual activity (question 4) on the Psychosexual Daily Questionnaire (PDQ-Q4) range from 0 to 12, with higher scores indicating more activity.
‖ Scores on the sexual-desire domain of the Derogatis Interview for Sexual Functioning in Men–II (DISF-M-II) range from 0 to 33, with higher scores indicating greater desire.
** Scores on the erectile-function domain of the International Index of Erectile Function (IIEF) range from 0 to 30, with higher scores indicating better function.
†† The outcomes for all Testosterone Trials participants are exploratory outcomes.

Table 2. Physical Function Trial Outcomes.*

Cohort and Outcome	No. of Men	Baseline Value	No. of Participants or Change from Baseline Value				Treatment Effect (95% CI)†	Effect Size (95% CI)‡	P Value§
			Month 3	Month 6	Month 9	Month 12			
Men enrolled in Physical Function Trial									
Primary outcome: increase of ≥50 m in 6-min walk test — no./total no. (%)									
Testosterone	191		20/179 (11.2)	24/174 (13.8)	28/172 (16.3)	35/172 (20.3)	1.42 (0.83 to 2.45)		0.20
Placebo	196		14/179 (7.8)	23/171 (13.5)	22/159 (13.8)	20/165 (12.1)			
Secondary outcomes									
6-Min walking distance — m									
Testosterone	191	347.7±69.1	10.2±35.8	8.2±41.5	5.3±50.3	14.3±45.9	4.09 (–3.00 to 11.18)	0.06 (–0.04 to 0.16)	0.28
Placebo	196	344.9±68.5	4.6±35.2	7.8±41.4	3.2±52.4	5.5±46.4			
Increase of ≥8 in PF-10 score — no./total no. (%)¶									
Testosterone	184		77/176 (43.8)	72/171 (42.1)	77/172 (44.8)	66/173 (38.2)	1.34 (0.90 to 2.00)		0.15
Placebo	181		59/171 (34.5)	73/159 (45.9)	60/159 (37.7)	58/167 (34.7)			
PF-10 score¶									
Testosterone	184	65.4±20.0	5.6±15.2	6.5±16.7	5.9±19.4	5.8±17.5	2.75 (0.20 to 5.29)	0.13 (0.01 to 0.26)	0.03
Placebo	181	64.8±21.3	4.2±13.7	4.8±17.0	3.3±18.9	2.4±17.3			
All Testosterone Trials participants‖									
Increase of ≥50 m in 6-min walk test — no./total no. (%)									
Testosterone	392		40/368 (10.9)	52/358 (14.5)	54/348 (15.5)	71/346 (20.5)	1.76 (1.21 to 2.57)		0.003
Placebo	389		25/356 (7.0)	39/339 (11.5)	37/320 (11.6)	41/326 (12.6)			

6-Min walking distance — m									
Testosterone	392	387.0±81.7	10.9±45.1	11.0±40.2	6.7±45.1	13.6±43.4	6.69 (1.80 to 11.57)	0.08 (0.02 to 0.14)	0.007
Placebo	389	387.0±83.7	1.6±41.9	5.7±45.1	3.2±47.4	6.4±45.8			
Increase of ≥8 in PF-10 score — no./total no. (%)¶									
Testosterone	309		111/285 (38.9)	113/281 (40.2)	115/276 (41.7)	103/281 (36.7)	1.50 (1.08 to 2.09)		0.02
Placebo	305		87/275 (31.6)	103/263 (39.2)	89/260 (34.2)	82/272 (30.1)			
PF-10 score¶									
Testosterone	309	71.2±20.2	5.0±14.7	6.1±16.7	5.3±18.5	4.3±16.9	3.06 (1.18 to 4.94)	0.15 (0.06 to 0.24)	0.002
Placebo	305	69.7±21.2	3.9±12.8	3.4±16.2	2.3±17.9	1.3±16.9			

* Plus–minus values are means ±SD.
† The treatment effect for dichotomous outcomes is the odds ratio for achieving the outcome versus not achieving the outcome among men assigned to testosterone versus those assigned to placebo. For continuous outcomes, the treatment effect is the mean difference in the outcome among men assigned to testosterone versus those assigned to placebo. All analyses are adjusted for balancing factors: baseline total testosterone level (≤200 or >200 ng per deciliter), age (≤75 or >75 years), trial site, participation in the main trials, use or nonuse of antidepressants, and use or nonuse of phosphodiesterase type 5 inhibitors.
‡ For continuous outcomes, the effect size is the treatment effect divided by the baseline standard deviation.
§ The P value for the treatment effect was determined with the use of a logistic mixed model with a random effect for participant for dichotomous outcomes and a linear mixed model with a random effect for participant for continuous outcomes.
¶ Scores on the physical-function scale (PF-10) of the Medical Outcomes Study 36-Item Short-Form Health Survey range from 0 to 100, with higher scores indicating better function.
‖ The outcomes for all Testosterone Trials participants are exploratory outcomes.

Table 3. Vitality Trial Outcomes.*

Cohort and Outcome	No. of Men	Baseline Value	No. of Participants or Change from Baseline Value				Treatment Effect (95% CI)†	Effect Size (95% CI)‡	P Value§
			Month 3	Month 6	Month 9	Month 12			
Men enrolled in Vitality Trial									
Primary outcome: increase of ≥4 in FACIT–Fatigue score — no./total no. (%)¶									
Testosterone	236		148/219 (67.6)	144/217 (66.4)	148/206 (71.8)	147/203 (72.4)	1.23 (0.83 to 1.84)		0.30
Placebo	238		138/207 (66.7)	126/196 (64.3)	127/188 (67.6)	120/191 (62.8)			
Secondary outcomes									
FACIT–Fatigue score¶									
Testosterone	236	31.6±6.4	7.7±8.4	7.4±9.1	8.6±9.1	8.0±8.4	1.21 (–0.04 to 2.46)	0.19 (0.01 to 0.38)	0.06
Placebo	238	31.3±6.4	7.2±8.8	5.9±9.2	7.2±9.2	6.7±9.4			
SF-36 vitality score‖									
Testosterone	208	50.6±13.8	7.4±13.6	7.2±14.6	8.4±14.4	8.2±15.3	2.41 (0.31 to 4.50)	0.18 (0.02 to 0.34)	0.03
Placebo	196	49.4±12.6	5.9±11.1	4.5±11.2	5.7±12.3	6.1±13.8			
PANAS positive affect score**									
Testosterone	229	15.3±3.2	0.7±3.2	0.9±3.8	0.9±3.4	0.7±3.9	0.47 (0.02 to 0.92)	0.14 (0.01 to 0.27)	0.04
Placebo	234	15.4±3.5	0.3±3.3	0.0±3.3	0.4±3.4	0.2±3.2			
PANAS negative affect score**									
Testosterone	229	7.5±2.7	–0.2±2.5	–0.4±2.4	–0.2±2.3	–0.6±2.1	–0.49 (–0.79 to –0.19)	–0.18 (–0.29 to –0.06)	<0.001
Placebo	234	7.4±2.8	0.3±2.4	0.4±2.6	–0.1±2.6	–0.1±2.6			

PHQ-9 depression score††									
Testosterone	230	6.6±4.0	−1.3±3.8	−1.7±3.8	−1.9±4.0	−1.8±3.7	−0.72 (−1.20 to −0.23)	−0.18 (−0.30 to −0.06)	0.004
Placebo	234	6.6±4.0	−0.8±3.5	−0.5±3.7	−1.2±4.2	−1.1±3.8			
All Testosterone Trials participants‡‡									
Increase of ≥4 in FACIT–Fatigue score — no./total no. (%)¶									
Testosterone	394		176/351 (50.1)	181/350 (51.7)	178/337 (52.8)	174/333 (52.3)	1.23 (0.89 to 1.70)		0.22
Placebo	394		166/337 (49.3)	151/329 (45.9)	154/317 (48.6)	152/316 (48.1)			
FACIT–Fatigue score¶									
Testosterone	394	37.0±8.6	4.7±8.5	4.8±8.7	5.2±9.1	4.7±8.8	1.27 (0.37 to 2.16)	0.15 (0.04 to 0.25)	0.006
Placebo	394	36.8±8.8	4.1±9.0	2.8±9.0	3.7±9.2	3.6±9.5			

* Plus–minus values are means ±SD.
† The treatment effect for dichotomous outcomes is the odds ratio for achieving the outcome versus not achieving the outcome among men assigned to testosterone versus those assigned to placebo. For continuous outcomes, the treatment effect is the mean difference in the outcome among men assigned to testosterone versus those assigned to placebo. All analyses are adjusted for balancing factors: baseline total testosterone level (≤200 or >200 ng per deciliter), age (≤75 or >75 years), trial site, participation in the main trials, use or nonuse of antidepressants, and use or nonuse of phosphodiesterase type 5 inhibitors.
‡ For continuous outcomes, the effect size is the treatment effect divided by the baseline standard deviation.
§ The P value for the treatment effect was determined with the use of a logistic mixed model with a random effect for participant for dichotomous outcomes and a linear mixed model with a random effect for participant for continuous outcomes.
¶ Scores on the Functional Assessment of Chronic Illness Therapy (FACIT)–Fatigue scale range from 0 to 52, with higher scores indicating less fatigue.
‖ Scores on the vitality scale of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) range from 0 to 100, with higher scores indicating more vitality.
** Scores for positive affect and for negative affect on the Positive and Negative Affect Schedule (PANAS) scales range from 5 to 50, with higher scores indicating a greater intensity of the affect.
†† Scores on the Patient Health Questionnaire 9 (PHQ-9) depression scale range from 0 to 27, with higher scores indicating greater intensity of depressive symptoms.
‡‡ The outcomes for all Testosterone Trials participants are exploratory outcomes.

Table 4. Adverse Events during the First Year (Treatment Period) of the Testosterone Trials.*

Event	Placebo (N = 394)	Testosterone (N = 394)
	no. of participants	
Prostate-related event		
Increase in PSA level by ≥1.0 ng/ml	8	23
Prostate cancer	0	1
IPSS >19†	26	27
Hemoglobin ≥17.5 g/dl	0	7
Cardiovascular event‡		
Myocardial infarction (definite or probable)	1	2
Stroke (definite or probable)	5	5
Death from cardiovascular causes	1	0
Myocardial infarction, stroke, or death from cardiovascular causes	7	7
Serious adverse events		
Death	7	3
Hospitalization	78	68
Other§	6	7

* PSA denotes prostate-specific antigen.
† The International Prostate Symptom Score (IPSS) questionnaire is used to identify symptoms of benign prostatic hyperplasia. Scores range from 0 to 35, with higher scores indicating more severe symptoms. A score of more than 19 indicates moderately severe lower urinary tract symptoms.
‡ Data on cardiovascular adverse events were collected with the use of a specific questionnaire administered at each visit and also identified from the adverse-event log and the form for reporting serious adverse events (see the protocol). Myocardial infarction, stroke, and death from cardiovascular causes were assessed by two adjudicators.
§ Other serious adverse events were defined as congenital anomaly, disability, a life-threatening event, or an event that may not be immediately life-threatening but is clearly of major clinical significance.

of a difference in risk with respect to the other cardiovascular adverse events (Table S4 in the Supplementary Appendix). No significant between-group differences were observed in cardiac adverse events defined according to *Medical Dictionary for Regulatory Activities* classification (Tables S5 and S6 in the Supplementary Appendix).

DISCUSSION

Increasing the serum testosterone concentrations of men 65 years of age or older from moderately low to the mid-normal range for men 19 to 40 years of age had significant effects on all measures of sexual function and some measures of physical function, mood, and depressive symptoms — all to small-to-moderate degrees, consistent with the degree of testosterone deficiency.

Men who received testosterone reported better sexual function, including activity, desire, and erectile function, than those who received placebo. Although the effect sizes were low to moderate, men in the testosterone group were more likely than those in the placebo group to report that their sexual desire had improved, which suggests that this effect was of clinical relevance. The effect of testosterone on erectile function was less than that reported with phosphodiesterase type 5 inhibitors.[28]

The percentage of men whose 6-minute walking distance increased by at least 50 m did not differ significantly between the two study groups in the Physical Function Trial but did differ significantly when men in all three trials were included, although the effect sizes did not differ markedly (1.42 vs. 1.76). Furthermore, men who received testosterone were more likely than those who received placebo to report that their walking ability was better, which suggests that the effect, although small in magnitude, might be clinically relevant.

Testosterone had no significant benefit with respect to vitality, as assessed by the FACIT–Fatigue scale, except as a continuous outcome when men in all three trials were included. However, testosterone was associated with small but significant benefits with respect to mood and depressive symptoms. Men in the testosterone group were also more likely than those in the placebo group to report that their energy was better.

We observed four cases of prostate cancer, three of which were in men treated with testosterone, and there was no significant difference in urinary symptoms (as assessed by means of the IPSS) between the study groups. The generalizability of these results is limited, however, because we excluded men with a high risk of prostate cancer and men with moderately severe urinary tract symptoms. Furthermore, the sample size was inadequate to assess reliably the effect of testosterone on the risk of these conditions.

Some studies have suggested that testosterone treatment is associated with increased cardiovascular risk,[29-32] although others have not.[6,33,34] We did not observe a pattern of increased risk, but this trial was too small to exclude other than a large increase.

Our three trials had certain strengths, including enrollment of men with an unequivocally

low mean testosterone concentration, adequate sample sizes, a double-blind, placebo-controlled design, an increase in serum testosterone concentration to the normal range for young men, and excellent participant retention. A major limitation, albeit an intentional one, is that the results apply only to men 65 years of age or older whose testosterone levels averaged less than 275 ng per deciliter.

Results of the primary outcomes in our three trials showed that testosterone treatment had a moderate, significant benefit with respect to sexual function but no significant benefit with respect to walking distance (among participants in the Physical Function Trial) or vitality. Testosterone treatment also had a significant benefit with respect to other prespecified outcomes, including walking distance when men in all three trials were included and mood and depressive symptoms. These results, together with those of the other four trials (now completed), should inform decisions about testosterone treatment for men 65 years of age or older whose levels are low for no apparent reason other than age. Such decisions will also require knowing the risks of testosterone treatment, which will necessitate larger and longer trials.

The Testosterone Trials were supported by a grant (U01 AG030644) from the National Institute on Aging (NIA), National Institutes of Health. The trials also received funding from the National Heart, Lung, and Blood Institute, National Institute of Neurological Disorders and Stroke, and National Institute of Child Health and Human Development. AbbVie (formerly Solvay and Abbott Laboratories) provided funding and donated AndroGel and placebo gel. The Boston site was partially supported by a grant (P30-AG013679) from the Boston Claude D. Pepper Older Americans Independence Center. The Yale Field Center was partially supported by a grant (P30-AG021342) from the Yale Claude D. Pepper Older Americans Independence Center and a grant (UL1 TR000142) from the Yale Center for Clinical Investigation. Dr. Cauley was supported by a grant (R01 AG37679) from the NIA. Dr. Gill was the recipient of a Midcareer Investigator Award in Patient-Oriented Research (K24-AG021507) and is the recipient of an Academic Leadership Award (K07AG043587), both from the NIA. Dr. Lewis was supported by a grant (DK079626) from the National Institute of Diabetes and Digestive and Kidney Diseases to the University of Alabama at Birmingham Diabetes Research and Training Center. Dr. Resnick was supported by the Intramural Research Program, NIA.

Dr. Snyder reports receiving consulting fees from Watson Laboratories. Dr. Bhasin reports receiving fees for serving on advisory boards from Eli Lilly and Sanofi, consulting fees from AbbVie, and grant support from Regeneron Pharmaceuticals, Eli Lilly, AbbVie, and Novartis; he also reports holding pending patents related to an algorithm for free testosterone determination based on a model of testosterone binding to sex hormone–binding globulin (provisional patent number 61/772,054) and the use of testosterone plus ornithine decarboxylase inhibitor as a prostate-sparing anabolic therapy (provisional patent number 13/679,889). Dr. Cunningham reports receiving fees for serving as an advisor to AbbVie, Apricus Biosciences, Clarus Therapeutics, Endo Pharmaceuticals, Ferring Pharmaceuticals, Eli Lilly, Purdue Pharma, and Repros Therapeutics and grant support from Ardana. Dr. Matsumoto reports receiving consulting fees from AbbVie, Eli Lilly, Endo Pharmaceuticals, and Clarus Therapeutics, study medication from AbbVie, and grant support from GlaxoSmithKline. Dr. Swerdloff reports receiving consulting fees from TesoRx and grant support from Clarus Therapeutics, Eli Lilly, Novartis, and Antares Pharma. Dr. Wang reports receiving fees for serving on an advisory board from TesoRx and grant support from Clarus Therapeutics, Lipocine, and Antares Pharma. Dr. Ensrud reports receiving fees for serving on a data and safety monitoring committee from Merck Sharp & Dohme. Dr. Farrar reports receiving fees for serving on a data and safety monitoring board from Cara Therapeutics, consulting fees from Bayer, Biogen, Mallinckrodt, the Campbell Consortium, Janssen, Daiichi-Sankyo, and Novartis, and grant support from Pfizer and Depomed. Dr. Cella reports receiving consulting fees from Pfizer and being the president of FACIT.org. Dr. Rosen reports receiving grant support from Bayer HealthCare, Eli Lilly, and Besins Healthcare. Dr. Molitch reports receiving consulting fees from AbbVie, Eli Lilly, and Pfizer. Dr. Anton reports receiving fees for serving as a scientific advisor to Reserveage Organics. Dr. Basaria reports receiving consulting fees from Eli Lilly and grant support from AbbVie. Dr. Mohler reports receiving consulting fees from AbbVie and Clarus Therapeutics. Dr. Parsons reports receiving grant support from Actavis. Dr. Ellenberg reports receiving grant support from AbbVie. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Drs. Evan Hadley and Sergei Romashkan of the NIA for their support throughout the trials.

APPENDIX

The authors' full names and academic degrees are as follows: Peter J. Snyder, M.D., Shalender Bhasin, M.D., Glenn R. Cunningham, M.D., Alvin M. Matsumoto, M.D., Alisa J. Stephens-Shields, Ph.D., Jane A. Cauley, Dr.P.H., Thomas M. Gill, M.D., Elizabeth Barrett-Connor, M.D., Ronald S. Swerdloff, M.D., Christina Wang, M.D., Kristine E. Ensrud, M.D., M.P.H., Cora E. Lewis, M.D., M.S.P.H., John T. Farrar, M.D., Ph.D., David Cella, Ph.D., Raymond C. Rosen, Ph.D., Marco Pahor, M.D., Jill P. Crandall, M.D., Mark E. Molitch, M.D., Denise Cifelli, M.S., Darlene Dougar, M.P.H., Laura Fluharty, M.P.H., Susan M. Resnick, Ph.D., Thomas W. Storer, Ph.D., Stephen Anton, Ph.D., Shehzad Basaria, M.D., Susan J. Diem, M.D., M.P.H., Xiaoling Hou, M.S., Emile R. Mohler III, M.D., J. Kellogg Parsons, M.D., M.H.S., Nanette K. Wenger, M.D., Bret Zeldow, M.S., J. Richard Landis, Ph.D., and Susan S. Ellenberg, Ph.D., for the Testosterone Trials Investigators

The authors' affiliations are as follows: the Division of Endocrinology, Diabetes, and Metabolism (P.J.S.), the Department of Biostatistics and Epidemiology (A.J.S.-S., J.T.F., X.H., B.Z., J.R.L., S.S.E.), the Center for Clinical Epidemiology and Biostatistics (D. Cifelli, D.D., L.F.), and the Division of Cardiovascular Disease, Section of Vascular Medicine, Department of Medicine (E.R.M.), Perelman School of Medicine, University of Pennsylvania, Philadelphia; Research Program in Men's Health: Aging and Metabolism, Brigham and Women's Hospital, Harvard Medical School, Boston (S. Bhasin, T.W.S., S. Basaria), and New England Research Institutes, Watertown (R.C.R.) — both in Massachusetts; the Departments of Medicine and Molecular and Cellular Biology, Division of Diabetes, Endocrinology, and Metabolism, Baylor College of Medicine and Baylor St. Luke's Medical Center, Houston (G.R.C.); Geriatric Research, Education, and Clinical Center, Department of Veterans Affairs (VA) Puget Sound Health Care System, and the Division of Gerontology and

Geriatric Medicine, Department of Internal Medicine, University of Washington School of Medicine — both in Seattle (A.M.M.); the Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh (J.A.C.); the Division of Geriatric Medicine, Yale School of Medicine, New Haven, CT (T.M.G.); the Department of Internal Medicine and Division of Epidemiology, Department of Family Medicine and Public Health, University of California, San Diego, School of Medicine, La Jolla (E.B.-C.), the Division of Endocrinology, Harbor–UCLA Medical Center (R.S.S., C.W.), and Los Angeles Biomedical Research Institute (R.S.S., C.W.), Torrance, and the Department of Urology, Moores Comprehensive Cancer Center, University of California, San Diego (J.K.P.) — all in California; the Department of Medicine, Division of Epidemiology and Community Health, University of Minnesota (K.E.E., S.J.D.), and Minneapolis VA Health Care System (K.E.E.) — both in Minneapolis; the Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham, (C.E.L.); the Department of Medical Social Sciences (D. Cella) and the Division of Endocrinology, Metabolism, and Molecular Medicine (M.E.M.), Feinberg School of Medicine, Northwestern University, Chicago; the Department of Aging and Geriatric Research, University of Florida, Gainesville (M.P., S.A.); the Divisions of Endocrinology and Geriatrics, Albert Einstein College of Medicine, Bronx, NY (J.P.C.); the Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore (S.M.R.); and the Division of Cardiology, Emory University School of Medicine, Atlanta (N.K.W.).

REFERENCES

1. Wu FC, Tajar A, Pye SR, et al. Hypothalamic-pituitary-testicular axis disruptions in older men are differentially linked to age and modifiable risk factors: the European Male Aging Study. J Clin Endocrinol Metab 2008;93:2737-45.
2. Harman SM, Metter EJ, Tobin JD, Pearson J, Blackman MR. Longitudinal effects of aging on serum total and free testosterone levels in healthy men. J Clin Endocrinol Metab 2001;86:724-31.
3. Snyder PJ, Peachey H, Hannoush P, et al. Effect of testosterone treatment on body composition and muscle strength in men over 65 years of age. J Clin Endocrinol Metab 1999;84:2647-53.
4. Page S, Amory J, Bowman F, et al. Exogenous testosterone (T) alone or with finasteride increases physical performance, grip strength, and lean body mass in older men with low serum T. J Clin Endocrinol Metab 2004;90:1502-10.
5. Travison TG, Basaria S, Storer TW, et al. Clinical meaningfulness of the changes in muscle performance and physical function associated with testosterone administration in older men with mobility limitation. J Gerontol A Biol Sci Med Sci 2011;66:1090-9.
6. Srinivas-Shankar U, Roberts SA, Connolly MJ, et al. Effects of testosterone on muscle strength, physical function, body composition, and quality of life in intermediate-frail and frail elderly men: a randomized, double-blind, placebo-controlled study. J Clin Endocrinol Metab 2010;95:639-50.
7. Gray PB, Singh AB, Woodhouse LJ, et al. Dose-dependent effects of testosterone on sexual function, mood, and visuospatial cognition in older men. J Clin Endocrinol Metab 2005;90:3838-46.
8. Vaughan C, Goldstein FC, Tenover JL. Exogenous testosterone alone or with finasteride does not improve measurements of cognition in healthy older men with low serum testosterone. J Androl 2007;28:875-82.
9. Liverman C, Blazer D, eds. Testosterone and aging: clinical research directions. Washington, DC: National Academies Press, 2004.
10. Snyder PJ, Ellenberg SS, Cunningham GR, et al. The Testosterone Trials: seven coordinated trials of testosterone treatment in elderly men. Clin Trials 2014;11:362-75.
11. Cauley JA, Fluharty L, Ellenberg SS, et al. Recruitment and screening for the Testosterone Trials. J Gerontol A Biol Sci Med Sci 2015;70:1105-11.
12. Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst 2006;98:529-34.
13. Derogatis LR. The Derogatis Interview for Sexual Functioning (DISF/DISF-SR): an introductory report. J Sex Marital Ther 1997;23:291-304.
14. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the Six Minute Walk test in chronic lung disease patients. Am J Respir Crit Care Med 1997;155:1278-82.
15. Cella D, Eton DT, Lai JS, Peterman AH, Merkel DE. Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales. J Pain Symptom Manage 2002;24:547-61.
16. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. Clin Pharmacol Ther 1974;15:443-53.
17. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 1975;31:103-15.
18. Lee KK, Berman N, Alexander GM, Hull L, Swerdloff RS, Wang C. A simple self-report diary for assessing psychosexual function in hypogonadal men. J Androl 2003;24:688-98.
19. Rosen RC, Cappelleri JC, Gendrano N III. The International Index of Erectile Function (IIEF): a state-of-the-science review. Int J Impot Res 2002;14:226-44.
20. Ware J, Snow K, Kosinski M. SF-36 Health Survey: manual and interpretation guide. Lincoln, RI: Quality Metric, 2000.
21. Ware J, Kosinski M. SF-36 Physical and Mental Health Summary Scales: a manual for users of Version 1, Second Edition. Lincoln, RI: Quality Metric, 2005.
22. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54:1063-70.
23. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16:606-13.
24. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990;9:811-8.
25. Little RJA. Pattern-mixture models for multivariate incomplete data. J Am Stat Assoc 1993;88:125-34.
26. Follmann D, Wu M. An approximate generalized linear model with random effects for informative missing data. Biometrics 1995;51:151-68.
27. Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. J Health Econ 2008;27:531-43.
28. Müller A, Smith L, Parker M, Mulhall JP. Analysis of the efficacy and safety of sildenafil citrate in the geriatric population. BJU Int 2007;100:117-21.
29. Vigen R, O'Donnell CI, Barón AE, et al. Association of testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. JAMA 2013;310:1829-36.
30. Finkle WD, Greenland S, Ridgeway GK, et al. Increased risk of non-fatal myocardial infarction following testosterone therapy prescription in men. PLoS One 2014;9(1):e85805.
31. Basaria S, Coviello AD, Travison TG, et al. Adverse events associated with testosterone administration. N Engl J Med 2010;363:109-22.
32. Xu L, Freeman G, Cowling BJ, Schooling CM. Testosterone therapy and cardiovascular events among men: a systematic review and meta-analysis of placebo-controlled randomized trials. BMC Med 2013;11:108.
33. Corona G, Maseroli E, Rastrelli G, et al. Cardiovascular risk associated with testosterone-boosting medications: a systematic review and meta-analysis. Expert Opin Drug Saf 2014;13:1327-51.
34. Sharma R, Oni OA, Gupta K, et al. Normalization of testosterone level is associated with reduced incidence of myocardial infarction and mortality in men. Eur Heart J 2015;36:2706-15.

Exhibit D

Author's Accepted Manuscript

Association between use of exogenous testosterone therapy (eTT) and risk of venous-thrombotic-events among eTT-treated and untreated men with hypogonadism

Hu Li , Karin Benoit , Wei Wang , Stephen Motsko

PII: S0022-5347(15)05157-5
DOI: 10.1016/j.juro.2015.10.134
Reference: JURO 13050

To appear in: *The Journal of Urology*
Accepted Date: 24 October 2015

Please cite this article as: Li H, Benoit K, Wang W, Motsko S, Association between use of exogenous testosterone therapy (eTT) and risk of venous-thrombotic-events among eTT-treated and untreated men with hypogonadism, *The Journal of Urology*® (2015), doi: 10.1016/j.juro.2015.10.134.

Association between use of exogenous testosterone therapy (eTT) and risk of venous-thrombotic-events among eTT-treated and untreated men with hypogonadism

Hu Li,[1,*] Karin Benoit,[1] Wei Wang,[1] Stephen Motsko [1]

Author affiliations

[1] Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana, USA

*** Correspondence:**

Hu Li, MBBS, PhD
Pharmacoepidemiologist, Global Patient Safety
Eli Lilly and Company
Indianapolis, IN, USA 46225
Phone: (317) 601-8777
FAX: (317) 433-5372
LI_HU_HL@lilly.com

Running title: VTE in eTT-Treated vs Untreated Hypogonadal Men

STATISTICAL SUMMARY

Abstract Text	**Manuscript Text (Introduction through Conclusions)**	**References**	**Figures and Tables**	**Supplementary Material**
N = **245**	N = **2497**	N = **23**	N = **7** (5 tables, 2 figures)	16 paragraphs of text; 5 supplementary tables

Abstract.

Purpose: Limited information exists about whether exogenous testosterone therapy (eTT) is associated with risk of venous thrombotic events (VTE). Here, we investigate via cohort and nested-case-control analyses whether eTT administration is associated with risk of VTE in men with hypogonadism.

Materials and Methods: Databases were reviewed to identify men prescribed eTT and/or men with a hypogonadism diagnosis. Propensity-score 1:1 matching was used to select patients for the cohort analysis. *Cases* (men with VTE) were matched 1:4 with *controls* (men without VTE) for the nested-case-control analysis. Primary outcome was defined as incident *idiopathic* VTE; Cox regression and conditional-logistic regression were used to assess hazard ratios (HRs) and odds ratios (ORs), respectively. Sensitivity analyses were also performed.

Results: 102,650 eTT-treated patients and 102,650 untreated patients were included in the cohort analysis after matching; 2785 *cases* and 11,119 *controls* were included in the case-control analysis. Cohort analysis revealed an HR of 1.08 for all eTT-treated patients (95% CI: 0.91, 1.27; p=0.378). Case-control analysis resulted in OR=1.02 (95% CI: 0.92, 1.13; p=0.702) for current eTT exposure and 0.92 (95% CI: 0.82, 1.03; p=0.145) for past eTT exposure. These results remained non statistically significant after stratifying by eTT-administration-route and age category. Results from most of the sensitivity analyses yielded results that were consistent.

Conclusions: No significant association was found between eTT and incidents of *idiopathic* VTE, as well as *overall* VTE in men with hypogonadism; however, some discrepant findings exist for the association between injectable formulations and *overall* VTE risk.

Keywords: venous thrombosis, hypogonadism, pulmonary embolism, testosterone.

Introduction

Venous thrombotic events (VTE) often manifest as deep vein thrombosis (DVT) or pulmonary embolism (PE). Major exogenous risk factors for VTE are surgery, hospitalization, and prolonged immobility, and endogenous risk factors are cancer, obesity, and hypercoagulation disorders.[1-3]

Exogeneous testosterone therapy (eTT) is administered to treat male hypogonadism in order to restore serum testosterone levels and relieve patient symptoms; however, several publications have suggested that eTT may be linked to increased hematocrit, polycythemia, and VTE.[4-7] In contrast, some studies have demonstrated that increases in endogenous estradiol or testosterone levels are not associated with increased risk of VTE.[8,9] Furthermore, a recently published study did not find a significant association between eTT and VTE.[10] However, based on postmarket spontaneous reports and published case reports,[11,12] the Food and Drug Administration (FDA) in 2014 required a change to drug labeling of all approved testosterone products, which included a general warning regarding a potential increased risk of VTE.[13]

The present study aimed to further examine whether eTT was associated with an increased risk of VTE in men with hypogonadism in both retrospective cohort and nested-case-control settings.

Materials and Methods

Data source

Medical claims data, pharmacy data, and healthcare enrollment information were obtained from Truven Health MarketScan® Databases[14] from December 2004 to December 2012 (see supplementary materials).

Patient population

Eligibility criteria included 1) men ≥18 years of age with continuous enrollment in a healthcare plan for ≥12 months; 2) had hypogonadism (eTT prescription and/or a hypogonadism diagnosis code per International Classification of Diseases, Ninth Revision [ICD-9]) (see supplementary Table 1). Patients who had a VTE during this period were excluded.

Study design

The study employed both a retrospective cohort and a nested-case-control study, to ensure consistent findings across different designs. For the retrospective cohort analysis, a propensity-score matching method was used to form cohorts of eTT-treated and untreated men with hypogonadism based on baseline demographics, comorbid conditions, concomitant medications, and resource utilization. Index date was defined as the first prescription date for eTT-treated men; and randomly assigned diagnosis date for untreated men in order to account for immortal time bias (see supplementary materials).[15] The baseline period was defined as 12-month period before patient's index date.

For the nested-case-control analysis, men with hypogonadism with VTE were selected from the original (prematched) cohort population to be *cases*. For each *case*, 4 patients without VTE were randomly selected to be *controls* and matched on index date and age.

Study variables

The exposure variable was any eTT exposure, further stratified by prespecified routes of eTT administration (topical/gel, injection, transdermal, or other/nonspecified). The exposure window was defined as duration of the prescription plus a 90-day wash out period. In the nested-case-control analysis, current eTT exposure was defined if VTE occurred during the exposure window, and past eTT was defined if VTE occurred at least 90 days after the end of the last prescription (i.e., outside exposure window).

The study outcome variables were incident *idiopathic* VTE (not associated with proxy risk factors of stroke, injury, paralysis/immobility, hospitalization >3 days, lower-limb fracture, major surgery, oxygen therapy, or anticoagulant use), as well as incident *overall* VTE in a sensitivity measure, defined via ICD-9 codes. These codes have been validated from a FDA mini-sentinel project with the highest positive predictive value (65%-90%).[16] Additionally, an adjudication process was employed to classify *idiopathic* VTE cases, although misclassification may still exist (concordance rate = 70%, 95% CI: 61.8% to 78.20%) due to limitations of the data source and lack of non-prescription information (see supplementary materials).

The other study variables including baseline characteristic variables (comorbidities, VTE risk factors, resource utilization, and medication use) were defined via ICD-9 or product codes.

Statistical analyses

Baseline characteristics and VTE risk factors were described for the patient populations. Between-cohort differences in these characteristics were calculated by t-test (continuous variables) or Pearson's chi-square test (categorical variables) with a 0.05 significance level.

For cohort analyses, a propensity score method was used. The propensity score for each patient was defined as the predicted probability of eTT initiation based on an assessment of measurable baseline characteristics.[17-19] A high-dimension propensity-score method developed by the Observational Medical Outcomes Partnership (OMOP) identified a comparison group with regard to elevated risk of drug-induced VTE by incorporating additional baseline variables to include in the propensity score model.[19,20] The propensity score generated for the entire population was applied to subcohorts.[19,21] Using a time-to-event analysis, VTE incidence rates (IRs) were calculated among eTT-treated and untreated patient cohorts (per person years). A Cox regression model was employed to determine hazard ratios (HRs) with 95% confidence intervals (CIs) and p-values. The proportionality assumption for the Cox regression models was checked;

no violations were observed.

For nested-case-control analyses, conditional stepwise logistic regression models adjusting for baseline characteristics were used to account for changes in drug exposure and time-varying confounding factors. The association between eTT exposure patterns and VTE risk was reported as an adjusted odds ratio (OR; with a 95% CI) after controlling for key VTE risk factors. In addition, stepwise criteria of variable selection applied a p-value of 0.20 for model entry and 0.10 for retaining variables. To be conservative, correction of multiple comparisons/Type I errors was not considered for multiple comparisons involving different hypothesis.

Sensitivity analyses were performed to assess the impact of different eTT exposure windows (60, 90, or 120 days), *overall* VTEs, and variations in study design (intent-to-treat [ITT] versus as-treated analysis) (see supplementary materials). Analyses were conducted with SAS Version 9.2 software (SAS Institute Inc., Cary, USA).

Results

Cohort analysis population

Figure 1 depicts the selection process for the cohort analysis population with 533,223 patients (306,507 eTT treated, and 226,716 untreated patients). After applying 1:1 propensity-score matching, 102,650 eTT-treated and 102,650 untreated men with hypogonadism with well-balanced baseline characteristics (Table 1) were selected for the primary analysis.

Case-control analysis population

A total of 2785 patients with incident *idiopathic* VTE were selected as *cases*, and 11,119 matched controls were randomly selected from the treated and untreated hypogonadal population. Demographics and baseline characteristics are summarized in Table 1.

Cohort analysis results

The incidence rate (IR) of *idiopathic* VTE for the treated versus untreated cohort was 3.70 (95% CI: 3.23 to 4.16) versus 3.20 (95% CI: 2.92 to 3.47) per 1000 patient-years respectively (Table 2).

The adjusted HRs from the cohort analysis demonstrated no significant differences in the incidence of VTE among eTT-treated and untreated men with hypogonadism (Table 2, Fig. 2). The adjusted HR for the entire retrospective cohort was 1.08 (95% CI: 0.91 to 1.27; p=0.378). Upon stratification by routes of eTT administration: the adjusted HR for the topical/gel route was 1.07 (95% CI: 0.88 to 1.29; p=0.496) and for injectable eTT was 1.32 (95% CI: 0.89 to 1.96; p=0.164). The adjusted HR among patients ≤65 and those >65 years old was 1.09 (95% CI: 0.91 to 1.29; p=0.350) and 0.96 (95% CI: 0.59 to 1.56; p=0.883), respectively.

Case-control analysis results

The adjusted OR from the case-control analysis was 1.02 (95% CI: 0.92 to 1.13; p=0.702) for current eTT exposure, and 0.92 (95% CI: 0.82 to 1.03; p=0.145) for past eTT exposure (Table 3). None of the analyses by age stratification, routes of administration, and interactions between eTT exposure status and routes of eTT administration reached statistical significance (Table 3).

Sensitivity analyses results

None demonstrated a significant association between eTT and VTE (Table 4). The adjusted HRs were consistent with those observed for the cohort analysis of the *idiopathic* VTE population (Table 2).

In the case-control analysis, despite the consistency, a few sensitivity analyses reached statistical significance: (1) the adjusted OR for ‘any past eTT exposure’ and ‘past exposure to topical/gel eTT’ was 1.08 (95% CI: 1.02 to 1.15; p=0.010) and 1.09 (95% CI: 1.02 to 1.16;

p=0.011) respectively among the entire population; (2) the adjusted OR for any injectable eTT exposure was 1.10 (95% CI: 1.01 to 1.19; p=0.023), further, for current injectable exposure, the adjusted OR was 1.15 (95% CI: 1.04 to 1.26; p=0.006) (Table 5).

Additional sensitivity analyses including using PS stratification methods (HR = 1.02, 95% CI: 0.90 to 1.15; p=0.799; Supplementary Table 4), applying intent-to-treat analysis (HR = 0.96, 95% CI: 0.85 to 1.08; p=0.461; Supplementary Table 5) yielded nonsignificant results.

Discussion

This real-world study, which utilized an incident user design, found no significant association between eTT and incident risk of *idiopathic* VTE via both retrospective cohort (HR of 1.08 [95% CI: 0.91, 1.27]; p=0.378) and nested-case-control designs (current eTT exposure (OR=1.02 [95% CI: 0.92, 1.13]; p=0.702) and past eTT exposure (OR=0.92 [95% CI: 0.82, 1.03]; p=0.145). Furthermore, none of the eTT routes of administration (injectable, gel, patch) were associated with increased *idiopathic* VTE risk. These findings are considered to be robust and consistent, because of the new user study design, large sample size, 2 complimentary study designs and analytical methods to control for confounding factors, and the inclusion of various sensitivity analyses. In order to address this public health issue that concerns patients and physicians, the study results can be extrapolated to the general population, but not to high-risk patients (e.g., pre-existing thrombophilia) because patients with baseline VTE were excluded from both analyses.

The results from the sensitivity analyses were generally consistent with the results from the primary analyses. Some exceptions existed when studying the *overall* VTE population. Specifically, in contrast to the cohort analysis, the nested-case-control analysis found some statistically significant findings among past eTT exposures and any injectable eTT users. These findings could be attributed to several factors. Compared to the *idiopathic* VTE population, patients in the *overall* VTE population were more likely to experience other proxy risk factors

for VTE, such as prolonged immobility, trauma, and injury. Although the fully adjusted statistical model included many of these terms as covariates, it is possible that other unmeasured confounding factors may exist. Additionally, this significant association was observed among patients who were not current eTT users and did not take any eTT (topical solution) at least 90 days prior to onset of VTE, thus it is unlikely that these VTEs were associated with eTT. Lastly, the dissimilar findings for any injectable exposure vs. other routes may be due to a difference in pharmacokinetics as suggested by a recently published study;[22] safety profiles may vary for different testosterone delivery mechanisms with altered pharmacokinetics (i.e., injections cause spikes in testosterone levels and transdermal patches and gels cause subtle but sustained increases).

Notably, our analyses replicated a recent case control analysis by Baillargeon et al, which reported that filling a prescription for eTT was not associated with an increased risk of VTE in nearly 31,000 middle aged and elder men.[10] Our findings further support their observations using a retrospective cohort study design, which strengthen the findings via examining the temporal relationship analysis. Both of these studies add value in that they are large general population-based comparative safety studies and offer a superior opportunity of evaluating drug safety compared to post-marketing cases and previously published case series reports.[11,12]

Further, the study examined both *idiopathic* VTE and *overall* VTE study outcomes. The reason *idiopathic* VTE was chosen as the primary outcome was to preclude confounding factors (independent of drug use) related to VTE, such as trauma, injury, and hospitalization, which are strong predictors, possibly diluting the association with the drug.

The study findings were contrary to an assumed link between eTT and incident VTE, which was thought to be mediated via increases in hematocrit and/or polycythemia, based on evidence for increased thromboembolic events in patients with primary polycythemia vera.[4-7] It is also theorized that the risk of testosterone-induced polycythemia (increased hematocrit value)

may increase blood viscosity, leading to an increased risk for thromboembolic events.[23] One study hypothesized that men with previously undiagnosed familial thrombophilias (Factor V Leiden) developed VTE while on eTT, due to peripheral conversion of testosterone to estradiol.[11] Although the results from the present study do not support the mechanisms above,[11] the study results are consistent with other studies that did not find an association between endogenous sex hormone levels and a 10-year risk of VTE in middle-aged and older men,[8] nor any significant association between endogenous testosterone or estradiol levels and risk of VTE, DVT, or PE in a study of 9331 men and women in the Copenhagen City Heart Study.[9]

Nevertheless, the results from this study should be interpreted with consideration of its limitations. While claims data are valuable for the effective examination of disease outcomes and treatment patterns, claims data are collected for the purpose of payment and not research. The presence of a claim for a filled prescription does not indicate that the medication was consumed or taken as prescribed. The presence or absence of disease may not be accurate, as the diagnostic code may be incorrectly coded or included as rule-out criteria rather than actual disease. The observational nature of this study precluded the ability to employ treatment randomization; thus, findings may be subject to changes due to residual confounding factors. Several important covariates were missing in the claims database, including (but not limited to) body weight and genotypes for inheritable hypercoagulation conditions (e.g., protein C deficiency, protein S deficiency). The study outcomes were not validated through chart validation, but rather 1) adapting the FDA recommended/validated algorithm, which yielded a positive predictive value (PPV) between 65% and 95%; 2) adjudicating some patients' claims to classify *idiopathic* VTE cases because misclassification cases may exist; and 3) combining both PE and DVT as the study outcome to improve PPV.[16] Although the IR of the study outcome was higher than reported in current literature, there is no evidence suggesting that the false-positive cases would be distributed unevenly between the study groups. Therefore, the drug-event association was

assumed to be unchanged. Further, the comparison group was defined as an inactive comparator group that was not treated, but the untreated cohort was formed to match eTT-treated patients based on baseline characteristics through a propensity score model[19] with the purpose of improving the comparability and reducing the confounding. Due to the lack of a specific ICD-9 code, the most frequently used codes were chosen; although others may exist, these would not appreciably change the size of the study population. Finally, this claims database lacked comprehensive laboratory data to further substantiate exposure (through serum testosterone levels) or potential mechanisms for possible increased VTE risk (through elevated hematocrit) or clinical presentation of symptoms (e.g. fatigue). While testosterone deficiency among the treated cohort was unconfirmed, we approached the research question with the assumption that adult males prescribed eTT were considered by their physician to have testosterone deficiency. Furthermore, baseline endogenous total serum testosterone level is not needed because untreated hypogonadism is not a well-established predictor of VTE as suggested by previous literature.[1-3] Therefore, the lack of laboratory measures should not confound the association between eTT and VTE.

Conclusion

In conclusion, the results from the analyses of this study using the Truven Health Analytics MarketScan Databases showed no significant association between eTT administration and incidents of idiopathic VTE, as well as overall VTE, although the two different study designs yielded discrepant findings for the association between injectable formulations and overall VTE risk.

Addendum

H. Li, K. Benoit, W. Wang, and S. Motsko contributed to the design of this study, the collection and analysis of the data, revision of the intellectual content of the manuscript, and approval of the final version to be published. H. Li, and S. Motsko were responsible for the concept of the study, and critically writing of the intellectual content of the manuscript.

Acknowledgments

The design and conduct of this study, as well as the preparation of this manuscript, were sponsored and funded by Eli Lilly and Company. The authors would like to thank Dr. John Seeger (Harvard Medical School/ Brigham and Women's Hospital), Dr. Shehzad Sultan Basaria (Harvard Medical School), and Dr. Susan Jick (Boston University School of Public Health) for their consultation in the study design, analysis plan and review of the study report; and thank Susan Kindig MD, Jason Sol Zielonka MD, and Thomas Ferry MD for their work in reviewing and adjudicating VTE to determine whether these events could be considered idiopathic. The authors would also like to thank Scott J. Burke MS and Deborah D'Souza, PhD (inVentiv Health Clinical) for their assistance in preparing this manuscript.

Disclosure of Conflicts of Interest

H. Li, K. Benoit, W. Wang, and S. Motsko are full-time employees of and minor shareholders in Eli Lilly and Company. Eli Lilly and Company is the market authorization holder of Axiron, an exogenous testosterone therapy.

References

1 Cushman M. Epidemiology and risk factors for venous thrombosis. *Semin Hematol* 2007; **44**: 62.

2 Goldhaber SZ. Risk factors for venous thromboembolism. *J Am Coll Cardiol.* 2010; **56**: 1.

3 Stein PD, Matta F. Epidemiology and incidence: the scope of the problem and risk factors for development of venous thromboembolism. *Clin Chest Med.* 2010; **31**: 611.

4 Cundy J. The perioperative management of patients with polycythaemia. *Ann R Coll Surg Engl.* 1980; **62**: 470.

5 Stohlawetz PJ, Dzirlo L, Hergovich N, et al. Effects of erythropoietin on platelet reactivity and thrombopoiesis in humans. *Blood.* 2000; **95**: 2983.

6 Jick SS, Hagberg KW. The risk of adverse outcomes in association with use of testosterone products: a cohort study using the UK-based general practice research database. *Br J Clin Pharmacol.* 2013; **75**: 260.

7 Bachman E, Travison TG, Basaria S, et al. Testosterone induces erythrocytosis via increased erythropoietin and suppressed hepcidin: evidence for a new erythropoietin/hemoglobin set point. *J Gerontol A Biol Sci Med Sci.* 2014; **69**: 725.

8 Svartberg J, Braekkan SK, Laughlin GA, et al. Endogenous sex hormone levels in men are not associated with risk of venous thromboembolism: the Tromsø study. *Eur J Endocrinol* 2009; **160**: 833.

9 Holmegard HN, Nordestgaard BG, Schnohr P, et al. Endogenous sex hormones and risk of venous thromboembolism in women and men. *J Thromb Haemost.* 2014; **12**: 297.

10 Baillargeon J, Urban RJ, Morgentaler A, et al. Risk of venous thromboembolism in men receiving testosterone therapy. *Mayo Clin Proc*. 2015; **90**: 1038.

11 Glueck CJ, Goldenberg N, Budhani S, et al. Thrombotic events after starting exogenous testosterone in men with previously undiagnosed familial thrombophilia. *Transl Res.* 2011; **158**: 225.

12 Glueck CJ, Richardson-Royer C, Schultz R, et al. Testosterone therapy, thrombophilia-hypofibrinolysis, and hospitalization for deep venous thrombosis-pulmonary embolus: an exploratory, hypothesis-generating study. *Clin Appl Thromb Hemost*. 2014; **20** :244.

13 Axiron® (testosterone) topical solution. Prescribing information. Copyright © 2010, 2015, Eli Lilly and Company.

14 Adamson DM, Chang S, Hansen LG. Health Research Data for the Real World: The MarketScan Databases. Thomson Medstat, Ann Arbor (2006).

15 Zhou Z, Rahme E, Abrahamowicz M, Pilote L (2005) Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. *Am J Epidemiol* 2005; **162**: 1016.

16 Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. *Pharmacoepidemiol Drug Saf.* 2012; **21**: Suppl. 1:154.

17 D'Agostino RB Jr, D'Agostino RB Sr. Estimating treatment effects using observational data. *JAMA.* 2007; **297** :314.

18 Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med.* 2007; **26**: 20.

19 Schneeweiss S, Rassen JA, Glynn RJ, et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiol.* 2009; **20**: 512.

20 [IMEDS] Innovation in Medical Evidence Development and Surveillance. 2013. Available at: http://imeds.reaganudall.org/. Accessed 09 July 2014.

21 Rassen JA, Glynn RJ, Rothman KJ, et al. Applying propensity scores estimated in a full cohort to adjust for confounding in subgroup analyses. *Pharmacoepidemiol Drug Saf.* 2012; **21**: 697.

22 Layton JB, Meier CR, Sharpless JL, et al. Comparative safety of testosterone dosage forms. *JAMA.* 2015; doi: 10.1001/jamainternmed.2015.1573 [Epub ahead of print].

23 Surampudi PN, Wang C, Swerdloff R. Hypogonadism in the aging male diagnosis, potential benefits, and risks of testosterone replacement therapy. *Int J Endocrinol.* 2012;2012:625434.

FIGURE LEGENDS

Fig. 1 – Flow chart describing the selection process for the study's patient cohort.

N, number of patients in group; eTT, exogenous testosterone therapy; VTE, venous thrombotic event(s).

Fig. 2 – Adjusted hazard ratios (HR) (with 95% confidence intervals [CI]) from Cox proportional regression analysis for retrospective cohort population: exogenous testosterone therapy (eTT) and *idiopathic* venous thrombotic events stratified by route of eTT administration. CI, confidence interval; HR, hazard ratio; LCL, lower confidence limit (low value of 95% CI); eTT, exogenous testosterone therapy; UCL, upper confidence limit (high value of 95% CI).

Table 1 Baseline Characteristics for Populations Selected for Retrospective Cohort Analysis (1:1 Propensity-Score-Matched) and Nested-Case-Control Analysis

	Population for Cohort Analysis			Population for Case-Control Analysis*		
Category	**Treated**	**Untreated**	**p-Value**	***Cases***	***Controls***	**p-Value**
(Characteristic)	**n (%)**	**n (%)**		**n (%)**	**n (%)**	
Total Number of Patients	***102,650***	***102,650***	---	***2894***	***11,576***	---
Mean Age (SD) at Index	51.50 (11.51)	51.51 (12.36)	0.793	54.18 (11.07)	54.18 (11.07)	1.000
VTE Risk Factors						
History of VTE	143 (0.14)	141 (0.14)	0.906	21 (0.73)	8 (0.07)	0.000
Genetic/congenital	239 (0.23)	238 (0.23)	0.963	--- (---)	19 (0.16)	0.021
Cancer	9306 (9.07)	9430 (9.19)	0.342	11 (0.38)	1195 (10.32)	0.000
Hypertension	45480 (44.31)	45375 (44.20)	0.641	1444 (49.90)	5136 (44.37)	0.000
Hypercholesterolemia	50351 (49.05)	50430 (49.13)	0.727	1401 (48.41)	5507 (47.57)	0.420
Diabetes	20737 (20.20)	20815 (20.28)	0.668	793 (27.40)	2491 (21.52)	0.000
Obesity	6748 (6.57)	6603 (6.43)	0.194	228 (7.88)	483 (4.17)	0.000
Renal disease	3065 (2.99)	3150 (3.07)	0.274	134 (4.63)	317 (2.74)	0.000
Myocardial infarction	1503 (1.46)	1545 (1.51)	0.443	47 (1.62)	236 (2.04)	0.150
Ischemic stroke	2575 (2.51)	2654 (2.59)	0.268	105 (3.63)	319 (2.76)	0.013
Congestive heart failure	4218 (4.11)	4235 (4.13)	0.850	182 (6.29)	535 (4.62)	0.000
Varicose vein(s)	1511 (1.47)	1538 (1.50)	0.622	95 (3.28)	154 (1.33)	0.000
Rheumatoid arthritis	1185 (1.15)	1155 (1.13)	0.533	49 (1.69)	112 (0.97)	0.001
Infection	8917 (8.69)	9094 (8.86)	0.167	355 (12.27)	995 (8.60)	0.000
Inflammatory bowel disease	824 (0.80)	791 (0.77)	0.410	34 (1.17)	87 (0.75)	0.025
Fracture(s)	836 (0.81)	851 (0.83)	0.714	36 (1.24)	95 (0.82)	0.032

Major trauma	334 (0.33)	342 (0.33)	0.758	8 (0.28)	24 (0.21)	0.479
Injury	2633 (2.57)	2625 (2.56)	0.911	96 (3.32)	281 (2.43)	0.007
Surgery	3178 (3.10)	3120 (3.04)	0.458	109 (3.77)	391 (3.38)	0.306
Hospitalization >3 days	3208 (3.13)	3233 (3.15)	0.752	112 (3.87)	363 (3.14)	0.047
Other Comorbidities						
Charlson comorbidity (SD)	0.97 (1.64)	0.98 (1.64)	0.126	1.09 (1.58)	1.02 (1.63)	0.050
Hypogonadism	85145 (82.95)	98919 (96.37)	0.000	1805 (62.37)	6999 (60.46)	0.060
Sexual dysfunction	98122 (95.59)	98658 (96.11)	0.000	1867 (64.51)	7365 (63.62)	0.373
Klinefelter's syndrome	211 (0.21)	295 (0.29)	0.000	5 (0.17)	10 (0.09)	0.199
Sleep disturbance	18132 (17.66)	17837 (17.38)	0.087	565 (19.52)	1795 (15.51)	0.000
Malaise/fatigue	32636 (31.79)	32966 (32.11)	0.118	808 (27.92)	2983 (25.77)	0.019
Pituitary disorders	3125 (3.04)	2986 (2.91)	0.071	90 (3.11)	346 (2.99)	0.734
Testicular cancer	433 (0.42)	429 (0.42)	0.891	2 (0.07)	34 (0.29)	0.034
Prostate disease	15224 (14.83)	15170 (14.78)	0.737	468 (16.17)	1800 (15.55)	0.410
Prostate cancer	2478 (2.41)	2514 (2.45)	0.606	21 (0.73)	322 (2.78)	0.000
Concomitant Medications						
Antihyperlipidemics	40844 (39.79)	40556 (39.51)	0.194	1213 (41.91)	5022 (43.38)	0.154
Antihypertensives	48016 (46.78)	47615 (46.39)	0.076	1528 (52.80)	5749 (49.66)	0.003
Diabetes medications	16225 (15.81)	16189 (15.77)	0.828	630 (21.77)	1975 (17.06)	0.000
Erectile dysfunction meds	14912 (14.53)	14700 (14.32)	0.183	472 (16.31)	1904 (16.45)	0.858
Hematological agents	6812 (6.64)	6821 (6.64)	0.936	116 (4.01)	876 (7.57)	0.000
Opiates	41548 (40.48)	41239 (40.17)	0.165	1381 (47.72)	4651 (40.18)	0.000
Psychotropics	35047 (34.14)	34453 (33.56)	0.006	1076 (37.18)	3938 (34.02)	0.001
Sleep medications	11469 (11.17)	11204 (10.91)	0.062	356 (12.30)	1312 (11.33)	0.145

Abbreviations: N, number of patients in population; n, number of patients exhibiting characteristic; SD, standard deviation; VTE, venous thrombotic event.
*Upon review of patient claim records, 109 of the selected *cases* and 457 of the selected *controls* were identified to have been treated with eTT via multiple administration routes. These patients were excluded from their respective groups, leaving a total of 2785 *cases* and 11,119 *controls*.

Table 2 Retrospective Cohort Analysis of Testosterone Use and *Idiopathic* Venous Thrombotic Events – Crude Incidence Rates and Adjusted Hazard Ratios (with 95% CIs) Among Testosterone-Treated and 1:1 Matched Untreated Hypogonadal Men – Stratified by Age Category and Route of Administration

	eTT-Treated Hypogonadal Men			**Untreated Hypogonadal Men**					
		Crude			**Crude**		**Adjusted**		
eTT Administration Route	**Total Patients (N)**	**Incidence Rate (IR)***	**95% CI**	**Total Patients (N)**	**Incidence Rate (IR)***	**95% CI**	**Hazard Ratio †**	**95% CI**	**p-Value**
Any	102,650	**3.70**	3.23, 4.16	102,650	**3.20**	2.92, 3.47	**1.08**	0.91, 1.27	**0.378**
≤65 years of age	93,292	**3.60**	3.12, 4.08	93,057	**3.17**	2.87, 3.46	**1.09**	0.91, 1.29	**0.350**
>65 years of age	9358	**4.67**	2.94, 6.40	9593	**3.42**	2.55, 4.28	**0.96**	0.59, 1.56	**0.883**
Topical/gel	71,095	**3.71**	3.17, 4.24	71,095	**3.26**	2.93, 3.60	**1.07**	0.88, 1.29	**0.496**
Injection	21,260	**4.20**	3.01, 5.39	21,260	**2.69**	2.08, 3.30	**1.32**	0.89, 1.96	**0.164**
Transdermal	6949	**1.57**	0.19, 2.95	6949	**3.60**	2.58, 4.62	**0.39**	0.15, 1.06	**0.065**
Other/nonspecified	3346	**3.99**	1.38, 6.59	3346	**3.46**	1.94, 4.98	**1.14**	0.46, 2.77	**0.781**

Abbreviations: CI, confidence interval; eTT, exogenous testosterone therapy; IR, incidence rate; N, number of patients in group; VTE, venous thrombotic event.

* Per 1000 patient-years.

† Adjusted for treatment and baseline characteristics that are commonly associated with VTE risk, including (but not limited to) the following: age, infection(s), previous VTE, obesity, cardiovascular disorders, cancer, and use of medications for diabetes and/or hematologic disorders.

Table 3 Adjusted Odds Ratios (with 95% CIs) from Nested-Case-Control Analysis (Conditional Logistic Regression Analysis) of Exogenous Testosterone Use and *Idiopathic* Venous Thrombotic Events – Stratified by Exposure Status and Route of eTT Administration

			Entire Population			Population ≤65 Years Old			Population >65 Years Old		
Category	***Cases**** **(N)**	***Controls*** **(N)**	**Odds Ratio†**	**95% CI**	**p-Value**	**Odds Ratio†**	**95% CI**	**p-Value**	**Odds Ratio†**	**95% CI**	**p-Value**
Current eTT vs no eTT	2785	11119	**1.02**	0.92, 1.13	0.702	**1.05**	0.94, 1.17	0.408	**0.83**	0.60, 1.13	0.235
Past eTT vs no eTT			**0.92**	0.82, 1.03	0.145	**0.93**	0.82, 1.05	0.230	**0.86**	0.62, 1.19	0.356
Topical/gel eTT vs no eTT	1223	4977	**0.95**	0.86, 1.05	0.300	**0.98**	0.88, 1.09	0.672	**0.77**	0.58, 1.03	0.078
Injectable eTT vs no eTT	316	1077	**1.11**	0.96, 1.29	0.163	**1.13**	0.96, 1.33	0.133	**1.01**	0.68, 1.52	0.954
Transdermal eTT vs no eTT	123	505	**0.91**	0.73, 1.13	0.389	**0.94**	0.74, 1.19	0.611	**0.78**	0.44, 1.40	0.406
Other/nonspec eTT vs no eTT	65	255	**1.00**	0.75, 1.33	0.983	**0.95**	0.69, 1.31	0.749	**1.28**	0.62, 2.68	0.505
Current ***topical/gel*** eTT vs no use/exposure	707	2778	**1.00**	0.89, 1.12	0.990	**1.03**	0.92, 1.17	0.590	**0.78**	0.54, 1.12	0.170
Current ***injection*** eTT vs no use/exposure	214	708	**1.15**	0.97, 1.38	0.120	**1.18**	0.98, 1.43	0.080	**0.94**	0.55, 1.61	0.820
Current ***transdermal*** eTT vs no use/exposure	49	229	**0.80**	0.57, 1.12	0.200	**0.89**	0.62, 1.26	0.500	**0.36**	0.12, 1.13	0.080
Current ***other/nonspec*** eTT vs no use/exposure	33	122	**0.99**	0.66, 1.48	0.960	**0.90**	0.58, 1.40	0.650	**2.21**	0.74, 6.57	0.160
Past ***topical/gel*** eTT vs no use/exposure	516	2199	**0.88**	0.78, 1.00	0.060	**0.90**	0.79, 1.04	0.140	**0.77**	0.53, 1.11	0.150
Past ***injection*** eTT vs no use/exposure	102	369	**1.03**	0.81, 1.32	0.790	**1.03**	0.79, 1.35	0.830	**1.08**	0.62, 1.88	0.800
Past ***transdermal*** eTT vs no use/exposure	74	276	**0.99**	0.74, 1.31	0.920	**0.98**	0.72, 1.34	0.890	**1.12**	0.57, 2.21	0.740
Past ***other/nonspec*** eTT vs no use/exposure	32	133	**1.00**	0.66, 1.50	0.990	**1.00**	0.63, 1.56	0.980	**0.86**	0.32, 2.34	0.770

Abbreviations: CI, confidence interval; eTT, exogenous testosterone therapy; N, number of patients in group; nonspec, nonspecified; VTE, venous thrombotic event; vs, versus.

* *Cases* with multiple routes of administration were excluded from the analysis.

† Conditional logistic regression, adjusted for imbalances in baseline characteristics that are commonly associated with VTE risk among the population subgroups under investigation. Baseline characteristics that were selected as covariates included those from the following categories: age, infection(s), previous VTE, obesity, cardiovascular disorders, cancer, diabetes medication use, and use of medications for hematologic disorders. Specific covariates varied with the particular patient subgroup under investigation.

Table 4 Sensitivity Analysis/Retrospective Cohort Analysis of Testosterone Use and *Overall* Venous Thrombotic Events – Crude Incidence Rates and Adjusted Hazard Ratios (with 95% CIs) Among Testosterone-Treated and 1:1 Matched Untreated Hypogonadal Men – Stratified by Age Category and Route of Administration

	eTT-Treated Hypogonadal Men			Untreated Hypogonadal Men					
		Crude			Crude		Adjusted		
eTT Administration Route	**Total Patients (N)**	**Incidence Rate (IR)***	**95% CI**	**Total Patients (N)**	**Incidence Rate (IR)***	**95% CI**	**Hazard Ratio†**	**95% CI**	**p-Value**
Any	102,637	**11.07**	10.26, 11.87	102,637	**11.42**	10.89, 11.95	**0.93**	0.85, 1.03	**0.151**
Topical/gel	71,110	**10.90**	9.97, 11.82	71,110	**11.54**	10.90, 12.17	**0.93**	0.83, 1.03	**0.161**
Injection	21,228	**12.48**	10.42, 14.53	21,228	**10.82**	9.59, 12.06	**1.05**	0.84, 1.31	**0.662**
Transdermal	6,965	**9.12**	5.80, 12.45	6,965	**11.60**	9.77, 13.43	**0.65**	0.42, 1.02	**0.061**
Other/nonspecified	3,334	**10.32**	6.10, 14.53	3,334	**11.54**	8.76, 14.32	**1.11**	0.65, 1.90	**0.704**

Abbreviations: CI, confidence interval; eTT, exogenous testosterone therapy; IR, incidence rate; N, number of patients in group; VTE, venous thrombotic event.

* Per 1000 patient-years.

† Adjusted for treatment and baseline characteristics that are commonly associated with VTE risk, including (but not limited to) the following: age, infection(s), previous VTE, obesity, cardiovascular disorders, cancer, and use of medications for diabetes and/or hematologic disorders.

Table 5 Sensitivity Analysis/Adjusted Odds Ratios (with 95% CIs) from Nested-Case-Control Analysis (Conditional Logistic Regression Analysis) of Exogenous Testosterone Use and *Overall* Venous Thrombotic Events – Stratified by Exposure Status and eTT Route of Administration

			Entire Population			**Population ≤65 Years Old**			**Population >65 Years Old**		
Category	***Cases**** **(N)**	***Controls*** **(N)**	**Odds Ratio†**	**95% CI**	**p-Value**	**Odds Ratio†**	**95% CI**	**p-Value**	**Odds Ratio†**	**95% CI**	**p-Value**
Current eTT vs no eTT	10205	40989	**1.03**	0.97, 1.09	0.300	**1.04**	0.97, 1.11	0.250	**0.98**	0.87, 1.10	0.670
Past eTT vs no eTT			**1.08**	1.02, 1.15	0.010	**1.06**	0.99, 1.14	0.110	**1.13**	1.01, 1.26	0.030
Topical/gel eTT vs no eTT	4549	18004	**1.05**	0.99, 1.10	0.087	**1.04**	0.98, 1.10	0.232	**1.07**	0.97, 1.18	0.203
Injectable eTT vs no eTT	1118	3947	**1.10**	1.01, 1.19	0.023	**1.13**	1.03, 1.24	0.012	**1.00**	0.86, 1.17	0.989
Transdermal eTT vs no eTT	525	1878	**1.07**	0.95, 1.19	0.262	**1.03**	0.90, 1.17	0.670	**1.11**	0.91, 1.36	0.314
Other/nonspec eTT vs no eTT	237	1079	**0.97**	0.83, 1.13	0.664	**0.93**	0.78, 1.12	0.463	**1.03**	0.78, 1.35	0.849
Current ***topical/gel*** eTT vs no use/exposure	2319	9575	**1.01**	0.95, 1.08	0.750	**1.01**	0.94, 1.09	0.725	**0.98**	0.86, 1.12	0.762
Current ***injection*** eTT vs no use/exposure	708	2410	**1.15**	1.04, 1.26	0.006	**1.20**	1.07, 1.34	0.001	**0.94**	0.76, 1.16	0.547
Current ***transdermal*** eTT vs no use/exposure	216	781	**0.98**	0.83, 1.15	0.775	**0.94**	0.78, 1.15	0.564	**1.00**	0.72, 1.38	0.995
Current ***other/nonspec*** eTT vs no use/exposure	109	518	**0.91**	0.73, 1.13	0.381	**0.86**	0.66, 1.11	0.239	**1.03**	0.68, 1.57	0.875
Past ***topical/gel*** eTT vs no use/exposure	2230	8429	**1.09**	1.02, 1.16	0.011	**1.07**	0.99, 1.15	0.088	**1.14**	1.01, 1.29	0.028
Past ***injection*** eTT vs no use/exposure	410	1537	**1.02**	0.90, 1.15	0.739	**1.00**	0.86, 1.16	0.954	**1.07**	0.87, 1.33	0.514
Past ***transdermal*** eTT vs no use/exposure	309	1097	**1.13**	0.99, 1.31	0.079	**1.10**	0.93, 1.30	0.278	**1.18**	0.92, 1.52	0.184
Past ***other/nonspec*** eTT vs no use/exposure	128	561	**1.03**	0.84, 1.26	0.812	**1.02**	0.79, 1.31	0.899	**1.03**	0.73, 1.46	0.860

Abbreviations: CI, confidence interval; eTT, exogenous testosterone therapy; N, number of patients in group; nonspec, nonspecified; VTE, venous thrombotic event; vs, versus.

* *Cases* with multiple routes of administration were excluded from the analysis.

† Conditional logistic regression, adjusted for imbalances in baseline characteristics that are commonly associated with VTE risk among the population subgroups under investigation. Baseline characteristics that were selected as covariates included those from the following categories: age, infection(s), previous VTE, obesity, cardiovascular disorders, cancer, diabetes medication use, and use of medications for hematologic disorders. Specific covariates varied with the particular patient subgroup under investigation.

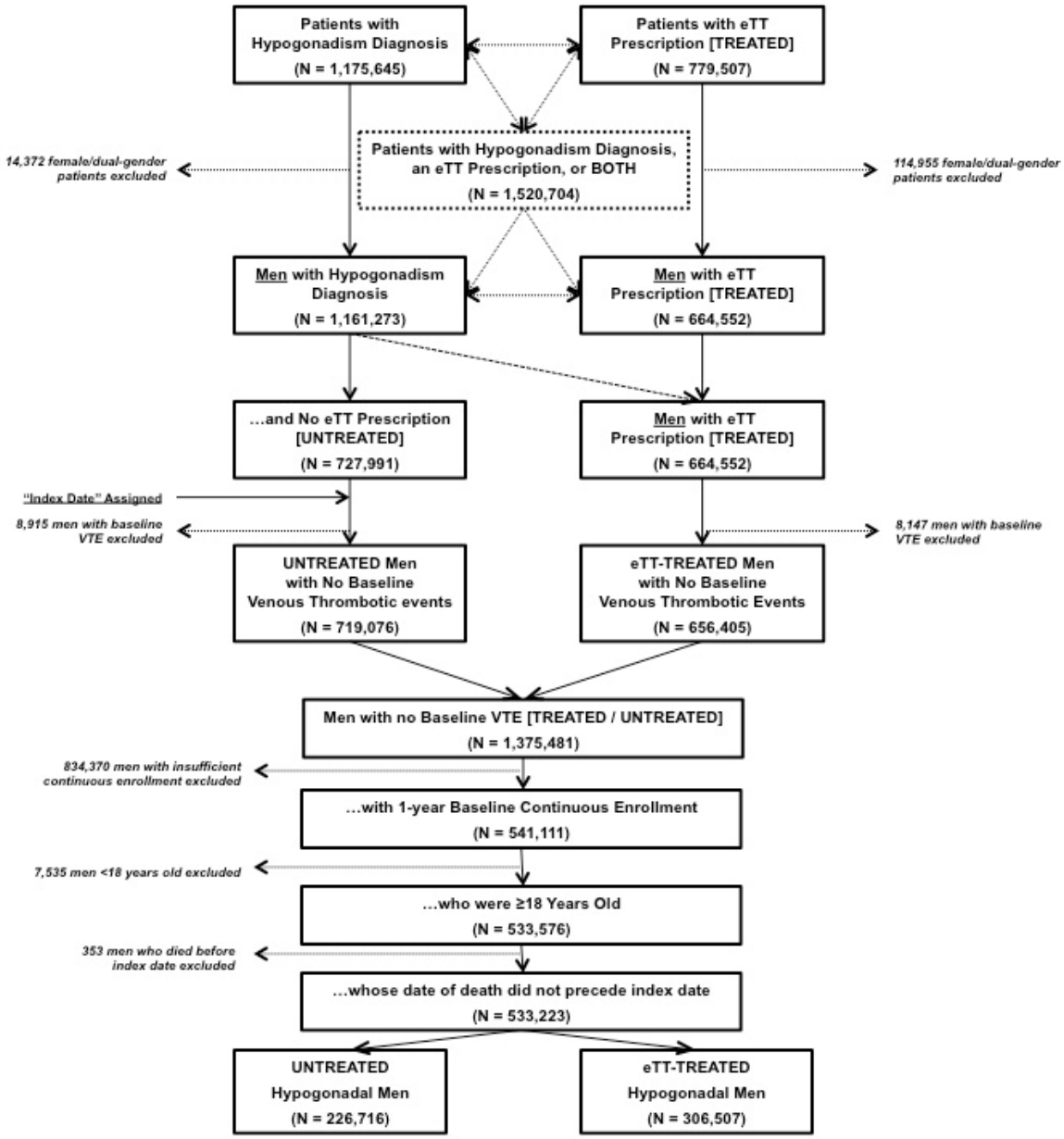
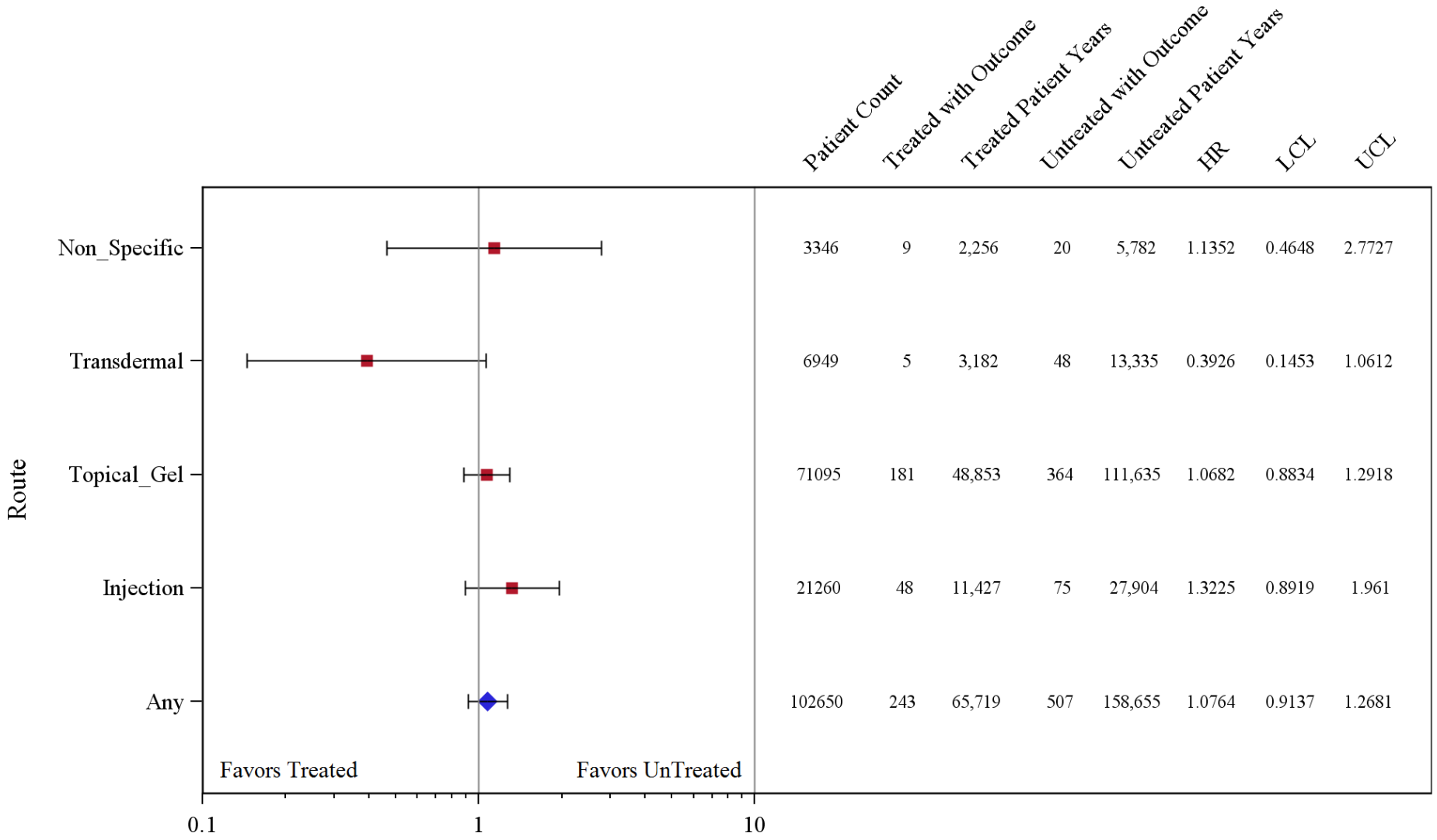

Patients with Hypogonadism Diagnosis (N = 1,175,645)
Patients with eTT Prescription [TREATED] (N = 779,507)
Patients with Hypogonadism Diagnosis, an eTT Prescription, or BOTH (N = 1,520,704)
14,372 female/dual-gender patients excluded
114,955 female/dual-gender patients excluded
Men with Hypogonadism Diagnosis (N = 1,161,273)
Men with eTT Prescription [TREATED] (N = 664,552)
...and No eTT Prescription [UNTREATED] (N = 727,991)
Men with eTT Prescription [TREATED] (N = 664,552)
"Index Date" Assigned
8,915 men with baseline VTE excluded
8,147 men with baseline VTE excluded
UNTREATED Men with No Baseline Venous Thrombotic events (N = 719,076)
eTT-TREATED Men with No Baseline Venous Thrombotic Events (N = 656,405)
Men with no Baseline VTE [TREATED / UNTREATED] (N = 1,375,481)
834,370 men with insufficient continuous enrollment excluded
...with 1-year Baseline Continuous Enrollment (N = 541,111)
7,535 men <18 years old excluded
...who were ≥18 Years Old (N = 533,576)
353 men who died before index date excluded
...whose date of death did not precede index date (N = 533,223)
UNTREATED Hypogonadal Men (N = 226,716)
eTT-TREATED Hypogonadal Men (N = 306,507)

Figure 2



Abbreviations

CI, confidence interval(s)

DVT, deep vein thrombosis

eTT, exogenous testosterone therapy

FDA, Food and Drug Administration

HR, hazard ratio(s)

IR, incidence rate(s)

ICD-9, International Classification of Diseases, 9th Edition

ITT, intent-to-treat

LCL, lower confidence limit

OMOP, Observational medical Outcomes Partnership

OR, odds ratio(s)

PE, pulmonary embolism

PPV, positive predictive value

SD, standard deviation

UCL, upper confidence limit

VTE, venous thrombotic event(s)

Exhibit E

Accepted Manuscript

Association between Testosterone Replacement Therapy and the Incidence of Deep Vein Thrombosis and Pulmonary Embolism: A Retrospective Cohort Study of the Veterans Administration Database

Rishi Sharma, Olurinde A. Oni, Guoqing Chen, Mukut Sharma, Buddhadeb Dawn, Ram Sharma, Deepak Parashara, Virginia J. Savin, Rajat S. Barua, Kamal Gupta

PII: S0012-3692(16)49000-8

DOI: 10.1016/j.chest.2016.05.007

Reference: CHEST 466

To appear in: *CHEST*

Received Date: 26 January 2016

Revised Date: 7 March 2016

Accepted Date: 2 May 2016

Please cite this article as: Sharma R, Oni OA, Chen G, Sharma M, Dawn B, Sharma R, Parashara D, Savin VJ, Barua RS, Gupta K, Association between Testosterone Replacement Therapy and the Incidence of Deep Vein Thrombosis and Pulmonary Embolism: A Retrospective Cohort Study of the Veterans Administration Database, *CHEST* (2016), doi: 10.1016/j.chest.2016.05.007.

Association between Testosterone Replacement Therapy and the Incidence of Deep Vein Thrombosis and Pulmonary Embolism: A Retrospective Cohort Study of the Veterans Administration Database

Rishi Sharma, † Olurinde A. Oni, † Guoqing Chen, § Mukut Sharma, † Buddhadeb Dawn, ‡ Ram Sharma, † Deepak Parashara, * Virginia J. Savin,Ψ Rajat S. Barua, * Kamal Gupta‡

† Kansas City VA Medical Center, Kansas City, MO, USA, (Rishi Sharma, Research Associate, Division of Cardiovascular Research), † Kansas City VA Medical Center, Kansas City, MO, USA (Olurinde Oni, Research Associate, Research Services), § University of Kansas Medical Center, Kansas City, KS, USA (Guoqing Chen, Director, Office of Scholarly, Academic & Research Mentoring), † Kansas City VA Medical Center, Kansas City, MO, USA, (Mukut Sharma, Scientist, Research and Development), ‡ University of Kansas Medical Center, Kansas City, KS, USA (Buddhadeb Dawn, Director, Division of Cardiovascular Diseases), † Kansas City VA Medical Center, Kansas City, MO, USA (Ram Sharma, Associate Chief of Staff, Research Service), *Kansas City VA Medical Center, Kansas City, MO, USA (Deepak Parashara, Chief, Division of Cardiovascular Medicine), Ψ Kansas City VA Medical Center, Kansas City, MO, USA (Virginia J. Savin, Professor, Division of Nephrology), *Kansas City VA Medical Center, Kansas City, MO, USA (Rajat S. Barua, Assistant Professor, Division of Cardiovascular Medicine), ‡ University of Kansas Medical Center, Kansas City, KS, USA (Kamal Gupta, Professor, Division of Cardiovascular Diseases

Correspondence to:

Kamal Gupta, MD

Professor, Department of Internal Medicine

Division of Cardiovascular Diseases

University of Kansas Medical Center

3901 Rainbow Blvd, Kansas City, KS 66160

Email: kgupta@kumc.edu

Conflict of Interest:

None of the authors have any conflict of interest to report.

ABBREVIATIONS:

DVT: Deep Vein Thrombosis

PE: Pulmonary Embolism

TRT: Testosterone Replacement Therapy

VINCI: Veterans Administrations Informatics and Computing Infrastructure

SIPTW: Stabilized inverse probability of treatment weights.

ABSTRACT:

Background: Testosterone Replacement Therapy (TRT) prescriptions have increased several-fold in the last decade. There have been concerns regarding a possible increased incidence of deep venous thrombosis (DVT) and pulmonary embolism (PE) with TRT. There is a paucity of data to support the association between TRT and DVT/ PE. We evaluated the incidence of DVT and PE in men who were prescribed TRT for low serum total testosterone levels (sTT).

Methods: This is a retrospective cohort study, conducted using data obtained from the Veterans Administrations Informatics and Computing Infrastructure (VINCI). We compared the incidence of DVT/PE between those who received TRT and subsequently had normal on-treatment sTT levels (Gp1), those who received TRT but continued to have low on-treatment sTT (Gp2), and those who did not receive TRT (Gp3). Those with prior history of DVT/PE, cancer, hypercoagulable state and chronic anticoagulation were excluded.

Results: The final cohort consisted of 71,407 subjects with low baseline sTT. Of these, 10,854 did not receive TRT (Gp3) and 60,553 received TRT. Of those who received TRT, 38362 achieved normal sTT (Gp1) while 22191 continued to have low sTT (Gp2). The incidence of DVT/PE was 0.5%, 0.4% and 0.4% in Gp1, Gp2 and Gp3 respectively. Univariate, Multivariate, and Stabilized inverse probability of treatment weights (SIPTW) analyses showed no statistically significant difference in DVT/PE free survival between different groups.

Conclusion: This study did not detect a significant association between testosterone replacement therapy and risk of DVT/PE in adult men with low sTT who were at low-moderate baseline risk of DVT/PE.

INTRODUCTION

With an aging population and increased awareness of adult testosterone deficiency, the use of testosterone replacement therapy (TRT) for age-related hypogonadism in the USA has increased several fold in the last decade[1]. Between 2010 and 2013, the number of people receiving TRT increased from 1,299,846 to 2,291,266 with 60% being prescribed by primary care physicians [2]. Several studies have evaluated the association of TRT with acute myocardial infarction and stroke [3-6]. Results of these studies have been conflicting and no consensus has emerged on TRT and the risk of myocardial infarction or stroke[7-9]. A recently published large population-based retrospective study from our center did find that normalization of serum total testosterone levels (sTT) after TRT in subjects with no prior history of CAD/MI/Stroke is associated with a decreased incidence of MI, Stroke and all-cause mortality[9].

More recently, it has been suggested that TRT may be associated with an increased risk of deep vein thrombosis and pulmonary embolism (DVT/ PE)[10,11]. These concerns are based on a limited number of case reports and case series. A recent study did not find an association between DVT/PE and TRT[12]. There is an urgent need for evidence as this is an important health issue as highlighted by the issuance of a recent US FDA and Health Canada warning regarding the possible association of DVT/ PE with TRT[13,14].

We conducted this retrospective study utilizing data from a large patient cohort to investigate the incidence of DVT/ PE in adult men with low sTT who were prescribed TRT.

METHODS

The Institutional Review Board of Kansas City Veterans Affairs Medical Center, USA, approved the study. We conducted this study using de-identified data from Veterans Administrations Corporate Data Warehouse (CDW) through the VINCI. Data on our study subjects was retrieved from December 1999 to May 2014.

Study Design

The aim of this study was to determine the incidence of DVT/PE in subjects who were prescribed TRT for low sTT.

Determination of sTT: A subject was considered to have low testosterone, when the measured sTT levels were below the lower limit of normal laboratory reference range (NLRR) for each reported test result. Position statement from the Endocrine Society and other sources suggest that testosterone levels can vary significantly between different laborateries, even when they use same commercial kits[15,16]. Moreover, because of assay ambiguities and biological variations, no single determination of testosterone can distinguish hypogonadism from eugonadism [17,18]. Consequently, we classified each test result as low or normal based on the respective laboratory reference range reported with the test result. This approach minimized the effect of changes due to multiple assay methodologies and provided sTT values in the context of the method used.

Outcome Measures: The primary outcome measure was the incidence of DVT/PE (ICD-9 codes for DVT included 453.40, 453.41, 453.42, 453.50, 453.51, 453.52, 453.72, 453.73, 453.82, 453.83, 451.11, 451.19, 451.81, 451.83; and for PE 415.11, 415.12, 415.19)

Ascertainment of TRT Exposure: Subjects who were prescribed any forms of TRT (injection, gel or patch) were considered treated with TRT. Treatment exposure and duration were ascertained from the prescriptions in the medical records.

Confounding factor measures**:** We controlled for the effects of confounding factors such as diabetes mellitus (DM), chronic kidney disease (CKD), congestive heart failure (CHF) captured by ICD-9-CM codes. Confounding measures also included patient age and baseline body mass index.

Start day and censoring: The first day of the study for all enrolled subjects was the date of their first reported low sTT. Those who did not develop DVT/PE during the period of follow-up were censored on the last day of utilization of VHA care.

Study Population

Inclusion criteria: Individuals with baseline sTT levels lower than NLRR were included in this study. In order to account for known variability in sTT levels in individuals, the Endocrine Society recommends rechecking levels to confirm true low sTT state prior to initiating TRT[19]. Accordingly, we included only those subjects who showed low sTT levels in a repeat sample prior to TRT. Those showing normal sTT levels in a replicate test before starting TRT were excluded from the final cohort in this study (Figure 1).

Exclusion criteria: We excluded, (i) females, (ii) those who received TRT without baseline (i.e pre-treatment) sTT test results (iii) those who had DVT/PE before the first day of study, (iv) those whose sTT levels normalized without documented TRT, (v) those with a diagnosis of any type of cancer, (vi) those receiving warfarin, and (vii) subjects with history of coagulopathy, ICD-9 code 289.81 (including Protein C and S

deficiency, Lupus anticoagulant, Factor V Leiden mutation, Antithrombin deficiency, prothrombin gene mutation, and activated protein C resistance).

There were no patients on Xa and IIa inhibitors in Gp1 and Gp2, while only 3 patients received these medications in Gp3. Thus, we did not include this parameter on the exclusionary list for this study.

Study Groups:

Normal on-treatment sTT group (Gp1): This group comprised of subjects with low baseline sTT levels who received TRT and in whom subsequent on-treatment testing demonstrated improvement in TT levels to normal i.e., levels within the NLRR. These subjects were considered adequately treated.

Low on-treatment sTT group (Gp2): This group comprised of subjects who continued to have low sTT levels while on TRT i.e. levels lower than NLRR, despite receiving TRT. These were considered to be inadequately treated and include those who may have been non-compliant with TRT use.

Untreated subjects (Gp3): This was the control group and comprised of subjects who had low baseline sTT but did not receive any TRT during the study period.

Statistical analysis

Continuous variables were reported as means (± standard deviation, SD) and categorical variables as percentages. Chi square test and Student's t-test were used to compare normally distributed baseline characteristics of patients. Non-parametric tests were used for non-normally distributed variables. To examine the relationship between

TRT and incident DVT/PE, univariable and multivariable Cox proportional hazard regression analyses were performed.

Propensity score and Stabilized Inverse Probability of Treatment Weights (SIPTW) analyses: To control for the potential selection bias for using TRT, a propensity score approach was used during data analyses. SIPTW analysis was performed which allowed us to retain most subjects in the study while using the weighted propensity scores to achieve balance between each pair of sub-groups compared[20-22]. The covariates used to compute propensity scores were age, body mass index, diabetes mellitus, congestive heart failure, and chronic kidney disease.

Statistical analysis was performed using SAS 9.4 and Stata 12 was used to plot Kaplan-Meier (KM) curves. Stabilized IPTW KM survival curves were plotted with TRT as a time-varying exposure variable to compare event-free survival time between the groups. Log-rank P-values were computed.

RESULTS

Cohort description

Figure 1 describes the study cohort enrollment. We identified a total of 117,094 patients in the database who had low sTT. We excluded 332 patients who had prior DVT/PE because the focus of this study was on incidence of DVT/PE. Also excluded were 28,081 patients without baseline sTT results before initial TRT and 7,388 with diagnosis of primary hypercoagulable state, cancers or use of warfarin. In order to enhance the effectiveness of SIPTW matching, 2374 patients who had missing data with regards to baseline covariates were also excluded from the final analysis. The cohort now comprised

of 78,919 subjects who were categorized into treated (60,553) and untreated (18,366) groups. Among the untreated group we further excluded 7,512 subjects whose testosterone levels normalized after repeat testing. There was no record of treatment for these people, and we could not rule out the possibility of non-VA prescriptions which could have been responsible for this finding. Thus, to prevent misclassification bias, these individuals with spuriously normalized testosterone levels were excluded. Thus the final study cohort comprised of 71,407 subjects with 10,854 untreated patients (Gp3) in the final analysis. In the treated arm, TRT achieved normal sTT levels in 38,362 subjects (Gp1) while the rest of this group i.e. 22,191 subjects continued to have low testosterone (Gp2).

Baseline Characteristics of the Patients

Table 1 presents the demographic and other baseline variables in detail. Mean ages at enrolment were 64.0 years (SD) 11.2), 63.9 years (SD 11.9), and 66.6 years (SD 13.1) for Gp1, Gp2 and Gp3 respectively. Mean BMI at enrollment were 33.0 kg/m^2 (SD 6.5), 33.7 kg/m^2 (SD 6.9), and 32.9 kg/m^2 (SD 6.8) for Gp1, Gp2 and Gp3 respectively. Mean follow-up time was significantly longer in Gp1 (6.1 years, SD 3.1) as compared to Gp2 (4.5 years, SD 2.9) or Gp3 (4.6 years, SD 2.9). Using SIPTW matching, we controlled for discrepancies in the baseline characteristics between the groups with regard to age, BMI and comorbidities. Thus, prior to performing Cox proportional hazard regression analyses, each pair of groups was well matched ($P>0.05$) with regard to these covariates as presented in Table 1.

Relationship between TRT and DVT/PE

The incidence of DVT/PE was 0.5% (207/38362) in Gp1, 0.4% (90/22191) in Gp2 and 0.4% (41/10854) in Gp3. Of these, the PE accounted for 28 of 207 events in Gp1, 15 of 90 events in Gp2 and 6 of 41 events in Gp3. Table 2 presents results of Cox proportional hazard regression analysis. Univariate, multivariate and SIPTW analyses did not detect a significant difference in the risk of incident DVT/PE between the groups. Results from SIPTW analysis were as follows: Gp1 vs. Gp3 (HR=1.1; CI=0.78-1.54, P=0.6); Gp1 vs. Gp2 (HR=0.96; CI=0.75-1.24, P=0.77); and Gp2 vs. Gp3 (HR=1.14; CI=0.78-1.65, P=0.50). KM curves (Figures 2-4) also show that the probability of DVT/PE free survival was not significantly different in the groups compared (logrank P>0.05). Data on race/ ethnicity were not available for a large number of subjects and thus we did not include this variable in the analysis. Where this data were available, subjects were predominantly white. In Gp1, data on race/ ethnicity was available for 12,272 patients and 82.2% were white; in Gp2, 6529 patients had race/ethnicity data and 81.5% were white ; in Gp3, race/ ethnicity data were available for 2840 patients and 77.3% were white.

DISCUSSION

To our knowledge, this is the largest study to evaluate the association between TRT and risk of DVT/ PE. The main finding of the study is that TRT is not associated with an increased incidence of DVT/PE in men with low baseline sTT and no known pre-existing hypercoagulable state.

The bulk of current evidence in regards to TRT and DVT/PE is derived from a series of reports from a single center [11,23-26]. The authors described 42 patients who developed a

venous thrombotic event while on TRT[11,25,26]. Of these, 27 had DVT/ PE . None had polycythemia/ cancer/ prior known DVT/PE or a hypercoagulable state. These 42 patients were compared to 105 healthy normal controls and 42 patients on TRT but no venous thrombosis. A significantly higher prevalence of one or more thrombophilia was observed in the TRT with DVT/PE group compared to the other groups. The authors hypothesized that TRT selected out subjects with previously occult, underlying thrombophilia and resulted in DVT/ PE.

In the recently published case-control study, Baillargeon et al investigated the association of TRT and DVT/PE using administrative health data from a large commercial health program. Similar to our study, the authors did not find any association between TRT and DVT/ PE[12]. The authors identified 7643 cases who were diagnosed with DVT or PE and had either at least one prescription of an anticoagulant filled or had a vena cava filter placed. There were 22,929 controls. The study found that having filled a prescription for TRT was not associated with an increased risk of VTE. Even though the overall study population was large, the numbers of subjects taking testosterone were small. In the study group 158 were on testosterone therapy and in the control group 505. [12]

From a plausibility standpoint, there are several potential mechanisms through which TRT may increase the risk of venous thrombosis. Erythrocytosis or polycythemia is a known side effect of TRT and may contribute to an increased risk of venous thrombosis[27,28]. In animals, platelet aggregation has been shown as the common pathway resulting in increased thrombotic events with androgens[29-36]. There is also evidence that exogenously administered testosterone regulates the expression of Platelet Thromboxane A2 receptors in humans [37]. Another proposed mechanism is that TRT may increase the

levels of circulating estrogens by aromatization of testosterone to estradiol (E2) and this may play a role in increasing thrombotic risk[38-40]. These findings have justifiably raised concern as parallels have been drawn between TRT in men and hormone replacement therapy (estrogen) in post-menopausal women where large well-conducted studies demonstrated that exogenous estrogen replacement is associated with an increased risk of DVT/ PE [41,42].

Our study provides strong epidemiologic evidence that in the population studied, TRT is not associated with an increased risk of DVT/ PE. Even though retrospective in design, the study results are strengthened by the large sample size and long-term follow up in a U.S. veteran population who received healthcare at the VA facilities thus minimizing the likelihood of missing outcome events due to patients receiving care elsewhere. Further, the study confirmed improvement in sTT levels for those on TRT and thus significantly decreases the chance that widespread non- compliance may have biased the results.

This is the first study with a large population on TRT to investigate the association between TRT and DVT/PE. More research, especially with large prospective trials will be needed to confirm these findings and until those are done, we recommend caution as recommended in the recent FDA advisory[13].

Study Limitations

This is a retrospective study and thus subject to the usual limitations including that of unidentified confounders. Despite using SIPTW method to match the groups before analysis, the impact of unidentified confounders cannot be completely eliminated. One of the factors affecting the occurrence of DVT/PE is recent surgery and immobilization. We

did not assess the study groups for these confounders and this may have potentially affected the results. In a retrospective study such as this, it is difficult to assess with certainty if patients in the treatment arm actually took the medications and in an adequate dose. Our study overcomes this limitation by assessing follow up sTT levels to confirm compliance and adequate dosing.

We did not specifically evaluate the prevalence of DVT/PE in individuals with sTT levels higher than the upper range of normal. Although the overall incidence of DVT/PE was very low, the effect of overdosing by TRT might have been missed. Since information on race/ ethnicity was not available for a majority of subjects, it is not possible to assess if effects of TRT vary with ethnicity. The study did not differentiate between formulations and routes of administration of TRT and was not designed to detect differences in the effects of various TRT preparations, if any. Since the study excluded patients with prior DVT/PE or know hypercoagulable state, this study cannot comment on the effect of TRT in this patient population.

From the data available to us for this analysis, the TRT group had longer follow-up time. Though this finding indicates a difference between groups, longer follow-up time provides the opportunity to observe more events in the TRT group if there were a positive association between TRT and DVT/PE. However, despite a longer follow up time, increased incidence of DVT/PE was not observed in the TRT group.

Despite these limitations our study does help fill the knowledge gap by studying the association of TRT with incidence of DVT/PE.

CONCLUSION

This large retrospective study shows that testosterone replacement therapy use in adult men with low sTT and low-moderate risk of DVT/PE is not associated with an increased risk of DVT/PE. This finding was consistent across all subgroups of those receiving TRT (those with normal on-treatment sTT and those with persistent low on-treatment sTT) when compared to untreated subjects. These findings should be confirmed in future studies, starting with prospective cohort studies and if needed randomized controlled trials.

FUNDING and ACKNOWLEDGEMENT

The KCVAMC, and the Midwest Biomedical Research Foundation (MBRF), Kansas City supported this study. Also received support from the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development (VA BX001037 awarded to V.J.S.).

Authors gratefully acknowledge the help received with the study design from Dr. Sharad C. Mathur, M.D., Chief, Pathology and Laboratory Medicine, Kansas City VA Medical center, Kansas City, MO.

GUARANTOR DECLARATION

Corresponding author (Kamal Gupta) is manuscript's guarantor, and affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted; and any discrepancies from the study as planned (and, if relevant, registered) have been explained.

AUTHOR CONTRIBUTIONS

Kamal Gupta, Rishi Sharma, Mukut Sharma, Rajat Barua, Olurinde Oni and Guoqing Chen all had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis, and contributed substantially to the study design, and the writing of the manuscript.

Ram Sharma, Buddhadeb Dawn, Deepak Parashara, and Virginia Savin contributed substantially to the study design, data analysis and interpretation, and writing of the manuscript.

REFERENCES

1 Baillargeon J, Urban RJ, Ottenbacher KJ, et al. Trends in androgen prescribing in the United States, 2001 to 2011. JAMA Intern Med 2013; 173:1465-1466

2 Mohamoud MA CG, Staffa J. Drug use review. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ReproductiveHealthDrugsAdvisoryCommittee/UCM412536.pdf

3 Shores MM, Smith NL, Forsberg CW, et al. Testosterone treatment and mortality in men with low testosterone levels. J Clin Endocrinol Metab 2012; 97:2050-2058

4 Basaria S, Coviello AD, Travison TG, et al. Adverse events associated with testosterone administration. N Engl J Med 2010; 363:109-122

5 Vigen R, O'Donnell CI, Baron AE, et al. Association of testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. JAMA 2013; 310:1829-1836

6 Xu L, Freeman G, Cowling BJ, et al. Testosterone therapy and cardiovascular events among men: a systematic review and meta-analysis of placebo-controlled randomized trials. BMC Med 2013; 11:108

7 Garnick MB. Testosterone replacement therapy faces FDA scrutiny. JAMA 2015; 313:563-564

8 Grech A, Breck J, Heidelbaugh J. Adverse effects of testosterone replacement therapy: an update on the evidence and controversy. Ther Adv Drug Saf 2014; 5:190-200

9 Sharma R, Oni OA, Gupta K, et al. Normalization of testosterone level is associated with reduced incidence of myocardial infarction and mortality in men. Eur Heart J 2015; 36:2706-2715

10 Glueck CJ, Richardson-Royer C, Schultz R, et al. Testosterone therapy, thrombophilia-hypofibrinolysis, and hospitalization for deep venous thrombosis-pulmonary embolus: an exploratory, hypothesis-generating study. Clin Appl Thromb Hemost 2014; 20:244-249

11 Glueck CJ, Richardson-Royer C, Schultz R, et al. Testosterone, thrombophilia, and thrombosis. Clin Appl Thromb Hemost 2014; 20:22-30

12 Baillargeon J, Urban RJ, Morgentaler A, et al. Risk of Venous Thromboembolism in Men Receiving Testosterone Therapy. Mayo Clin Proc 2015; 90:1038-1045

13 FDA adding general warning to testosterone products about potential for venous blood clots. http://www.fda.gov/Drugs/DrugSafety/ucm401746.htm

14 Summary safety review. Testosterone replacement products and cardiovascular risk. Health Canada, 2014. http://www.hc-sc.gc.ca/dhp-mps/medeff/reviews-examens/testosterone-eng.php

15 Lazarou S, Reyes-Vallejo L, Morgentaler A. Wide variability in laboratory reference values for serum testosterone. J Sex Med 2006; 3:1085-1089

16 Rosner W, Auchus RJ, Azziz R, et al. Position statement: Utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. J Clin Endocrinol Metab 2007; 92:405-413
17 Wang C, Catlin DH, Demers LM, et al. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry. J Clin Endocrinol Metab 2004; 89:534-543
18 Vesper HW, Botelho JC. Standardization of testosterone measurements in humans. J Steroid Biochem Mol Biol 2010; 121:513-519
19 Bhasin S, Cunningham GR, Hayes FJ, et al. Testosterone therapy in men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2010; 95:2536-2559
20 Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000; 11:550-560
21 Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res 2011; 46:399-424
22 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods 2010; 15:234-249
23 Freedman J, Glueck CJ, Prince M, et al. Testosterone, thrombophilia, thrombosis. Transl Res 2015; 165:537-548
24 Glueck CJ, Bowe D, Valdez A, et al. Thrombosis in three postmenopausal women receiving testosterone therapy for low libido. Womens Health (Lond Engl) 2013; 9:405-410
25 Glueck CJ, Freiberg RA, Wang P. Heritable thrombophilia-hypofibrinolysis and osteonecrosis of the femoral head. Clin Orthop Relat Res 2008; 466:1034-1040
26 Glueck CJ, Friedman J, Hafeez A, et al. Testosterone, thrombophilia, thrombosis. Blood Coagul Fibrinolysis 2014; 25:683-687
27 Fernandez-Balsells MM, Murad MH, Lane M, et al. Clinical review 1: Adverse effects of testosterone therapy in adult men: a systematic review and meta-analysis. J Clin Endocrinol Metab 2010; 95:2560-2575
28 Coviello AD, Kaplan B, Lakshman KM, et al. Effects of graded doses of testosterone on erythropoiesis in healthy young and older men. J Clin Endocrinol Metab 2008; 93:914-919
29 Pilo R, Aharony D, Raz A. Testosterone potentiation of ionophore and ADP induced platelet aggregation: relationship to arachidonic acid metabolism. Thromb Haemost 1981; 46:538-542
30 Johnson M, Ramey E, Ramwell PW. Androgen-mediated sensitivity in platelet aggregation. Am J Physiol 1977; 232:H381-385
31 Rosenblum WI, el-Sabban F, Nelson GH, et al. Effects in mice of testosterone and dihydrotestosterone on platelet aggregation in injured arterioles and ex vivo. Thromb Res 1987; 45:719-728

32 Nakao J, Change WC, Murota SI, et al. Testosterone inhibits prostacyclin production by rat aortic smooth muscle cells in culture. Atherosclerosis 1981; 39:203-209
33 Gonzales RJ, Ghaffari AA, Duckles SP, et al. Testosterone treatment increases thromboxane function in rat cerebral arteries. Am J Physiol Heart Circ Physiol 2005; 289:H578-585
34 Cembrano J, Lillo M, Val J, et al. Influence of sex difference and hormones on elastine and collagen in the aorta of chickens. Circ Res 1960; 8:527-529
35 Wolinsky H. Effects of androgen treatment on the male rat aorta. J Clin Invest 1972; 51:2552-2555
36 Greenberg S, George WR, Kadowitz PJ, et al. Androgen-induced enhancement of vascular reactivity. Can J Physiol Pharmacol 1974; 52:14-22
37 Ajayi AA, Mathur R, Halushka PV. Testosterone increases human platelet thromboxane A2 receptor density and aggregation responses. Circulation 1995; 91:2742-2747
38 Mooradian AD, Morley JE, Korenman SG. Biological actions of androgens. Endocr Rev 1987; 8:1-28
39 Glueck CJ, Goldenberg N, Budhani S, et al. Thrombotic events after starting exogenous testosterone in men with previously undiagnosed familial thrombophilia. Transl Res 2011; 158:225-234
40 Cherrier MM, Matsumoto AM, Amory JK, et al. The role of aromatization in testosterone supplementation: effects on cognition in older men. Neurology 2005; 64:290-296
41 Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. JAMA 1998; 280:605-613
42 Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA 2002; 288:321-333

TABLES

	UNMATCHED COHORT			SIPTW MATCHED COHORT		
	NORMAL ON-TREATMENT sTT VERSUS UNTREATED GROUPS					
	Normal on-treatment N=38362	Untreated N=10854	P-value	Normal on-treatment N=38362	Untreated N=10855	P-value
Age ≥50 years n (%)	34212 (89.2)	9713 (89.5)	0.3639	34236 (89.2)	9682 (89.2)	0.8862
Age, Mean (SD) Years	64.0 (11.2)	66.6 (13.1)		64.0 (11.2)	66.4 (13.2)	
Body mass index ≥30 kg/m^2	25383 (66.2)	6978 (64.3)	0.0003	25226 (65.8)	7141 (65.8)	0.9588
Body mass index kg/m^2, mean (SD)	33.0 (6.5)	32.9 (6.8)		33.0 (6.6)	33.0 (6.7)	
Follow up time (years), mean (SD)	6.1 (3.1)	4.6 (2.9)		6.1 (3.1)	4.6 (2.9)	
Chronic kidney disease n (%)	311 (0.8)	157 (1.5)	<.0001	364 (1.0)	102 (0.9)	0.9440
Diabetes mellitus n (%)	11649 (30.4)	3467 (31.9)	0.0017	11782 (30.7)	3333 (30.7)	0.9913
Congestive heart failure n (%)	483 (1.3	206 (1.9)	<.0001	536 (1.4)	151 (1.4)	0.9514
	NORMAL ON-TREATMENT sTT VERSUS LOW ON-TREATMENT sTT GROUPS					
	Normal on-treatment N=38362	Low on-treatment N=22191	P-value	Normal on-treatment N=38362	Low on-treatment N=22191	P-value
Age ≥50 years n (%)	34212 (89.2)	19578 (88.2)	0.0003	34080 (88.8)	19716 (88.9)	0.9673
Age, Mean (SD) Years	64.0 (11.2)	63.9 (11.9)		63.9 (11.2)	64.1 (11.9)	
Body mass index ≥30 kg/m^2	25383 (66.2)	15449	<.0001	25873 (67.4)	14967 (67.5)	0.9901
Body mass index kg/m^2, mean (SD)	33.0 (6.5)	33.7 (6.9)		33.2 (6.6)	33.4 (6.9)	
Follow up time (years), mean (SD)	6.1 (3.1)	4.5 (2.9)		6.0 (3.1)	4.6 (3.0)	
Chronic kidney disease n (%)	311 (0.8)	296 (1.3)	<.0001	384 (1.0)	222 (1.0)	0.9924
Diabetes mellitus n (%)	11649 (30.4)	8000 (36.1)	<.0001	12449 (32.5)	7201 (32.5)	0.9969
Congestive heart failure n (%)	483 (1.3	426 (1.9)	<.0001	577 (1.5)	334 (1.5)	0.9974
	LOW ON-TREATMENT sTT VERSUS UNTREATED GROUPS					
	Low on-treatment N=22191	Untreated N=10854	P-value	Low on-treatment N=22191	Untreated N=10855	P-value
Age ≥50 years n (%)	19578 (88.2)	9713 (89.5)	0.0007	19669 (88.6)	9620 (88.6)	0.9800
Age, Mean (SD) Years	63.9 (11.9)	66.6 (13.1)		64.1 (11.9)	66.2 (13.2)	
Body mass index ≥30 kg/m^2	15449	6978 (64.3)	<.0001	15063 (67.9)	7371 (67.9)	0.9594
Body mass index kg/m^2, mean (SD)	33.7 (6.9)	32.9 (6.8)		33.5 (6.9)	33.3 (6.8)	
Follow up time (years), mean (SD)	4.5 (2.9)	4.6 (2.9)		4.5 (3.0)	4.6 (2.9)	
Chronic kidney disease n (%)	296 (1.3)	157 (1.5)	0.4084	304 (1.4)	149 (1.4)	0.9945
Diabetes mellitus n (%)	8000 (36.1)	3467 (31.9)	<.0001	7701 (34.7)	3767 (34.7)	0.9991
Congestive heart failure n (%)	426 (1.9)	206 (1.9)	0.8920	425 (1.9)	208 (1.9)	0.9928

Table 1: Baseline characteristics of all study subjects

Normal-on-treatment group comprised of patients with low pre-treatment total testosterone (sTT) levels that achieved normal sTT levels after testosterone replacement

therapy (TRT). Low-on-treatment group did not achieve normalization of their sTT after TRT. Untreated group comprised of patients who did not receive any form of TRT during the follow-up period. At first, the groups were unmatched ($p = 0.05$) with regard to baseline covariates. Prior to performing data analysis, each pair of groups were matched ($p > 0.05$) using stabilized inverse probability treatment weighted propensity scores.

Model	Normal on-treatment Versus Untreated			Normal on-treatment Versus Low on-treatment			Low on-treatment Versus Untreated		
	HR	95% CI	p-value	HR	95% CI	p-value	HR	95% CI	p-value
Univariate	1.048	0.749-1.466	0.7855	0.953	0.743-1.223	0.7066	1.095	0.757-1.584	0.6309
Multivariate	1.066	0.761-1.492	0.7116	0.970	0.756-1.245	0.8124	1.130	0.781-1.636	0.5169
Stabilized IPTW	1.096	0.778-1.543	0.5998	0.964	0.751-1.237	0.7732	1.138	0.784-1.652	0.4970

Table 2: Unadjusted and adjusted hazard ratios for DVT/PE

Multivariate and stabilized inverse probability treatment weighted cox regression models were fitted using baseline variables presented in Table 1 as covariates. Testosterone replacement therapy was fitted as a time-varying covariate in all models.

FIGURE LEGENDS

Figure 1: Methodology and patients selection process

Inclusion and exclusion criteria used for selecting the final sample cohorts comprising of three groups (normal-on-treatment total testosterone levels, low-on-treatment total testosterone levels and untreated patients.

Figure 2 : Kaplan-Meier curve depicting DVT/PE survival between untreated and normal-on-treatment groups:

DVT/PE-free survival probability was not significantly different between the groups.

Figure 3 : Kaplan-Meier curve depicting the DVT/PE survival between low-on-treatment and normal-on-treatment groups:

DVT/PE-free survival probability was not significantly different between the groups.

Figure 4 : Kaplan-Meier curve depicting the DVT/PE survival between untreated and the low on-treatment groups:

DVT/PE-free survival probability was not significantly different between the groups.

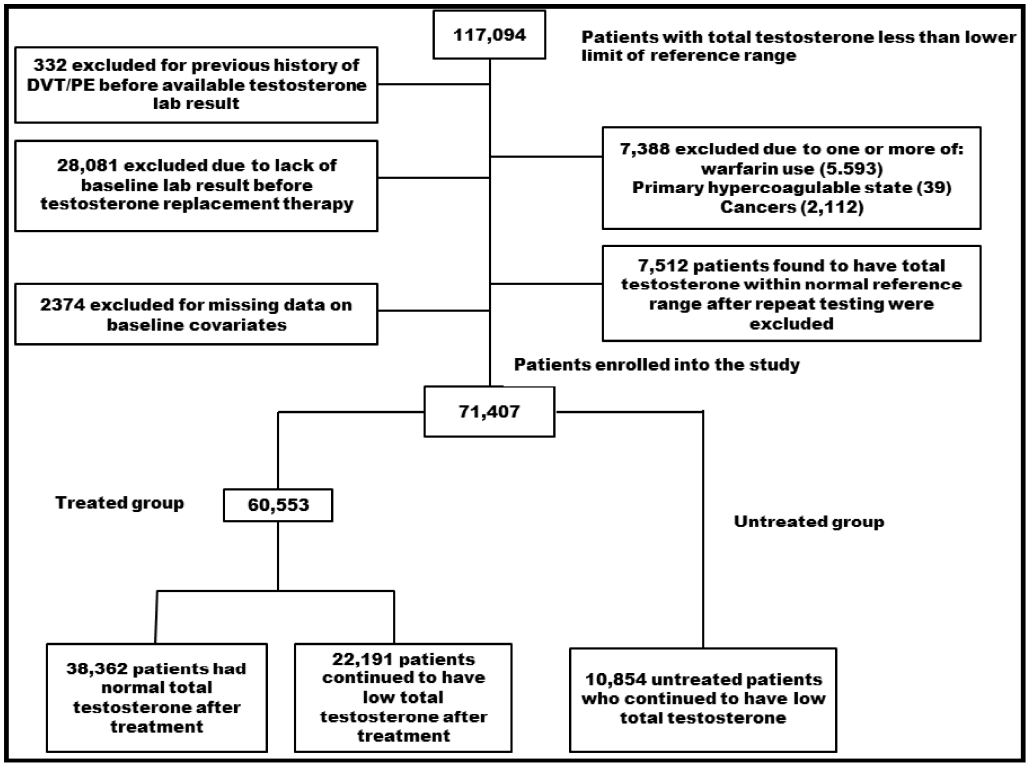


Figure 1: Methodology and patients selection process

Inclusion and exclusion criteria used to select final sample cohorts comprising of the three groups (normal on-treatment total testosterone levels, low on-treatment totoal testosterone levels and untreated patients. DVT/PE- Deep vein thrombosis/pulmonary embolism.

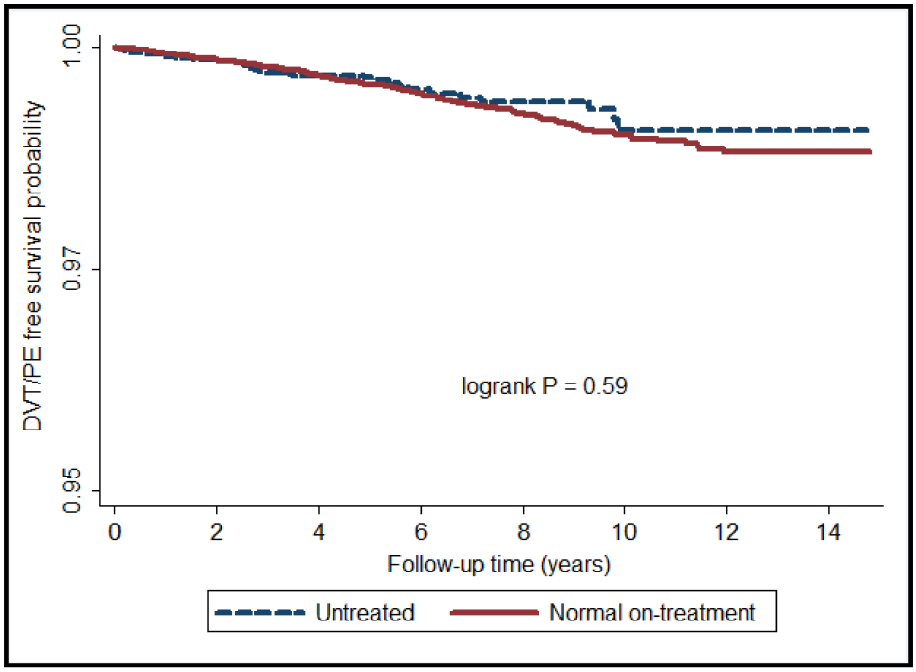


Figure 2 : Kaplan-Meier curve depicting deep vein thrombosis/pulmonary embolism free (DVT/PE) survival between untreated and normal on-treatment groups:

DVT/PE free survival probability was not significantly different between the groups.

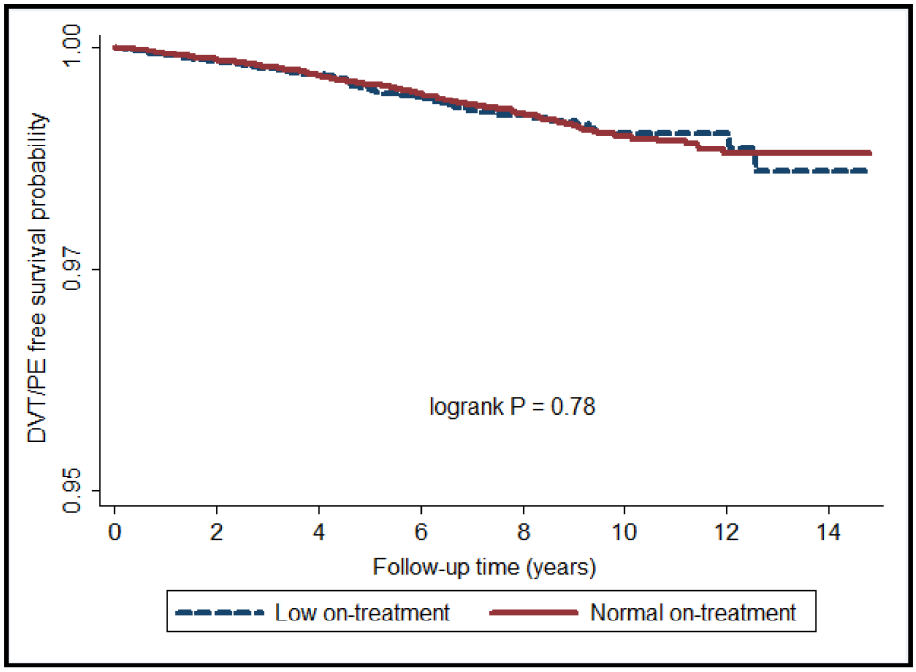


Figure 3 : Kaplan-Meier curve depicting the deep vein thrombosis/pulmonary embolism free (DVT/PE) survival between low on-treatment and normal on-treatment groups:

DVT/PE free survival probability was not significantly different between the groups.

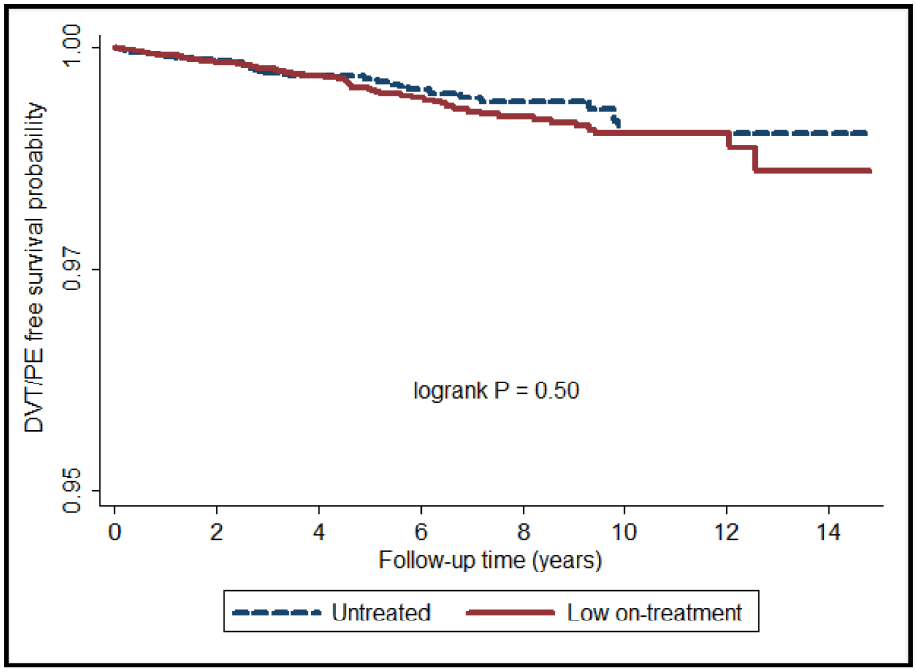


Figure 4 : Kaplan-Meier curve depicting the deep vein thrombosis/pulmonary embolism free (DVT/PE) survival between untreated and the low on-treatment groups:

DVT/PE free survival probability was not significantly different between the groups.