IN THE 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** **This Document Relates to:** **All Actions**

DEFENDANTS' MOTION TO EXCLUDE PLAINTIFFS' EXPERT TESTIMONY ON THE ISSUE OF CAUSATION, AND FOR SUMMARY JUDGMENT, AND MEMORANDUM OF LAW IN SUPPORT

TABLE OF CONTENTS

Page

TABLE OF CONTENTS
(cont'd)

Page

TABLE OF CONTENTS
(cont'd)

Page

TABLE OF CONTENTS
(cont'd)

Page

TABLE OF CONTENTS
(cont'd)

Page

TABLE OF CONTENTS
(cont'd)

Page

TABLE OF AUTHORITIES

Page(s)

CASES

TABLE OF AUTHORITIES
(cont)

Page

TABLE OF AUTHORITIES
(cont)

Page

TABLE OF AUTHORITIES
(cont)

Page

TABLE OF AUTHORITIES
(cont)

Page

TABLE OF AUTHORITIES
(cont)

Page

TABLE OF AUTHORITIES
(cont)

Page

OTHER AUTHORITIES

Defendants AbbVie Inc. and Abbott Laboratories (collectively "AbbVie")[1] respectfully move, pursuant to Federal Rules of Evidence 104(a), 702, 703 and 403, to exclude Plaintiffs' expert testimony that AndroGel causes arterial cardiovascular adverse events or venous thromboembolic events.

Specifically, Defendants request that the Court exclude the causation opinions of Drs. Ardehali (general CV causation: cardiology; specific causation: Plaintiff Mitchell), Cuculich (general CV causation: cardiology), Gerstman (general CV causation: epidemiology), Halushka (general causation: thromboxane mechanism), Rinder (general VTE causation: hematology; specific causation: Plaintiffs Myers, Nolte, Rowley), Setaro (specific causation: Plaintiff Cribbs), Wells (general CV causation: statistics), and Ziman (specific causation: Plaintiff Frost). Because Plaintiffs lack admissible expert testimony to establish general or specific causation, which are essential elements of their claims, AbbVie also moves, pursuant to Federal Rule of Civil Procedure 56, for summary judgment in its favor with respect to all of Plaintiffs' claims. In support of these motions, AbbVie states as follows:

PRELIMINARY STATEMENT

As the Court already is aware, the administration of testosterone replacement therapy ("***TRT***") has been an accepted, FDA-approved clinical practice for over 75 years, and AndroGel in particular has been on the market for over 15 years. This brief focuses on the scientific evidence published from 2010 to 2014 that actually prompted this litigation, as well as the significant additional research published since 2014.

---

[1] AbbVie Inc. was established in January 2013 as an independent, publicly traded company from the innovative pharmaceutical business of Co-Defendant Abbott Laboratories, which no longer sells AndroGel in the US.

The initiation of thousands of product liability cases on the heels of new scientific developments poses special problems for burgeoning litigation. There is no regulator on the filing of suits; and once filed, they must be addressed systematically without any procedural device specifically designed for the purpose. Compounding these challenges, the science which both prompts and predicates such litigation is often thin. Whether it supports any of the claims and, if so, which ones, becomes a crucial question. Over 20 years ago, the Supreme Court's *Daubert* decision made addressing that question a requirement. The years since have produced notable examples of what can happen if this is not done timely and effectively. In the Breast Implant litigation during the early 1990s, for example, thousands of claims were filed and litigated in state and federal court within a period of months. By the time *Daubert* review was undertaken in the federal courts years later, the litigation dispersed around the country and had cost billions of dollars.

In this case, the Court is in a far better position to undertake a meaningful *Daubert* review. First, the available science regarding TRT has grown substantially even since this litigation began. The issues of cardiovascular ("***CV***") and venous thromboembolic ("***VTE***") risks can be addressed with the benefit of that science. Second, courts are now far more experienced in understanding how the available science can answer the issues of causation that are key in a *Daubert* analysis.

For these reasons, the science and the law can readily be joined in this case. As discussed below, the established science of both cardiovascular disease and venous thromboembolisms requires statistically significant association in a well-designed study, replication through other such studies, and then analysis of scientifically accepted causal factors—such as the strength and

consistency of the association, temporality, dose-response, and biological plausibility—to determine whether there is an actual causal relationship.

The courts that have delved into the science in pharmaceutical cases have properly heeded these rigorous scientific requirements in reviewing expert testimony under *Daubert*. That is the only viable path for a court to discharge its gatekeeping function. The ultimate decision represents one of the primary values the MDL process can confer on the litigation before cases are remanded.

Taking that path is particularly crucial in this case, because Plaintiffs' experts uniformly seek to abrogate the established methods and standards for determining CV and VTE causal factors in favor of precisely the kind of subjective judgments that *Daubert* was designed to avoid. Tests for statistical significance are characterized as "misleading" and rejected in favor of non-statistical "estimates," "clinical judgment," and "gestalt" views of the evidence. A "Bayesian" computation is preferred to the statistical analyses employed by all of the researchers who have published epidemiological studies on TRT. This is true despite the fact that the Bayesian computation was done solely for this litigation, by an expert who has never published the algorithm he used, and who still uses traditional methods of significance testing in his own published research. Indeed, his own initial report in this case nowhere launched the attack on statistical significance that followed his receipt of Defendants' expert reports. Another expert—Plaintiffs' sole expert on VTE causation—vigorously applies strict tests of causation in reciting the established causes of VTE, admits that the science on TRT meets none of those same requirements, and even goes so far as to say that all of the VTE epidemiological studies are "flawed." Yet, he does not hesitate to conclude that TRT is a new VTE causal risk factor and

adopts legal terminology (which he admits has no meaning to him as a scientist) to conclude that TRT was a "substantial contributing factor" for each of Plaintiffs' injuries.

In this case, "leaving it to the jury" is certain to cause confusion and prejudice. No matter how technical-sounding, the evidence and methods relied upon by Plaintiffs' experts must be assessed rigorously. Indeed, the jury can only decide the facts based upon what the Court determines is reliable science. Because they flout the metrics of reliable science, the opinions of Plaintiffs' experts must be excluded. *See Rosen* v. *Ciba-Geigy Corp.*, 78 F.3d 316, 319 (7th Cir. 1996) ("Law lags science; it does not lead it.").

***

Moreover, Plaintiffs' experts are not writing on a clean slate when they opine on general causation. Experts at the FDA, employing scientifically sound methodology and rigorous standards, have repeatedly assessed most of the science proffered by Plaintiffs' experts. The same four CV studies upon which they still rely were immediately investigated by the FDA after the articles were published in 2010 and 2013–14, and more expanded reviews were conducted by the FDA thereafter.

Each time, the FDA found the available evidence insufficient to establish that TRT is even ***associated*** with increased CV risk. Rather, the FDA determined that the studies showed a "***possible risk***" or a "***weak signal***." The process and resulting findings are extensively documented. The conclusions were summarized by the FDA staff in an August 2015 article in the New England Journal of Medicine. They reported that an FDA Advisory Committee "acknowledged the limitations of the available data on adverse cardiovascular events but concluded that the totality of the evidence ***suggests a weak signal*** of cardiovascular risk and recommended updating drug labels to reflect this information." (emphasis supplied)

Of course, the FDA's mission is to protect public health, and it can act to do so even without scientific evidence establishing that a drug causes an adverse outcome. And it did so with respect to TRT by requiring in 2015 that the labeling be revised to warn doctors to tell their patients that some studies had reported a possible increased risk of CV events. But the predicates for regulatory action in service of public health are not the same as the requirements for scientific proof of causation in a court of law. What is crucial about the FDA's action, therefore, is its analysis of the relevant science on the merits and its determination that the science did not support a finding of increased risk, much less causation. Indeed, the FDA has never found that an increased risk exists or that TRT causes CV events.

The FDA also investigated whether TRT increases the risk of VTE. After initially proposing a label which said that TRT "***may*** increase the risk" of such events, upon completing its investigation the FDA decided not to include this language. Instead, the label was revised in June 2014 to state, in part: "There have been postmarketing reports of venous thromboembolic events, including deep vein thrombosis (DVT) and pulmonary embolism (PE), in patients using testosterone products such as AndroGel 1%."

Very substantial science has been published during and since the FDA's investigations. There are no fewer than 25 epidemiological TRT studies on CV risk and 5 studies on VTE risk, all of them published in peer-reviewed journals and authored by a wide variety of investigators. These studies have analyzed the potential risk to TRT patients generally and to specific subgroups. The 30 studies report scores of relevant findings. Only three of the ***new*** studies report statistically significant associations between TRT and CV events, and these associations are traced to similar subgroups of patients: men who are over 65 years of age who used TRT for a short period of time. All of the other reported findings (including many in the same studies)

are either not statistically significant or show that TRT patients have a ***reduced risk*** of CV events. So, Plaintiffs' experts continue to rely upon the ***earlier*** studies that the FDA found too flawed to demonstrate association.

This means that, even using Plaintiffs' own selection of research, the general causation issues come down to a handful of studies. A summary chart of CV studies is attached to this brief as **Figure 3**. Also, Exhibit 68 includes each of those studies and relevant excerpts from AbbVie's expert reports. In a nutshell, Plaintiffs' experts select and rely upon studies that report ***statistically significant risks*** for the following particular ***groups***:

- **Risk to TRT Users Overall:** Only one study, Xu (2013), found a statistically significant risk to all TRT users. It was criticized by the FDA after it was published. Plaintiffs' expert epidemiologist Dr. Gerstman continues to cite Xu despite the fact that he also relies upon a more recent study using the same methodology but including more recent data (praising it for providing a "best estimate). That study found no statistically significant risk of CV events for users overall (Albert (2016)). Several additional studies report the same conclusion as Albert (2016).

- **Users Over the Age of 65:** As noted, the only statistically significant findings for this subgroup are traced to a sub-subgroup of those with **short-term treatment**. The research reports no statistically significant association for those with longer treatment. As reflected in the expert materials in Exhibit 68, the studies of short term treatment have important limitations, including some identified by the FDA. None of the bellwether Plaintiffs in this case is in this subgroup.

- **Users Under the Age of 65:** Plaintiffs' experts single out one earlier study specific to this population, Finkle (2014). Finkle's finding of a statistically significant risk is limited to **short duration exposure** (90 days), and, further, to those **with a history of CV disease.** Finkle found no such association for patients without such a history. Not only was this study found by the FDA to be of limited value, but it has never been replicated. Other studies focused on patients under the age of 65, and on patients with history of CV disease, have not found statistically significant associations. Plaintiffs' experts erroneously opine that one bellwether Plaintiff, Mr. Konrad, is in this Finkle sub-subgroup.

With respect to VTE, a total of five epidemiological studies have been published on TRT use. All five have reported no statistically significant increased risk to large groups of TRT users. One study (Martinez (2016)) did a subgroup analysis and reported a statistically

significant association limited to a "transient risk" (***fewer than 6 months*** of TRT use) for patients ***without a history of other VTE risk factors***. The authors of that study further acknowledge that their research merely "***suggests*** a transient increase in the risk of venous thromboembolism that peaks during the first three to six months" and that "[f]urther research is needed to confirm this temporal increase in the risk." None of the bellwether Plaintiffs is in this subgroup. Another of the five reports is actually a letter describing some *ad hoc* meta-analysis work on limited data selected from the earlier Xu (2013) CV study referenced above. Re-analyzing some of that data to focus on VTE rather than CV yielded a non-statistically significant result. The authors of the letter then added data from another source and found a significant result (further details on this letter are provided below). Each of the epidemiological studies on VTE risk is reflected in ***Figure 4*** and contained in Exhibit 69. (Ex. 69).

In reviewing this body of scientific evidence, it is also essential to compare it to the methods and standards which have been adopted by the scientific community for decades in determining the ***non-TRT causal factors*** for the diseases at issue. There is no dispute in this case that, for these traditional CV and VTE risk factors, science has required (1) the demonstration and replication of statistically significant association in well-designed studies, (2) evidence satisfying other criteria needed to support a finding of causation, such as specificity, dose-response, temporality, and biological plausibility.

None of Plaintiffs' experts do, or reasonably can contend that the science for TRT meets these requirements. Instead, Plaintiffs' experts adopt the bold strategy of attacking the bedrock scientific methods and replacing them with subjective judgments based upon the "totality of the evidence." Depending upon the expert, even the basic tests of statistical significance are simply ignored, dismissed as misleading, or rejected in favor of "Bayesian" probability methods, non-

statistical estimates, or "clinical judgment." This dismissal of statistical significance stands in stark contrast to all of the TRT epidemiological studies (including those that Plaintiffs single out), which, without exception, apply the traditional tests of statistical significance. Further, while Plaintiffs' experts recognize that association is not the same thing as causation, many of them follow no accepted method or process for deciding causation. For them, it is sufficient to rely upon their own analyses of post-marketing adverse event reports, other case reports, non-statistically significant findings in epidemiological studies, and theories of biological mechanism that have not been tied to CV or VTE events in TRT users. Though it may be sufficient for these experts in their medical practices, it is not sufficient to satisfy the stringent *Daubert* standards, and they should be prohibited from testifying.

While the seven bellwether cases are treated individually only at the end of this admittedly lengthy brief, they involve case-specific causation issues that also are also dispositive. **Figure 1** shows which of the causation cross-cutting issues apply to each of the bellwether Plaintiffs. With respect to the three clotting cases, it is particularly notable that two of the cases involve Plaintiffs whose hematocrit levels were at or below normal lab limits (generally, at or a somewhat above 50%), and one Plaintiff had a reading above 51%, thus framing two of the cross-cutting causation issues listed in Plaintiffs' causation case. Moreover, Plaintiffs' sole VTE expert Dr. Rinder has admitted that, in all three of those VTE cases, he has no evidence the Plaintiff's hematocrit was even changed as a result of TRT. Indeed, he has no evidence linking any of the VTE Plaintiffs to any of his proffered theories of mechanism. Despite this lack of evidence, Dr. Rinder is still prepared to opine that TRT is a "substantial contributing factor" to their VTEs. These and other facts are explored further below.

STATEMENT OF FACTS

A. Cardiovascular Events and Venous Thromboembolisms

The CV events at issue in this case are myocardial infarctions ("MI," also known as heart attacks) and strokes. The kind of heart attacks and strokes alleged are the result of atherosclerosis, a process that occurs when the inner layer of an ***artery*** is injured by activities like smoking, high blood pressure, high cholesterol, or diabetes. Cellular waste products traveling through the blood stream near the site of injury accumulate into a plaque (made primarily of cholesterol), and platelets clump around the plaque to try to repair the injury. Eventually, the clumps of platelets either grow and obstruct blood flow or can embolize (or separate) from the arterial wall, causing a clot (or thrombus) that prevents blood flow downstream. When this clot blocks an artery, an MI occurs. When the clot interrupts blood flow to a certain area of the brain, a stroke occurs. *See* Expert Report of William J. French, M.D. at 7–9 (Ex. 1)[2]; Expert Report of Phillip Cuculich, M.D. at 9–11 (Ex. 2); Expert Report of Hossein Ardehali, M.D., Ph.D. at 15–20 (Ex. 3).

In contrast, *venous* thromboembolisms are clots of red blood cells that can form anywhere in the body, typically due to one of three mechanisms (also known as "Virchow's triad"): injury to the blood vessel wall (similar to atherosclerosis but in a vein, not an artery),

[2] AbbVie's causation experts are: William French (general CV causation: cardiology; specific causation: Plaintiffs Frost and Konrad), Sucha Nand (general VTE causation: hematology; specific causation: Plaintiff Rowley), Barbara Bierer (general VTE causation: hematology), Jacques Baillargeon (general causation: epidemiology), Laurentius Marais (general causation: biostatistics), William White (specific causation (cardiology): Plaintiffs Cribbs and Mitchell), Kenneth Bauer (specific causation (hematology): Plaintiffs Myers and Nolte), Howard Kirshner (specific causation (neurology): Plaintiffs Frost and Nolte), Mohit Khera (general causation and prescribing of TRT: urology; specific prescribing of TRT to Plaintiffs Nolte and Rowley), Adrian Dobs (prescribing of TRT to Plaintiff Cribbs: endocrinology), Andrew Heyman (prescribing of TRT to Plaintiff Frost: family medicine), Martin Miner (prescribing of TRT to Plaintiffs Konrad and Mitchell: family medicine), Helena Rodbard (prescribing of TRT to Plaintiffs Myers and Konrad: endocrinology).

changes in the pattern of blood flow, and changes in the red blood cell concentration of the blood. A deep vein thrombosis ("***DVT***") usually occurs when a clot forms in the lower extremities (e.g., the legs) and causes pain, swelling, and redness. A pulmonary embolism ("***PE***") occurs when a clot that formed in one part of the body embolizes from the venous wall, travels to the right side of the heart, and becomes lodged in a pulmonary blood vessel. PEs most commonly cause chest pain and shortness of breath, but they can be fatal. *See* Expert Report of Sucha Nand, M.D. at 6–8 (Ex. 4); Expert Report of Henry Rinder, M.D. at 5–7 (Ex. 5).

B. General Causation of CV Events

The ultimate issues of CV general causation in this case are whether there is an association between TRT and CV events and whether that association rises to the level of a causal relationship. *See, e.g.*, Gerstman Dep 9:7–15, Jan. 4, 2017 (Ex. 6). But the Plaintiffs' experts reject outright the scientifically recognized tests for what constitutes a reliable association with TRT and fail to apply the scientific methods and criteria for demonstrating causation.

1. The Accepted Causal Risk Factors for MI and Stroke

It is undisputed that the types of CV events that Plaintiffs attribute to TRT are the same as those that are caused by a variety of well-known inherited and acquired casual risk factors. These risk factors include: increased age, male gender, family history of CV disease, smoking, obesity, elevated blood cholesterol levels (hyperlipidemia), high blood pressure (hypertension), lack of physical activity, and diabetes. *See generally* Ardehali Dep. 97:9–98:16, 123:14–124:24, Jan. 6, 2017 (Ex. 7); Cuculich Dep. 71:16–72:20, Jan. 13, 2017 (Ex. 8); French Dep. 301:6–22, Jan. 18, 2017 (Ex. 9); French Rpt. at 10–18 (Ex. 1).

The scientific methods that were used to establish these accepted causes also are not disputed. Ardehali Dep. 101:10–102:7 (Ex. 7); French Dep. 462:14–464:8 (Ex. 9). The ideal

way to research a causal relationship between a drug and an outcome is to conduct a large randomized controlled trial ("***RCT***"). Expert Report of B. Burt Gerstman, Ph.D. at 30 (Ex. 10); Marais Dep. 489:4–490:3, Jan. 17, 2017 (Ex. 13). Randomization "maximizes the possibility that exposures other than the one under study are evenly distributed between the exposed and the control cohorts," and it therefore increases the probability that any observed effect is due to the exposure being studied rather than bias or confounding factors inherent in the data sample. Fed. Judicial Ctr., Nat'l Research Counsel of the Nat'l Acads., *Reference Manual on Scientific Evidence* at 592 (3d ed. 2011) [hereinafter "Reference Manual"] (Ex. 14). RCTs are "prospective" studies because the subjects are selected before any exposure occurs, are exposed to the drug or placebo, and then followed. *Reference Manual,* at 582. Increasing the number of subjects then can increase the statistical "power" of the study to detect a hypothesized risk. *Reference Manual,* at 626 (Ex. 14). At the conclusion of the RCT, the results are analyzed statistically to determine whether the subjects in the treatment group experienced a higher or lower rate of disease compared to the placebo group. *Reference Manual,* at 555 (Ex. 14).

Where the goal is to determine risk factors for the development of diseases that occur over long periods of time, RCTs can take years and large RCTs to determine all adverse events for drugs are not conducted as a routine matter. In the absence of such RCTs, observational studies are next best. They too gather data for treatment and control groups and statistically compare them to determine whether the treatment group experienced a different incidence of disease than the control group. Observational studies can be conducted prospectively or retrospectively, the latter meaning that the researchers use already-existing data collected by (or being collected by) government organizations, other medical researchers, or healthcare providers. *See Reference Manual,* at 555-60 (Ex. 14).

Meta-analysis is another epidemiological research method. It combines data already collected from other (usually smaller) RCTs and statistically analyzes the data as a whole to generate "the single, more accurate result that would have been obtained from a hypothetical single, larger study" with greater statistical power. Expert Report of M. Laurentius Marais, Ph.D. at 44–45, Dec. 6, 2016 (Ex. 16); *Reference Manual,* at 581, n.89 (Ex. 14). When there is a limited body of RCTs, separate meta-analyses examining the same issue will rely on the same general body of underlying data, although more recent meta-analyses will have access to the newest RCT data. *See* Gerstman Rpt. at 65–67, Table 4.4 (Ex. 10) (showing that meta-analyses related to CV risk of TRT incorporate many of the same underlying studies).

As noted, these epidemiological studies look for differences in the incidence rates of CV events among the groups being compared. In this case, people taking TRT are the "treatment" or "exposed" group and people who are not taking TRT are the "control" group. Comparing the frequency of CV events in the treatment group to the frequency in the control group enables researchers to determine whether the observed CV events are "associated" with exposure to the drug or potential risk factor being studied, that is, whether the CV events and exposure occur together. This comparison is usually expressed as a ratio, usually an odds ratio ("OR") or relative risk ("RR")). *Reference Manual,* at 566–69 (Ex. 14). If the incidence rates of CV events in the two groups is the same, the ratio will be 1.0, indicating that the treatment is not associated with CV outcomes. *See generally Reference Manual,* at 567 (Ex. 14). Thus, an OR or RR of 1.0 means there is no difference between the treatment and control samples in terms of CV events. An OR or RR greater (or less) than 1.0 signifies that the treatment or risk factor is associated with an increase (or decrease), respectively, in the incidence rate of CV events. *Id.*

Any association found must be tested for its statistical significance. Statistical significance testing measures the likelihood that the observed association could be due to chance variation among samples. Scientists evaluate whether an observed effect is due to chance using p-values and confidence intervals. The prevailing scientific convention requires that there be 95% probability that the observed association is not due to chance (expressed as a ***p-value < 0.05***) before reporting a result as "statistically significant." *See Reference Manual,* at 576–81 (Ex. 14); Marais Dep. 224:1–14, 440:18–23, 451:17–452:5 (Ex. 13); Gerstman Dep. 104:1–105:1 (Ex. 6); (Ex. 70) (explaining statistical significance). This process guards against reporting false positive results by setting a ceiling for the probability that the observed positive association could be due to chance alone, assuming that no association was actually present. Thus, if a p-value lies below 0.05, the result is statistically significant because there is at least a 95% probability that it was not due to chance. If the p-value is greater than 0.05, the result is not statistically significant because one cannot rule out with at least 95% certainty that the observed result was due to chance. *See Reference Manual,* at 576–77 (Ex. 14); Marais Dep. 441:7–16 (Ex. 13).

The determination of statistical significance can be described equivalently in terms of the ***confidence interval*** calculated in connection with the association. A confidence interval indicates the level of uncertainty that exists around the measured value of the association (i.e., the OR or RR). A confidence interval defines the range of possible values for the actual OR or RR that are compatible with the sample data, at a specified ***confidence level***, typically 95% under the prevailing scientific convention. *Reference Manual,* at 580 (Ex. 14) ("If a 95% confidence interval is specified, the range encompasses the results we would expect 95% of the time if samples for new studies were repeatedly drawn from the same population."). If a point

estimate indicates increased risk (i.e., OR or RR > 1.0) but the confidence interval includes results equal to or less than 1.0 (e.g., 0.2-2.3), it is impossible to be 95% confident that the observed effect was not due to chance. If the confidence interval crosses 1.0, this means there may be no difference between the treatment group and the control group, therefore the result is not considered statistically significant.

It is undisputed that the accepted causes of MIs and strokes have been established in multiple, consistent studies showing ***statistically significant*** associations with the various causal factors. *See* Ardehali Dep. 101:10–102:7 (Ex. 7); Gerstman Dep. 274:17–275:7 (Ex. 6); *see generally* French Rpt. at 10–18 (Ex. 1). The magnitude of the effect of each of these causal factors also has been quantified. Ardehali Dep. 112:18–113:23 (Ex. 7) (discussing Framingham Study risk calculator); Cuculich Dep. 283:23–284:2, 301:1–22 (Ex. 8) (discussing American Heart Association risk calculator); Cuculich Rpt. at 10-11 (Ex. 2). The same methodologies also have been used to establish that certain factors are associated with reduced incidence of CV events and therefore are protective.

The acceptance of the established CV risk factors as "**causal**" has required more than a well-designed study showing a statistically significant association. The basic criteria for causation were established over 50 years ago in the 1964 Surgeon General's Report on Smoking and Health and a paper by Sir Austin Bradford Hill from 1965. (Ex. 17); Sir Austin Bradford Hill, *The Environment and Disease: Association or Causation?*, 58 Proc. Royal Soc'y Med. 295–300 (1965) (Ex. 18); *see also Reference Manual,* at 600 (Ex. 14). The factors outlined by the Surgeon General are consistency and strength of the observed association, specificity of the outcome to the risk factor being studied, temporality (meaning the effect occurs after someone is exposed to the risk factor), and coherence of the scientific evidence in support of a causal

relationship. (Ex. 17 at 20). The Bradford Hill factors are similar: (i) consistency and strength of relationship, (ii) temporal relationship between exposure and outcome, (iii) dose-response relationship, (iv) replication of findings, (v) biological plausibility,[3] (vi) alternative explanations, (vii) specificity, and (viii) coherence. (Ex. 18 at 295–59). Certain of Plaintiffs' experts acknowledge that demonstration of causation is a process that should follow these criteria. Cuculich Dep. 118:4-20 (Ex. 8); Cuculich Rpt. at 4 (Ex. 2); Gerstman Rpt. at 89 (Ex. 10); Gerstman Dep. 336:23–337:7, 341:18–343:8 (Ex. 6).

2. Science Regarding Whether TRT Can Cause MI or Stroke

To determine whether TRT is a **new** causal risk factor for CV events, AbbVie's experts apply the same methods and metrics used to determine the **established** risk factors for CV events. In doing so, they assure both that their work is reliable and that their conclusions fit the ultimate legal issue of general causation. Critically, those same methods are used by the researchers who generated the relevant body of science on TRT and also by the FDA when it assesses research.

(a) Epidemiological TRT Research Through the FDA Investigation

Scientific and medical knowledge of the safety of TRT arose from long-standing approval and use of these drugs dating back to 1940 and from clinical trials focused on whether the products were effective in raising deficient serum testosterone levels to the normal range. The FDA played a central role in designing these trials. The clinical trials conducted for AndroGel 1%, and later for 1.62%, followed this same basic design: efficacy was determined by measuring AndroGel's effect on serum testosterone levels in men with testosterone below

[3] Specifically, biological plausibility requires evidence of the "biological mechanism" by which the risk factor could cause the disease. *See* 1964 Surgeon General's Report, at 185 (Ex. 17); Hill, *Environment and Disease*, at 298 (Ex. 18). This is discussed in further detail below.

300ng/dL. Safety was monitored by evaluating any adverse events. The trials also measured subjects' estradiol levels (estradiol is a form of the female hormone estrogen and a metabolite of testosterone; it is associated with CV events in women on hormone replacement therapy ("HRT"). But the TRT clinical trials were not large or long enough (meaning they lacked the statistical "power") to gather the data necessary to look for a statistical association between AndroGel and CV risk. This was the result of the prevailing regulatory standards for testing TRT, *see* Kessler (a regulatory expert for Plaintiffs) Dep. 98:6–99:4, 114:21–24, Jan. 12, 2017 (Ex. 19), and the fact that no cardiovascular "signal" had been observed in the decades of TRT use before AndroGel was tested, *id.* 158:13–18. As the FDA's Guidance on the subject explains, conducting such large and highly-powered trials for long-term chronic disease is not feasible or necessary for every drug, even when a potential safety signal has arisen. *See* FDA Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment Mar. 2005 [hereinafter "FDA Pharmacovigilance Guidance"] (Ex. 20 at 12–13). The FDA followed this approach in regulating TRT until it called for a large RCT to evaluate CV risk in 2015.

Some of the clinical trials that were conducted on TRT were published. The results were unremarkable, none of them reporting an elevated CV risk. None were seen as a "signal" of adverse events by the drug companies or by the FDA. Three meta-analyses were done in 2005, 2007, and 2010. These studies aggregated the results of many published RCTs in order to specifically focus on whether there was evidence of an association between TRT and increased CV risk. None of the meta-analyses reported a statistically significant increased risk of CV events. *See* Ex. 68 (Calof (2005), Haddad (2007), Fernandez-Balsells (2010)).

The first trial to report increased CV risk among TRT users was published in July 2010 by Dr. Shehzad Basaria, *et al*. The paper reported on the Testosterone in Older Men with

Mobility Limitations ("TOM") Trial. This was a small RCT designed to examine the effects of TRT on physical function and leg strength in 209 frail men over the age of 65 (average age of 74) with mobility limitations and low serum testosterone levels between 100-350 ng/dL. Although it demonstrated that men who received TRT experienced improvements in strength and functional status, the study was stopped early because a higher proportion of subjects in the treatment group experienced a broad variety of adverse CV events than did subjects in the placebo group. A statistical analysis of these results revealed that the odds ratio of increased CV risk in the treatment group was 5.4 (95% CI = 2.0-14.9). As this result shows, the association was statistically significant (the lower bound of the range is above 1.0) (Ex. 68) (Basaria (July 2010)).

The Basaria authors did not claim that their study demonstrated that TRT caused CV events; rather, they reached the limited conclusion that: "In this population of older men with limitations in mobility and a high prevalence of chronic disease, the application of a testosterone gel was ***associated*** with an increased risk of cardiovascular adverse events." (Ex. 68) ((emphasis added)). Dr. Basaria later noted that "these findings should be viewed as ***hypothesis-generating*** and need confirmation in prospective randomized trials. (Ex. 22) (emphasis added).

After the study was halted, the FDA conducted an investigation of the literature on TRT and CV available at that time and concluded that "the results regarding overall cardiovascular risk associated with TRT compared to placebo are consistent and ***do not support an association between TRT and increased risk of cardiovascular events in men.***" (Ex. 23 at 6). The agency further examined the issue in December of that year, and again determined that "one cannot make the conclusion. . . that testosterone therapy increases the risk of cardiovascular disease." (Ex. 24).

The agency's Division of Epidemiology ("DEPI") again concluded a few months later that Basaria and the other epidemiological evidence available as of January 2011 "did not demonstrate that TRT is associated with an increased risk of CV events" and "that there was ***insufficient evidence of a cardiovascular risk associated with TRT to warrant a regulatory action***." Ad Comm. Docs., at E13-23 (Ex. 25) (emphasis added). Indeed, DEPI noted that TRT was associated with "a general trend towards improved lipid profiles, with the potential to decrease the risk of adverse CV events." *Id*.; *see also* (Ex. 26 at 76); (Ex. 27).

No further TRT study reported an association with CV risk until the middle of 2013. In the intervening years, the clinical trials for AndroGel 1.62% were completed, and the drug was approved without any finding of CV risk. *See* Ex. 28 (approving 1.62% with no changes to the label).

From April 2013 through January 2014, three new TRT studies were published, all reporting statistically significant associations with CV risk. As a result, the FDA again began to explore this issue in January 2014. The investigation advanced in steps, with consistent findings at each step.

The review was done to respond to a petition filed by Public Citizen, a public advocacy group, requesting that three separate measures be undertaken to warn of an increased CV risk for TRT users. The group contended that four studies (Basaria (2010), Xu (2013), Vigen (2013), and Finkle (2014), provided sufficient evidence to demonstrate such a risk. Pub. Citizen Pet., Feb. 25, 2014 (Ex. 29). In responding to the request, the FDA undertook a systematic review of the four studies, together with additional research reported in the literature. All four studies were found to have significant limitations. FDA Resp. to Pub. Citizen, July 16, 2014 (Ex. 30).

Re-reviewing Basaria, described above, the FDA found the following:

- **Overbroad definition of CV events led to difference between groups.** The FDA explained that "[t]he majority of the cardiovascular events reported were not major adverse cardiac events (MACE), MI, stroke, and deaths due to stroke or MI, and represented diverse pathophysiology. When we evaluated only the MACE [events]. . ., we found only a very small numerical imbalance between the testosterone and placebo groups." (Ex. 30 at 6). For example, the Basaria study considered "peripheral edema" as a cardiovascular event.

- **Designed to evaluate benefits, not CV safety.** The study was not designed to assess cardiovascular risk and therefore "the testosterone and placebo groups were not balanced for cardiovascular risk factors, which could explain the imbalance in MACE between the two groups." *Id.*

- **Small number of subjects from highly specific population.** The FDA also found it "questionable whether the study results are applicable to the population for whom testosterone therapy is indicated." "[T]he indicated population for testosterone therapy is men of all ages with confirmed hypogonadism," but the Basaria researchers examined only frail elderly men with mobility limitations who already faced "a high risk of cardiovascular disease and [had] low-normal testosterone levels." *Id.*

Further analysis of Basaria (2010) in this litigation has confirmed these limitations and found others. The study and the additional findings made by AbbVie's experts can be found at Exhibit 68. (Ex. 68).

The FDA also critically reviewed the Xu meta-analysis published in April 2013. The study combined the data from 27 small RCTs (including Basaria) in which TRT was administered to subjects and a large variety of CV events were recorded. It reported that the odds ratio of any CV event among overall TRT users was 1.54 (95% CI = 1.09-2.18), and of a serious CV event was 1.61 (95% CI = 1.01-2.56). The Xu authors did not characterize their study as establishing a causal link between TRT and increased CV risk, but stated only that they "examined the overall risk of cardiovascular-related events ***associated*** with testosterone therapy." (Ex. 68) (Xu (Apr. 2013)) (emphasis supplied).

The FDA found important limitations in the Xu meta-analysis as well:

- **Heterogeneity of component studies.** The FDA stated that a "major concern" with this study "is the heterogeneity of the trials and their suitability for

integration," as they differed "in almost all aspects of study design—age of participants, inclusion and exclusion criteria, study duration, drug formulation, [ ] dose. . . [and] in the baseline cardiovascular risk of participants." (Ex. 30 at 7–8).

- **Overly broad definition of "CV event."** The FDA also criticized the Xu authors' use of a composite cardiovascular events outcome, which tracked over twenty types of outcomes, "because they anticipated too few events for a robust assessment by cardiovascular event type." The agency noted that, "[w]hile combining these clinically heterogeneous events with widely varying severity and biological mechanisms may provide the necessary power to detect a difference between treatment arms, the aggregated outcome is difficult to interpret and may mask or distort the signal for the most clinically important cardiovascular outcomes. . . The limited interpretability of such broad composite outcomes is one reason that FDA typically uses MACE to assess cardiovascular risks." *Id*. at 8.

- **Excluded studies.** The Xu authors noted "that 138 out of 169 potentially relevant studies (82%) were excluded for 'no cardiovascular-related events reported by study arm.'" The authors also "excluded trials under 12 weeks' duration to 'assess long-term rather than acute effects of testosterone therapy." The FDA explained that these decisions could have substantially shaped the outcome of the trial, and specifically stated "that systematically excluding adverse events that occur early in treatment could create bias." *Id.*at 8–9.

Again, analysis in this litigation has confirmed these findings and shown additional limitations and flaws. The study and the additional limitations of Xu (2013) can be found at Exhibit 68. (Ex. 68).

The third study, Vigen, *et al.*, was published in November 2013. It examined the incidence of death, MIs, and strokes in hypogonadal men who underwent coronary angiography (meaning they already had cardiovascular disease) in the Veterans Administration healthcare system. The TRT group experienced 123 events while the non-TRT group experienced 1,587 events, but after the authors applied a novel statistical method (stabilized inverse probability of treatment weighting or "SIPW") that uses over 50 variables (none of which are testosterone levels), they concluded that the absolute risk increase for TRT users compared to controls was 1.29 (95% CI = 1.05-1.58, p = 0.02). This was the **sole** statistically significant finding from the Vigen study. **That finding was reached only after applying the SIPW method to data that**

showed statistically <u>insignificant</u> results for overall risk. The Vigen authors did not claim to have established that TRT causes increased CV risk, stating instead that "[u]se of testosterone therapy in this cohort of veterans with significant medical comorbidities was ***associated*** with increased risk of mortality, MI, or ischemic stroke" and "[f]uture studies including randomized controlled trials are needed to properly characterize the potential risks of testosterone therapy in men with comorbidities." (Ex. 68) (Vigen (2013)) (emphasis added).

Again, the FDA found important limitations including:

- **<u>Lower incidence of CV events in TRT group reversed by "unclear" statistical method.</u>** The raw data collected from the database by the Vigen investigators "show[ed] no significant difference between the testosterone and no-testosterone groups in the incidence of MI, stroke, and death." The FDA criticized the study on the ground that the authors were only able to reach the opposite conclusion after applying "a complex statistical model using a weighting scheme involving approximately 50 variables," which excluded "[a]t least one key variable. . . : the significantly lower baseline testosterone level in the testosterone group, which was viewed by the authors' peers as a significant oversight." (Ex. 30 at 12).

- **<u>Increased risk possibly due to inadequately treated underlying hypogonadism.</u>** On average, subjects in the Vigen study began with a baseline serum testosterone level of 175.5 ng/dL that was raised to 322.2 ng/dL, just above the level of being hypogonadal and below the range generally used for treatment (400-500 ng/dL). Based on this data, the FDA concluded that "[t]he adequacy of testosterone treatment seen in this study does not appear to reflect the recommended clinical guidelines for testosterone replacement. Due to these treatment uncertainties, it is difficult to attribute the increased risk for the composite outcome to testosterone therapy alone. It is important to consider that the study subjects might have simply remained hypogonadal and thus at higher risk for cardiovascular events, regardless of treatment." *Id.* at 11.

- **<u>Bias from excluded patients.</u>** The FDA joined the larger medical and scientific communities in criticizing the Vigen authors for excluding "128 patients who experienced MI or stroke before initiating testosterone. . . . These patients should have been included in the analysis and their events included in the no-testosterone group, which would have raised the event rate in the no-testosterone group by 71%. Their exclusion biased the results by reducing the number of events in the no-testosterone group." *Id.* at 12.

In this litigation, Plaintiffs' expert Dr. Gerstman agrees with the FDA's assessment and does not list Vigen among the studies he finds more reliable. Gerstman Dep. 67:9–68:16, 200:18–201:2

(Ex. 6) (testifying that the FDA reached "a reasonable conclusion with respect to Vigen."). AbbVie's experts have reached the same conclusion for these and additional reasons. *See* Ex. 68.

The fourth study, Finkle, *et al.*, was published in January 2014. This study analyzed data in a health insurance database to determine the rate of non-fatal MIs ("NfMI") **90 days following an initial prescription**. The authors did not use a control group, but instead compared the incidence of NfMIs in TRT users one year before and 90 days after filling their initial prescription, and found a statistically significant increased risk (RR 1.36 (95% CI = 1.03-1.81)). They then "stratified" their data in order to study **subgroups** based upon age and history of CV disease. They found other statistically significant findings during the 90-day exposure window for men over the age of 65 (RR 2.19 (95% CI = 1.27–3.77)), men over 65 without history of CV disease (RR 2.21 (95% CI = 1.09–4.45)), and men under the age of 65 with history of CV disease (RR 2.90 (95% CI = 1.49–5.62)). The Finkle authors concluded only that "the evidence supports an ***association*** between testosterone therapy and risk of serious, adverse cardiovascular-related events–including non-fatal myocardial infarction–in men." (Ex. 68) (Finkle (Jan. 2014)).

The Finkle study also suffers from several methodological flaws that the FDA determined limit its value:

- **Increased risk possibly due to inadequately treated underlying hypogonadism because TRT exposure unknown.** TRT exposure was determined solely on the basis of a patient filling a single TRT prescription, and the study's authors made no attempt to measure how much—if at all—those patients actually used TRT. The FDA explained that this fact "makes it impossible to determine if the testosterone levels in these treated males had reached therapeutic range. Due to these limitations, it is difficult to completely attribute the increased risk for [NfMI] to testosterone treatment." (Ex. 30 at 13–14).

- **Lack of control group and limitation to short-term use.** The Finkle authors' choices to use self-control analysis (i.e. incidence of non-fatal MIs before and

after TRT prescription in the same patient) instead of a placebo group and to limit follow-up to only 3 months of therapy were questioned by the FDA, which doubted that they were "adequate to capture the relevant outcomes" in view of the fact that "[n]ormally, testosterone is prescribed for chronic use." *Id.* at 14.

- **Improperly limited CV endpoint.** The authors chose to evaluate only NfMIs, meaning that any men who died from an MI during the study period were not included in the analysis. The FDA noted that the lack of clarity on how inclusion of fatal MIs would have affected the result was a significant limitation of this study. *Id.*

AbbVie's experts have confirmed these observations and made many others regarding the problems with Finkle (2014). *See* Ex. 68.

Based upon its review of the four studies and additional research, the FDA rejected each of the requests made by Public Citizen. Specifically, in July 2014, the FDA "considered the studies submitted in support of [the] requests. . . and conclude[d] that, at this time, there is insufficient evidence of a causal link between testosterone therapy and adverse cardiovascular outcomes to support the regulatory actions requested in your Petition. . . . The Agency [has] determined that each study had significant limitations." (Ex. 30 at 5). **And the FDA directly addressed the question of whether an increased risk had been found**: "FDA believes that the publication of these studies warrants further exploration of a ***possible safety signal*** regarding testosterone and cardiovascular risk." *Id.* at 16 (emphasis added).

A "signal" is a predicate for the FDA to conduct further investigation. In the words of Plaintiffs' epidemiologist Dr. Gerstman, a signal is an "indication." Gerstman Dep. 230:22–231:12 (Ex. 6); *see also* Ardehali Dep. 177:24–178:9 (Ex. 7) (defining signal as "enough evidence out there to look into and consider that a drug ***may*** have adverse effects." (emphasis added)); Cuculich Dep. 164:7–18 (Ex. 8).

Where the potential signal comes from case reports or adverse event reports of events that occurred after a product is marketed, the FDA's process and metrics for evaluating that kind of

signal is called "pharmacovigilance." *See* FDA Pharmacovigilance Guidance (Ex. 20). But the CV issue concerning TRT in 2013–14 was based on epidemiological studies, **not** case reports or adverse event reports. The FDA can use epidemiology to assess the strength or validity of a potential signal and the analysis of the studies is termed "pharmacoepidemiology." This follows general epidemiological methods. FDA Pharmacovigilance Guidance (Ex. 20). The FDA's specific assessment of the four epidemiological studies in 2014 applied these methods and standards. **The FDA's 2014 evaluation is, in this way, directly relevant to the assessment of the same evidence in this case.**

This background clarifies the FDA's ultimate conclusion in July 2014 that the studies supported only a "***possible safety signal,***" a "possible" indication. FDA Resp. to Pub. Citizen, at 16 (Ex. 30) (emphasis added). The possibility of an indication is far from the actual increased risk advocated by the petitioning group, as further evidenced by the FDA's Guidance: "Signals generally indicate the need for further investigation, which ***may or may not*** lead to the conclusion that the product ***caused*** the event." (Ex. 20 at 4) (emphasis added). Determining whether a signal needs further investigation requires the analysis of further criteria, which overlap with the criteria for causation.[4] *See id.* at 4–12. Thus, a "possible safety signal"—before it has even been investigated according to FDA Guidance—is not equivalent to a finding of increased risk, or causation, or even a reliable, statistically significant association.

Moving on to the next step of the FDA's 2014 investigation, the FDA decided to convene an Advisory Committee in September 2014 to discuss various issues concerning TRT, including

[4] While these FDA criteria are designed to produce a "risk/benefit" assessment rather than a finding of scientific causation, they include several criteria similar to the Bradford Hill factors (along with assessment of benefits). *See* FDA Pharmacovigilance Guidance, at 4–12; Hill, *Environment and Disease* (Ex. 18).

CV risk. In preparation for this meeting, FDA Staff from various divisions worked to compile a Briefing Book. The Staff's Briefing Book reported that, "[b]ecause of important limitations, the available epidemiological studies do not provide convincing evidence that TRT is associated with adverse CV outcomes" and "the evidence for increased risk of cardiovascular events with TRT is not conclusive." (Ex. 31 at E31–31) [hereinafter "Advisory Comm. Briefing Book"] (emphasis added). Crucially, the Staff again concluded that the available studies "***suggested a possible cardiovascular safety signal*** associated with testosterone therapy," but that "each of the studies had major limitations, precluding the ability to draw definitive conclusions." (Ex. 31 at E31–154) (emphasis added). *See also* Gerstman Dep. 219:5–17, 219:24–220:14 (Ex. 6) (testifying that the FDA Staff's conclusion was "very similar" to the conclusion reached by the FDA in rejecting the Public Citizen petition).

The Advisory Committee agreed with these conclusions when it met in September 2014: ***"a weak signal of cardiovascular risk had emerged from results of recent large epidemiologic studies."*** (Ex. 32 at 7) [hereinafter "Ad. Comm. Mins."] (emphasis added). "Overall the committee agreed that the ***potential signal*** for cardiovascular risk should be added to the labeling." *Id.* (emphasis added). Thus, the Advisory Committee agreed that the available evidence showed only a weak indication, not actual CV risk.

Following the Advisory Committee meeting, the FDA took action on all TRT labelling and required the addition of new warning language. The warning was consistent with the Committee's conclusion, as it did not state that TRT causes CV events, but rather said that the existing studies "have been ***inconclusive*** for determining the risk of major adverse cardiovascular events" and instructed doctors to inform patients "of this ***possible risk*** when

deciding whether to use or to continue to use" TRT. AndroGel 1.62% Label § 5.5, May 2015 (Ex. 33) (emphasis added).

In an August 2015 article, referenced above, the FDA Deputy Director for Safety, Division of Bone, Reproductive and Urologic Products and her colleagues summarized the findings of the Advisory Committee. (Ex. 34 at 689–91). They wrote that the Advisory Committee had "acknowledged the limitations of the available data on adverse cardiovascular events but concluded that the totality of the evidence ***suggests a weak signal*** of cardiovascular risk and recommended updating drug labels to reflect this information." *Id.* at 691 (emphasis added). The Committee also "commented that only a controlled clinical trial—not observational studies—will be able to definitively determine the effects of testosterone therapy on cardiovascular outcomes." *Id.* "The FDA agreed with the advisory committee's recommendations and subsequently required revisions to the labels of all testosterone products." *Id.*

(b) Epidemiological TRT Research Since the FDA Investigation

Today, a total of 25 epidemiological studies have been published on the potential association between TRT and CV events. Currently, at least 18 studies—many of which report multiple findings—have reported no statistically significant association between TRT and increased CV risk. Six have found that TRT is associated with protection against CV events. *See* **Figure 3**.

In contrast, seven studies report that TRT is significantly associated with an increase in CV events. Four of the seven studies are older, already have been assessed by the FDA, and have been found wanting. *See supra* Section B.2.a. The remaining three studies—Etminan (2015), Wallis (2016), and Albert (2016)—have looked broadly for associations and also traced associations down to specific sub-groups of people. In doing so, they have found statistically

significant results for patients over 65 and only for short periods of treatment. *See* Gerstman Dep at 253:19–254:8 (Ex. 6). Importantly, these same studies and many others have examined other associations between CV and TRT among overall users and in many subpopulations, and they have reported non-statistically significant results. *See* **Figure 3**.

Three general points that characterize this body of evidence should be emphasized before reviewing the state of the epidemiology today in more detail below:

First, all of the researchers have tested if associations are statistically significant by determining whether p-values are <0.05 and/or whether the 95% confidence intervals are above a point estimate of 1.0, thereby excluding the possibility that TRT treatment is not different from controls (i.e., no TRT treatment). None of the researchers bridle at the notion of testing for statistical significance—none criticize these tests, adopt different tests for determining the significance of the reported results, or refuse to report the test results. *See generally* Ex. 68.

Second, like the FDA, all of the researchers have used epidemiological analysis and methods to assess the limitations of their studies. Experts on both sides agree that this should be done and have found limitations in many studies. Ardehali Dep. 69:14–70:1 (Ex. 7); Baillargeon Dep. 205:3–16 (Ex. 11); French Dep. 424:11–425:11 (Ex. 9); Gerstman Dep. 31:16–24 (Ex. 6).

Third, while the three new articles that report statistically significant findings conclude that TRT may be associated with increased CV risk in some subgroups, none say that they have established that TRT is a new causal factor for CV events, much less do they analyze the criteria necessary to reach such a conclusion. Indeed, no such researcher has ever reported such a result in the peer-reviewed literature. *See generally* Ex. 68.

The following are the principal subgroups of TRT patients studied in the articles. Again, Exhibit 68 contains the studies and AbbVie's experts' assessments of them. To help keep the big

picture in view, **Figure 3** organizes the most significant study findings by the group studied, and where a finding has been followed down to a subgroup level, that finding appears on the chart. This process is described in further below.

TRT Users Overall

For the broadest patient group, only one study by Xu (2013) has reported a statistically significant association between TRT and increased CV risk. Gerstman Dep. 308:20–309:3 (Ex. 6); French Dep. 451:18–452:1 (Ex. 9); **Figure 3.** As described above, this was one of the studies reviewed by the FDA and found too flawed to demonstrate increased CV risk.

Several studies before and since Xu (2013) have looked for the same risk and not found statistically significant evidence to support it. Among them is a more recent meta-analysis by Albert, *et al.* (2016), which evaluated more data from newer RCTs. Albert is regarded by Plaintiffs' epidemiologist, Dr. Gerstman, as one of the more reliable studies. Gerstman Dep. 68:16–69:12 (Ex. 6). That study reported no statistically significant association between TRT and increased CV risk among overall users. (Ex. 68) (Albert (2016)). Apart from Albert, eight other meta-analyses show no statistically significant association for users overall, and four observational studies (Shores (2012), Anderson (2015), Sharma (2015), Tan (2015)) report a statistically significant association between TRT and ***decreased*** CV risk, meaning that TRT is ***protective*** against CV events. **Figure 3.**

Subgroups of TRT Users

Looking beyond TRT users overall, researchers have increasingly sought to define and study specific subpopulations for whom TRT use might increase CV risk. While these same subjects are included among TRT users as a whole, the investigation of subgroups analyzes data that is specific to those particular groups and determines whether each subgroup faces a risk that the overall population of TRT users does not. Epidemiological studies can be designed to study

these subpopulations from the outset. For example, researchers could decide that the inclusion criteria for their study requires subjects to be over 65 years of age; therefore, only patients meeting that requirement will be enrolled at the outset of the study. Similar limitations can be imposed by using exclusion criteria.

Subgroups can also be evaluated as a later step in studies that begin by looking at the overall population. For example, a study of TRT users overall could then divide the data according to distinct populations, e.g. users aged 65 and older and users under the age of 65, and then analyze each group separately.[5] This further step might reveal that an association found in the larger group actually reflects an association that exists only in a subgroup within the population.[6]

Several subgroups of TRT users have now been studied, including some subpopulations that are defined by multiple characteristics:

TRT Users with Pre-Existing Disease ("Comorbidities"): The only study that has reported a statistically significant increased risk of CV events among all TRT users who have a

---

[5] **Figure 2** diagrams these steps in the Albert (2016) study, discussed above.

[6] However, the practice of studying subpopulations *post-hoc* (as opposed to pre-specifying them at the outset of the study) is problematic for two primary reasons. First, when subgroups are not pre-specified, the investigators cannot balance potential confounding factors between the exposure and control subgroups being considered. **This is even more problematic when researchers proceed to stratify subgroups from a larger group in which there was no statistically significant association (for example, finding no statistically significant association among overall TRT users, but proceeding to stratify according to age or duration of use).** *See* French Dep. 284:18–286:4 (Ex. 9) (criticizing Albert (2016) meta-analysis). Second, performing too many comparative statistical tests on the same set of data complicates the interpretation of p-values associated with those tests because, using a 95% level of evidence, one out of every twenty tests will be statistically significant due to chance alone. *See Reference Manual*, at 290 ("For example, if 20 divisions of a company are examined, and one division is found to have a disparity significant at the 5% level, the result is not surprising; indeed, it would be expected"). Experts on both sides have recognized that this "data mining" raises additional issues of reliability beyond those already associated with retrospective studies. Baillargeon Dep. 232:6–234:4, 236:2–8 (Ex. 11); Marais Rpt. at 16–18 (Ex. 16); *see also* Gerstman Rpt. at 70 (Ex. 10).

history of CV events is Vigen (2013). **Figure 3.** As described above, Vigen was an observational study that found a statistically significant increased incidence of CV events in those patients whose records included a prescription for TRT. The FDA had several criticisms of the Vigen study and, in this litigation, both Plaintiffs' experts and AbbVie's experts have testified that Vigen is not reliable. Gerstman Dep. 200:18–201:2 (Ex. 6); Marais Dep. 345:2–347:22 (Ex. 13); French Dep. 265:24–267:7, 269:1–270:10 (Ex. 9). Other studies also make different findings for TRT users with comorbidities. Two meta-analyses (Corona (2014), Alexander (2016)) and three observational studies (Shores (2012), Baillargeon (2014), and Etminan (2016)) failed to find any association between TRT and increased CV risk for comorbid groups, and one observational study (Baillargeon (2014)) reported that TRT was protective against mortality in diabetic patients. **Figure 3.**

Longer–Term and Shorter–Term Use Generally: No study finds statistically significant increased risk of CV events for long-term users overall. Marais Rpt. at App'x. F, App'x G (Ex. 16); French Dep. 439:14–18 (Ex. 9); **Figure 3.** Two meta-analyses have failed to find a statistically significant increase in CV risk among long-term TRT users (Corona (2014) and Albert (2016)), and one observational study reported that TRT was significantly associated with a protective effect against CV events (Wallis (2016)). **Figure 3**; *see also* Gerstman Rpt. at 52–53, 81–84 (Ex. 10) (relying on Albert and Wallis); Gerstman Dep. 54:23–55:5, 68:16–69:12 (Ex. 6) (testifying that Corona and Albert are among the more reliable meta-analyses). For short-term or first-time users, one meta-analysis reports a statistically significant increase in risk among TRT users (Albert (2016)),[7] but one meta-analysis (Corona (2014)) and two observational studies (Baillargeon (2014) and Etminan (2015)) have failed to find a statistically significant

[7] This finding is discussed in further detail below.

association between TRT exposure and increased CV events in this subgroup. **Figure 3**; *see also* Gerstman Rpt. at 52-53 (relying on Wallis as a reliable study) (Ex. 10); Gerstman Dep. 54:23–55:5, 68:16–69:12 (Ex. 6).

Users Over the Age of 65 Generally: No epidemiological study reports a statistically significant association between TRT exposure and increased CV events in users over the age of 65 generally. **Figure 3**; Marais Rpt. at 121 Fig. 15, App'x F, App'x G (Ex. 16); French Dep. 442:7–10 (Ex. 9). Two meta-analyses (Corona (2014) and Albert (2016)) and six observational studies (Shores (2012), Baillargeon (2014), Etminan (2015), Ramasamy (2015), Wallis (2016), and Maggi (2016), failed to find any statistically significant association with increased CV risk. Indeed, two of these observational studies reported that TRT was associated with a statistically significant reduction in mortality in this population (Shores (2012) and Wallis (2016)). One long-term RCT studying the elderly, albeit powered to measure benefits rather than risk, showed no safety signal (Snyder (2016)). *See* **Figure 3**. (noting another RCT (Basaria 2015) that found no statistically significant increased risk).

Short-Term Users Over the Age of 65: Moving to **subgroups of subgroups**, one meta-analysis (Albert (2016)) and three observational studies (Finkle (2014), Etminan (2015), and Wallis (2016)) report statistically significant association with increased CV risk among patients over the age of 65 **during short-term use of TRT**. Marais Rpt. at 127 Fig. 19, App'x. F, App'x. G (Ex. 16); French Dep. 442:16–443:16, 444:21–445:7 (Ex. 9); Gerstman Dep. 164:19–165:25, 254:9–255:6, 256:16–24 (Ex. 6); **Figure 3.** The durations of use range from less than three months to less than twelve months.

Specifically, Finkle (2014), referenced above, reported that there was a statistically significant increased risk of CV events among men over the age of 65 who had filled a

prescription for a TRT product in the 90 days prior to their CV event and had no history of CV disease. Again, this older study has been criticized by the FDA. *See supra* Section B.2.a.; App'x A at A-251–60.

The second study is the Albert (2016) meta-analysis, which was designed to examine data gathered in RCTs on TRT users overall. As noted above, Albert found no statistically significant association for users overall. Notwithstanding this finding, the Albert investigators proceeded to define and study subgroups based on duration of use.[8] After finding a statistically significant association between short-term TRT use (defined as less than twelve months) and increased CV outcomes, the investigators went to a third level of subgroup analysis and found that ***the statistically significant association among short-term users persisted only in individuals over the age of 65***. *See* Ex. 68 (Albert (2016); French Dep. 452:8–453:5 (Ex. 9) (criticizing Albert study for proceeding to further subgroup analyses of data that was not statistically significant initially).

Two recent observational studies have found statistically significant associations in similar subpopulations. Etminan (2015) reported that there was a statistically significant association with a small increased risk of MI among first-time TRT users (defined as men whose first TRT prescription was within 90 days of having an MI). While the study was not specifically stratified to separate those who were below and above 65, the patients were an "older" group, average age 70 years old. The authors concluded only that this data supported "***a potential MI safety signal***" among this subgroup. *See* Ex. 68 (Etminan (2015)).

Finally, Wallis (2016) is a recent observational study that found no association between TRT exposure and CV events among TRT users over the age of 65 generally. However, after

[8] For the reasons explained above, this practice increases the unreliability of the result. *See* n.7, *supra*.

performing a *post-hoc* analysis to stratify the data according to duration of use, the authors found a statistically significant association with increased CV risk among elderly men exposed to TRT for up to 120 days (average of two months). The Wallis authors acknowledged several limitations of their study, including the inability to match cases and controls on the basis of underlying hypogonadism or testosterone levels and lack of data on the adequacy of TRT in patients, which could have driven the observed increased risk of CV events in the short-duration group. *See* Ex. 68 (Wallis (2016)).[9]

Long-Term Users Over the Age of 65: No study reports a statistically significant association for this population. Marais Rpt. at 127 Fig. 19, App'x F, App'x G; French Dep. 442:8–10 (Ex. 9); Gerstman Dep. 314:14–315:6 (Ex. 6). One observational study (Maggi (2016) reported no statistically significant association, and one reported that long-term TRT exposure is statistically significantly associated with decreased CV risk (Wallis (2016)). **Figure 3.**

Users Under the Age of 65 Generally: No study reports a statistically significant association. Marais Rpt. at 121 Fig. 15, App'x F, App'x G (Ex. 16); French Dep. 439:19–440:1 (Ex. 9); Gerstman Dep. 310:18-311:4 (Ex. 6). One meta-analysis (Albert (2016)) and one observational study (Maggi (2016) report no statistically significant association. One observational study reports a statistically significant association with ***decreased CV risk*** (Shores (2012)). **Figure 3.**

Short-Term Users <u>Under</u> the Age of 65, with <u>Pre-Existing Disease</u>: Finkle (2014), again, reports a statistically significant association in short duration use (less than six months)

[9] As discussed above, the Basaria (2010) study also examined frail men with mobility limitations who were over the age of 65 and exposed to high doses of TRT for six months and found a statistically significant association, but both the authors and the FDA identified significant limitations of this study. (Ex. 68) (Basaria (2010))

among users under the age of 65. Gerstman Dep. 271:6–10; 310:18–311:4 (Ex. 6). This is the only study to make such a finding. Looking further, Finkle ultimately found increased risk for the subgroup with a prior history of CV events, but **no** such increased risk for the subgroup **without** prior history. Gerstman Dep. 271:6–272:15 (Ex. 6); French Dep. 442:3–6 (Ex. 9). But, in contrast to Finkle, Albert (2016) reported no statistically significant association for users under the age of 65 who were exposed to TRT for less than twelve months. **Figure 3**. (Finkle (2014) and Albert (2016)).

(c) Moving from Association to Causation: Application of Causal Criteria

Association is not the same as causation, and without association there cannot be causation. *Reference Manual* at 598. In sharp contrast to the research that was required to establish the many **other causal risk factors for CV events**, the body of **TRT** science set out above fails even to report a statistically significant association, excepting the four early studies that the FDA found insufficiently reliable to establish increased risk and the three more recent studies that reported risk only for a sub-subgroup of patients. *Cf.* Ardehali Dep. 101:10–102:7 (Ex. 7) (accepted causal factors are established based on "extensive epidemiological" evidence); French Dep. 462:14–464:8 (Ex. 9) (same). This lack of consistent and reliable evidence of a statistically significant association ends the causation inquiry. *See Reference Manual,* at 598–99 (Ex. 14) ("We emphasize that these guidelines are employed only *after* a study finds an association to determine whether that association reflects a true causal relationship"); *id.* at 599 n.141 (applying causal criteria in the absence of epidemiological evidence of association "does not reflect accepted epidemiologic methodology").

The following briefly sets out additional facts about the body of science that are relevant to the additional criteria necessary to scientifically—and therefore, legally—demonstrate

causation. All of the other accepted causes of CV events have met such criteria. But TRT use does not. *See* Ardehali Dep. 104:1–12 (Ex. 7) ("risk factors for atherosclerotic cardiovascular disease are a set of already identified risk factors, and testosterone is not one of those risk factors"); Cuculich Dep. 69:2–12 (Ex. 8); French Dep. 462:14–465:1 (Ex. 9).

(i) ***Consistency and Strength of the Relationship:*** *Any Association Between TRT Use and the CV Events Is Weak and Inconsistent*

When science establishes only lower relative risks, "the epidemiologist will scrutinize such associations more closely because there is a greater chance that they are the result of uncontrolled confounding or biases." *Reference Manual,* at 602 (Ex. 14). Most of the studies reporting any association between TRT and increased CV risk report relatively ***low risk estimates***. For example, the Xu study concluded that the relative increased risk of CV events due to TRT is **1.54**, and the Vigen study reported a RR of only **1.29**. *See* App'x A at A-131–47, A-189–99 (Xu (2013), Vigen (2013)). The FDA took note of this fact in concluding that these studies have limited value in determining whether there was an association. *See* FDA Resp. to Pub. Citizen, at 15 (Ex. 30) (noting "Overall effect estimates are small and may be due, in part, to residual confounding (Vigen, Xu, Finkle).").

Finkle also finds a significant association for men over 65, with no history of CV disease, who were exposed for a short duration of 90 days or less (RR: **2.21**) for only non-fatal MIs. Albert reports a significant RR of **2.90** for men over age 65 and TRT use for less than 12 months, but those results do not evidence a **consistent** association because these studies measured different outcomes and defined "short term" as different durations. There are important differences among these studies and the two others reporting similar associations that make it difficult to conclude that they are consistent with one another. First, Finkle was an observational study that examined only non-fatal MIs. In contrast, Albert is a meta-analysis that

examined a wide variety of CV events and reported an association in subgroups defined differently from Finkle's (Finkle looked at men with history of CV disease and limited exposure to 90 days, but Albert stratified by duration of under 12 months or 12 months and longer and did not stratify by disease history). Second, Finkle and Etminan (which also reported a low risk estimate of **1.41**) are subject to the criticism that they did not obtain data on subjects' underlying testosterone levels, making it difficult to determine whether the observed increased short-term risk was due to TRT exposure or, instead, to persisting hypogonadism. (Ex. 68 (Finkle (2014) Etminan (2015)). *Third*, as reflected immediately below, research on longer term use by patients over the age of 65—including from these studies—shows no statistically significant association with increased CV risk. *See* **Figure 3.** The same is true of other research on men with pre-existing disease. *Id*. *Fourth*, no study has established a reason for the different results reached for men under 65 with short-term use. *Fifth, and finally*, none of these studies concludes that they have demonstrated that TRT can, in fact, cause CV events.

(ii) ***Temporality:*** *The Requirement of A Temporal Relationship Between Exposure and Outcome Is Not Met*

A causal relationship requires that exposure to the allegedly harmful drug precede the adverse outcome; however, such a temporal relationship is not sufficient on its own to establish causation. *See Reference Manual,* at 601 (Ex. 14). "[W]hen latency periods are lengthy, variable, or not known and a substantial proportion of the disease is due to unknown causes, temporal relationship provides little beyond satisfying the requirement that cause precede effect." *Id.* at 602.

Here, the type of CV events at issue take time to develop because they result from subclinical injury to the arterial wall and plaque that builds up over time into a thrombus that either occludes an artery or embolizes to block blood flow and causes an adverse event. This

latency period weighs against drawing any reliable conclusions as to causation based on timing. Moreover, the apparently strong temporal evidence of associations among subgroups of short-term users is tempered by the fact that low testosterone levels are a known risk factor for CV events (Ex. 30 at 13-14) and TRT does not raise serum testosterone levels immediately and often requires dose adjustment over the course of weeks or months before testosterone levels reach the normal range (*see* French Dep. 277:19–279:4 (Ex. 9)). Thus, apparent short-term risk could be due to underlying hypogonadism unaffected by short-term TRT exposure, and therefore the temporal connection between TRT exposure and a CV event could be the result of confounding.

(iii) ***Dose-Response:*** *There Is No Evidence of Any Dose-Response Relationship*

"A dose–response relationship means that the greater the exposure, the greater the risk of disease." *Reference Manual*, at 603 (Ex. 14). Simple exposure to a substance is insufficient to support a causal relationship; instead, the dose at which the substance exhibits harmful effects must be established and there should be a direct relationship between increasing dose and the incidence or severity of the harm. This is particularly relevant to TRT because, as all human beings naturally produce some level of testosterone necessary for survival, it cannot be true that just *any* dose is toxic.

Scientific research on TRT has not produced any data on a dose-response relationship between adverse CV events and TRT generally or AndroGel specifically. The method of TRT administration is crucial to considering any dose-response relationship because transdermal, injection, and oral methods of administration deliver testosterone at different rates and have different pharmacokinetic impact.[10] No studies have shown that higher doses of testosterone gel

[10] *See* A. Dobs, *et al.*, *Pharmacokinetics, Efficacy, and Safety of a Permeation-Enhanced Testosterone Transdermal System in Comparison with Bi-Weekly Injections of Testosterone Enanthate for the*

within the therapeutic range are associated with greater incidence or severity of CV events than lower doses.[11] Plaintiffs' own experts do not contend otherwise. *See* Ardehali Dep. 109:9–110:12 (Ex. 7) (only evidence of dose was outside context of TRT); Cuculich Dep. 191:2–22 (Ex. 8) ("We just don't have any dose finding studies"); Halushka Dep. 220:19–221:11 (Ex. 21).

(iv) ***Replication:*** *No Associations Have Been Replicated*

"Rarely, if ever, does a single study persuasively demonstrate a cause–effect relationship." *Reference Manual,* at 164 (Ex. 14). Therefore, "[i]t is important that a study be replicated in different populations and by different investigators before a causal relationship is accepted by epidemiologists and other scientists." *Id.*

Here, there has been no replication of the studies reporting a statistically significant risk. As described above, Xu's meta-analysis finding of an increased risk for TRT users overall is effectively superseded by Albert's later meta-analysis. Gerstman Dep. 308:20–309:3 (Ex. 6); French Dep. 451:18–452:1 (Ex. 9). And Albert finds no statistically significant association for the same group. (Ex. 68 (Albert (2016)). Eight other meta-analyses report the same result as Albert. **Figure 3.** Finkle's finding of increased risk for a subgroup of short-term (90 days or less) TRT users under 65 and with a history of CV disease has never been replicated. Gerstman Dep. 310:18–311:4 (Ex. 6). The same lack of replication is true of Vigen's finding of increased

---

Treatment of Hypogonadal Men, J. Clin. Endocrinology & Metab. 84(10) 3469–78 (Ex. 35) (1999) (studying disparate pharmacokinetic responses in men given testosterone injections compared to transdermal testosterone).

[11] Plaintiffs' expert Dr. Ardehali relies on literature reporting adverse CV events in athletes who abuse testosterone supplements. Ardehali Dep. 109:9–110:12 (Ex. 7). Dr. Ardehali acknowledges that the doses are significantly higher than what TRT patients receive and are used under circumstances that constitute abuse. Ardehali Dep. 88:17–22, 110:2–12, 214:6–215:10 (Ex. 7); *see also* Cuculich Dep. 87:16–88:10 (Ex. 8) (explaining that steroid doses were "extremely high"). As importantly, these are only case reports, not epidemiological studies. And extrapolation is necessary in order for this evidence to be relevant to establishing a dose-response for TRT. But Dr. Ardehali does not even have a method for extrapolating from those high doses to TRT, (Ardehali Dep. 88:17–22, 110:2–12, 214:6–215:10 (Ex. 7)).

risk among men with pre-existing comorbidities, which Plaintiffs' expert Dr. Gerstman does not even rely on. *See* **Figure 3**; Gerstman Dep. 67:9–68:15 (Ex. 6).

Nor have the findings of increased CV risk in patients over 65 with short-term TRT use been replicated. As discussed above, *supra* Section B.2.c.i. ("Consistency and Strength"), while four studies have reported significant associations, they are too different from one another to be considered replicated. None of the authors conclude that they have replicated previous findings.

(v) ***Biological Plausibility:*** *No Significant TRT-Specific Research*

This criterion "depends on existing knowledge about the mechanisms by which the disease develops." *Reference Manual,* at 605 (Ex. 14). This element has been satisfied for the other causal risk factors for CV events. By contrast, the TRT literature has hypothesized but not researched a number of mechanisms by which TRT could cause CV events. *See, e.g.*, (Ex. 68) (Baillargeon (2014) (mentioning platelet aggregation as a proposed mechanism that has never been confirmed)).

Plaintiffs' experts advance four principal mechanisms (*see generally* Ardehali Rpt. at 67 (Ex. 3); Cuculich Rpt. at 9–16 (Ex. 2); Halushka Rpt. at 25–28 (Ex. 44)):

- **Increasing serum testosterone increases estradiol:** Estradiol is a form of estrogen, a female hormone, and a metabolite of testosterone. Estradiol is associated with CV events in women on HRT. None of the experts produce evidence linking elevated estradiol levels in men to clinical disease data. Rather, the HRT experience is said to be a "parallel" to TRT. Gerstman Rpt. at 24 (Ex. 10). Moreover, actual studies of estradiol given to men show no pattern of increased risk. Marais Rpt. at 166–171, Figure 11 (Ex. 16); *see also* Ardehali Dep. at 191:11–19 (Ex. 7); Cuculich Dep. 199:21–200:3 (Ex. 8); *infra* Section I.C.3.d. (regarding estradiol as proposed mechanism for VTEs).

- **Increased thromboxane:** Thromboxane A2 and its receptors are associated with platelet activation. Plaintiffs' experts opine that this increases the risk of arterial clots among those already at risk. Halushka Dep. 4:3–-10 (Ex. 21); Ardehali Dep. at 191:6–10 (Ex. 7); Cuculich Dep. at 199:2–10 (Ex. 8). Yet the one 1995 TRT study they cite that evaluated human subjects did not look for any association with CV

events and involved healthy young men with normal testosterone levels.[12] TXA2 is discussed further below in connection with VTE. *See infra* Section C.3.d.

- **Increased hematocrit:** That testosterone increases hematocrit is well known, but no study has ever demonstrated that hematocrit increases from gel TRT (which have a lower dose rate than injection) are statistically associated with CV events. Again, this is discussed further below in connection with VTE. *See infra* Section C.3.d.

- **Reactive Oxygen Species ("ROS"):** Like hematocrit, Plaintiffs' experts produce no evidence linking TRT to CV events via ROS in human men, and instead rely on dated studies of animals and isolated cells that were administered supraphysiological doses of testosterone to produce results. *See* Ardehali Dep. at 191:25–192:5 (Ex. 7); French Rpt. at 109-10 (Ex. 1).

3. Conclusion Reached by AbbVie's Experts Based Upon the Science

By referring to the methods used to establish the accepted CV causal risk factors, assessing whether TRT is a new causal risk factor is straightforward. The first step of the causal analysis is determining whether there is a reliable statistically significant association between TRT and CV events, specifically MI and stroke. Today, even after many studies have supplemented the evidence found by the FDA to be insufficient, no such reliable association has been shown for overall TRT users or for subpopulations of TRT users. This ends the causal analysis. Application of the other causal criteria does nothing other than to highlight the absence of causal evidence.

C. General Causation of VTEs

The scientific evidence relating to the generally accepted causes of VTE can be summarized in the same way: generally accepted risk factors have been identified as causes of VTEs through consistent, replicated findings of statistically significant associations from multiple studies. By contrast, the five epidemiological studies which have evaluated whether TRT is associated with an increased risk of VTE present no such picture.

[12] A.L. Ajayi, *et al.*, *Testosterone Increases Human Platelet Thromboxane A_2 Receptor Density and Aggregation Responses*, Circulation 91(11) 2742–47 (1995) (Ex. 36).

1. The Accepted Causal Risk Factors for VTEs

As with the causes of CV events, there are many accepted causal risk factors for VTE. Plaintiffs' sole general causation expert on clotting risk, Dr. Rinder, carefully distinguishes the risk factors that have been established scientifically from those that have not. There are inherited and acquired risk factors. The inherited risk factors include: Factor V Leiden mutation, prothrombin gene mutation, protein C and S deficiencies, antithombin3 deficiency, and elevated factor levels. Nand Rpt. at 9 (Ex. 4); Rinder Dep. 83:11–84:17 (Ex. 37). There are also acquired risk factors for VTE: previous thromboembolism, antiphospholipid antibodies, malignancy, recent major surgery, trauma, long period of immobilization, pregnancy, and polycythemia vera. Nand Rpt. at 12–16; Rinder Dep. 84:18–85:15 (Ex. 37).

Dr. Rinder acknowledges that all of these risk factors have been demonstrated through well-designed epidemiological studies which find statistically significant associations between the risk factor and increased incidence of VTE. Rinder Dep. 27:12–21, 92:8–23, 111:22–113:2, 133:9–134:2 (Ex. 37). In most cases, these findings have been replicated, the magnitude of each risk factor has been quantified, and strong evidence of biological mechanism has been developed for each risk factor. *Id*. 91:7–13, 94:24–95:22. Dr. Rinder distinguishes these risk factors from other, yet unproven VTE risk factors on two principal grounds: (1) the epidemiological studies for the unproven risk factors have greater uncertainty because their designs do not sufficiently isolate the effect of the risk factor at issue (for example, he finds this true of certain risk factors that have been established for cardiovascular disease but not VTE); and (2) the evidence of mechanism is wanting (e.g., inflammatory bowel disease, smoking, high cholesterol). *Id*. 120:13–124:6.

2. Science Regarding Whether TRT Can Cause VTEs

Comparing the science establishing the accepted risk factors for VTE to the ***science*** that Dr. Rinder relies upon to establish TRT as a ***new causal risk factor*** is, in his words, like comparing "apples and oranges." Rinder Dep. 269:17–270:6 (Ex. 37) ("Q. Give me another risk factor today where the epidemiology is that limited; just as we've talked about for TRT. . . A. I think it's apples and oranges. . . they're not comparable situations").

(a) VTE Research on Testosterone Through the FDA Investigation

It has been reported in the literature for decades that TRT, particularly injections, can stimulate the production of red blood cells, a condition known as erythropoeisis. This increase in red blood cells is measurable, and the results are reported as a percentage of hematocrit ("Hct"). The hematocrit percentage refers to the ratio of the volume of red blood cells to the total volume of blood. Increased red blood cell production, in turn, increases blood viscosity. Because elevated blood viscosity is one factor that can contribute to clotting, clotting has been reported as an adverse reaction to testosterone injections. Starting in 1981, the class label for all TRT said that patients should be monitored for increases of hematocrit while on TRT.

The mode of administration of TRT makes a difference. Injectable TRT introduces testosterone into the blood stream directly, and serum testosterone levels rise very quickly as a result. Transdermal TRT, such as patches and gels, deliver testosterone much more gradually and result in lower peak testosterone levels. This was verified in research published in 1999, before the introduction of AndroGel. Dobs, *Pharmcokinetics*, 3469–78 (Ex. 35). The same research found that transdermal TRT also reduced the effect on increases in hematocrit, as mentioned above.

The effect of AndroGel on hematocrit was specifically studied in the AndroGel 1% trial. Likewise, the approved label for AndroGel included the class labeling recommendation that doctors monitor hematocrit. AndroGel 1% Label, at 1, 4, Mar. 2000, (Ex. 38) (cautioning that androgens may "stimulate the production of red blood cells by stimulating erythropoietin production," and instructing physicians to monitor "[h]emoglobin and hematocrit levels. . . to detect polycythemia in patients on long-term androgen therapy."). In December 2007, the FDA added the following language to the label for AndroGel and other TRT: "Increases in hematocrit, reflective of increases in red blood cell mass, may require lowering or discontinuation of testosterone. Increase in red blood cell mass may increase the risk for a thromboembolic event." (Ex. 67 at 6).

As set forth in more detail in AbbVie's failure to warn brief, additional epidemiological data is found in post-marketing surveillance on AndroGel, which included reports of venous thromboembolism. *See* Arrowsmith Rpt. at 104–05 (Ex. 39). The one such adverse event reported in a clinical study of AndroGel 1% was also included in the label. AndroGel 1% Label, at 11, Aug. 2005 (Ex. 64) ("Two patients reported serious events considered possibly related to treatment: deep vein thrombosis (DVT) and prostate disorder requiring a transurethral resection of the prostate (TURP).").

Beginning in 2010, a series of case reports were published by a research group led by Dr. Glueck which studied patients at a single health care center who experienced VTEs while on TRT and also had a genetic pre-disposition to VTEs due to the Leiden Factor V mutation. (Ex. 69). This work never led to Dr. Glueck publishing a full epidemiological study; indeed, Dr. Glueck and colleagues referred to their work as an "exploratory, hypothesis-generating

study." (Ex. 69). These cases of VTE were reported by AbbVie to the FDA as part of its pharmacovigilance process. (Ex. 40 at 347–354, 358–367).

In 2014, the FDA advised AbbVie it had "become aware of new safety information related to the serious risk of venous thromboembolic events associated with testosterone use." (Ex. 65 at 1). The FDA originally proposed amending the label to state that TRT "may increase the risk" of VTEs. *Id*. In response, AbbVie conducted and submitted its analysis of the issue to the FDA, including an assessment of adverse event reports. After further consideration of the issue and AbbVie's analysis, the FDA rejected its own proposed language and agreed to add: "There have been postmarketing reports of venous thromboembolic events, including [DVT] and [PE], in patients using [TRT] such as AndroGel." AndroGel 1.62% Label, § 5.4, June 2014 (Ex. 48). The FDA explained that this revision was necessary "***to better reflect the level of uncertainty of drug causality***." (Ex. 49) (emphasis added).

(b) Epidemiological TRT Research on VTE Since the FDA Investigation

The most important development since the FDA added additional VTE warnings to the AndroGel label has been the publication of additional epidemiological studies. There are now five such studies. All of these studies are provided in Exhibit 69, and a summary chart is provided as **Figure 4**.

Three observational studies have found that TRT was not associated with increased risk for idiopathic VTE (meaning among individuals without histories of VTE risk factors). (Ex. 69). Sharma's observational study of over 70,000 men with low testosterone levels compared the incidence of VTEs in three groups: men who received TRT and achieved a normal testosterone level, men who received TRT and did not achieve a normal testosterone level, and men who did not receive TRT at all. Each group experienced nearly identical rates of VTEs, thus

demonstrating that there is no association between using TRT and increased VTE risk. Ex. 69 (Sharma (Sept. 2016)).

Li studied over 200,000 hypogonadal men, half of whom were on TRT, for the incidence of idiopathic VTE (meaning the subjects had no other VTE risk factors). This design permitted the Li authors to isolate the effect due to TRT alone, which mitigates against confounding from non-testosterone causes. *See* Baillargeon Dep. 425:6–426:12. The Li study concluded that there is no statistically significant association between VTE and TRT (OR 1.08, 95% CI = 0.91–1.27, p = 0.378). Ex. 69 (Li (Apr. 2016)).

Finally, Dr. Baillargeon's VTE study of over 30,000 middle-aged and older men also found no association between TRT and VTE. He and his colleagues identified 7,643 men who were diagnosed with DVT or PE, 158 of whom were exposed to some form of TRT (inferred from filing a TRT prescription) in 15 days before their VTE. No statistically significant association was found (OR 0.90, 95% CI = 0.73–1.12). This finding persisted when the exposure window (i.e. date of last prescription) was extended to 30 days and 60 days before the VTE, and was consistent across means of administration of testosterone (topical (OR 0.80, 95% CI = 0.61–10.4), transdermal (OR 0.91, 95% CI = 0.38–2.16), and intramuscular (OR 1.15, 95% CI = 0.80–1.64)). Ex. 69 (Baillargeon, *et al.* (Aug. 2015)).

Xu & Schooling (2015) was a meta-analysis performed using selected studies from the dataset Xu had earlier analyzed for CV effects and published. *See* Ex. 68 (Xu (2013)). The Xu & Schooling dataset included three prior RCTs which collectively reported only five VTEs among TRT users. Xu & Schooling used the same data to look for any VTE effects. None of the three RCTs individually found a statistically significant association. Combined, they examined a total of 283 men exposed to TRT and 233 controls, and reported a total of 5 adverse VTE events, all

in the testosterone groups. After statistically analyzing this data, the result was not statistically significant because the lower bound of the confidence interval was 1.0[13] Ex. 69 (Xu & Schooling Ltr. (2015)).

A recent paper by Martinez (2016) is regarded by Dr. Rinder as the most complete of these studies, **all** of which he still regards as "flawed." Rinder Dep. 187:23–190:2, 334:13–335:3 (Ex. 37); (Ex. 69). Martinez found no statistically significant risk of VTE for TRT users overall. The researchers then analyzed various subgroups. The authors reported low but statistically significant associations for short-term users (defined as less than six months) (1) without history of other VTE risk factors (RR 1.91 (95% CI = 1.13–3.23)) or (2) without "pathological" hypogonadism (RR 1.88 (95% CI = 1.02–3.45)).[14] The finding related to short-term users without VTE history was based on only seventeen men. Similarly, the second finding related to short-term users without pathological hypogonadism was based on only eleven men. It cannot be determined from the study to what extent the groups overlap or to what extent that overlap may have affected the statistical analysis. The authors could not explain their findings,

---

[13] As noted earlier, the authors then added to the existing data one more small RCT that did not describe adverse events in the placebo group but did report another single DVT in the TRT group. By adding the data, the additional DVT in the TRT group would clearly affect the statistics. And it did, producing statistical significance that did not exist before. Further, the three of the six total VTE events reported by the four RCTs were from the Copenhagen study, whose subjects all had cirrhosis of the liver and were administered extremely high doses of testosterone via an unapproved oral version of TRT. Dr. Marais performed a sensitivity analysis of the Xu letter, and concluded that its finding was due almost entirely to the particularities of the underlying studies and did not persist when other RCTs that would have been eligible for inclusion were added to the mix. *See* Marais Rpt. at 116–17 (Ex. 16); Bierer Rpt. at 7–8 (Ex. 42).

[14] Again, this is a subgroup analysis, where the ultimate finding is made for a specifically defined group within the larger group of patients. Thus, the authors here begin with patients overall (finding no statistically significant association), then stratify by current/recent users, and then by type of hypogonadism (finding a statistically significant association among all current TRT users without pathological hypogonadism). But, once the authors further stratified according to duration of use this finding persisted only among the current *short-term* users. The current users exposed for longer than six months were not tied to a statistically significant association. Thus, the short-term users drive the result for the larger group.

acknowledging that their study only "***suggests*** a *transient* increase in the risk of venous thromboembolism," and that "[f]urther research is needed to confirm this temporal increase in the risk." Ex. 69 (Martinez (2016) (emphasis supplied). The authors also recognized the limitations of their study, including the fact that all of their findings were based on only 69 men exposed to TRT and that residual confounding, hidden bias, or chance alone could be responsible for their observations. *Id*.

Additional limitations of this study have been noted by AbbVie's experts, including:

- **Tyranny of small numbers.** The "recent user" finding was based on only 36 VTE cases out of over 19,000 patients. Moreover, the difference between the groups with and without VTE risk factors was *two* individuals, meaning the result could have been the opposite if only two more people with VTE risk factors were found in the database. *See* Bierer Rpt. at 13–14 (Ex. 42).

- **Arbitrary definition of "recent" use.** The authors provided no reason for stratifying their data among men who used TRT for less than six months compared to six months and longer. This time period is greater than what has been used in other studies to evaluate short-term risk of TRT use. *See* Baillargeon Rpt. at 56 (Ex. 15).

- **Risk factor results inexplicable.** The relative risk was higher in recent users who did not have preexisting risk factors for VTE than those who did. This finding casts serious doubt on the reliability of the Martinez study, particularly because the authors and Plaintiffs' experts have failed to properly explain it. *See* Baillargeon Rpt. at 56–57 (Ex. 15).

- **Multiplicity of comparisons.** The subgroup finding was one of many post-hoc subgroup analyses performed by the Martinez authors. When multiple comparisons are performed on the same set of data, using the regular 95% level of evidence means that one out of every twenty comparisons will appear to be statistically significant but will, in fact, be due to chance. *See* Marais Rpt. at 16 (Ex. 16); Baillargeon Rpt. at 56–57 (Ex. 15).

3. Conclusion Reached by AbbVie's Experts

Based upon the methods and criteria which have been used to establish other casual risk factors, science does not establish TRT as a causal risk factor for VTE. As in the case of CV events, the causal analysis is simple: all five epidemiological studies consistently found no

statistically significant association between VTE and broad groups of TRT users, and one study reported associations in two highly-specified, related subgroups. The authors of that study try to explain these findings but do not come to any conclusion about it. They say no more than it "suggests" a transient increase in risk and needs to be confirmed by a RCT, which it has not. This makes further causal analysis unnecessary.

Nor would undertaking such analysis make a difference because there is no evidence to support the causal factors:

- Strength and consistency of relationship does not exist because no strong associations are reported;
- Temporality has little significance because VTE frequently results from underlying genetic predisposition and chronic conditions (e.g. alcohol abuse, smoking);
- Dose-response has never been studied, much less confirmed;
- No association between TRT and increased VTE risk in any group has been replicated; and
- While high levels of hematocrit have been linked to VTEs, no study ties increased hematocrit due to TRT to any disease outcome and thus biological plausibility in the context of TRT is lacking.

D. Analyzing the Science Regarding Specific Causation

General causation is a threshold requirement that Plaintiffs must satisfy before offering evidence of specific causation. "[A]n agent cannot be considered to cause the illness of a specific person unless it is recognized as a cause of that disease in general." *Reference Manual*, at 611 (Ex. 14) (citation omitted).

The results of general causation studies determine the causal risk factors for CV events and VTEs among various groups of people and can be applied to individuals who fall within those groups. Thus, the relative risk posed by each of these factors can be quantified and then applied to individuals to assist in determining the most likely causal factors for those events. Science on subpopulations works in a similar way: if epidemiological evidence has established that a risk factor can cause CV events or VTEs in a subgroup of people, then that science can be applied to determine whether that risk applies to an individual Plaintiff in that subgroup.

However, "[o]nly if the study subjects and the Plaintiff are similar with respect to other risk factors will a risk estimate from a study or studies be valid when applied to an individual." *Reference Manual,* at 613 (Ex. 14). For instance, "if all of the subjects in a study are participating because they were identified as having a family history of heart disease, the magnitude of risk found in a study of smoking on the risk of heart disease cannot validly be applied to an individual without such a family history." *Id.*[15]

[15] As explained further in detail below, *infra* Section D.1.b., the method for applying generally accepted and scientifically reliable risk factors established by epidemiological evidence to individuals is a form of differential diagnosis (also called differential etiology). *Reference Manual.* at 617 (Ex. 14). In a differential etiology, established causes and the agent in question must first be "ruled in" as potential causes and alternative causes must be "ruled out" as the triggers of an individual Plaintiff's injury. A differential etiology "is only valid if general causation exists and a substantial proportion of competing causes are known. . . . And, like any scientific methodology, it can be performed in an unreliable manner." *Id.* at 618 (Ex. 14).

ARGUMENT

I. THE COURT SHOULD EXCLUDE PLAINTIFFS' EXPERT TESTIMONY ON GENERAL AND SPECIFIC CAUSATION

A. Legal Standard

1. Law Governing the Admission of Expert Testimony Generally

Federal Rule of Evidence 702 sets forth the standard governing the admission of expert testimony. In particular, Rule 702 provides:

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based on sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

This rule incorporates principles established by the Supreme Court in *Daubert* v. *Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993). Significantly, *Daubert* was a pharmaceutical case. Plaintiffs there alleged that Bendectin, a prescription anti-nausea drug marketed by defendant, caused birth defects. *See id.* at 582. In that case, as here, the defendant moved for summary judgment and sought to exclude the proffered testimony of the Plaintiffs' experts on the issue of causation. *See id.* The district court granted the defendant's motion and the Ninth Circuit affirmed. *See id.* at 583-85. The Supreme Court "granted certiorari. . . in light of sharp divisions among the courts regarding the proper standard for the admission of expert testimony." *Id.* at 585. Indeed, the *Daubert* case was one of a series of cases in various trial and appellate courts involving the issue of whether Bendectin caused birth defects, and although Plaintiffs proffered similar evidence by many of the same experts in those various cases, there was disagreement among the courts as to whether such expert testimony was admissible. *See, e.g., Daubert* v. *Merrell Dow Pharm., Inc.*, 727 F. Supp. 570, 572-73 (S.D. Cal. 1989) (discussing the "two

schools of thought governing expert testimony in these Bendectin cases"). In at least one of those cases, Plaintiff was allowed to proceed to trial even in the absence of "a single original or recalculated epidemiological study which showed a statistically significant association between Bendectin and birth defects." *Id.* at 573 (citing *Oxendine* v. *Merrell Dow Pharm., Inc.*, 506 A.2d 1100 (D.C. App. 1986).

The Supreme Court's decision in *Daubert* effectively resolved the disagreement among the lower courts by imposing stringent reliability and relevance requirements in order for expert scientific evidence to be admissible, and by giving the judge the "gatekeeping" responsibility to decide whether proffered expert testimony meets these strict standards. *Daubert*, 509 U.S. at 589, 597 ("the trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable").

(a) The Exacting Standards of Reliability

"The reliability analysis [required by *Daubert*] applies to all aspects of an expert's testimony: the methodology, the facts underlying the expert's opinion, [and] the link between the facts and the conclusion." *ZF Meritor, LLC* v. *Eaton Corp.*, 696 F.3d 254, 291 (3d Cir. 2012) (internal quotation omitted). Accordingly, courts must exclude expert evidence that is not "the product of reliable principles and methods," that is not "based on sufficient facts or data," or that has not "reliably applied the principles and methods to the facts of the case." Fed. R. Evid. 702. "[A]ny step that renders the analysis unreliable under the *Daubert* factors renders the expert's testimony inadmissible. This is true whether the step completely changes a reliable methodology or merely misapplies that methodology." *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717,745 (3d Cir. 1994).

While the *Daubert* reliability inquiry is "a flexible one," the Supreme Court identified several factors that may assist a district court in determining whether an expert's testimony is

reliable. *Daubert*, 509 U.S. at 593-94. The first factor is whether "the theory. . . can be (and has been) tested." *Id.* at 593. The second factor is whether the theory has been subjected to evaluation by peer review and publication. *See id.* The third factor is the known or potential rate of error and the existence and maintenance of standards controlling the technique's operation. *See id.* at 594. And the fourth factor is whether the theory has been generally accepted in the scientific community. *See id.* These factors are "neither definitive nor exhaustive, however, and some factors may be more pertinent than others depending on the nature of the issue, the expert's particular expertise, and the subject of his testimony." *Newman* v. *Motorola, Inc.*, 218 F. Supp. 2d 769, 773 (D. Md. 2002), aff'd, 78 F. App'x 292 (4th Cir. 2003). *See Daubert*, 509 U.S. at 593–94. *Daubert*'s flexible four-pronged analysis supplanted the longstanding "austere" *Frye* standard, which allowed the admission of expert testimony when it was generally accepted in the relevant scientific community. *See Frye* v. *United States*, 293 F. 1013, 1013 (D.C. Cir. 1923). Since *Daubert*, "general acceptance" alone is not sufficient.

These "exacting standards of reliability," *Weisgram* v. *Marley Co.*, 528 U.S. 440, 455 (2000), provide that "conjecture, hypothesis, subjective belief, or unsupported speculation are impermissible grounds on which to base an expert opinion," *Wehling* v. *Sandoz Pharm. Corp.*, 162 F.3d 1158, 1998 WL 546097, at *5 (4th Cir. Aug. 20, 1998) (per curiam). Courts must ensure the expert "employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field." *Kumho Tire Co.* v. *Carmichael*, 526 U.S. 137, 152 (1999); *see also Braun* v. *Lorillard Inc.*, 84 F.3d 230, 234 (7th Cir. 1996). "[T]he courtroom is not the place for scientific guesswork, even of the inspired sort. Law lags science; it does not lead it." *Rosen* v. *Ciba-Geigy Corp.*, 78 F.3d 316, 319 (7th Cir. 1996).

Though *Daubert* addresses methods, "[a]s the Supreme Court has recognized, 'conclusions and methodology are not entirely distinct from one another,'" *Cooper* v. *Smith & Nephew, Inc.*, 259 F.3d 194, 202 (4th Cir. 2001) (quoting *Gen. Elec. Co.* v. *Joiner*, 522 U.S. 136, 146 (1997)), and the difference "has only limited practical import," *Paoli*, 35 F.3d at 746. "When a judge disagrees with the conclusions of an expert, it will generally be because he or she thinks that there is a mistake at some step in the investigative or reasoning process of that expert." *Id.* "[T]rial judges may evaluate the data offered to support an expert's bottom-line opinions to determine if that data provides adequate support to mark the expert's testimony as reliable." *Brown* v. *Nucor Corp.*, 785 F.3d 895, 936 (4th Cir. 2015).

"Because expert witnesses have the potential to be both powerful and quite misleading, it is crucial that the district court conduct a careful analysis into the reliability of the expert's proposed opinion." *United States* v. *Fultz*, 591 F. App'x. 226, 227 (4th Cir. 2015). "[C]lose judicial analysis of expert testimony is necessary 'because expert witnesses are not necessarily always unbiased scientists.'" *Nelson* v. *Tenn. Gas Pipeline Co.*, 243 F.3d 244, 252 (6th Cir. 2001) (quoting *Turpin* v. *Merrell Dow Pharms., Inc.*, 959 F.2d 1349, 1352 (6th Cir. 1992)).

(b) Relevance and "Fit"

The court also "must determine whether the evidence or testimony assists the trier of fact in understanding the evidence or in determining a fact in issue." *Cummins* v. *Lyle Indus.*, 93 F.3d 362, 370 (7th Cir. 1996) (citing *Daubert*, 590 U.S. at 592). "Stated differently, the scientific testimony must 'fit' the issue to which the expert is testifying." *Id.* As the Supreme Court held in *Daubert*, "scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes." 509 U.S. at 591. "Rule 702's 'helpfulness' standard requires a valid scientific connection to the pertinent inquiry as a precondition to admissibility." *Id.*

A district court therefore is not required "to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert." *Joiner*, 522 U.S. at 146. In those circumstances, the "court may conclude that there is simply too great an analytical gap between the data and the opinion proffered." *Id.*

These "stricter" and "more stringent" standards of reliability and fit "are necessary because of the potential impact on the jury of expert testimony." *Allison* v. *McGhan Med. Corp.*, 184 F.3d 1300, 1310 (11th Cir. 1999); *see also Daubert* v. *Merrell Dow Pharm., Inc.*, 43 F.3d 1311, n.17 (9th Cir. 1995) ("In elucidating the 'fit' requirement, the Supreme Court noted that scientific expert testimony carries special dangers to the fact-finding process because it 'can be both powerful and quite misleading because of the difficulty in evaluating it.'") (quoting *Daubert*, 590 U.S. at 595 (internal quotation omitted)).

(c) Plaintiffs Bear the Burden of Proof

Plaintiffs, as the proponents of the expert testimony, bear the burden of showing by a preponderance of the evidence that it is admissible. *Lewis* v. *CITGO Petroleum Corp.*, 561 F.3d 698, 705 (7th Cir. 2009) (citing Fed. R. Evid. 702 advisory committee's note 2000 amends. ("[T]he admissibility of all expert testimony is governed by the principles of Rule 104(a). Under that Rule, the proponent has the burden of establishing that the pertinent admissibility requirements are met by a preponderance of the evidence.")). AbbVie does not bear the burden of demonstrating its inadmissibility. *See Soldo* v. *Sandoz Pharms. Corp.*, 244 F. Supp. 2d 434, 534 (W.D. Pa. 2003).

2. Criteria for Association and Causation in Pharmaceutical Cases

Pharmaceutical cases have extensively explored the use of science in connection with *Daubert* analyses on association and causation. Earlier sections of this brief have reviewed the roles of association and causal criteria in scientific research (and as explained in the Reference

Manual on Scientific Evidence). The following demonstrates that courts have rigorously applied these scientific criteria in *Daubert* rulings.

"[T]o carry the burden of proving a Plaintiff's injury was caused by exposure to a specified substance, the Plaintiff must demonstrate the levels of exposure that are hazardous to human beings generally as well as Plaintiff's actual level of exposure." *Zellers* v. *NexTech Ne., LLC*, 533 F. App'x 192, 196 (4th Cir. 2013). "These two levels of causation are known as 'general causation' and 'specific causation.'" *Id.* at 196 n.6.

Statistically Significant Association. To establish general causation in a pharmaceutical products liability case, "the generally accepted method" begins, but does not end, by "look[ing] for statistically significant associations between medication exposure and [the injury], which are consistent and replicated across epidemiological studies." *In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 26 F. Supp. 3d 449, 455 (E.D. Pa. 2014); *see also In re Lipitor (Atorvastatin Calcium) Mktg., Sales Practices & Prod. Liab. Litig.*, 174 F. Supp. 3d 911, 914 (D.S.C. 2016) ("Epidemiology provides the primary generally accepted methodology for demonstrating a causal relation between a chemical compound and a set of symptoms or disease." (internal quotation marks omitted)). As explained above, epidemiological studies report this association as a ratio of the incidence of adverse events observed in the exposure and placebo groups, and statistical significance testing measures the likelihood that any observed association—the relative risk or the odds ratio—is due to chance by calculating p-values and confidence intervals consistent with a 95% level of confidence. *See supra* Section B.1.

Reliability. If a study reports a statistically significant association, the next step of the methodology is to evaluate whether the result is reliable. *In re Zoloft*, 26 F. Supp. 3d 449, 454 (E.D. Pa. 2014) (explaining that even when a statistically significant result is found, scientists

"will not draw firm conclusions from a single study, as apparent associations may reflect flaws in methodology, including multiple comparisons, bias, or confounding."). This analysis considers whether the observed association could be due to anything other than exposure to the medication being studied, such as bias, confounding, or some other flaw in the design of the study. *See Reference Manual,* at 583-96. Bias refers to systematic, non-random error—for example, information bias occurs where the available records for one group are more likely to include relevant information than the records for another. *See id.* at 585. Confounding refers to related factors that may be the true or contributory cause of an observed association—for example, if the group taking the study medication has risk factors that may also be the cause of the disease being studied. *See id.* at 591.

Application of Causation Methodology. Once a reliable statistically significant association has been shown, the final step in this accepted methodology is to evaluate other causal criteria, such as the Bradford-Hill criteria or the elements outlined by the Surgeon General's 1964 report, to determine whether the association rises to the level of a causal relationship. *See, e.g., In re Lipitor*, 174 F. Supp. at 925-26 ("Courts exclude expert testimony that attempts to start at step two, applying the Bradford Hill criteria without adequate evidence of an association."); *Mathews* v. *Novartis Pharm. Corp.*, No. 3:12–CV–314, 2013 WL 5780415, at *27 (S.D. Ohio Oct. 25, 2013) ("Unless there is a statistically significant association between the drug and the disease, the Bradford-Hill analysis to determine causation is inapplicable."); *Reference Manual,* at 598–99 (emphasizing that the Bradford-Hill factors "are employed only after a study finds an association to determine whether that association reflects a true causal relationship." (emphasis in original)). As noted above, these factors include, *inter alia*, the consistency and strength of the observed association, temporality, dose-response relationship,

replication, biological plausibility, alternative explanations, and coherence. *See* Hill, *Environment and Disease* (Ex. 18); *supra* Section B.2.c.

While the ultimate determination of "[w]hether an established association is causal is a matter of scientific judgment," *Milward* v. *Acuity Specialty Products Grp., Inc.,* 639 F.3d 11, 18 (1st Cir. 2011), courts routinely exclude experts who have failed to produce reliable evidence supporting these causal factors. For example, experts have been excluded for lacking evidence of a strong association (*see, e.g.*, *Daubert*, 43 F.3d at 1320-21 (requiring epidemiological data to report sufficiently strong association with relative risk exceeding 2.0 to "to show causation under a preponderance standard"); *Allison,* 184 F.3d at 1315 n. 16 (11th Cir. 1999) (relative risk greater than 2.0 allows inference that agent caused illness; risk of 1.24 insufficient)), temporality (*see Moore* v. *Ashland Chem. Inc.*, 151 F.3d 269, 278 (5th Cir. 1998) ("In the absence of an established scientific connection between exposure and illness, or compelling circumstances. . ., the temporal connection between exposure to chemicals and an onset of symptoms, standing alone, is entitled to little weight in determining causation.")), and a dose-response relationship (*In re Prempro Prods. Liab. Litig.,* 738 F. Supp. 2d 887, 895 (E.D. Ark. 2010) (rejecting Plaintiffs' experts' reliance "on a study of a high dose to determine adverse effects of a lower dose, without supplying a method to substantiate this inference."); *In re Rezulin*, 369 F. Supp. 2d 398, 427 (S.D.N.Y. 2005) (excluding experts in part because they "admitted that they have no information on whether therapeutic doses" of the defendant's drug could cause the injury alleged by Plaintiffs)).

3. Further Guidance from the Pharmaceutical Cases

The results reported in significant decisions in pharmaceutical cases reflect three basic principles: (1) *Daubert* motions are routinely granted when experts fail to satisfy the requirements association and (subsequently) causation; (2) experts must account for the total

body of evidence available, particularly when the majority of that evidence contradicts their opinions; and (3) the evidence underlying experts' opinions must be scrutinized to ensure that it matches the expert's conclusions, is not cherry-picked from the broader universe, and is relevant to the issue under consideration.

(a) Demonstration of a Reliable, Statistically Significant Association Is a Threshold Requirement

In pharmaceutical cases, many courts have excluded expert testimony offered by Plaintiffs to prove general causation on the ground that the expert, by failing to produce reliable evidence showing that the defendant's drug was statistically significantly associated with the Plaintiff's injury, did not follow the scientific methodology explained above:

- In *Norris* v. *Baxter Healthcare Corp.*, 397 F.3d 878, 885–87 (10th Cir. 2005), the Tenth Circuit affirmed the district court's decision to exclude Plaintiff's expert testimony on general causation where Plaintiff's experts failed to show "a consistent, statistically significant association between breast implants and systemic disease," and instead "relied solely on differential diagnosis and case studies [instead of epidemiological evidence] to support their belief that silicone gel breast implants can cause systemic disease."

- In *In re Lipitor (Atorvastatin Calcium) Mktg., Sales Practices & Prod. Liab. Litig.*, 174 F. Supp. 3d 911, 926 (D.S.C. 2016), the district court excluded the testimony of Plaintiffs' expert that lower doses of defendant's drug could cause diabetes where he failed to produce any data demonstrating a statistically significant association between those doses and increased risk of diabetes, and construed his admission that "a lack of statistical significance means that either a study has 'low power' or 'no risk exists'" as "demonstrat[ing] that studies without statistical significance are insufficient to support a causation opinion."

- In a two decisions involving prescription antidepressant Zoloft, the district court excluded Plaintiffs' expert testimony on general causation because, among other things, Plaintiffs' expert failed to explain why the studies on which he relied reflected true associations between maternal Zoloft use and cardiac birth defects, and were not the result of methodological flaws (*see In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, MDL No. 2342, 2015 WL 7776911, at *16 (E.D. Pa. Dec. 2, 2015)), and Plaintiffs' epidemiology expert attempted to rely on non-statistically significant trends, instead of replicated statistically significant associations, which was not accepted within her field (*see In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 26 F. Supp. 3d 449, 456, 460–61 (E.D. Pa. 2014)).

- In *Frischhertz* v. *SmithKline Beecham Corp.,* No. CIV.A. 10–2125, 2012 WL 6697124, at *3 (E.D. La. Dec. 21, 2012), the court excluded an expert's general causation opinion as unreliable on the ground that, "[b]ecause there [wa]s no data showing an association between Paxil and limb defects, no association existed for Dr. Goldstein to apply the Bradford–Hill criteria."

- In *Soldo* v. *Sandoz Pharm. Corp.*, 244 F.Supp.2d 434, 569 (W.D. Pa.2003), the district court excluded the general causation opinions of Plaintiffs' experts that relied on the Bradford Hill criteria because they "ha[d] not demonstrated any statistically significant epidemiologic study showing an increased risk of postpartum stroke in women using Parlodel, [and thus] application of the Bradford–Hill criteria [wa]s unwarranted."

- Finally, in *Happel* v. *Walmart Stores, Inc.*, 602 F.3d 820 (7th Cir. 2010), the Seventh Circuit affirmed the district court's decision to exclude the testimony of Plaintiff's causation expert who, *inter alia*, relied on articles that "stop[ped] short of reaching the same conclusion" as the experts that the observed data supported a causal relationship on the ground that "[i]t is axiomatic that causation testimony is inadmissible if an expert relies upon studies for publications, the authors of which were themselves unwilling to conclude that causation had been proven." *Id.* at 826 (quoting *Huss* v. *Gayden*, 571 F.3d 442, 459 (5th Cir. 2009)).

(b) Experts Must Reconcile Their Opinions with the Larger Universe of Epidemiological Evidence

Many courts have also excluded expert testimony from pharmaceutical cases where the expert not reconciled their opinions with the larger body of evidence contrary to their opinion:

- In *Norris* v. *Baxter Healthcare Corp.*, 397 F.3d 878, 885–87 (10th Cir. 2005), the Tenth Circuit affirmed the district court's decision to exclude Plaintiff's expert testimony on general causation in part because Plaintiff's experts failed to reconcile their opinions with "the volume of contrary medical opinion."

- In *Allison* v. *McGhan Med. Corp.*, 184 F.3d 1300, 1315 (11th Cir. 1999), the Eleventh Circuit affirmed the district court's decision to exclude Plaintiff's expert testimony on general causation because, among other things, the four studies on which Plaintiff's expert relied "were in direct contrast to over twenty other epidemiological studies which found no statistical correlation between silicone breast implants and systemic disease, strong evidence that a consensus exists in the general scientific community that no correlation exists."

- In *Daubert* v. *Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1314, 1320–22 (9th Cir. 1995), on remand from the Supreme Court, the Ninth Circuit affirmed the district court's decision to exclude Plaintiff's expert testimony on general causation and grant summary judgment because the large body of science did not support a statistically significant association between Bendectin and birth defects.

- In *In re Lipitor (Atorvastatin Calcium) Mktg., Sales Practices & Prod. Liab. Litig.*, 174 F. Supp. 3d 911, 929–32 (D.S.C. 2016), the district court excluded Plaintiffs' expert testimony on general causation in part because the experts failed to account for evidence that contradicted their opinions that Lipitor could cause diabetes.

- In a series of decisions involving prescription antidepressant Zoloft, the district court excluded Plaintiffs' expert testimony on general causation because, among other things: Plaintiffs' expert failed to reconcile his opinion with other studies showing no increased risk of cardiac birth defects (*see In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, MDL No. 2342, 2015 WL 7776911, at *16 (E.D. Pa. Dec. 2, 2015)); Plaintiffs' experts failed to reconcile inconsistent epidemiological research with their opinions that Zoloft causes birth defects (*see In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 26 F. Supp. 3d 466, 475–476 (E.D. Pa. 2014)); and Plaintiffs' expert failed to account adequately for contrary epidemiological evidence (*see In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 26 F. Supp. 3d 449, 456, 460–61 (E.D. Pa. 2014)).

- In *In re Rezulin Prod. Liab. Litig.*, 369 F. Supp. 2d 398 (S.D.N.Y. 2005), the district court excluded Plaintiffs' expert testimony that Rezulin causes liver damage because, *inter alia*, Plaintiffs' experts effectively ignored the larger body of contrary evidence. *See id.* at 425–426.

(c) Experts Cannot Cherry-Pick Studies That Support Their Opinions or Rely on Irrelevant Data That Does Not Involve the Drug at Issue

Finally, general causation experts cannot selectively rely only on those studies that support their opinions. Nor can they use those studies to reach stronger conclusions than those stated by the study authors themselves. Relatedly, the data underlying the experts' opinions must pertain to the drug and injury under consideration in order for their opinion to be considered reliable. This latter principle also overlaps with *Daubert*'s "fit" requirement. Courts have excluded experts for using these flawed techniques in other pharmaceutical cases:

- In *Daubert* v. *Merrell Dow Pharm., Inc.*, 43 F.3d 1311 (9th Cir. 1995), on remand from the Supreme Court, the Ninth Circuit affirmed the district court's decision to exclude Plaintiff's expert testimony on general causation and grant summary judgment in part because Plaintiffs' experts sought to extrapolate a causal link between Bendectin and birth defects based on animal studies and an analysis of other drugs with a similar chemical structure, which the court held failed the Supreme Court's "fit" requirement. *See id.* at 1314, 1320–22.

- Similarly, in *Caraker* v. *Sandoz Pharms. Corp.*, 188 F. Supp. 2d 1026, 1031–1040 (S.D. Ill. 2001), the district court excluded expert testimony that Parlodel could cause ICH because the experts' conclusion "require[d] too many extrapolations from dissimilar data, too many analytical leaps and involve[d] a loose application of purportedly objective scientific causation standards."

- In *In re Rezulin Prod. Liab. Litig.*, 369 F. Supp. 2d 398 (S.D.N.Y. 2005), the district court excluded Plaintiffs' expert testimony that Rezulin causes liver damage because, *inter alia*, Plaintiffs' experts selectively chose their support from a large body of scientific evidence. *See id.* at 425–426.

- In *Soldo* v. *Sandoz Pharms. Corp.*, 244 F. Supp. 2d 434, 542, 546–50, 567–72 (W.D. Pa. 2003), the district court excluded testimony that Parlodel could cause intracerebral hemorrhage ("ICH") because the testimony, which was based on anecdotal case reports, animal studies, other drugs, and studies of patients with pathologies other than ICH, flunked all of the *Daubert* criteria.

- In *Hall* v. *Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1407–11 (D. Or. 1996), the district court excluding testimony of Plaintiffs' general causation experts that silicone is capable of causing certain symptoms because the experts were making "too great a leap of faith;" animal studies could not be reliably extrapolated to humans without explanation, case reports were insufficient to prove causation, and studies of crystalline silica were irrelevant because Plaintiffs had not shown that silicone breast implants were associated with the presence of crystalline silica.

Here, as discussed below, the testimony of all of Plaintiffs' general causation experts fails to meet all of these standards, and therefore their opinions should be excluded.

B. General Causation of CV Events: Plaintiffs' General CV Causation Evidence Does Not Meet the Stringent *Daubert* Standards

In line with the *Daubert* decisions discussed above, Defendants below answer three questions regarding the admissibility of Plaintiffs' experts' general CV causation testimony:

- Do they show that the threshold requirement of a statistically significant association has been met by well-designed epidemiological research and, additionally, that casual criteria are met?

- Do they reconcile their opinions with the body of epidemiological evidence, particularly the majority of studies that contradict their view?

- Do they use reliable evidence that reports conclusions commensurate with their opinions, is not cherry-picked, and is related to TRT and the CV events at issue?

1. Plaintiffs' Experts Disavow the Established Methods for Demonstrating a Statistically Significant Association and Fail To Satisfy The Additional Criteria of Causation

Plaintiffs CV causation experts *first* dismiss the requirement of establishing reliable proof of a statistically significant association, which should end the causal analysis. *Second*, they fail to produce reliable evidence demonstrating that any of the additional causal criteria are satisfied in the context of TRT and CV risk. *Third*, the methodologies that Plaintiffs' experts fall back on are unreliable and do not pass scrutiny under *Daubert* and Rule 702.

(a) Plaintiffs' Experts Outright Reject The Established Methods for Demonstrating Statistically Reliable Associations and Do Not Demonstrate Such an Association Between AndroGel and CV Events

The threshold issue in any causation analysis is whether there is a "statistically significant association[ ] between medication exposure and [the injury], which [is] consistent and replicated across epidemiological studies." *In re Zoloft*, 26 F. Supp. 3d 449, 455. As described above, all of the accepted causal risk factors for CV events were established using this same methodology. Reliable epidemiological studies reported a statistically significant association, later reliable studies replicated the association, and only then did scientists evaluate the association to determine whether it rose to the level of causation. Cuculich Dep. 56:3–11, 71:16–72:10 (Ex. 8); Ardehali Dep. 101:10–102:7 (Ex. 7); Gerstman Dep. 274:17–275:7 7 (Ex. 6); *see generally* French Rpt. at 10–18 (Ex. 1). Plaintiffs' experts have provided no scientific basis for failing to follow the same approach in determining whether AndroGel is a causal risk factor for CV events. They instead outright reject statistical significance as a requirement for finding an association.

In particular, Dr. Gerstman testifies that he believes the standard statistical method of relying on p-values less than 0.05 and 95% confidence intervals is misused and improperly

excludes important results.[16] (Ex. 43 at 12-13); Gerstman Dep. 93:17–94:5; 394:2–16 (Ex. 6). Indeed, he even fights the validity of what existing research has found using the established tests. *E.g.*, *id.* at 144:12–17 (refusing to acknowledge that Albert (2016) study had statistically insignificant results, and instead contending that they are "statistically meaningful" if the 95% level of evidence is abandoned by increasing the acceptable alpha threshold), 254:25-255:6 (admitting that Albert (2016) finding was statistically insignificant at a 95% level of evidence only after "ask[ing] the founding fathers of [his] discipline to forgive" him). But Dr. Gerstman's position in this litigation contradicts the accepted standards in his field and his practice as an epidemiologist, as he uses traditional significance testing in his publications and his opening expert report in this case. *Id.* 120:18–121:3; *see generally* Gerstman Rpt. (including and analyzing confidence intervals and p-values reported by TRT studies) (Ex. 6). The same is true of Plaintiffs' other general causation experts who reject the importance of statistical significance. *See, e.g.*, Cuculich Dep. 131:23–133:24 (Ex. 8); Rinder Dep. 25:15–26:13 (Ex. 37).

Because these experts reject a necessary step of the accepted methodology for determining general causation from epidemiological evidence and because this rejection is contrary to the practices of these experts outside of this litigation, the Court should exclude their opinions as unreliable. *See Kumho Tire Co.* v. *Carmichael*, 526 U.S. 137, 152 (1999) (requiring

[16] The statistical community is currently debating whether scientists who lack expertise in statistics misunderstand p-values and overvalue significance testing. *See, e.g.*, R. Wasserstein & N. Lazar, *The ASA's Statement on p-Values: Context, Process, and Purpose*, The American Statistician, 70:2 192-33 (2016). The fact that there is a debate among professional statisticians on this narrow issue does not validate Dr. Gerstman's rejection of the importance of statistical significance testing, or undermine Defendants' reliance on accepted methods for determining association and causation. *See* Marais Dep. 462:2-463:3. Regardless of this debate, courts have routinely found the traditional epidemiological method—including bedrock principles of significance testing—to be the most reliable and accepted way to establish general causation. *See, e.g.*, *In re Zoloft*, 26 F. Supp. 3d 449, 455; *see also Rosen* v. *Ciba-Geigy Corp.*, 78 F.3d 316, 319 (7th Cir. 1996) ("The law lags science; it does not lead it.").

that expert "employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field."); *Norris*, 397 F.3d at 885-87; *In re Lipitor*, 174 F. Supp. 3d at 926; *In re Zoloft*, 26 F. Supp. 3d 449, 456, 460-61.

Plaintiffs' experts' failure to establish a statistically significant association ends the analysis of causation. *See, e.g., In re Lipitor*, 174 F. Supp. at 925-26 ("Courts exclude expert testimony that attempts to start at step two, applying the Bradford Hill criteria without adequate evidence of an association."); *Mathews* v. *Novartis Pharm. Corp.*, No. 3:12–CV–314, 2013 WL 5780415, at *27 (S.D. Ohio Oct. 25, 2013) ("Unless there is a statistically significant association between the drug and the disease, the Bradford-Hill analysis to determine causation is inapplicable."); *Frischhertz* v. *SmithKline Beecham Corp.*, No. CIV.A. 10–2125, 2012 WL 6697124, at *3 (E.D. La. Dec. 21, 2012), (excluding general causation expert opinion on the ground that, "[b]ecause there [wa]s no data showing an association between Paxil and limb defects, no association existed for Dr. Goldstein to apply the Bradford–Hill criteria."); *Soldo* v. *Sandoz Pharm. Corp.*, 244 F. Supp. 2d 434, 569 (W.D. Pa. 2003) (same); *Reference Manual,* at 598–99 (emphasizing that the Bradford-Hill factors "are employed only after a study finds an association to determine whether that association reflects a true causal relationship." (emphasis in original)).

(b) Plaintiffs' Experts Fail to Systematically Apply the Scientific Criteria on Causation

Plaintiffs' expert testimony also should be excluded because they fail to properly apply the causal criteria that at least some of them concede must be met. Gerstman Rpt. at 88–91 (Ex. 10) (discussing "global principles" of causation based on the Surgeon General's report); Cuculich Rpt. at 19 (Ex. 2). In reaching their conclusions that AndroGel causes CV events, none of Plaintiffs' experts fully analyze these causal criteria:

- Dr. Gerstman purports to offer testimony on only one of the Bradford-Hill factors (strength of the association) and fails to apply any of the others.
- Dr. Ardehali purports to offer testimony on three of the Bradford-Hill factors (temporality, dose-response relationship and biological plausibility) but not the others.
- Dr. Cuculich purports to offer testimony on two of the Bradford-Hill factors (biological plausibility and temporality) but not the others.
- Dr. Wells does not purport to offer testimony on any of the Bradford-Hill factors.

None of Plaintiffs' experts can actually meet their *Daubert* burden on any of the specific causal criteria.

Strength and Consistency of Association. "For an epidemiological study to show causation under a preponderance standard, the relative risk . . . arising from the epidemiological data will, at a minimum, have to exceed 2." *Daubert*, 43 F.3d at 1320-21; *see also Allison,* 184 F.3d at 1315 n. 16 (11th Cir. 1999) (relative risk greater than 2.0 allows inference that agent caused illness; risk of 1.24 insufficient); *Pozefsky* v. *Baxter Healthcare Corp.,* No. 92–314, 2001 WL 967608 (N.D.N.Y. Aug. 16, 2001) (breast implant Plaintiff must show relative risk greater than 2.0); *Hall* v. *Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1403 (D. Or. 1996) (same).

As explained above, most of the studies that find any statistically significant association between TRT and increased CV risk report relatively low point estimates. *See supra* Section B.2.c.i. (noting Xu (RR 1.54), Vigen (1.29), Etminan (RR among first-time users: 1.41) and Wallis (HR among short-term users: 1.26)); *see also* App'x A at A-131–47, A-189–99, A-388–97, A-526–36. As also explained above, the studies that do report a relative risk above 2.0 (Finkle (2014) and Albert (2016)) are very limited because they are the result of *post-hoc* analyses of subgroups that are statistically questionable and clearly inapplicable to the entire population of TRT users. *See supra* Section B.2.b.; App'x A at A-562–72.

Plaintiffs' solution is to rely on findings that were not statistically significant, *see, e.g.*, Gerstman Dep. 323:9–23(Ex. 6) (testifying that non-statistically significant findings, while uncertain, "are more consistent with an elevation in risk than not"); *see also* Ardehali Dep. 69:14–70:1 (Ex. 7) (relying on the "totality of the evidence" or a "combination of the evidence that is out there"); Cuculich Dep. 129:15-130:3 (Ex. 8) (relying on the "totality" of the evidence), but these cannot satisfy the criteria of a strong and consistent association.

Replication of Findings. Statistically significant associations must be "replicated across epidemiological studies" before they can qualify as evidence of a causal relationship. *In re Zoloft,* 26 F. Supp. 3d 449, 455 (E.D. Pa. 2014).

In this case, no study has ever replicated the Xu authors' finding for the overall population of men exposed to TRT. Indeed, nearly every study since Xu has come to the opposite conclusion. Further, of all of the subgroups of TRT users, more than one study has reported a statistically significant association only for one sub-sub group for "short term" use in men over 65, *see* Marais Rpt. at 127 Fig. 19, App'x F, App'x G (Ex. 16); French Dep. 442:16–443:16, 444:21–445:7 (Ex. 9); Gerstman Dep. 164:19–165:25, 254:9–255:6, 256:16–24 (Ex. 6). And these studies cannot be considered replications because they are not consistent and the authors themselves do not make any claim of replication. *See supra* Section B.2.c.i., iv. Plaintiffs' experts' failure to produce evidence that any statistically significant association has been replicated prevents them from satisfying this element that is necessary to establish a causal relationship between AndroGel and CV events, making their opinions on general causation unreliable under *Daubert*.

Temporal Relationship. Plaintiffs' experts essentially contend that, because some individuals suffered MIs or strokes after taking TRT, TRT must have been the cause. This type

of reasoning is not sufficient to support a causal inference. *See Glastetter* v. *Novartis Pharm. Corp.*, 252 F.3d 986, 990 (8th Cir. 2001) (holding that evidence "demonstrat[ing] a temporal association between [drug] and stroke, or stroke-precursors," was "not scientifically valid proof of causation."). Moreover, Plaintiffs' experts' reasoning is undermined by the nature of the CV events at issue in this case, which take time to develop, the nature of hypogonadism being related to CV events, and the nature of TRT as requiring time to become effective in raising serum testosterone levels to offset the underlying CV risk from hypogonadism. Plaintiffs' experts' failure to address this latency period and sources of confounding is another example of their inability to satisfy the criteria necessary to demonstrate causation, and therefore their general causation opinions should be excluded. *See Moore* v. *Ashland Chem. Inc.*, 151 F.3d 269, 278 (5th Cir. 1998) ("In the absence of an established scientific connection between exposure and illness, or compelling circumstances. . ., the temporal connection between exposure to chemicals and an onset of symptoms, standing alone, is entitled to little weight in determining causation.").

Dose-Response Relationship. Plaintiffs' experts cannot produce evidence of a dose-response relationship between TRT and increased incidence or severity of CV events because there is no published study establishing such a relationship. *See* Cuculich Dep. 191:2–10, 192:8–14 (Ex. 8) (testifying that he has no evidence of dose-response and that his opinion does not depend on dose); *see also* Baillargeon Rpt. at 63 (Ex. 15). Nor have they produced any evidence that establishes a dose-response relationship between CV events and gels, specifically, which is necessary due to the different pharmacokinetic reactions to the various methods of TRT administration. *E.g.,* Ardehali Dep. 109:9–110:12 (Ex. 7) (testifying that only evidence of dose was outside context of TRT); Cuculich Dep. 191:2–22 ("We just don't have any dose finding studies"), (Ex. 8); Gerstman Dep. 136:24–137:4 (Ex. 6) (doubting that there is sufficient

evidence to distinguish between gels and injections); Halushka Dep. 182:4–185:22 (testifying that he has not studied the pharmacokinetic differences between testosterone injections and gels and has not evaluated whether his injection study is applicable to AndroGel), 220:19–221:11 (Ex. 21). Thus, this is another element of a causal relationship that Plaintiffs' experts have failed to fulfill, and for this reason their opinions that AndroGel causes CV events should be excluded. *Cf. In re Lipitor* 174 F. Supp. 3d at 926 (excluding experts who lacked evidence that adverse event occurred at when exposed to each of the various doses of defendant's drug); *In re Rezulin*, 369 F. Supp. 2d at 426 (excluding experts who admitted they had no evidence that the Plaintiffs' alleged injuries could occur at therapeutic doses of the defendant's drug).[17]

Biological Plausibility/Mechanism. As described above, *supra* Section C.2.c.v., Plaintiffs' experts rely on evidence that is old and largely from studies outside the context of TRT.

(c) The Alternative Approaches Used by Plaintiffs' Experts Are Not Reliable Methodologies and Are Not Helpful to the Jury

Instead of following the basic generally accepted methods, Plaintiffs' experts proffer substitutes such as "considering the totality of the evidence" or "risk benefit analysis." *See, e.g.*, Ardehali Rpt. at 9 (Ex. 3); Cuculich Dep. 70:11–14, 128:25–130:3 (Ex. 8); Gerstman Dep. 76:19–21 (Ex. 6). For the reasons explained below, none of these alternative methods reliably address the question of whether AndroGel can cause CV events, and they provide additional grounds for excluding the opinions of Plaintiffs' experts.

[17] Instead, Plaintiffs' experts attempt to satisfy this criterion by relying on evidence of different drugs or studies conducted on animals or cells. For the reasons explained in §§ I.B.3.c-d., *infra*, Plaintiffs' experts' failure to explain how that evidence can be extrapolated to the population of TRT users makes their opinions unreliable and inadmissible under *Daubert*.

(i) *Bayesian Statistical Methods*

Dr. Gerstman purports to rely on Bayesian statistical methods. Gerstman Response Rpt. at 12–13 (Ex. 43); Gerstman Dep. 93:17–94:5; 394:2–16 (Ex. 6). In contrast to traditional statistical significance testing, Bayesian statistical methods evaluate the "posterior" probability of a hypothesis being true based on the effect of data on subjective underlying assumptions. *See Reference Manual,* at 529 (Ex. 14) ("The frequentist view of statistics is more conventional; subjective Bayesians are a well–established minority."). These underlying assumptions mean that Bayesian computations are "necessarily subjective." *Id.*

Dr. Gerstman's rejection of significance testing in favor of Bayesian techniques is deeply self–conflicted. Dr. Gerstman has never published an article on Bayesian statistical methods, does not consider himself an expert in the field, and has never used this methodology as an expert in litigation. Gerstman Dep. 97:3–12, 112:4–18, 123:24–124:16 (Ex. 6). As noted above, he has continued to use traditional significance testing in his own publications, and did so in his opening expert report in this case. Gerstman Dep. 97:3–12 (Ex. 6). Even Dr. Gerstman's assertion that the tide has turned in favor of Bayesian methods[18] is wrong: both the Reference Manual on Scientific Evidence and Dr. Gerstman's own *Basic Biostatistics* textbook introduce p-values and confidence intervals as foundational tools of statistical inference. (Ex. 41, at 204);

[18] Finally, Dr. Gerstman's attempt to characterize Bayesian methods as growing in acceptance among some epidemiologists is irrelevant to whether his opinion is reliable. Regardless of whether this method is new or old, Rule 702 demands that it is reliably applied and fits the issue under consideration. Fed. R. Evid. 702; *see, e.g.*, *Daubert*, 43 F.3d at 1319 n. 11 ("[M]ethods accepted by a minority in the scientific community may well be sufficient. However, the party proffering the evidence must explain the expert's methodology and demonstrate in some objectively verifiable way that the expert has both chosen a reliable scientific method and followed it faithfully."). "[T]he courtroom is not the place for scientific guesswork, even of the inspired sort. Law lags science; it does not lead it." *Rosen* v. *Ciba-Geigy Corp.*, 78 F.3d 316, 319 (7th Cir. 1996). Thus, Dr. Gerstman's experimental application of Bayesian methods to determine risks associated with TRT has no place in a court of law and should be excluded.

Reference Manual, at 576–581 (Ex. 14). Bayesian methods were never applied in the many studies on TRT Nor do they appear in other studies designed to evaluate the risks associated with a pharmaceutical. Gerstman Dep. 98:18–99:2, 321:1–10 (Ex. 6); Baillargeon Dep. 444:7–21 (Ex. 11); Marais Dep. 444:7–13, 460:9–461:20 (Ex. 13). Defendants have found no authority concluding that an expert's use of Bayesian methodology to opine on general causation in a products liability case is admissible. *See also Reference Manual,* at 259 (Ex. 14) ("subjective Bayesians are a well-established minority" of scientists whose methods "have rarely been used in court").

Dr. Gerstman has concluded on the basis of Bayesian calculations that there is a 85% chance that TRT is associated with some increased risk of CV events (i.e., RR > 1.0), but he acknowledged that he has no way to determine the size of that risk between 1.0 and infinity. Gerstman Response Rpt. at 14 (Ex. 43); Gerstman Dep. 409:5–410:7 (Ex. 6). This opinion is unacceptable under *Daubert*. *First,* like any Bayesian analysis in this context, Dr. Gerstman simply performed a subjective thought experiment that produced no actual measurement of the risk purportedly presented by TRT. *See Wehling* v. *Sandoz Pharm. Corp.*, 162 F.3d 1158, 1998 WL 546097, at *5 (4th Cir. Aug. 20, 1998) (per curiam) ("conjecture, hypothesis, subjective belief, or unsupported speculation are impermissible grounds on which to base an expert opinion."). *Second,* as Dr. Marais testified, Dr. Gerstman's estimate that TRT is associated with increased CV risk meets only a 70% level of evidence, which is far below the 95% level required by the scientific community. Marais Dep. 459:16–23 (Ex. 13). Thus, Dr. Gerstman has failed to apply his method with "the same level of intellectual rigor that characterizes" his own practice or that of other epidemiologists that require 95% confidence that the result was not due to chance

alone. *Kumho Tire Co.* v. *Carmichael*, 526 U.S. 137, 152 (1999). This provides yet another ground for the exclusion of his opinion under *Daubert*.

(ii) *The "Totality of the Evidence" and "Gestalt" Judgment Are Not Reliable Methodologies*

Drs. Ardehali and Cuculich base their opinions that TRT causes CV events on their standard-less judgement about "totality of the evidence." Ardehali Rpt. at 9 (Ex. 3); Ardehali Dep. 68:21–69:12 (Ex. 7); Cuculich Dep. 70:11–14, 128:25–130:3 (Ex. 8). These experts cannot provide the information necessary for a court to conclude that this methodology is reliable under *Daubert*. *E.g.* Ardehali Dep. 200:9-24 (Ex. 7) (responding that it was "strange" to ask how someone else could tell whether he had properly applied this method), 203:16-204:8 (responding "that's not how science works" when asked for an error rate); Cuculich Dep. 128:10-130:3 (Ex. 8) ("[T]here's no name for [my method.] It is called medical training."). Indeed, they acknowledged that they were using their judgment alone and nothing that could be tested or replicated by another scientist. Ardehali Dep. 200:9–24, 203:16–204:8 (Ex. 7); Cuculich Dep. 130:18–24 (Ex. 8). This clearly does not satisfy the of *Daubert*'s requirements that scientific methodologies used by experts must be testable, subject to evaluation by peer review and publication, have a known or potential rate of error and standards controlling its operation, and be generally accepted in the scientific community. *Daubert*, 509 U.S. at 593-94. Simply reading studies and reaching a conclusion is the definition of "conjecture, hypothesis, subjective belief, or unsupported speculation [that] are impermissible grounds on which to base an expert opinion," *Wehling* v. *Sandoz Pharm. Corp.*, 162 F.3d 1158, 1998 WL 546097, at *5 (4th Cir. Aug. 20, 1998) (per curiam).

Dr. Gerstman's brand of reliance on subjective judgment is captured by his "gestalt" approach, which he used to reach his opinion that the epidemiological data is "consistent" with a

30-50% increased risk of CV events from TRT among overall users. Dr. Gerstman explained that he got to this estimate using point estimates from studies in the range of 1.3-1.5. Gerstman Dep. 354:20–355:23, 357:11–358:23, 360:9–361:19 (Ex. 6). These contained statistically insignificant results. *Id.* at 354:20–355:23, 357:11–358:23, 360:9–361:19; Gerstman Rpt. at 61, 83. In selecting these studies, he performed no analysis that considered whether these point estimates were statistically significant, whether the study or finding was limited to a subpopulation, or whether the study was too flawed to be reliable. Instead, the analysis was just "[b]ased upon the totality of the evidence." *Id.* at 358:21–23, 364:8–365:6 (explaining his method was to make a "qualitative judgment" based on the "quantifications," or risk ratios, reported in the study). Further, despite numerous opportunities, Dr. Gerstman could not explain the methodological steps he followed after these studies were chosen to reach his 30-50% increased risk opinion. *Id.* at 354:20–24, 364:8–365:7. Instead, he testified that he used his "gestalt" judgment to conclude from these studies that TRT increases CV risk by 30-50%. *Id.* at 360:13–21.

"Gestalt" is not an acceptable method of applying objective epidemiological or statistical principles; it is simply Dr. Gerstman's subjective judgment based on a handful of apparently similar data points in certain studies. *Rosen* v. *Ciba–Geigy Corp.*, 78 F.3d 316, 318 (7th Cir. 1996) ("Under the regime of *Daubert* a district judge asked to admit scientific evidence must determine whether the evidence is genuinely scientific, as distinct from being unscientific speculation offered by a genuine scientist."). The "gestalt" method fails every *Daubert* factor for determining whether a methodology is reliable: there can be no scientific testing of an inherently subjective approach, it has no known or potential rate of error or any standards to control its operation, and certainly has not been generally accepted in the epidemiological

community that relies on the system of analyzing evidence of potential associations for statistical significance, bias, and confounding. *See Daubert*, 509 U.S. at 593-94; *Zenith Elecs. Corp.* v. *WH-TV Broad. Corp.*, 395 F.3d 416, 419 (7th Cir. 2005) ("'[E]xpert intuition,' is neither normal among [ ] scientists nor testable—and conclusions that are not falsifiable aren't worth much to either science or the judiciary."). For this reason, Dr. Gerstman's opinion that TRT increases CV risk by 30-50% is unreliable and must be excluded. *See Zenith*, 395 F.3d at 419 ("A witness who invokes 'my expertise' rather than analytic strategies widely used by specialists is not an expert as Rule 702 defines that term.").

(iii) *"Risk Benefit" Analysis Does Not Fit*

Equally unavailing is Dr. Gerstman's purported reliance on what he calls the "risk-assessment" methodology of determining causation, citing FDA practice. Gerstman Dep. 76:19–21. Although Dr. Gerstman's opening report made no mention of this method, he appears to be referencing the public health approach of determining potential risk that is used by the FDA. Gerstman Dep. 18:15–23. As many courts have decided, however, this is not a valid methodology for determining legal causation. *See Hollander* v. *Sandoz Pharmaceuticals Corp.*, 289 F.3d 1193, 1215 (10th Cir. 2002) (determining that FDA rulings could not create a factual dispute as to whether drug caused Plaintiff's injury); *In re Neurontin Mktg., Sales Practices, & Prod. Liab. Litig.*, 612 F. Supp. 2d 116, 137 (D. Mass. 2009) ("[T]he FDA often uses a different standard than a court does to evaluate evidence of causation in a products liability action."). In any event, after analyzing the very same epidemiological studies on which Dr. Gerstman relies, the FDA determined that it was insufficient to establish an association or causal relationship between TRT and CV risk, and that it constituted only a "weak signal" or "possible risk." *See supra* Section B.2.a.

Furthermore, Dr. Gerstman testified at his deposition that he interprets "weak signal" to mean that the FDA was communicating that it was 60-80% confident that TRT causes CV events. Gerstman Dep. 242:7–19 (Ex. 6). Dr. Gerstman provided no evidence to support his interpretation, nor can he because it is entirely inconsistent with how the FDA uses the term "signal." The FDA does not use the term "signal" to describe a relationship that it has determined to be causal to some lesser degree of certainty. Rather, the FDA's Guidance clearly states that "[s]ignals generally indicate the need for further investigation, which may or may not lead to the conclusion that the product caused the event." FDA Pharmacovigilance Guidance (Ex. 20). This understanding of the term is consistent with the way that the 2014 Advisory Committee used the phrase "weak signal": "[I]n general, the available studies informing the cardiovascular safety signal with testosterone therapy are limited in scope, quality, design, and size. Nonetheless. . . a weak signal of cardiovascular risk had emerged from results of recent large epidemiological studies. Given this signal. . . committee members believed that the need for additional studies was critical." Ad. Comm. Mins. at 6 (Ex. 32).

(iv) *Analysis of Adverse Event Reports*

Dr. Ardehali's report contains an Appendix that evaluates seventeen adverse event reports ("AERs") from AbbVie's Period Safety Update Reports ("PSURs"), the FDA's Adverse Event Reporting system (FAER), Medwatch, and CIOMS. Ardehali Rpt., App'x C, at 1 (Ex. 3). These spontaneous AERs describe MIs and strokes that occurred in patients exposed to AndroGel between 2000 and 2007. *Id.* at 1-8.

The AERs reviewed by Dr. Ardehali were generated as part of AbbVie's post-market pharmacovigilance efforts, which attempt to evaluate potential "safety signals" of risks associated with AndroGel by collecting data on adverse events from healthcare providers and

patients. FDA Pharmacovigilance Guidance, at 3–4 (Ex. 20); *see also* Ardehali Dep. 70:2–10 (Ex. 7).

The FDA's Guidance clearly explains that spontaneous AERs are evaluated according to a specific method, which involves (a) evaluating the quality of the reports, (b) searching for similar reports, (c) collecting and analyzing those reports as a case series, (d) comparing the reported number of adverse events to a rough estimate of the expected number of adverse events, and (e) determining whether an increased rate of reported events is a signal that requires further investigation. FDA Pharmacovigilance Guidance, at 4–12 (Ex. 20). Finally, the FDA recommends that any detected signal be placed in context by comparing it to the background rate of the adverse event being investigated and whether that rate is different for certain subpopulations of patients exposed to the drug. *Id.* at 10–11.

Dr. Arrowsmith applied this methodology in reviewing AbbVie's and the FDA's AER databases, and confirmed that there was no signal of increased CV events among AndroGel users. Arrowsmith Rpt. at 136–37 (Ex. 39) (noting additionally that the increased stimulated reporting beginning in late 2013 to early 2014 harmed the usefulness of the data set for signal generation), Attachment D; *see also id*. at 133-34 (describing frequent gaps in information reported in AERs, which can make them "inaccurate or misleading in significant ways").

In contrast, Dr. Ardehali had no method for critically reviewing these AERs. He simply read the AERs in chronological order and stopped at the last AER from 2007 because he "had enough to write [his] report" and "[t]here were enough cases" to reach his conclusion that "by 2007 there was reasonable evidence of a causal association between AndroGel therapy and heart attack and stroke." Ardehali Rpt., App'x C, at 11 (Ex. 3); Ardehali Dep. 175:22–176:14, 328:19–23 (Ex. 7). Dr. Ardehali also testified that the AERs he reviewed were part of the

"totality of the evidence" on which he made a subjective judgment that AndroGel can cause CV events. Ardehali Dep. 70:11–17; *see supra* Section B.1.d.ii. (arguing that "totality of the evidence" is not a real method for determining causation and requires Dr. Ardehali's opinions to be excluded under *Daubert*). Nor does Dr. Ardehali have the training or experience to conduct a proper pharmacovigilance evaluation of these AERs to determine whether they support a potential safety signal. Ardehali Dep. 198:16–199:2, 293:20–294:10 (Ex. 7); *see also* Fed. R. Evid. 702 (requiring expert to be qualified in order for opinion to be admissible).

Additionally, Dr. Ardehali's use of AERs amounts to reliance on unpublished, incomplete case reports, which is plainly impermissible under *Daubert*. It is well-established that such case reports are not a scientifically reliable basis for a general causation opinion because they are simply anecdotal evidence, and "the events were not observed in such a way as to rule out coincidence or other potential causes." *In re Accutane Prods. Liab.*, 511 F. Supp. 2d 1288, 1298 (M.D. Fla. 2007); *Hall* v. *Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1411 (D. Or. 1996) ("[C]ase reports and case studies are universally regarded as an insufficient scientific basis for a conclusion regarding causation because case reports lack controls."); *Casey* v. *Ohio Med. Prods.*, 877 F. Supp. 1380, 1385 (N.D. Cal. 1995) (finding case reports do not provide reliable scientific evidence of causation); *see also Lennon* v. *Norfolk & Western Ry. Co.*, 123 F. Supp. 2d 1143, 1152 (N.D. Ind. 2000) (collecting cases).

2. Plaintiffs' Experts Fail to Reconcile their Opinions with the Larger Body of Scientific Evidence

As explained above, the body of evidence on TRT and CV risk has greatly expanded in the last few years, and most studies have not found a statistically significant association between TRT and increased CV risk, or have suggested that TRT is protective against CV events. *See*

supra Section B.2.b. Plaintiffs' experts therefore must reconcile their opinions with this large body of epidemiological research in order for their opinions to be reliable under *Daubert*.

Instead, some of Plaintiffs' experts attempt to dismiss it entirely. For example, Plaintiffs' experts contend that the epidemiological studies that have repeatedly found no association between TRT and CV events are predominantly observational studies and therefore subject to many limitations. *See* Ardehali Rpt. at 85–87 (Ex. 3); Cuculich Rpt. at 16–18 (Ex. 2); Gerstman Rpt. at 7, 24, 30 (Ex. 10). But this criticism applies to all observational studies as a rule, *see Reference Manual,* at 556 (Ex. 14) ("We emphasize that the Achilles' heel of observational studies is the possibility of differences in the two populations being studied with regard to risk factors other than exposure to the agent.") Thus, it does not help Plaintiffs meet their burden of proving that TRT *does* cause CV events. *See Norris*, 397 F.3d at 886. Dr. Gerstman's extensive criticism of design limitations in studies that report no significant association, Gerstman Rpt. (Ex. 10), are unavailing for the same reason—the same can be and has been said (both by the FDA and by AbbVie's experts) about the studies he prefers.

Plaintiffs' experts next insist that the existing research on TRT and CV risk is "underpowered" and therefore cannot rule out the possibility of increased risk. *E.g.* Ardehali Rpt. at 9, 100 (Ex. 3) ("The safety of the administration of exogenous testosterone to middle-aged and older men has never been adequately studied through appropriately rigorous, prospective, randomized, placebo-controlled, sufficiently powered, double-blind investigations that include specific cardiovascular endpoints."); Gerstman Rpt. at 55 (Ex. 10) ("Individually, these RCTs do not have sufficient power to discern differences in cardiovascular event rates."). Indeed, Plaintiffs' expert Dr. Wells is offered by Plaintiffs only to opine that the Corona and Albert meta-analyses are underpowered and that the Basaria (2010) result is statistically

significant.[19] From this premise, he concludes that "[t]hese studies are essentially unable to rule out an increased risk of a MACE due to testosterone therapy and thus fail to prove, or provide assurance about the safety of testosterone therapy as it pertains to MACE risk." Wells Rpt. at 2. (Ex. 46).

All of these opinions of Drs. Ardehali, Cuculich, Wells, and Gerstman should be excluded. Not only is it legally insufficient to attempt to prove causation by merely criticizing the larger body of data, *Norris* 397 F.3d at 886, but any opinion that the existing data is insufficient because it does not "rule out" all potential risk (whether due to the limitations of observational studies or the lack of statistical power) is unhelpful and confusing to a jury. That logic is clearly at odds with the standard that applies in a court of law: it is ***Plaintiffs'*** burden to affirmatively prove that AndroGel causes increased CV risk, not Defendants' burden to produce evidence "ruling out" that possibility. *See, e.g.*, *Zellers* v. *NexTech Ne., LLC*, 533 F. App'x 192, 196 (4th Cir. 2013); *Norris*, 397 F.3d at 881; *In re Lipitor (Atorvastatin Calcium) Mktg., Sales Practices & Prod. Liab. Litig.*, 185 F. Supp. 3d 786, 790 (D.S.C. 2016).

3. The Evidence Relied Upon by Plaintiffs' Experts Fails to Support a Conclusion of Causation

Finally, as explained above, courts rigorously scrutinize the evidence underlying the opinions of general causation experts in pharmaceutical cases to determine whether r it supports those opinions. This scrutiny results in the exclusion of experts who: (i) go beyond the conclusions of the studies they rely upon (*see Happel* v. *Walmart Stores, Inc.*, 602 F.3d 820, 826 (7th Cir. 2010)); (ii) rely on cherry-picked data (*see, e.g.*, *In re Zoloft*, 26 F. Supp. 3d at 460-61,

[19] These facts are obvious based on the texts of the studies themselves, and therefore Dr. Wells's opinions will not be helpful to a jury. Thus, Dr. Wells's opinion should also be excluded because it will not "assist[] the trier of fact in understanding the evidence or in determining a fact in issue." *Cummins* v. *Lyle Indus.*, 93 F.3d 362, 370 (7th Cir. 1996) (citing *Daubert*, 590 U.S. at 592).

In re Rezulin, 369 F. Supp. 2d at 425); (iii) assert opinions based on no evidence at all (*see, e.g.*, Fed. R. Evid. 702 (requiring an expert's testimony be "based on sufficient facts or data")); and (iv) assert opinions based on evidence irrelevant to the issue under consideration without providing reasoning to support extrapolation from that evidence (*see, e.g.*, *Daubert* v. *Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1314, 1320-22 (9th Cir. 1995)). For the reasons explained below, Plaintiffs' general causation experts commit all of these flaws.

(a) Plaintiffs' Experts Cherry-Pick Only Those Studies Which Favor the Plaintiffs' Position

Plaintiffs' experts rely on selected findings in a selective group of seven studies—Basaria (2010), Xu (2013), Vigen (2013), Finkle (2014), Etminan (2015), Wallis (2016), and Albert (2016)—to support their opinions that TRT causes CV events. *E.g.*, Gerstman Dep. 68:8–20 (Ex. 6) (identifying Finkle, Etminan, Xu, and Albert as the most reliable studies). On their face, their selections have a plain and simple common denominator: they are precisely those which report statistically significant associations. Plaintiffs' experts go even further with this blinkered approach when they only use those particular findings among all of the other statistical findings the same studies report. Because Plaintiffs' experts rely on selective studies to the exclusion of the larger body of evidence, their opinions should be excluded as unreliable. *In re Zoloft*, 26 F. Supp. 3d at 460-61 (excluding expert who selectively discussed evidence favoring her opinion to the exclusion of larger body of contrary evidence as applying unreliable methodology); *In re Rezulin,* 369 F. Supp. 2d at 425 (excluding expert testimony "where the expert selectively chose his support from the scientific landscape").

(b) Even Considered Alone, the Evidence Relied Upon by Plaintiffs' Experts Is Insufficient to Demonstrate Causation

Though *Daubert* addresses methods, "[a]s the Supreme Court has recognized, 'conclusions and methodology are not entirely distinct from one another,'" *Cooper* v. *Smith & Nephew, Inc.*, 259 F.3d 194, 202 (4th Cir. 2001) (quoting *Joiner*, 522 U.S. at 146), and the difference "has only limited practical import." *Paoli*, 35 F.3d at 746. "Rule 702 also requires courts to verify that expert testimony is 'based on sufficient facts or data.' Thus, trial judges may evaluate the data offered to support an expert's bottom-line opinions to determine if that data provides adequate support to mark the expert's testimony as reliable." *In re Lipitor*, 174 F. Supp. 3d at 920.

In this case, even taking the TRT studies relied upon by Plaintiffs' experts at face value, ignoring their flaws and limitations, and setting aside the body of contrary science, they cannot get to a conclusion that comports with *Daubert* .

Only the Xu study has found a statistically significant association with increased CV risk among overall TRT users. As explained above, this finding was superseded by the Albert study, upon which Plaintiffs experts also rely. Similar problems prevent Plaintiffs' experts from relying solely on the Finkle study to support causation among short-term users under the age of 65 with prior CV history. In addition to being the only study to report a statistically significant association in this group, the FDA concluded that this study could not support anything more than a "weak signal" or a "possible risk." *See supra* Section B.2.a. Finally, while four studies have reported statistically significant findings for men over age 65 who used TRT for some short period of time, these studies also are insufficient to demonstrate general causation for this subgroup because they are not consistent with one another for the reasons explained above. *See supra* Section B.2.c.i.; App'x A [CV chart]. Specifically, a statistically significant association

among men over age 65 who were exposed to TRT for 90 days (i.e. Finkle (2014)) is not replicated by a statistically significant association among men over age 65 who were exposed to TRT for 12 months or less (i.e. Albert (2016)). The four studies addressing this subgroup are too heterogeneous to constitute evidence of a replicated association between TRT and increased CV risk, let alone sufficient evidence of causation. *Cf. In re Lipitor*, 174 F. Supp. 3d at 924 ("To be sure, it is possible for the entirety of the evidence to support an opinion even when individual pieces of evidence are not sufficient in isolation, but it is also possible that multiple pieces of insufficient evidence add up to insufficient evidence.").

Nor do any of the authors reach any conclusion that causation has been established. Plaintiffs' experts' "rel[iance] on studies for publication, the authors of which were themselves unwilling to conclude that causation had been proven" makes their testimony too unreliable to be admissible. *Happel* v. *Walmart Stores, Inc.*, 602 F.3d 820, 826 (7th Cir. 2010). Thus, if these seven studies constitute the total body of the epidemiological evidence supporting Plaintiffs' experts' opinions, then there is no basis for them to conclude that AndroGel causes MIs and strokes because the studies' authors determined that the data supported only an association or signal.

(c) Plaintiffs' Experts Assert Opinions Based on No Evidence

Dr. Ardehali seeks to testify to an opinion not based on any data. Rule 702 requires that an expert's opinion be based on "sufficient facts or data," Fed. R. Evid. 702, and *Daubert* provides that expert testimony must be "supported by appropriate validation—i.e., 'good grounds,' based on what is known." *Daubert*, 509 U.S. at 590. The Seventh Circuit has affirmed the exclusion of experts who merely hypothesized without scientific proof of their conclusions. *Porter* v. *Whitehall Labs., Inc.*, 9 F.3d 607, 614-15 (7th Cir. 1993) ("If experts cannot tie their assessment of data to known scientific conclusions, based on research or studies,

then there is no comparison for the jury to evaluate and the expert's testimony is not helpful to the jury."); *see also Wehling*, 162 F.3d 1158, 1998 WL 546097, at *5 ("unsupported speculation [is] impermissible" as a basis for an expert opinion); *In re Lipitor*, 174 F. Supp. 3d at 920 (explaining that a trial judge should "evaluate the data offered to support an expert's bottom-line opinions to determine if that data provides adequate support to mark the expert's testimony as reliable.").

In this case, Dr. Ardehali opines that TRT increases the risk of MI and stroke only in men who are "off-label," which he uses to refer to patients with non-classical hypogonadism. Ardehali Rpt. at 9, 40,56, 63, 98, 101 (Ex. 3); Ardehali Dep. 149:25–150:6, 150:6–17 (Ex. 7). But he cites absolutely no research to support this distinction. Indeed, as explained in Dr. Khera's report, there is no data that isolates the risk of MI and stroke associated with TRT exclusively in classical or non-classical patients; rather, the clinical trials of AndroGel and every other TRT product, as well as all independent observational studies and RCTs, included subjects with low serum testosterone levels regardless of etiology. *See* Expert Report of Mohit Khera, M.D., M.B.A., M.P.H. at 13, Dec. 6, 2016 (Ex. 47); Arrowsmith Rpt. at 88 (Ex. 39). Therefore, Dr. Ardehali's opinion that TRT is unsafe in "off-label" (non-classical) patients is without basis and should be excluded as unreliable and unhelpful to the jury under *Daubert*. Fed. R. Evid. 702; *Porter*, 9 F.3d at 614-15.

(d) Plaintiffs' Experts Rely on Studies That Are Irrelevant to AndroGel or CV Events

The opinions of Drs. Ardehali and Cuculich that TRT causes CV events are based on research on steroid abuse in athletes. Ardehali Rpt. at 28–29 (Ex. 3); Ardehali Dep. 88:17–22, 110:2–12; 214:6–215:10 (Ex. 7); Cuculich Rpt. at 5-6, 14, 19 (Ex. 2). But studies evaluating (i) another class of drugs, (ii) ingested at levels so high to be considered abuse, (iii) in a different

population of people are not studies on the product at issue. *See also* Khera Rpt. at 86-87 (Ex. 47) (explaining that anabolic steroids abused by athletes are more dangerous because they can have up to 30 times the anabolic:androgenic ratio to testosterone products).

Use of such studies is impermissible under *Daubert* unless Plaintiffs' experts have a clear methodology for extrapolating the results of these studies to the context of TRT. *See Daubert*, 43 F.3d 1311, 1314, 1320-22 (9th Cir. 1995) (excluding Plaintiffs' expert who sought to extrapolate a causal link between Bendectin and birth defects based on studies of other drugs with a similar chemical structure for failure to meet "fit" requirement). They do not. Drs. Ardehali and Cuculich supply no reasoning or explanation of how this data has any relevance to determining the association between AndroGel and increased risk of MIs and strokes. *See* Ardehali Rpt. at 29 (Ex. 3); Cuculich Rpt. at 5, 14–15 (Ex. 2). At his deposition, Dr. Cuculich testified that the fact that anabolic steroids and testosterone are in "the same families" of drugs should be sufficient grounds for comparing the two. This is the very contention that was rejected in the Ninth Circuit's *Daubert* decision, affirming the exclusion of an expert who relied on data concerning similar chemicals to the one at issue. *See Daubert*, 43 F.3d at 1314, 1320-22. Dr. Cuculich did not know or account for the vast differences in dose (testifying only that steroid abusers used "considerably higher" doses) and identified no methodology for drawing a direct comparison between steroid-abusing athletes and middle-aged TRT patients. Cuculich Dep. 221:17–223:8 (Ex. 8).

Similarly, Dr. Gerstman has an entire section of his report dedicated to "Parallels with studies of hormone therapy safety in women." Gerstman Rpt. at 24–26 (Ex. 10). Nowhere does his description contain any reasoning for why the use of estrogen and progestin in post-

menopausal women to alleviate symptoms of menopause should be directly comparable to the use of TRT in men who suffer from an illness due to pathologically low testosterone levels.[20]

The foregoing opinions are "connected to existing data only by [the] *ipse dixit*" of Plaintiffs' experts. *See Joiner*, 522 U.S. at 146. Their failure to provide any support for extrapolating data on the use of different chemicals by different populations demonstrates that their opinions inherently rely on a logical leap that is impermissible under *Daubert*. *Id.*; *Porter* v. *Whitehall Labs., Inc.*, 9 F.3d 607, 614-15 (7th Cir. 1993). Therefore, any of Dr. Ardehali's and Dr. Cuculich's opinions that rely on steroid abuse by athletes, and any of Dr. Gerstman's opinions that rely on the adverse events associated with HRT in women, should be excluded. *See Joiner*, 522 U.S. at 146; *Rosen*, 78 F.3d at 319 (7th Cir. 1996) ("[T]he courtroom is not the place for scientific guesswork, even of the inspired sort.").

(e) Plaintiffs' Experts Rely on Animal Studies and *In Vitro* Studies

Drs. Ardehali and Halushka rely heavily on animal and *in vitro* studies as evidence of the mechanism by which TRT increases CV risk in human beings. "Federal courts have consistently cautioned against extrapolation of human effects from animal studies. In addition to the biological differences between species, most animal studies involve significantly higher

[20] Dr. Gerstman also opines that there are "*striking* parallels between the study of hormone therapy (HT) in women and in men." Gerstman Rpt. at 24 (Ex. 10). But nothing could be further from the truth: the dangers of HT in women were studied and confirmed after the publication of large, highly-powered, robust RCTs demonstrated and replicated statistically significant associations between HT and certain adverse events. As Dr. Gerstman readily points out, no such study has been completed for TRT yet. Therefore, the only readily applicable lesson from the studies of HT in women appears to be that the risk of adverse events must be confirmed in a large, highly-powered RCT, a proposition with which experts on both sides agree. *See, e.g.*, Gerstman Rpt. at 30 (Ex. 10); Baillargeon Dep. 113:24–115:1, 147:23–148:3 (Ex. 11); French Dep. 422:22–424:10 (Ex. 9); Wells Dep. 83:5–10 (Ex. 12); Cuculich Dep. 109:8–9, 237:20–238:1 (Ex. 8); Marais Dep. 489:4–490:3 (Ex. 13).

concentrations of a substance than would ever be present in humans." *In re Prempro Prod. Liab. Litig.*, 738 F. Supp. 2d 887, 894 (E.D. Ark. 2010) (collecting cases).

"These considerations have equal force for *in vitro* studies. Even where human tissue is studied, an expert must still explain how laboratory results will reliably predict effects in living humans." *Id.* (expert did not explain how animal and in vitro studies could be reasonably applied to humans); *Allison*, 184 F.3d at 1313 (exclusion of Plaintiff's expert who failed to "explain why the results of these animal studies should trump more than twenty controlled epidemiological studies of breast implants in humans which have found no valid increased risk of autoimmune disease").[21]

In the case of increased CV risk, Dr. Ardehali proposes four mechanisms: increased estradiol, increased blood viscosity, increase of thromboxane ("TxA2") receptors that contribute to plaque disruption, and increased reactive oxygen species. Ardehali Rpt. at 67 (Ex. 3). Dr. Ardehali consistently cites studies of animals and cells to conclude that these mechanisms are demonstrated as the means by which AndroGel causes CV events. Dr. Halushka is offered solely to explain how the TxA2 and its receptors contribute to increased CV risk from TRT

[21] *See also, e.g., Fabrizi* v. *Rexall Sundown, Inc.*, No. 01-289, 2004 WL 1202984, at *6-8 (W.D. Pa. June 2, 2004), Rep & Rec. adopted June 24, 2004 (excluding testimony of a physician that ingestion of St. John's wort could cause eye cataracts on grounds that *in vitro* studies of animal eye tissue could not be extrapolated to live humans); *Soldo* v. *Sandoz Pharms. Corp.*, 244 F. Supp. 2d 434, 542, 546-50, 567–72 (W.D. Pa. 2003) (excluding testimony that Parlodel can cause intracerebral hemorrhage ("ICH") because the testimony, which was based on anecdotal case reports, animal studies, other drugs, and studies of patients with pathologies other than ICH, flunked all of the *Daubert* criteria); *Caraker* v. *Sandoz Pharms. Corp.*, 188 F. Supp. 2d 1026, 1031–1040 (S.D. Ill. 2001) (excluding testimony that Parlodel can cause ICH because the experts' conclusion "requires too many extrapolations from dissimilar data, too many analytical leaps and involves a loose application of purportedly objective scientific causation standards"); *Hall* v. *Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1407-11 (D. Or. 1996) (excluding testimony that silicone is capable of causing certain symptoms because the experts were making "too great a leap of faith;" animal studies could not be reliably extrapolated to humans without explanation, case reports were insufficient to prove causation, and studies of crystalline silica were irrelevant because Plaintiffs had not shown that silicone breast implants were associated with the presence of crystalline silica).

exposure, and he too relies on two *in vitro* cell studies, two small animal studies in which rats and guinea pigs were administered massive overdoses of testosterone. Halushka Rpt. at 19–20 (Ex. 44); Halushka Dep. 116:15–117:6 (Ex. 21). Neither of these experts produce research showing how the results from these studies can be extrapolated to living, human men. Their inability to do so renders their opinions on mechanism inadmissible. *See Allison*, 184 F.3d at 1313; *In re Prempro*, 738 F. Supp. 2d at 894 (E.D. Ark. 2010).

C. The Testimony of Plaintiffs' Sole General Causation Experts Regarding VTE Risk, Drs. Rinder and Halushka, Does Not Satisfy the Stringent *Daubert* Standards

In reaching his opinions regarding the risk of VTE, Dr. Rinder has developed a distinctive and highly flawed approach in which he essentially dismisses all but one of the epidemiological studies as defective, dismisses the concept of statistical significance, and improperly relies on notoriously unreliable sources (case reports) to support his opinion. His opinions are unreliable and should be dismissed.

1. Dr. Rinder Fails to Demonstrate a Statistically Significant Association Between AndroGel and VTEs

There are now five epidemiological studies which report that there is no statistically significant association between TRT and VTEs for large groups of users, one of which finds an association for two related sub-subgroups. *See* **Figure 4**.

Sharma's observational study of over 70,000 men with low testosterone levels compared the incidence of VTEs in three groups: men who received TRT and achieved a normal testosterone level, men who received TRT and did not achieve a normal testosterone level, and men who did not receive TRT at all. Ex. 69 (Sharma (2016)). Each group experienced nearly

identical rates of VTEs, thus demonstrating that there is no association between using TRT and increased VTE risk.

Li studied over 200,000 hypogonadal men, half of whom were on TRT, for the incidence of idiopathic VTE (meaning the subjects had no other VTE risk factors). Ex. 69 (Li (2016)). This design permitted the Li authors to isolate the effect due to TRT alone, which militates against confounding from non-testosterone causes. *See* Baillargeon Dep. 194:7–195:5 (Ex. 11). The Li study concluded that there is no statistically significant association between VTE and TRT (OR 1.08, 95% CI = 0.91-1.27, p = 0.378). Ex. 69 (Li (2016)).

Dr. Baillargeon's VTE study of over 30,000 middle-aged and older men also found no association between TRT and VTE. Ex. 69 (Baillargeon (2015)). He and his colleagues identified 7,643 men who were diagnosed with DVT or PE, 158 of whom were exposed to some form of TRT (inferred from filing a TRT prescription) in 15 days before their VTE. *Id.* at 1042 Table 2. No statistically significant association was found (OR 0.90, 95% CI = 0.73-1.12). This finding persisted when the exposure window (i.e., date of last prescription) was extended to 30 days and 60 days before the VTE, and was consistent across means of administration of testosterone (topical (OR 0.80, 95% CI = 0.61-10.4), transdermal (OR 0.91, 95% CI = 0.31-2.16), and intramuscular (OR 1.15, 95% CI = 0.80-1.64)). *Id.* at 1038, 1042.

The fourth study, Martinez (2016), examined the medical records of 19,000 UK men, aged 20-89, who suffered a VTE between January 2001 and May 2013, 69 of whom received TRT. Ex. 69 (Martinez (2016)). All 19,000 subjects were matched to over 90,000 control subjects (1,251 of whom used TRT) on the basis of age, VTE risk factors in the 90 days preceding VTE, and history of "pathological" primary or secondary hypogonadism. The authors stratified this data and statistically compared various groups.

The study failed to find any association between VTE and TRT users overall (RR 1.25, 95% CI = 0.94-1.66). It found a statistically increased risk of VTE among current users of TRT, who had taken TRT for 6 months or less (RR 1.63, 95% CI = 1.12-2.37), but this finding persisted only among recent TRT users *without* known VTE risk factors (RR 1.91, 95% CI = 1.13-3.23), and not short-term TRT users *with* known VTE risk factors (RR 1.41, 95% CI = 0.82-2.41). The study also found statistically significant increased incidence of VTE among men without pathological hypogonadism (RR 1.69, 95% CI = 1.09-2.63) but after these results were stratified between short-term (≤ 6 months) users and longer-term (> 6 months) users, the results presented only for short-term users (RR 1.88, 95% CI = 1.02-3.45). When the authors stratified the data according to method of TRT administration, the authors found no statistically significant incidence of VTE associated with any single method (intramuscular, gel, oral). *Id.*

The Martinez authors acknowledged that this study "*suggests* a transient increase in the risk of venous thromboembolism that peaks during the first three to six months and declines gradually thereafter," but called for further research before to confirm such an association. *Id.* (emphasis added). Thus, reliance on Martinez to support an opinion that an association, let alone a causal relationship, has been established goes beyond the conclusion that the authors themselves were willing to reach and is impermissible under *Daubert*. *See Happel*, 602 F.3d at 826.

In addition, the findings of this study are significantly flawed for a number of reasons.

- **Tyranny of small numbers.** The "recent user" finding was based on only 36 VTE cases out of over 19,000 patients. Moreover, the difference between the groups with and without VTE risk factors was *two* individuals, meaning the result could have been the opposite if only two more people with VTE risk factors were found in the database. *See* Bierer Rpt. at 13–14. (Ex. 42)

- **Arbitrary definition of "recent" use.** The authors provided no reason for stratifying their data among men who used TRT for less than six months compared to six months and longer. This time period is greater than what has been used in other studies to evaluate short-term risk of TRT use. *See* Baillargeon Rpt. at 56 (Ex. 15).

- **Risk factor results inexplicable.** The relative risk was higher in recent users who did not have preexisting risk factors for VTE than those who did. This finding casts serious doubt on the reliability of the Martinez study, particularly because the authors and Plaintiffs' experts have failed to properly explain it. *See* Baillargeon Rpt. at 56–57 (Ex. 15).

- **Multiplicity of comparisons.** The subgroup finding was one of many post-hoc subgroup analyses performed by the Martinez authors. When multiple comparisons are performed on the same set of data, using the regular 95% level of evidence means that one out of every twenty comparisons will appear to be statistically significant but will, in fact, be due to chance. *See* Marais Rpt. at 16 (Ex. 16); Baillargeon Rpt. at 56–57 (Ex. 15).

Dr. Rinder regards all of these studies as "flawed." Rinder Dep. at 187:18–188:7 (Ex. 37). Yet he reaches out to rely on a **letter** by Xu describing VTE meta-analysis work using three small RCTs. Ex. 69 (Xu & Schooling (Jan. 2015)). As referenced earlier, none of the RCTs individually found a statistically significant association between TRT and VTE risk, and Xu and Schooling's original analysis reports a statistically insignificant odds ratio of 5.94 (95% CI = 1.0–35.3). By including one more small RCT that reported another single DVT in the TRT group that did not describe adverse events in the placebo group, Xu and Schooling got to a statistically significant odds ratio of 5.62 (95% CI = 1.39–22.8).

Not only is this second calculation clearly based on a *post hoc* analysis, but three of the six total VTE events reported by the four RCTs were from the Copenhagen study, whose subjects all had cirrhosis of the liver and were administered extremely high doses of testosterone via an unapproved oral version of TRT. Dr. Marais performed a sensitivity analysis of the Xu letter, and concluded that its finding was due almost entirely to the particularities of the underlying studies and did not persist when other RCTs that would have been eligible for

inclusion were added to the mix. Marais Rpt. at 116–17 (Ex. 16); *see also* Bierer Rpt. at 7–8 (Ex. 42). For these reasons, that portion of the letter reflecting the *post-hoc* re-analysis is too unreliable to provide a basis for Plaintiffs' experts' opinions.

* * *

Taking Dr. Rinder's insistence that all of the epidemiology today is "flawed," his opinions cannot meet the threshold requirement of statistically significant association. This ends the causal analysis. Taking them at face value, they reach no opinion that TRT increases the risk of VTE and, with one exception for a sub-group of a subgroup, report no statistically significant associations. And that exception applies to none of the bellwether Plaintiffs. Rinder Dep. 332:7–335:17 (Myers), 365:5–366:20 (Nolte), 372:21–374:16 (Rowley) (Ex. 37).

2. Case Reports Cannot Yield an Association That Is not Established by the Epidemiology

Dr. Rinder is left to rely on case reports described in articles by Charles Glueck, *et al.* regarding VTEs, which also suffer from serious statistical and methodological flaws.[22]

After learning of 20 total patients who took TRT and subsequently developed a variety of venous thrombosis (blood clot) related events, including deep vein thrombosis

[22] Ex. 69 (Charles J. Glueck *et al*., Thrombotic events after starting exogenous testosterone in men with previously undiagnosed familial thrombophilila, 158(4) TRANS. RES. 225 (2011); Charles J. Glueck *et al*., Testosterone, thrombophilia, and thrombosis, 20(1) CLIN. APPL. THROMB. HEMOST. 22 (2014). In addition, Glueck *et al*. have published a "review" with similar conclusions (Charles J. Glueck *et al*., Testosterone therapy, thrombosis, thrombophilia, cardiovascular events, 63(8) METABOLISM 989 (2014)) and an "exploratory, hypothesis-generating study" (Charles J. Glueck *et al*., Testosterone therapy thrombophilia-hypofibrinolysis and hospitalization for DVT-PE: an exploratory, hypothesis-generating study, 20(3) CLIN. APPL. THROMB. HEMOST. 244 (2014)). In the latter, the authors tracked down 139 men hospitalized for (non-cancer-related) DVT and PE. Seven of these had been taking TRT at admission. Five of the seven were found to have previously undiagnosed familial or acquired thrombophilia or hypofibrinolysis. Glueck *et al*. compared these seven men to eight non-TRT controls and found that seven of the eight controls also had measures of underlying thrombophilia. Nevertheless, Glueck *et al*. reiterate the possible "interaction" effect of TRT with underlying thrombophilia.

("DVT") and pulmonary embolism ("PE"), the authors decided to investigate and screen the men for familial thrombophilia (an inherited abnormality of blood coagulation that increases the risk of blood clots). They found previously undiagnosed familial thrombophilia in nearly all of the men. According to the articles, this led the authors to "speculate that when exogenous testosterone is aromatized to [estradiol], and [estradiol]-induced thrombophilia is superimposed on familial thrombophilia, thrombosis occurs." The authors did not do anything to rule out the underlying familial thrombophilia as the sole cause of the events, but instead apparently relied entirely on the "temporal association" of thrombosis following TRT in time. These shortcomings render the Glueck articles unreliable as evidence of VTE risk.

The Glueck articles also are unreliable evidence because they are merely reports of individual cases. As explained above, it is well established that such case reports are not a scientifically reliable basis for a general causation opinion because they are simply anecdotal evidence, and "the events were not observed in such a way as to rule out coincidence or other potential causes." *In re Accutane Prods. Liab.*, 511 F. Supp. 2d 1288, 1298 (M.D. Fla. 2007); *Hall* v. *Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1411 (D. Or. 1996) ("[C]ase reports and case studies are universally regarded as an insufficient scientific basis for a conclusion regarding causation because case reports lack controls."); *Casey* v. *Ohio Med. Prods.*, 877 F. Supp. 1380, 1385 (N.D. Cal.1995) (finding case reports do not provide reliable scientific evidence of causation); *see also Lennon* v. *Norfolk & Western Ry. Co.*, 123 F. Supp. 2d 1143, 1152 (N.D. Ind. 2000) (collecting cases).

The Glueck articles also cannot be squared with the available epidemiological data, which suggests that there is no increased risk of VTEs. While "[t]o date, there are no prospective studies that have evaluated the risk of VTE in men receiving exogenous

[testosterone] supplementation,"[23] two prospective studies of *endogenous* testosterone have found that "high endogenous concentrations of estradiol and testosterone in . . . men in the general population are not associated with increased risk of VTE, DVT or PE."[24] Case reports that are contrary to the weight of epidemiological evidence are particularly unreliable. *See, e.g.*, *Norris*, 397 F.3d at 884 (10th Cir. 2005); *Allison*, 184 F.3d at 1316.

The reliance of Dr. Rinder on the Glueck articles is particularly unwarranted since the authors themselves acknowledge that the case reports are merely "exploratory," and thus not reliable evidence of an association between TRT and VTEs (much less evidence of causation). Not only do Dr. Glueck and colleagues refer to their work as "exploratory," but they expressly characterize it as "hypothesis-generating," and "speculation." Ex. 69 (Glueck (2011), Glueck (2014)). As noted above, such studies are insufficient as a matter of law to support an opinion of scientific causation. *Rosen* v. *Ciba–Geigy Corp.*, 78 F.3d 316, 318 (7th Cir. 1996) ("Under the regime of *Daubert* a district judge asked to admit scientific evidence must determine whether the evidence is genuinely scientific, as distinct from being unscientific speculation offered by a genuine scientist.").

3. Dr. Rinder Fails to Meet the Additional Criteria for Causation

In addition to failing to demonstrate any statistically significant association between AndroGel and VTE risk, Dr. Rinder also fails to address, much less meet additional criteria for causation.

---

[23] Anthony Grech *et al.*, Adverse effects of testosterone replacement therapy: an update on the evidence and controversy, 5(5) THER. ADV. DRUG SAF. 190 (2014) (Ex. 50).

[24] Haya N. Holmegard *et al.*, Endogenous sex hormones and risk of venous thromboembolism in women and men, 12(3) J. THROMB. HAEMOST. 297 (2014) (Ex. 66); *see also* Johan Svartberg *et al.*, Endogenous sex hormone levels in men are not associated with risk of venous thromboembolism: the Tromso study, 160(5) EUR. J. ENDOCRINOL. 833 (2009) (Ex. 51).

(a) There Is No Consistency or Strength of Association or Replication

This already has been shown above. The only consistency that exists in the VTE epidemiological studies is that repeated finding of no statistically significant association in TRT users overall.

(b) The Temporal Relationship Between the Exposure and the Outcome Does Not Establish Causation

As discussed above, the mere fact that someone experienced a VTE after using AndroGel is not sufficient to demonstrate a causal relationship under the Bradford-Hill framework. *See Reference Manual,* at 601 (Ex. 14). VTEs are the result of clots, which typically take time to form, break off, and travel to the site of injury, and thus generally involve some latency period between initiation and harm. As a result of this latency period, the fact that a VTE occurred subsequent to AndroGel exposure is entitled to little weight in evaluating general causation. *See id.* at 602 ("[W]hen latency periods are lengthy, variable, or not known and a substantial proportion of the disease is due to unknown causes, temporal relationship provides little beyond satisfying the requirement that cause precede effect."); *Glastetter*, 252 F.3d at 990 (8th Cir. 2001) (holding that evidence "demonstrat[ing] a temporal association between [drug] and stroke, or stroke-precursors," was "not scientifically valid proof of causation."); *Moore*, 151 F.3d at 278 (5th Cir. 1998) ("In the absence of an established scientific connection between exposure and illness, or compelling circumstances. . ., the temporal connection between exposure to chemicals and an onset of symptoms, standing alone, is entitled to little weight in determining causation.").

(c) Dr. Rinder Ignores the Dose-Response Relationship

It is a well-documented fact that testosterone injections cause a spike in hematocrit levels that does not occur with gel formulations. *See* Nand Rpt. at 26–27 (Ex. 4).

"Data exists to suggest that injectable androgen products have a different pharmacological profile than topically applied androgen; therefore, adverse events reported for injectable androgen products cannot be pooled with those of topical androgen products.[25] AndroGel is applied transdermally and adverse events observed with AndroGel administration must be differentiated from adverse events observed with other formulations." Bierer Rpt. at 5 (Ex. 42). "Thus, AndroGel, as compared to injectable forms of testosterone, is less likely to contribute to erythrocytosis." Nand Rpt. at 27 (Ex. 4). Dr. Rinder's claim that all TRT, regardless of dose or formulation, cause increases in hematocrit so severe as to trigger VTEs, but the body of scientific evidence does not support such a broad conclusion. Hematocrit also is addressed in the next section, in connection with biological mechanism.

(d) Dr. Rinder Fails to Demonstrate the Biological Plausibility of TRT Increasing VTE Risk

As noted above, this criterion "depends on existing knowledge about the mechanisms by which the disease develops." *Reference Manual*, at 605 (Ex. 14). Here, Plaintiffs' experts propose three mechanisms or physiological effects for how TRT purportedly causes VTE: (i) an increase in hematocrit resulting in polycythemia, (ii) an increase in platelet aggregation resulting from increased thromboxane A2 receptor density, and (iii) an increase in serum level of estradiol.

Increase in Hematocrit

As discussed above (*see supra* Section C.2.a), the impact of testosterone on the production of red blood cells has prompted precautions to monitor hematocrit in TRT users. The

[25] A.W. Pastuszak, *et al.*, Comparison of the Effects of Testosterone Gels, Injections, and Pellets on Serum Hormones, Erythrocytosis, Lipids, and Prostate-Specific Antigen, 3 SEX MED. 165 (2015) (Ex. 52).

effects of transdermal delivery on hematocrit have been shown to be smaller than injections, reflecting the fact that this pathway delivers testosterone without the sharper spikes associated with injection (Dobs, *Pharmacokinetics*, 3477 (Ex. 35)). In such studies and in medical practice, the upper boundary of normal lab range for hematocrit is generally 50% or somewhat more. (Dobs, *Pharmacokinetics*, 3473–74 (Ex. 35)). Doctors monitoring TRT patients for hematocrit, as advised by the AndroGel label, can make their own judgment about exactly what level is appropriate.

Dr. Rinder's opinion seeks to overturn all of this and find causation even for patients whose hematocrit while on TRT not only did not exceed 50% but did not even change from levels while not on TRT. These opinions simply cannot be reconciled with science or medicine. Not only do they lack substantiation in the TRT epidemiological literature, but they say that current medical practice is just wrong.

Specifically, Dr. Rinder claims that "[m]en with hematocrits ≥ 46% had a more than 2-fold increase in VTE risk compared with men whose hematocrit was < 43%." Rinder Rpt. at 10 (Ex. 5). But the study on which Dr. Rinder relies for that opinion (Braekken), does not even involve TRT, much less purport to find any link between TRT and increased hematocrit, or, even further, to research the relationship between TRT and increased VTE risk. Moreover, the stratification of the data in the Braekken study, in which the "normal" hematocrit range (< 43%) was defined at levels that are in fact well below the actual normal hematocrit range for men, renders any opinion based on this study highly unreliable. Given this deficiency, it is not surprising that the findings in the Braekken study have not been replicated in any other study.

Dr. Rinder also seeks to rely on animal studies, but he fails to explain how the results of those studies can be extrapolated to humans. *See Daubert*, 43 F.3d at 1319-20; *Allison*,

184 F.3d at 1313 (noting that Plaintiff failed to "explain why the results of these animal studies should trump more than twenty controlled epidemiological studies of breast implants in humans which have found no valid increased risk of autoimmune disease"); *Conde* v. *Velsicol Chem. Corp.*, 24 F.3d 809, 814 (6th Cir. 1994) (finding animal studies inadequate for showing causation of disease in humans with chlordane exposure); *In re Prempro Prod. Liab. Litig.*, 738 F. Supp. 2d 887, 894 (E.D. Ark. 2010) ("Federal courts have consistently cautioned against extrapolation of human effects from animal studies. In addition to the biological differences between species, most animal studies involve significantly higher concentrations of a substance than would ever be present in humans."). Thus, the only connection between the animal studies on which Dr. Rinder relies and the potential that TRT increases the risk of VTE via increased hematocrit is his "*ipse dixit*," which is impermissible under *Daubert* and requires his opinion to be excluded. *Joiner*, 522 U.S. at 146.

Increase in Thromboxane A2 Receptor Density

All but one of the studies on which Plaintiffs' experts rely for their theory that TRT increases thromboxane A2 receptor density (which they claim in turn increases VTE risk) also were performed on animals, and thus are insufficient under *Daubert* for the same reasons. Indeed, Drs. Halushka and Rinder fail to provide any methodology or explanation for extrapolating the results of these animal studies to human men on appropriate doses of testosterone. *See* Nand Rpt. at 30–31 (Ex. 4); *see also* Bierer Rpt. at 5 (Ex. 42) (pointing out flaws in Dr. Rinder's analysis of TXA2 receptor mechanism). Dr. Rinder simply declares that "[p]latelet effects of testosterone have been confirmed in humans" without citing a shred of evidence supporting this proposition. Rinder Rpt. at 13 (Ex. 5).

Similarly, Dr. Halushka simply discusses how platelet aggregation plays a role in VTE development—which is obvious because VTEs are blood clots formed by platelets (as opposed to the white blood cells that clot and cause MIs and strokes). *See supra* Section A; Nand Rpt. at 6 (Ex. 4). Dr. Halushka's report discloses absolutely no evidence tying testosterone to VTE via activation of thromboxane A2 and its receptors. Halushka Rpt. at 33–35 (Ex. 44); Halushka Dep. 72:18–22, 235:2–16 (Ex. 21). Moreover, he testified that he has never studied the relationship between TRT and VTE and was unable to explain his opinion regarding VTE and how it had any relation to TRT. Halushka Dep. 235:11–236:24, 237:24–238:3 (Ex. 21). This glaring lack of evidence and analysis is plainly insufficient under *Daubert* requires Dr. Halushka's opinion on VTE risk to be excluded.

Increase in Estradiol Levels

Dr. Rinder's opinion that TRT increases the risk of VTE by raising estradiol levels is based on speculation. As Dr. Bierer explains in her report, Dr. Rinder relies on studies that have nothing to do with TRT to support this theory; rather, the "[a]dministration of direct estrogen treatment to men either as treatment for prostate cancer[26] or to determine if estrogen administration may be a beneficial impact on hyperlipidemia is not analogous to AndroGel® administration."[27] Bierer Rpt. at 5. Similarly, Dr. Rinder fails to explain how evidence from studies of estradiol in women on HRT can be directly applied to men on TRT. For these reasons, Dr. Rinder's opinions are nothing more than speculation, based on insufficient evidence related to the drug at issue in this case, and thus should be excluded.

---

[26] M. Van Hemelrijck, *et al.*, Risk of thromboembolic diseases in men with prostate cancer: results from the population-based PCBaSe Sweden, 11 LANCET ONCOL. 450 (2010) (Ex. 53).

[27] Coronary Drug Research Project: Findings Leading to Discontinuation of the 25-mg/day Estrogen Group, 226(6) JAMA 652 (1973) (Ex. 54).

And Dr. Rinder's focus on whether certain mechanisms are biologically plausible is also unsupportable because he has not established a statistically significant association between TRT and VTE. *In re Lipitor*, 174 F. Supp. at 924-25 (emphasis in original) (citing *Mathews* v. *Novartis Pharm. Corp.*, No. 3:12–CV–314, 2013 WL 5780415, at *27 (S.D. Ohio Oct. 25, 2013) ("[u]nless there is a statistically significant association between the drug and the disease, the Bradford–Hill analysis to determine causation is inapplicable."); *Soldo*, 244 F.Supp.2d at 461 ("[A]pplication of the Bradford Hill criteria depends first on an association by epidemiology between a disease and an exposure to an agent. The association must rule out chance.")). "Courts exclude expert testimony that attempts to start at step two, applying the Bradford Hill criteria without adequate evidence of an association." *Id*. at 925-26.

4. Dr. Rinder's Opinions on VTE Risk Suffer from Numerous, Additional Methodological Flaws

First and foremost, Dr. Rinder all but admits that he is applying a double standard. He has a clear-cut methodology for determining whether a potential non-TRT risk factor is causal, but he admits that he cannot compare that methodology to what he did in this case. He simply applied a different set of criteria to evaluating the epidemiological studies on TRT. Rinder Dep. at 103–04 (Ex. 37). When asked to provide an example of a risk factor that has such limited epidemiology as what he analyzed in this case, he explained, "to me, it's kind of an apples and oranges . . . they're not comparable situations." Rinder Dep. at 269 (Ex. 37).

The standards and methods he follows for assessing general causation is not science, but practical medicine. It is the same approach he applies "in the clinical setting when evaluating patients." Rinder Rpt. at 3 (Ex. 5); Rinder Dep. at 116:3–19 (Ex. 37). But the diagnosis of VTE in individual patients has no bearing on whether TRT is capable of causing VTEs in the general patient population. Many courts have excluded general causation opinions

reached by the method of differential diagnosis or clinical judgment. *See, e.g.*, *Hollander* v. *Sandoz Pharmaceuticals Corp.*, 289 F.3d 1193 (10th Cir. 2002) (affirming exclusion of opinion based on differential diagnosis offered to prove general causation); *In re Rezulin*, 2004 WL 2884327, at *4 (S.D.N.Y. Dec. 10, 2004) (excluding expert opinion proffered for general causation because based on inapplicable methodology of differential diagnosis); *Soldo* v. *Sandoz Pharmaceuticals Corp.,* 244 F.Supp.2d 434, 516 (W.D. Pa. 2003) ("differential diagnosis is not a reliable methodology for determining general causation"); *Hall* v. *Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1413-14 (D. Or. 1996) (excluding expert testimony based on differential diagnosis as evidence of general causation because "differential diagnosis *assumes* that general causation has been proven for the list of possible causes it eliminates"). Thus, Dr. Rinder's general causation opinions should be excluded on the ground that they are unreliable and based on a method does not fit the issues on which he opines.

Dr. Rinder also is another Plaintiffs' expert who denies the significance of statistical significance. While he disclaims facility with the tests of statistical significance, the results do not slow him down – he relies on all point estimates, statistically significant or not. Rinder Dep. at 28:19–29:1, 39:2–11, 40:11–41:7 (Ex. 37); Rinder Dep. at 11:24–12:19, 15:18–16:24, Oct. 5, 2011, *In re: Yaz/Yasmin/Ocella Prod. Liab. Litig*., Case No. 3:08-md-2100 (Ex. 71).

Dr. Rinder also joins with Plaintiffs' other experts in relying almost exclusively on case reports to demonstrate a link between TRT and VTE. And like Dr. Ardehali, this includes the incomplete and unpublished case reports known as adverse event reports. *See* Rinder Rpt., App'x A (Ex. 5). Dr. Rinder testified at his deposition that his methodology for conducting this review was to evaluate the subjects of the AERs as if they were his own patients and he was diagnosing them with VTE. Rinder Dep. 286:15–23 (Ex. 37). While he was familiar with the

reason for conducting post-marketing surveillance, Dr. Rinder admitted that he has no familiarity with the proper FDA methodology for evaluating AERs to determine potential safety signals, he did not review the FDA's Guidance on this topic, and he was not asked to evaluate whether or not there was a potential signal in this case. *Id.* at 283:17–285:21, 297:7–20. But the purpose of post-marketing AERs, according to the FDA Guidance that governs their use, is to evaluate potential safety signals according to the multi-step process outlined above. *See* FDA Pharmacovigilance Guidance at 3–4 (Ex. 20); *See supra* Section B.2.a (describing pharmacovigilance process for examining AERs). Dr. Arrowsmith applied this methodology to AbbVie's and the FDA's AER databases and concluded there was no VTE signal among AndroGel patients. Arrowsmith Rpt. at 136–37 (Ex. 39).

D. The Testimony of Plaintiffs' Specific Causation Experts Does Not Satisfy the Stringent *Daubert* Standard

In addition to proving general causation, Plaintiffs must prove that AndroGel actually caused each of their injuries. *Reference Manual,* at 609 (Ex. 14) ("The Plaintiff must establish not only that the defendant's agent is capable of causing disease, but also that it did cause the Plaintiff's injury.") "General causation is whether a substance is capable of causing a particular injury or condition in the general population and specific causation is whether a substance caused a particular individual's injury." *See, e.g.*, *Norris* v. *Baxter Healthcare Corp.*, 397 F.3d 878, 881 (10th Cir. 2005). Each Plaintiff must "demonstrate[ ] that [AndroGel] actually caused injury in h[is] particular case." *Guinn* v. *AstraZeneca Pharm. LP*, 602 F.3d 1245, 1248 n.1 (11th Cir. 2010).

Specifically, each Plaintiff must show that it is more likely than not that AndroGel caused or was a substantial contributing factor to his particular injury. *See In re Lipitor (Atorvastatin Calcium) Mktg., Sales Practices & Prod. Liab. Litig.*, 150 F. Supp. 3d 644, 649 (D.S.C. 2015).

In doing so, Plaintiffs' experts must first "rule in" AndroGel as a potential cause of their injuries by demonstrating a scientifically reliable causal connection between TRT and CV events or VTEs generally, and then must also "rule out" alternative potential causes for each of their specific injuries.

Plaintiffs offer case-specific testimony of five experts: three cardiologists (Drs. John Setaro, Phillip Cuculich, and Hossain Ardehali), a neurologist (Dr. Ronald Ziman), and a hematologist (Dr. Henry Rinder). Their opinions neither rule out the established non-TRT risk factors that apply to each of the Plaintiffs. Nor do they methodically rule in AndroGel.

1. The Specific-Causation Testimony of Plaintiffs' Case-Specific Experts Is Not the Product of Scientifically Reliable Principles, Methodologies, or Analyses

These experts vary to some degree in their approaches to case specific causation. But their analyses are uniformly flawed in that they fail to adhere to any of the basic steps of the methodologies have been found by the courts to be reliable under Daubert.

(a) <u>There Can Be No Specific-Causation Testimony Without Admissible Evidence of General Causation</u>

As an initial matter, because an "agent cannot be considered to cause the illness of a specific person unless it is recognized as a cause of that disease in general," *Reference Manual*, at 611 (Ex. 14), no expert can opine on specific causation without admissible evidence of general causation. *See, e.g.*, *Norris*, 397 F.3d at 887 (holding there was "no scientific basis" for any expert testimony on specific causation because the Plaintiffs' experts had failed to demonstrate general causation). Thus, if the Court excludes Plaintiffs' expert testimony on general causation due to its many flaws, the Court also must exclude Plaintiffs' expert testimony on specific causation.

(b) Plaintiffs' Specific-Causation Experts Use the Wrong Kind Of Differential Diagnosis

Plaintiffs' specific causation experts perform what they term a "differential diagnosis." Determining whether TRT, as opposed to other risk factors, caused a Plaintiff's particular injury is in fact an inherently comparative analysis. And the methodology used to make this comparison is sometimes referred to as "differential diagnosis." Beyond that, Plaintiffs' experts come to a fork in the road and take the wrong path. Plaintiffs' experts uniformly use their "clinical judgment" to make the differential diagnosis that predicates their opinions in this case.[28] But in the Seventh Circuit, the appropriate method is more properly referred to as "differential ***etiology,***" and it is nothing close to what Plaintiffs' experts perform in this case. *Myers* v. *Illinois Cent. R. Co.*, 629 F.3d 639, 644 (7th Cir. 2010) (emphasis supplied). "When a physician makes a differential diagnosis, he systematically compares and contrasts clinical findings from a patient's medical history to determine" the "identity of a specific ailment." *Id*. But where, as here, the question is not what a Plaintiff is "suffering from but what caused his ailments," a differential etiology is required. *Id*.

In a differential etiology, the expert "rules in all the potential causes of a patient's ailment and then by systematically ruling out causes that would not apply to the patient, the physician arrives at what is the likely cause of the ailment." *Id*. Whether a differential etiology is reliable under *Daubert* depends on "which potential causes should be 'ruled in' and which should be 'ruled out.'" *Id*.

[28] Setaro Dep. 130:5–24, 308:21–309:9 (Ex. 55); Ziman Dep. 269:21–271:9 (Ex. 56); Rinder Dep. 356:23–358:4, 364:6–21 (Ex. 37); Ardehali Dep. 92:19–944 (Ex. 7); Cuculich Dep. 128:25–130:24 (Ex. 8).

Plaintiffs' experts in this case therefore have made a basic category error by applying the methodologies used in their respective clinical practices, which they generally label as "differential diagnosis," to attempt to conduct a differential etiology. But whatever name they use, "[t]he ability to diagnose medical conditions is not remotely the same as the ability to deduce in a scientifically reliable manner, the causes of those medical conditions." *Tamraz*, 620 F.3d at 673. Moreover, "[s]imply claiming that an expert used the 'differential diagnosis' method is not some incantation that opens the *Daubert* gate." *Tamraz* v. *Lincoln Elec. Co.*, 620 F.3d 665, 674 (6th Cir. 2010). Other than Plaintiffs' experts' assertions that they performed a differential diagnosis, there is nothing in the record that suggests they actually did so, or if it was done, that it was reliable.

(c) Plaintiffs' Specific-Causation Experts Improperly "Rule In" AndroGel as a Potential Cause Because They Cannot "Connect the Dots" From the Science to Plaintiffs' Facts

Exposure to an agent alone is not a sufficient reason to rule it in as a potential cause. *Ervin v. Johnson & Johnson, Inc.*, 492 F.3d 901, 904-05 (7th Cir. 2007) ("The mere existence of a temporal relationship between taking a medication and the onset of symptoms does not show a sufficient causal relationship."). "[A]n expert must have a sufficient, reliable justification for ruling in each potential cause into the etiology in the first place." *Zimmer Prod. Liab. Litig.*, 2016 WL 6135685, at *12. One justification for ruling in a potential cause would be if the reliable scientific evidence "fit" the Plaintiff. This would require the Plaintiff to be among the group of people to which the science applies. "Only if the study subjects and the Plaintiff are similar with respect to other risk factors will a risk estimate from a study or studies be valid when applied to an individual." *Reference Manual,* at 613 (Ex. 14). This is equally true for individuals in sub-populations or with competing risk factors. *See id.* ("Thus, if all of the subjects in a study are participating because they were identified as having a family history of

heart disease, the magnitude of risk found in a study of smoking on the risk of heart disease cannot validly be applied to an individual without such a family history.").

The studies "fitting" the Plaintiff in question also must be evaluated using appropriate methodology and, if reliable, the application of those studies to Plaintiff must be evaluated. The application of epidemiological evidence to individuals requires consideration of many of the same issues relevant to general causation: whether the study's results are tainted by bias, whether the individual was exposed in a different manner or to a different dose, or whether the individual's history has different risk factors for the disease than the subjects of the study. *Id.* at 615-16.

In this case, Plaintiffs' specific causation experts fail to show that the Plaintiffs they assess actually fit into any subpopulations where a statistically significant association has been reported.

(d) Plaintiffs' Specific-Causation Experts Fail to Properly Compare and "Rule Out" Alternative Potential Causes

Once the science has been evaluated, and its fit to the Plaintiff has been established, then the causal factors must be compared to determine whether it is TRT, as opposed to an alternative cause, which caused the Plaintiff's injury. Like the decision to rule in a cause (here TRT), the decision to rule out causes must be systematic and Plaintiffs' experts must apply a reliable methodology. *Myers*, 629 F.3d at 644. Plaintiffs' experts must be able to say, based on a reliable, evidence-based methodology, that other potential causes alone could not have caused the Plaintiffs' injuries. *Clausen* v. *M/V NEW CARISSA*, 339 F.3d 1049, 1058 (9th Cir. 2003). Their experts must be able to support that conclusion with sufficient facts and data. The expert must provide reasons for rejecting alternative hypotheses "using scientific methods and procedures" and the elimination of those hypotheses must be founded on more than "subjective

beliefs or unsupported speculation." *Claar* v. *Burlington N. R.R. Co.*, 29 F.3d 499, 502 (9th Cir.1994).

As shown case-by-case below, Plaintiffs' experts examine other non-TRT risk factors that apply to the individual Plaintiffs and offer their clinical views, but they do nothing to rule out those factors scientifically. Yet an expert "cannot merely conclude that all risk factors for a disease are substantial contributing factors in its development. The fact that exposure to [a substance] may be a risk factor for [a disease] does not make it an actual cause simply because [the disease] developed." *Guinn*, 602 F.3d at 1255; *Viterbo* v. *Dow Chem. Co.,* 826 F.2d 420, 424 (5th Cir. 1987) (If a patient's "symptoms could have numerous causes," the expert cannot "simply pick[ ] the cause that is most advantageous to [the Plaintiff's] claim").

(e) Plaintiffs' Specific-Causation Experts Fail to Quantitatively Compare the Relative Contribution of Potential Causes

More specifically, none of the case specific experts report the quantified risk for those risk factors, as required to support any scientifically credible and reliable "more likely than not" opinion regarding AndroGel. *McDowell* v. *Brown*, 392 F.3d 1283, 1300-01 (1lth Cir. 2004) (excluding specific-causation testimony from experts who failed to "mete out" various contributors or "quantify" their effect on Plaintiff's injuries); *Magistrini* v. *One Hour Martinizing Dry Cleaning*, 180 F. Supp. 2d 584, 609–10 (D.N.J. 2002) (same).

An expert "cannot merely conclude that all risk factors for a disease are substantial contributing factors in its development. The fact that exposure to [a substance] may be a risk factor for [a disease] does not make it an actual cause simply because [the disease] developed." *Guinn*, 602 F.3d at 1255; *Viterbo* v. *Dow Chem. Co.,* 826 F.2d 420, 424 (5th Cir. 1987) (If a patient's "symptoms could have numerous causes," the expert cannot "simply pick[ ] the cause that is most advantageous to [the Plaintiff's] claim").

Nor are Plaintiffs' experts' claims that all of the risk factors somehow work together to produce the injuries alleged enough without quantifying the effect of AndroGel in comparison to the other factors. *See Haller* v. *AstraZeneca Pharm. LP*, 598 F. Supp. 2d 1271, 1306–07 (M.D. Fla. 2009) (mere reliance on a "straw that broke the camel's back" form of logic is improper under *Daubert*) Plaintiffs' experts' failure to meaningfully consider the combination of potential causes that alone "were each not likely the sole cause of his ailments," but together "could have been wholly responsible for" the Plaintiff's injuries renders Plaintiffs' experts' opinions as to specific causation inadmissible under Rule 702 and *Daubert*. *Brown* v. *Burlington N. Santa Fe Ry. Co*., 765 F.3d 765, 773–74 (7th Cir. 2014).

(f) Plaintiffs' Specific-Causation Experts' Opinions Amount to Conclusions Based Solely and Impermissibly on Temporal Relationship

Finally, an expert cannot reliably opine that AndroGel caused a bellwether Plaintiff's injury simply because he only suffered an injury *after* exposure to AndroGel. "[S]imply because a person takes drugs and then suffers an injury does not show causation. Drawing such a conclusion from temporal relationships leads to the blunder of the *post hoc ergo propter hoc* fallacy." *McClain* v. *Metabolife Int'l, Inc.,* 401 F.3d 1233, 1243 (11th Cir. 2005). "Temporal proximity is generally not a reliable indicator of a causal relationship [and] several factors make it especially unreliable in this case, including the facts that CV disease and VTE can be caused by many comorbid conditions and risk factors, and can take years to develop to the point where an individual suffers harm from a heart attack, stroke, or clot. *Guinn,* 602 F.3d at 1255. For the reasons explained below, many of Plaintiffs' experts' specific causation opinions are based solely on the temporal proximity of AndroGel exposure and a Plaintiff's injury, which requires the exclusion of their testimony as unreliable.

2. Plaintiffs' Experts Specific-Causation Testimony Should Be Excluded

(a) Dr. Setaro (Plaintiff Cribbs

(i) *Dr. Setaro Applies an Unreliable Differential Diagnosis Methodology*

Dr. Setaro describes his methodology as the same one he uses "in clinical practice, in treating patients and arriving at differential diagnoses." Expert Report of John F. Setaro, M.D. at 4, Oct. 31, 2016 (Ex. 57). As described above, clinical judgment used to treat patients is an unreliable application of differential etiology and for that reason alone Dr. Setaro's testimony should be prohibited. *See McClain*, 401 F.3d at 1253 (11th Cir. 2005).

(ii) *Dr. Setaro Improperly "Rules In" AndroGel as a Potential Cause Without a Connection to the Facts of Mr. Cribbs's Case*

Dr. Setaro fails to analyze the scientific evidence to determine whether any of it could actually apply to Mr. Cribbs. Dr. Setaro testified that, even though Mr. Cribbs was 61 years old at the time of his injury, general causation could be established for Mr. Cribbs by evidence that allegedly demonstrates increased CV risk in TRT users over the age of 65 because he was "close enough." Setaro Dep. at 166:17–167:13, 250:13–25, Jan. 18, 2017 (Ex. 55). But "close enough" is not sufficient under *Daubert*, particularly when the epidemiological evidence makes important distinctions in risk at the age of 65.

Additionally, Dr. Setaro rejects the diagnosis reached by Mr. Cribbs's own doctors that his MI was caused by vasospasm, but testified at his deposition that he never even read the transcripts of the doctors who treated Mr. Cribbs. Setaro Dep. at 62:25–35:3 (Ex. 55) But the "link between the facts and the conclusion" of expert testimony must be reliable, *ZF Meritor, LLC* v. *Eaton Corp.*, 696 F.3d 254, 291 (3d Cir. 2012). Upon examining Dr. Setaro's opinion, the link is nonexistent, let alone unreliable. This is unacceptable. Dr. Setaro has no factual basis

for reaching this opinion; it is merely *ipse dixit*. *See Joiner*, 522 U.S. at 146. *See also* Fed. R. Evid. 702 (expert testimony must be "based on sufficient facts or data" to be admissible).

(iii) *Dr. Setaro Fails to Properly Compare and "Rule Out" Alternative Potential Causes*

Dr. Setaro fails to properly apply the second step of differential diagnosis: ruling out the other obvious causes of Mr. Cribbs's MI. Dr. Setaro admits that Mr. Cribbs suffered from many cardiovascular risk factors, "includ[ing] family history, obesity, elevated lipids, hypertension, diabetes, and prior tobacco use," which "are known to promote vascular inflammation, and thereby to heighten the tendency toward athero-thrombotic events." Setaro Rpt. at 8–9 (Ex. 57). Yet, according to Dr. Setaro, TRT was the real cause of Mr. Cribbs's MI. Dr. Setaro performs no analysis of Mr. Cribbs's medical history and the risk factors therein. *See* Expert Report of William B. White, M.D., Re: Edward R. Cribbs at 17, Dec. 6, 2016 [hereinafter White Cribbs Rpt.] (Ex. 58 Failing to do so requires Dr. Setaro's opinion to be excluded under *Daubert* because he has not properly applied the methodology he claims to be using.

Instead of attempting to eliminate potential causes other than AndroGel, Dr. Setaro opines that "the adverse effects of testosterone therapy are proportionately greater in Mr. Cribbs compared with other individuals in the general population because of these pre-existing conditions and cardiovascular risk factors." Setaro Rpt. at 9 (Ex. 57). This is based on no scientific data whatsoever. No cardiovascular risk calculators accepted by the scientific community include testosterone as a risk factor that adds to the otherwise-present risk in a particular individual. Rather, these calculators rely *exclusively* on the risk factors that already plagued Mr. Cribbs: family history of cardiovascular events, obesity, diabetes, elevated lipids, and hypertension, to name only a few.

Indeed, Dr. Setaro has no response at all to Dr. White's calculation of Mr. Cribbs's 10-year risk of a CV event in the range of 35-40%. Nor does Dr. Setaro quantify the relative risk posed by testosterone as compared to the many other problems in Mr. Cribbs's history. *Cf.* White Cribbs Rpt. at 18–21 (Ex. 58) (calculating Mr. Cribbs's cardiovascular risk profile based on quantified measurements); *see, e.g. In re Lipitor*, 150 F. Supp. 3d at 660 (explaining that *Haller* district court properly excluded expert who "did not even attempt to quantify" the allegedly accelerated risk presented by defendant's product) (citing *Haller*, 598 F. Supp. 2d at 1279).

Dr. Setaro fails to cite to any study or evidence that testosterone *amplifies* the risk already presented by other well-accepted comorbid conditions. It is not a reliable methodology to "merely conclude that all risk factors for a disease are substantial contributing factors in its development." *Guinn*, 602 F.3d at 1255. Dr. Setaro's only explanation of his opinion is that "[t]he presence of these risk factors does not exculpate testosterone; rather testosterone therapy was a substantial factor in promoting long-term atherosclerotic arterial disease . . . by its effect on coagulation under conditions of generalized systemic vascular inflammation." Setaro Rpt. at 9 (Ex. 57). That testosterone is not "exculpate[d]," in Dr. Setaro's opinion, does not amount to "inculpating" it without further analysis. *Id*. Dr. Setaro's opinion is ultimately based on nothing other than temporality, which is insufficient as a methodology for the reasons explained above. *See, e.g.*, *McClain*, 401 F.3d 1233 at (11th Cir. 2005); *Guinn,* 602 F.3d at 1255.

For all of these reasons, Dr. Setaro's opinion should be excluded as unreliable.

(b) Dr. Ziman (Plaintiff Frost)

(i) *Dr. Ziman Applies an Unreliable Differential Diagnosis Methodology*

Dr. Ziman claims that he has "employed the same methodologies that [he uses] in clinical practice, namely deductive reasoning and differential diagnosis." Expert Report of Ronald B. Ziman, M.D., F.A.C.P., F.A.A.N. at 2, Oct. 31, 2016 (Ex. 59). The methodology used in clinical practice is not the same as the methodology required to come to a conclusion of causation. In essence, Dr. Ziman's opinion is based on: (1) his acceptance of Dr. Ardehali's opinion that testosterone can generally cause stroke and (2) the timing between Mr. Frost's alleged use of AndroGel and his stroke. Thus, temporality lies at the heart of Dr. Ziman's testimony on case specific causation

The primary conclusion that Dr. Ziman makes with respect to AndroGel is that "[t]he time frame and interval between re-initiation of the AndroGel therapy and Mr. Frost's ischemic stroke is consistent with the time frame of onset of testosterone therapy's physiological prothrombotic effects." Ziman Rpt. at 14. But, as previously discussed, temporal relationship alone is an unreliable basis for an expert opinion on specific causation under *Daubert*. *See, e.g.*, *Guinn,* 602 F.3d at 1255 ("The fact that exposure to [a substance] may be a risk factor for [a disease] does not make it an actual cause simply because [the disease] developed."); *McClain* v. *Metabolife Int'l, Inc.*, 401 F.3d 1233, 1243 (11th Cir. 2005) ("[S]imply because a person takes drugs and then suffers an injury does not show causation. Drawing such a conclusion from temporal relationships leads to the blunder of the *post hoc ergo propter hoc* fallacy."); *In re Lipitor*, 150 F. Supp. 3d at 654 ("In other words, according to Dr. Murphy, if general causation exists (i.e., Lipitor has the *potential* to cause diabetes), then specific causation exists for every patient that took Lipitor and thereafter developed diabetes. Again, this logic is flawed.").

(ii) *Dr. Ziman Fails to Properly Compare and "Rule Out" Alternative Potential Causes*

Additionally, Dr. Ziman does not address the other obvious potential causes of Mr. Frost's stroke, including smoking, hypertension, diabetes, hyperlipidemia, depression, alcohol abuse, and family history of heart disease. Expert Report of Dr. William J. French, M.D., Re: Cecile B. Frost at 12 [hereinafter French Frost Rpt.] (Ex. 60). Dr. Ziman does not rule out these risk factors. He opines merely that Mr. Frost's stroke was due to his "heightened prothrombotic state due to his multiple atherosclerotic risk factors," which was "enhance[ed]" by AndroGel. Ziman Rpt. at 14. This "enhancement" opinion, which is untested and hypothetical, is impermissible under *Daubert*. *See Joiner*, 522 U.S. at 139–40 (excluding expert opinion that chemical "promoted" cancer in Plaintiff" already at a heightened risk").

(iii) *Dr. Ziman's Opinion Is Not Connected to the Facts of Mr. Frost's Case*

Dr. Ziman's temporality opinion is expressly based on a factual assumption that find no support in the record. Specifically, Dr. Ziman assumes that Mr. Frost was continuing to use AndroGel at the time of his stroke because Mr. Frost had additional AndroGel left over from a prescription filled one year prior to his stroke and that he started on it in January when he was told to resume AndroGel. Ziman Rpt. at 6 (Ex. 59); Ziman Dep. 110:2–112:7, Jan. 26, 2017 (Ex. 56). But Dr. Ziman identified no evidence to support this assumption, and even Frost's testimony expressly contradicts it. Ziman Dep. 93:11–94:15 (Ex. 56); Frost Dep. 229:30-230:7, May 4, 2016 (Ex. 61). Thus, Dr. Ziman's opinion must be excluded under Rule 702 and *Daubert* because it lacks "sufficient facts or data" to be reliable, and does not fit the facts of Mr. Frost's case. Fed. R. Evid 702 (requiring an expert to "reliably appl[y] the principles and methods to the facts of the case" for testimony to be admissible); *Ervin*, 492 F.3d at 904; *Westberry*, 178 F.3d at 262; *Guinn* , 602 F.3d at 1253; *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d at 759.

(c) Dr. Cuculich (Plaintiff Konrad)

(i) *Dr. Cuculich Applies an Unreliable Differential Diagnosis Methodology*

Dr. Cuculich summarizes his specific causation opinion as follows:

> In the absence of other significant cardiovascular risks, and with the recent initiation of exogenous testosterone, in particular for short duration of therapy (approximately 6 weeks) exogenous testosterone was a substantial contributing factor in causing Mr. Konrad's myocardial infarction. This is consistent with studies discussed in my general opinion report, e.g. Finkle 2014; Wallis 2016.

Thus, Dr. Cuculich essentially ascribes the cause of Mr. Konrad's MI to AndroGel based on the fact that he was only on the drug for six weeks before he was injured. Cuculich Dep. 257:16–258:19 (Ex. 8). But temporality alone is an insufficient basis for a specific causation opinion. *See, e.g.*, *Guinn*, 602 F.3d at 1255; *McClain*, 401 F.3d at 1243 (11th Cir. 2005) ("[S]imply because a person takes drugs and then suffers an injury does not show causation."); *In re Lipitor*, 150 F. Supp. 3d at 654.

(ii) *Dr. Cuculich Improperly "Rules In" AndroGel as a Potential Cause Without a Connection to the Facts of Mr. Konrad's Case*

Dr. Cuculich's opinion is flawed for a number of additional reasons. Dr. Cuculich "rules in" AndroGel on the basis of the Finkle study. But he fails to consider two serious problems with the Finkle study that are relevant Mr. Konrad: (1) that Mr. Konrad's underlying hypogonadism could have caused his MI and (2) that Finkle is only applicable to the extent that Afib patients were included in the sub-group for which Finkle found an association and Mr. Konrad has Afib. But there is no dispute that Afib did not play any role in Mr. Konrad's MI. (cite).

Dr. Cuculich acknowledges that he cannot identify a mechanism by which AndroGel caused Mr. Konrad's injury, but he nonetheless rules it in as a potential cause. At his deposition,

he testified that there are no clinical measurements of thromboxane or estradiol that could trace an individual's MI to TRT exposure and that there was no specific level of hematocrit at which an individual's risk of MI is increased. Cuculich Dep. at 200:8–202:23 (Ex. 8).

(iii) *Dr. Cuculich Fails to Properly Compare and "Rule Out" Alternative Potential Causes*

Dr. Cuculich also testified that, in his opinion, Mr. Konrad did not have heart disease or cardiovascular risk factors. Cuculich Dep. at 264:23–265:17 (Ex. 8). This opinion is belied by the record: Mr. Konrad suffered from a number of conditions that are accepted causal factors for MI, including poor cholesterol levels, hypertriglyceridemia, and hypertension.. *See* Expert Report of William J. French, M.D., Re; Jeffrey Konrad at 8, Dec. 7, 2016 [hereinafter French Konrad Rpt.] (Ex. 62). Dr. Cuculich's failure to properly address these other risk factors renders his opinion inadmissible under *Daubert*. *See, e.g.*, *Westberry*, 178 F.3d at 262 ("differential diagnosis that fails to take serious account of other potential causes may be so lacking that it cannot provide a reliable basis for an opinion on causation.").

Asked about this, Dr. Cuculich testified at his deposition that he simply did not need to rule out other potential contributing factors to Mr. Konrad's MI because all of them, including testosterone, play a role. Cuculich Dep. at 298:13–300:10 (Ex. 8). Dr. Cuculich based this opinion on standard CV risk calculation which was not in his report and has not been produced. This type of calculation estimates the synergistic effect of various risk factors in a particular individual. *Id.* at 300:12–301:15. But Dr. Cuculich also testified that testosterone is not one of the risk factors considered by the calculator. *Id.* at 301:16–302:9. Therefore, Dr. Cuculich's opinion that testosterone contributed to the synergistic effect of Mr. Konrad's other risk factors is based on no calculation or analysis. This makes his opinion unreliable under *Daubert* and Rule

702. *See* Fed. R. Evid. 702 (requiring expert testimony to be based on "sufficient facts or data" to be admissible).

(d) Dr. Ardehali (Plaintiff Mitchell)

(i) *Dr. Ardehali Improperly "Rules In" AndroGel as a Potential Cause Without a Connection to the Facts of Mr. Mitchell's Case*

Dr. Ardehali testified at his deposition that he did not know which of the mechanisms he discussed in his general report was the pathway by which AndroGel caused Mr. Mitchell's MI. Ardehali Dep. at 248:3–21, 255:8–262:7 (Ex. 7). Dr. Ardehali testified that he does not know how consistently Mr. Mitchell was using AndroGel and that he cannot explain the fact that Mr. Mitchell's serum testosterone levels were inconsistently above or below the hypogonadal level. Ardehali Dep. at 264:2–271:17 (Ex. 7).

(ii) *Dr. Ardehali Fails to Properly Compare and "Rule Out" Alternative Potential Causes*

Dr. Ardehali admits that Mr. Mitchell suffered from many diseases that are risk factors for MI, but opines that "[t]hese risk factors increase the atherosclerotic burden and associated chronic inflammatory state on which testosterone acts to cause an acute coronary artery thrombotic event. Expert Report of Hossein Ardehali, M.D., Ph.D., Re: Jess Mitchell at 9, Oct. 26, 2016 [hereinafter Ardehali Mitchell Rpt.] (Ex. 63). He testified at his deposition that Mr. Mitchell's MI was substantially caused by his hypertension and many other risk factors. Ardehali Dep. at 238:5–241:17 (Ex. 7). But it is not a reliable methodology to "merely conclude that all risk factors for a disease are substantial contributing factors in its development." *Guinn*, 602 F.3d at 1255; *see Haller*, 598 F. Supp. 2d at 1306–07. For this reason, Dr. Ardehali's specific causation opinion must be excluded.

Dr. Ardehali also testified that, in the absence of AndroGel, Mr. Mitchell's risk of an MI in 10 years was 15%, but that his risk of an MI with AndroGel was 100% because that is what happened. Ardehali Dep. at 242:9–21 (Ex. 7). In essence, Dr. Ardehali opines that testosterone was the "straw that broke the camel's back" in the context of Mr. Mitchell's many comorbid conditions. But this type of reasoning was rejected as unreliable and unscientific in the *Haller* decision. *See Haller*, 598 F. Supp. 2d at 1306-07 (excluding specific causation expert's opinion that drug added to the Plaintiff's risk of developing diabetes according to a "straw/camel" analogy on reliability grounds). Thus, for this additional reason, Dr. Ardehali's specific causation opinion should be excluded under *Daubert*.

(e) Dr. Rinder (Plaintiffs Myers, Nolte, and Rowley)

(i) *Dr. Rinder Applies an Unreliable and Undefined Methodology*

Dr. Rinder has admitted that his conclusion that AndroGel was a substantial contributing factor in causing the VTE injuries of Mr. Myers, Mr. Nolte, and Mr. Rowley is not based on any objective test. Rinder Dep. at 347:12–351:14 (Ex. 37). He further acknowledged that another doctor could not reproduce his methodology to reach the same conclusion as to these three Plaintiffs. *Id*. at 356:10–358:4. He thus admits that there is no testable methodology underlying any of his specific causation opinions in this case. Rather, his "bare assertion that [AndroGel] was a 'substantial contributing factor' in [Plaintiffs'] development of [VTEs] amounts to nothing more than inadmissible *ipse dixit*, as 'the only connection between the conclusion and the existing data is the expert's own assertions.'" *Guinn*, 598 F. Supp. 2d at 1243 (citation omitted); *see also Haller*, 598 F. Supp. 2d at 1297 (excluding specific causation expert in part because he could not determine that defendant's drug was the but-for cause of Plaintiff's injury). Because

Dr. Rinder relies on no reliable methodology to reach his conclusions, his opinion should be excluded.

Moreover, Dr. Rinder acknowledged that "substantial contributing factor" is not a term he uses in medical practice and he failed to define its meaning in his report. Rinder Dep. at 351:15–352:14 (Ex. 37). Thus, his methodology in this case does not match the work he does outside of litigation, making it unreliable under *Daubert*. *Kumho Tire Co.* v. *Carmichael*, 526 U.S. 137, 152 (1999) (Courts must ensure the expert "employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.").

(ii) *Dr. Rinder Improperly "Rules In" AndroGel as a Potential Cause Without a Connection to the Facts of Any Plaintiff's Case*

Dr. Rinder has no basis to rule in TRT as a potential cause. Dr. Rinder testified that all of the epidemiological studies are flawed. Rinder Dep. at 187:18–188:7 (Ex. 37). He also testified that the available science on general causation of VTE is inapplicable to the question of specific causation. *Id.* at 328:4–330:4, 331:12–332:5. And, he testified that AndroGel exposure was a contributing factor to each of Plaintiffs' injuries, but that he had not ruled out other obvious potential causes. *Id.* at 363:16–365:4.

Yet, he also claims that TRT can be ruled in for each bellwether Plaintiff based on one of the subgroups analyzed in the Martinez (2016) study. Rinder Dep. 332:6–333:2, 365:5–366:7, 372:21–373:9 (Ex. 37). However, as Dr. Rinder admitted at his deposition, none of the subgroups to which Mr. Myers, Mr. Nolte, or Mr. Rowley belongs had a statistically significant finding of increased risk. *Id.* at 332:6–333:24, 365:5–366:20, 372:21–374:16. There are many additional methodological problems with Dr. Rinder's attempt to use the Martinez study: he goes beyond the conclusions reached by the authors (who found no statistically significant risk for any

of these subgroups), *Id.* at 251:16–253:10, and the Martinez study was highly flawed and does not provide a sufficient basis for ruling in testosterone as a potential cause of Plaintiffs' injuries.

Further, Dr. Rinder testified that he has no information about which mechanism was the biological pathway by which AndroGel could have caused any Plaintiff's VTE injury, let alone *did* cause a particular Plaintiff's injury. *Id.* at 343:18–345:14, 366:22–367:12, 371:24–372:20.

(iii) *Dr. Rinder Fails to Properly Compare and "Rule Out" Alternative Potential Causes*

Dr. Rinder admitted that he did not rule out risk factors other than AndroGel. Rinder Dep. at 363-64 (Ex. 37). He also testified that he cannot quantify how much risk is due to each factor, including AndroGel. Rinder Dep. at 345:16–25, 370:6–23 (Ex. 37). These are methodological flaws that are clear to anybody with even a passing understanding of *Daubert* law.

Inability to quantify the risk due to an allegedly harmful drug is a methodological flaw that has been cited in previous cases to exclude specific causation opinions as unreliable, particularly for diseases that have well-known and already quantified risk factors. *See Guinn*, 598 F. Supp. 2d at 1247 (excluding expert who testified "it [wa]s not possible" to quantify relative contribution of Plaintiff's background risk factors and defendant's drug as unreliable); *Haller*, 598 F. Supp. 2d at 1283, 1298-99 (excluding specific causation opinion from expert who testified that he could not quantify risk presented by defendant's drug in addition to Plaintiff's other risk factors as aspect of unreliable opinion that drug was the "straw that broke the camel's back").

Each of the Plaintiffs' cases addressed by Dr. Rinder involved multiple other risk factors for VTEs, each of which has been quantified by substantial epidemiological evidence. Dr. Rinder readily admitted that generally risk factors *can* be quantified, but that he simply did

not do so for any of the Plaintiffs he opined about. Rinder Dep. at 327, 329–31 (Ex. 37). He conceded that no objective test exists to determine whether AndroGel, as opposed to any of the Plaintiffs' other comorbidities, actually caused an injury. Rinder Dep. at 348 (Ex. 37). Dr. Rinder's failure to reconcile AndroGel exposure in relation to these other risk factors renders his opinion unreliable.

Additionally, Dr. Rinder admits that he did not apply the legally required standard for specific causation. It is **not** his opinion that the injuries alleged would have occurred but for the Plaintiffs' TRT use. Rinder Dep. at 349-50 (Ex. 37). Rather, he limited himself to a determination that AndroGel was a "substantial contributing factor," a standard that he argues cannot be quantitatively measured, has not been published in any peer-reviewed paper, and cannot be reproduced. Rinder Dep. at 351-54, 356-57 (Ex. 37). But an expert "cannot simply opine that all present risk factors are substantial contributing factors," without any methodology at all. *Norris* v. *Baxter Healthcare Corp.*, 397 F.3d 878, 881 (10th Cir. 2005). For all of these reasons, Dr. Rinder's testimony as to specific causation should be excluded as unreliable under *Daubert* and Rule 702.

(iv) *All Plaintiffs Save Nolte Always Had Hematocrit In Normal Lab Range (generally <50%) and There Is No Evidence of Mechanism*

As previously noted, having hematocrit above 50% is a cross-cutting issue connected with Plaintiffs' mechanism theories for clotting risk. Two of three of the bellwether Plaintiffs who allege VTE injuries, however, always had hematocrit in the normal range. Yet Dr. Rinder can cite no study indicating that patients with *normal* hematocrit levels have an increased risk of VTE. Rinder Dep. at 314 (Ex. 37). Nor does he think it is even possible to design a study that would support his hematocrit theory. *Id*.

Dr. Rinder also testified that he has no opinion on whether the hematocrit levels for any of the three Plaintiffs changed when they started treatment. Rinder Dep. at 342 (Ex. 37). This is a plain "failure of the expert[] to connect the dots from the studies to the illnesses," and cannot survive *Daubert*. *Wood* v. *Textron, Inc.*, 807 F.3d 827, 837 (7th Cir. 2015).

Moreover, Dr. Rinder's mechanism opinions are so theoretical he does not even attempt to apply them to the Plaintiffs whose allegations he is supporting. Estradiol is just one example. He has not studied the conversion rate from TRT to estradiol, admits that estradiol cannot be extrapolated from the mere use of TRT, and agrees that the estradiol levels of the Plaintiffs were not a data point he could consider. Rinder Dep. at 335, 339, 366, 370 (Ex. 37). But "a biological explanation without evidence of the mechanism by which it works is merely an unproven hypothesis, a theory," insufficient to be put before a jury. *In re Accutane Prods. Liab.*, 511 F. Supp. 2d 1288, 1295 (M.D. Fla. 2007).

Dr. Rinder admits that he has no data from which he can opine that a change in estradiol is enough to have a significant impact on VTE. Rinder Dep. at 310 (Ex. 37). But having no data does not stop him from opining that, via its effect on estradiol, TRT was a substantial factor in causing Plaintiffs' injuries. The Court should conclude that "there is simply too great an analytical gap between the data and the opinion proffered." *Gen. Elec. Co.* v. *Joiner*, 522 U.S. 136, 146 (1997).

II. SUMMARY JUDGMENT IS REQUIRED BECAUSE PLAINTIFFS LACK SUFFICIENT ADMISSIBLE EVIDENCE OF GENERAL AND SPECIFIC CAUSATION

"The court shall grant summary judgment if the movant shows that there is no genuine dispute as to any material fact and the movant is entitled to summary judgment as a matter of law." Fed. R. Civ. P. 56. The "plain language" of Rule 56 "mandates the entry of summary judgment . . . against a party who fails to make a showing sufficient to establish the existence of

an element essential to that party's case, and on which the party will bear the burden of proof at trial." *Celotex Corp.* v. *Catrett*, 477 U.S. 317, 322 (1986). A "failure of proof concerning an essential element of the nonmoving party's case necessarily renders all other facts immaterial." *Id.* at 323.

Causation is a required element of all of Plaintiffs' claims. *See, e.g.*, *Cooper* v. *Smith & Nephew, Inc.*, 259 F.3d 194, 203 (4th Cir. 2001). "To establish medical causation in a product liability case, a Plaintiff must show both general and specific causation." *In re Bausch & Lomb*, 693 F. Supp. 2d 515, 518 (D.S.C. 2010). "[C]ases involving pharmaceuticals, toxins or medical devices involve complex questions of medical causation beyond the understanding of a lay person," and thus expert testimony is required. *In re Baycol Prods. Litig.*, 321 F. Supp. 2d 1118, 1126 (D. Minn. 2004); *see also Silverstein* v. *Procter & Gamble Mfg. Co.*, 700 F. Supp. 2d 1312, 1316 (S.D. Ga. 2009) ("[I]f the inference that the defendant's product caused the Plaintiff's injury is not a 'natural inference that the juror could make through human experience . . . medical expert testimony [is] essential to prove causation.'") (alterations in original) (quoting *Allison* v. *McGhan Med. Corp.*, 184 F.3d 1300, 1320 (11th Cir. 1999)); *Zellers*, 533 F. App'x at 200 (proof of causation requires "relevant and reliable expert testimony, as the health effects of toxic exposure to chemicals are beyond the knowledge and experience of the average layperson").

Here, because Plaintiffs lack admissible expert testimony on both general and specific causation, "there is a complete failure of proof on the critical element of causation," and thus summary judgment is warranted. *Id.*; *see also C.W. ex rel. Wood* v. *Textron, Inc.*, 807 F.3d 827, 838 (7th Cir. 2015) ("With no experts to prove causation . . . the appellants cannot prove their toxic-tort case . . . [and] summary judgment . . . was proper.") (citation omitted); *Wells* v. *SmithKline Beecham Corp.*, 601 F.3d 375, 381 (5th Cir. 2010) ("[w]ithout the expert testimony,

Plaintiff "cannot prove general causation" and thus "judgment ***must*** be entered for" defendant) (emphasis added); *Rutigliano* v. *Valley Bus. Forms*, 929 F. Supp. 779, 783 (D.N.J. 1996) (where "expert opinion evidence regarding causation is inadmissible . . . summary judgment ***must*** be granted to defendants") (emphasis added), *aff'd*, 118 F.3d 1577 (3d Cir. 1997).

Indeed, courts around the country routinely grant summary judgment where, as here, Plaintiffs failed to pass the *Daubert* threshold on general causation. *See, e.g., Chapman* v. *Procter & Gamble Distrib., LLC*, 766 F.3d 1296, 1316-17 (11th Cir. 2014); *Wells*, 601 F.3d at 381; *Ruggiero* v. *Warner-Lambert Co.*, 424 F.3d 249, 254-55 (2d Cir. 2005); *Norris* v. *Baxter Healthcare Corp.*, 397 F.3d 878, 884-86 (10th Cir. 2005); *Miller* v. *Pfizer, Inc.*, 356 F.3d 1326, 1335-36 (10th Cir. 2004); *Hayes* v. *Raytheon Co.*, 23 F.3d 410 (7th Cir. 1994); *In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 2016 WL 1320799, at *10-11 (E.D. Pa. Apr. 5, 2016); *In re Bausch & Lomb Inc. Contacts Lens Sol. Prods. Liab. Litig.*, 2010 WL 1727807, at *1-3 (D.S.C. Apr. 26, 2010); *In re Bausch & Lomb*, 693 F. Supp. 2d at 517-19; *In re Viagra Prods. Liab. Litig.*, 658 F. Supp. 2d 950, 967-69 (D. Minn. 2009); *In re Human Tissue Prods. Liab. Litig.*, 582 F. Supp. 2d 644, 690-91 (D.N.J. 2008); *In re Rezulin Prods. Liab. Litig.*, 441 F. Supp. 2d 434, 577-78 (W.D. Pa. 2003); *Siharath* v. *Sandoz Pharms. Corp.*, 131 F. Supp. 2d 1347, 1373-74 (N.D. Ga. 2001); *Wade-Greaux* v. *Whitehall Labs., Inc.*, 874 F. Supp. 1441, 1485-86 (D.V.I. 1994), *aff'd*, 46 F.3d 1120 (3d Cir. 1994).

In the *Zoloft* MDL, for example, after excluding the opinions of Plaintiffs' general causation experts as unreliable and inadmissible, the court held that summary judgment was warranted because "[a]t the end of the day, Plaintiffs have failed to raise a jury question on the necessary predicate to success in any case: that Zoloft was capable of causing their injuries." *In re Zoloft*, 2016 WL 1320799, at *11.

Likewise, courts routinely grant summary judgment where, as here, Plaintiffs failed to pass the *Daubert* threshold on specific causation. *See, e.g., Wilson* v. *Taser Int'l, Inc.*, 303 F. App'x 708, 716 (11th Cir. 2008) (granting summary judgment against Plaintiff who failed to present an expert who could testify to a reasonable degree of medical certainty that defendant's product cause his injuries); *In re Zimmer Nexgen Knee Implant Prod. Liab. Litig.*, No. 11 C 5468, 2016 WL 6135685, at *15 (N.D. Ill. Oct. 21, 2016) (requiring expert testimony on specific causation in case where Plaintiff's injury could have been the result of many potential causes, and granting summary judgment when Plaintiffs failed to produce admissible expert testimony on specific causation); *Haller* v. *AstraZeneca Pharmaceuticals LP*, 598 F. Supp. 2d 1271, 1275 (M.D. Fla. 2009) ("Indeed, because establishing medical causation is critical to each of Haller's claims, AstraZeneca is entitled to judgment as a matter of law as to all claims set forth in the complaint if Haller fails to establish a genuine issue of material fact with regard to whether his ingestion of Seroquel caused him to develop diabetes."); *see also Oines* v. *Bridgestone/Firestone N. Am. Tire, LLC*, 93 F. App'x 80, 81 (7th Cir. 2004) (affirming grant of summary judgment against Plaintiff who, despite establishing general causation, failed to produce any evidence of specific causation).[29]

[29] *See also Chapman* v. *Procter & Gamble Distrib., LLC*, 766 F.3d 1296, 1316–17 (11th Cir. 2014) (affirming summary judgment for defendant on the ground that Plaintiffs lacked admissible expert testimony on specific causation); *Neal-Lomax* v. *Las Vegas Metro. Police Dep't*, 574 F. Supp. 2d 1193, 1207 (D. Nev. 2008) ("Plaintiffs have failed to produce any admissible evidence stating to a reasonable degree of medical certainty that the Taser caused or contributed to Lomax's death. Plaintiffs therefore have not raised a genuine issue of material fact as to causation, and the Court will grant Defendant TI's motion for summary judgment."), *aff'd,* 371 F. App'x 752 (9th Cir. 2010); *Smith* v. *Gen. Motors Corp.*, 376 F. Supp. 2d 664, 667–68 (W.D. Va. 2005) ("Modern case law requires expert medical testimony to establish causation in cases where the Plaintiff has suffered a complex injury."), *aff'd sub nom. Estate of Smith* v. *Gen. Motors Corp.*, 179 F. App'x 890 (4th Cir. 2006).

For example, in the *Seroquel* litigation, the district court granted summary judgment in two separate cases on the ground that Plaintiffs had failed to produce admissible expert testimony demonstrating specific causation. *See Guinn*, 598 F. Supp. 2d at 1246-47, *aff'd* 602 F.3d 1245 (11th Cir. 2010); *Haller*, 598 F. Supp. 2d at 1304-06. In those cases, Plaintiffs alleged that defendant's anti-psychotic medication caused them to develop diabetes. In the *Guinn* case, after concluding that Plaintiff's specific causation expert's testimony amounted to nothing more than speculation as to whether Seroquel was a substantial contributing factor in her development of diabetes, the court held that no reasonable juror could find that Seroquel was, more likely than not, a substantial contributing factor in causing the Plaintiff's injury. *See Guinn*, 598 F. Supp. 2d at 1247-48. This ruling was upheld on appeal. *See Guinn*, 602 F. 3d at 1257; *see also Haller*, 598 F. Supp. 2d at 1306-07 (granting summary judgment on the ground that no reasonable jury could find that Seroquel was a substantial contributing factor based on specific causation expert's opinion that drug added to the Plaintiff's risk of developing diabetes according to a "straw/camel" analogy).

The same reasoning applies here. Plaintiffs' specific causation experts are merely speculating that AndroGel, and not the many other risk factors that plagued Plaintiffs, was the cause of their injuries. They have no methodology for distinguishing or quantifying the risk presented by AndroGel alone, and any reliance on a "straw that broke the camel's back" form of logic is improper under *Daubert*. *See Haller*, 598 F. Supp. 2d at 1306-07. Plaintiffs must present expert testimony on this issue because a lay jury is not equipped to differentiate the risk presented by AndroGel from other complex potential causes of Plaintiffs' injuries. *See Wills* v. *Amerada Hess Corp.,* 379 F.3d 32, 46 (2d Cir. 2004) ("[W]here an injury has multiple potential etiologies, expert testimony is necessary to establish causation"). Because Plaintiffs have failed

to meet their burden of demonstrating that their experts' specific causation opinions are admissible, no reasonable jury could conclude that AndroGel was the specific cause of any of their injuries.

Accordingly, because Plaintiffs' experts cannot demonstrate reliable scientific evidence of general or specific causation, AbbVie is entitled to summary judgment and all of Plaintiffs claims should be dismissed.

Finally, even if the Court were to find that Plaintiffs' experts have raised a triable issue of fact with respect to certain subgroups of TRT users, the Court still should grant summary judgment with respect to those subgroups for which Plaintiffs have not even attempted to raise any triable issue of fact. In terms of CV risk, that includes patients under 65 with no preexisting CVE history, a subgroup for which no study has ever suggested that there is any increased risk from using TRT. With respect to VTE risk, that includes long-term users of TRT, a subgroup for which there is no study suggesting any increased risk from using TRT.

CONCLUSION

For the foregoing reasons, AbbVie respectfully requests that the Court exclude Plaintiffs' experts from testifying that AndroGel causes CVEs or VTEs, and grant summary judgment dismissing Plaintiffs' claims.

Dated: February 20, 2017

Respectfully submitted,

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

Attorneys for AbbVie Inc.

Nathan Hoffman
DECHERT LLP
35 West Wacker Drive, Suite 3400
Chicago, IL 60601
Tel: (312) 646-5827
Fax: (312) 646-5858
nathan.hoffman@dechert.com

Hope S. Freiwald
Friedrich-Wilhelm W. Sachse
DECHERT LLP
2929 Arch St., Cira Centre
Philadelphia, Pennsylvania 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, Michelle Hart Yeary, hereby certify that on February 20, 2017, 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/*Michelle Hart Yeary*
Michelle Hart Yeary

CROSS CUTTING ISSUES RE CAUSATION

CV General Causation

(Gerstman, Ardehali, Cuculich)

1. <65, short term use (90 days), w/ history of CV events
2. >65, short term use (<3, <6, <12 months)
3. All other Users (Overall Users)

VTE General Causation

(Rinder)

4. <6 months, without existing VTE risk factors
5. All other Users (Overall Users)
6. Hct ≤ 50% (Normal lab range)
7. Hct >50% (Above normal lab range)

Case Specific Causation

Issue

1. **Konrad** (Cuculich)
3. **Mitchell** (Ardehali)
3. **Cribbs** (Setaro)
3. **Frost** (Ziman)
5 6. **Myers** (Rinder)
5 6. **Rowley** (Rinder)
5 7. **Nolte** (Rinder)

ALBERT (2016): FINDING RE TRT USERS OVERALL

Figure 2

CV Studies Summary Chart

Red = statistically significant increased risk
Black = no statistically significant increased risk
Green = statistically significant decreased risk

Study Type	Overall	Age < 65	Age ≥ 65	Short / 1st Use (<3, <6, or ≤12 mos.)	Long-Term Use (>12 mos.)	Gel v. IM	Pre-Existing Disease
Meta-analysis	1. Calof (2005) (19 RCTs) 2. Haddad (2007) (6 RCTs) 3. Fernandez-Balsells (2010) (7 RCTs) 4. Xu (2013) (27 RCTs) CV Events & Serious CV Events 5. Xu (2013) (27 RCTs) CV-Related Death 6. Ruige (2013) (10 RCTs) 7. Corona (2014) (75 RCTs) 8. Borst (2014) (35 RCTs) 9. Corona & Maggi (2015) (24 RCTs) 10. Albert (2016) (45 RCTs) 11. Alexander (2016) (30 RCTs)	1. Albert (2016) (45 RCTs)	1. Corona (2014) (10 RCTs - MACE) 2. Albert (2016) (45 RCTs)	1. Corona (2014) (4 RCTs ≤12 wk) 2. Albert (2016) (45 RCTs) if <12 mos. or ≥65 and <12 mos. use or <12 mos. transdermal use 3. Albert (2016) (45 RCTs) if <65 and <12 mos. use or <12 mos. IM	1. Corona (2014) (22 RCTs >12 wk) 2. Albert (2016) (45 RCTs)	1. Borst (2014) (35 RCTs) 2. Albert (2016) (45 RCTs)	1. Corona (2014) (2 RCTs – CVD history; 4 RCTs – Metabolic Syndrome) 2. Alexander (2016) (30 RCTs)
RCTs			1. Basaria (2015) 2. Snyder (2016)	1. Basaria (2010)			1. Copenhagen (1986)

Figure 3

CV Studies Summary Chart

Study Type	Overall	Age < 65	Age ≥ 65	Short / 1st Use (<3, <6, or ≤12 mos.)	Long-Term Use (>12 mos.)	Gel v. IM	Pre-Existing Disease
Observational	1. Shores (2012) no diabetes (Mortality) 2. Shores (2012) Overall Mortality; Mortality with no CVD 3. Anderson (2015) Normal v. Low T: 1 yr MACE; 1 yr Death; 3 yr MI; 1 & 3 yr Stroke; 1 & 3 yr Coronary Death; 1 & 3 yr CV Death; 1 & 3 yr CV-MACE 4. Anderson (2015) Normal v. Low T: 3-yr MACE; 3 yr Death; 1 yr MI 5. Sharma (2015) Non-normalized v. no T (MI & Stroke); normalized v. no T (MI) 6. Sharma (2015) Normalized T v. no T (MI, Stroke & All-Cause Mortality); normalized T v. non-normalized T (MI, Stroke & All-Cause Mortality); non-normalized T v. no T (All-Cause Mortality) 7. Tan (2015) 8. Maggi (2016)	1. Shores (2012) 2. Maggi (2016)	1. Shores (2012) Mortality 2. Shores (2012) Mortality with diabetes 3. Baillargeon (2014) 4. Etminan (2015) 5. Ramasamy (2015) 6. Maggi (2016) 7. Wallis (2016) CV/VTE 8. Wallis (2016) Mortality	1. Finkle (2014) if age <65, CVD history and ≤3 mos. use or age ≥65 and ≤3 mos. use 2. Finkle (2014) if age <65 overall and <3 mos. use or age <65, no CVD history and <3 mos. use 3. Baillargeon (2014) 4. Etminan (2015) 1st use or 1st use gel/spray 5. Etminan (2015) 1st IM use 6. Wallis (2016) ≥65 with 2 mos. mean use (CV/VTE & Mortality) or ≥65 and 9 mos. mean use (CV/VTE) 7. Wallis (2016) ≥65 and 9 mos. mean use (Mortality)	1. Wallis (2016)	1. Etminan (2015)	1. Shores (2012) 2. Vigen (2013) 3. Baillargeon (2014) 1st – 3rd quartiles of MI risk 4. Baillargeon (2014) highest quartile of MI risk 5. Etminan (2015)

Figure 3

Study Type	Overall	Age < 65	Age ≥ 65	Short / 1st Use (<3, <6, or ≤12 mos.)	Long-Term Use (>12 mos.)	Gel v. IM	Pre-Existing Disease
Non-CV Outcome						1. Ajayi (1995) increase in thromboxane receptor density and platelet aggregation 2. Ajayi (1995) no change in thromboxane receptor affinity or thrombin-induced platelet aggregation	

VTE Studies Summary Chart

Red = statistically significant increased risk Black = no statistically significant increased risk Green = statistically significant decreased risk

Study Type	Overall	Age < 65	Age ≥ 65	Short / 1st Use (<3, <6, or ≤12 mos.)	Long-Term Use (>12 mos.)	Gel v. IM	Pre-Existing Disease
Meta-analysis	1. Xu & Schooling (2015)						
Observational	1. Baillargeon (2015) 2. Li (2016) 3. Sharma (2016)	1. Li (2016)	1.Li (2016)	1. Martinez (2016) if >6 mos. or <6 mos. with VTE risk factors 2 Martinez (2016) if <6 mos.		1. Li (2016) cohort 2. Li (2016) case-control if any or current IM use	1. Martinez (2016)
Case Series							1. Glueck (2011) 6 VTEs in 6 men w/ clotting factors 2. Glueck (Apr 2014) 7 VTEs in 7 men; 5 w/ clotting factors 3. Glueck (Oct 2014) 14 VTEs in 11 men w/ clotting factors 4. Glueck (2016) 84 VTEs in 67 men & women w/ clotting factors
Other Studies	1. Braekkan (2010) incremental increases in hemoglobin & hematocrit (total & unprovoked VTE) ***not TRT** 2. Braekkan (2010) incremental increases in red blood count (unprovoked VTE) or mean corpuscular volume (total & unprovoked VTE) ***not TRT**					1. Ajayi (1995) increase in thromboxane receptor density and platelet aggregation w/ T injections in young healthy men 2. Ajayi (1995) no change in thromboxane receptor affinity or thrombin-induced platelet aggregation w/ T injections in young healthy men	

Figure 4