IN THE UNITED STATES DISTRICT C OURT
FOR THE DISTRICT OF SOUTH CAROLINA
CHARLESTON DIVISION

IN RE: LIPITOR (ATORVASTATIN CALCIUM) MARKETING, SALES PRACTICES AND PRODUCTS LIABILITY LITIGATION	MDL No. 2:14-mn-02502-RMG CASE MANAGEMENT ORDER NO. 54 **This Order relates to all cases.**

In this MDL, Plaintiffs allege that Lipitor caused them to develop Type 2 diabetes. To carry their burden, Plaintiffs must prove both general and specific causation. *Westberry v. Gislaved Gummi AB*, 178 F.3d 257, 263 (4th Cir. 1999). Defendant has moved to exclude the testimony of Plaintiffs' general causation experts. (Dkt. No. 972). In its motion, Defendant also moves to exclude the testimony of Dr. Nicholas Jewell.[1] Dr. Jewell is a statistician. He does not opine on whether Lipitor causes diabetes but offers opinions related to whether particular data show a statistical association between Lipitor and new-onset diabetes.[2] All of Plaintiffs' general

---

[1] Some of Defendant's arguments regarding Dr. Jewell appear in its motion to exclude the testimony of Dr. Abramson. (Dkt. No. 974.)

[2] The parties agree that epidemiologists use a two-step process for establishing general causation. (Dkt. No. 972 at 27-28; Dkt. No. 1053 at 13); *see also Ambrosini v. Labarraque*, 101 F.3d 129, 136 (D.C. Cir. 1996). First, studies must establish an association or correlation between two variables, here, Lipitor and diabetes. If two variables correlate, the incidence of one variable (diabetes) changes with the incidence of another (Liptor). Once an association is established, epidemiologists apply the "Hill factors" to evaluate whether an association is causal. These factors are (1) strength of the association, (2) replication of the findings, (3) specificity of the association, (4) temporal relationship, (5) dose-response relationship (aka biological gradient), (6) biological plausibility, (7) consistency with other knowledge (aka coherence), (8) consideration of alternative explanations, and (9) cessation of exposure. Reference Manual on Scientific Evidence 600 (3d ed. 2011); *In re Zoloft (Sertraline Hydrochloride) Products Liab. Litig.*, 26 F. Supp. 3d 449, 454-55 (E.D. Pa. 2014), *recon. denied*, 2015 WL 314149 (E.D. Pa. Jan. 23, 2015).

causation experts have relied on Dr. Jewell's analysis to some extent in their initial expert reports.

In Dr. Jewell's initial report, he analyzes (1) data submitted with Lipitor's FDA New Drug Application (NDA), (2) data from the SPARCL trial, and (3) data from the IDEAL and TNT trials. (Dkt. Nos. 972-10, 1247-9, 1247-10). From the outset, Plaintiffs' counsel provided Dr. Jewell with data from the ASCOT trial, but Dr. Jewell did not consider or discuss the ASCOT trial in his initial report. (Dkt. No. 972-7 at 121-22; Dkt. No. 972-10). In his deposition, he testified that "chose not to study the data in ASCOT." (Dkt. No. 972-7 at 120, 123).

Defendant's experts criticized Dr. Jewell's statistical methodologies and analyses and specifically attacked Dr. Jewell's analysis on the fact that he did not consider the ASCOT trial data. Plaintiff then sought leave from the Court for Dr. Jewell to submit a supplemental and/or rebuttal report addressing Defendant's experts' criticisms and allowing Dr. Jewell to consider and analyze the ASCOT data. "In an abundance of caution, and to ensure this Court ha[d] the best information possible when addressing *Daubert* motions," the Court allowed Dr. Jewell to submit a supplemental report. (CMO 34, Dkt. No. 869 at 2). In this supplemental report, Dr. Jewell analyzes the ASCOT data, performs additional analysis on the NDA data, and performs some additional analysis of the SPARCL data. (Dkt. Nos. 972-34, 1247-11).

Defendant specifically attacks Dr. Jewell's analyses of the NDA data and ASCOT data as unreliable. (*See* Dkt. No. 972 at 14-15, 34-38, 44-46, 51; Dkt. No. 1247-2 at 27-35; Dkt. No. 1247-5 at 7-8, 12-15, 2; Dkt. No. 1247-6 at 18-19). In briefing, Defendant does not specifically attack Dr. Jewell's statistical analyses of the SPARCL, TNT and IDEAL data but argues that Dr. Jewell has cherry-picked these studies and that such cherry-picking cannot be the basis of a

general causation opinion. (*See* Dkt. No. 972 at 12-15, 38-43, 53). The Court takes each issue in

turn.

I. Legal Standard

Under Rule 104(a) and 702, "the trial judge must ensure that any and all scientific

testimony or evidence admitted is not only relevant, but reliable." *Daubert v. Merrell Dow*

Pharm., Inc., 509 U.S. 579, 589 (1993). Thus, the trial court must ensure that (1) "the testimony

is the product of reliable principles and methods," that (2) "the expert has reliably applied the

principles and methods to the facts of the case," and (3) that the "testimony is based on sufficient

facts or data." Fed. R. Evid. 702(b), (c), (d). "This entails a preliminary assessment of whether

the reasoning or methodology underlying the testimony is scientifically valid," *Daubert,* 509

U.S. at 592-93, and whether the expert has "faithfully appl[ied] the methodology to facts."

Roche v. Lincoln Prop. Co., 175 F. App'x 597, 602 (4th Cir. 2006)

Factors to be considered include "whether a theory or technique . . . can be (and has been)

tested," "whether the theory or technique has been subjected to peer review and publication," the

"known or potential rate of error," the "existence and maintenance of standards controlling the

technique's operation," and whether the theory or technique has garnered "general acceptance."

Daubert, 509 U.S. at 593-94; *accord United States v. Hassan*, 742 F.3d 104, 130 (4th Cir. 2014).

However, these factors are neither definitive nor exhaustive, *United States v. Fultz*, 591 F. App'x

226, 227 (4th Cir. 2015), *cert. denied*, 135 S. Ct. 2370 (2015), and "merely illustrate[] the types

of factors that will bear on the inquiry." *Hassan*, 742 F.3d at 130. Courts have also considered

whether the "expert developed his opinions expressly for the purposes of testifying," *Wehling v.*

Sandoz Pharm. Corp., 162 F.3d 1158 (4th Cir. 1998), or through "research they have conducted

independent of the litigation," *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1317 (9th

Cir. 1995) (on remand), and whether experts have "failed to meaningfully account for . . . literature at odds with their testimony." *McEwen v. Baltimore Washington Med. Ctr. Inc.*, 404 F. App'x 789, 791-92 (4th Cir. 2010).

Rule 702 also requires courts "to verify that expert testimony is 'based on sufficient facts or data.'" *E.E.O.C. v. Freeman*, 778 F.3d 463, 472 (4th Cir. 2015) (quoting Fed. R. Evid. 702(b)). 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." *Id.* The court may exclude an opinion if "there is simply too great an analytical gap between the data and the opinion offered." *Id.* "The proponent of the [expert] testimony must establish its admissibility by a preponderance of proof." *Cooper v. Smith & Nephew, Inc.*, 259 F.3d 194, 199 (4th Cir. 2001).

The Court is mindful that the *Daubert* inquiry involves "two guiding, and sometimes competing, principles." *Westberry v. Gislaved Gummi AB*, 178 F.3d 257, 261 (4th Cir. 1999). "On the one hand, . . . Rule 702 was intended to liberalize the introduction of relevant expert evidence," *id.*, and "the trial court's role as a gatekeeper is not intended to serve as a replacement for the adversary system." *United States v. Stanley*, 533 F. App'x 325, 327 (4th Cir. 2013) *cert. denied*, 134 S. Ct. 1002 (2014). On the other, "[b]ecause 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.) *cert. denied*, 135 S. Ct. 2370 (2015); *accord Westberry*, 178 F.3d at 261.

II. Dr. Jewell's Analysis of the NDA Data

A. Background on the NDA Data

Dr. Jewell's first opinion concerns data that Defendant submitted with Lipitor's FDA

New Drug Application (NDA). (Dkt. No. 1247-9 at ¶ 6).   With its application to the FDA,

Defendant submitted an Integrated Summary of Safety (ISS) that summarizes data from 31

completed clinical pharmacology studies and 21 completed clinical studies. (Dkt. No. 1063-8 at

6). Section 5.2.1 and Table 42 group all of the data from the placebo-controlled clinical trials.

(*Id*. at 118-19). There were seven such placebo-clinical trials, excluding one placebo-controlled

trial where all participants had pre-existing Type 2 diabetes. (*See id*. at 120; Dkt. No. 1247-9 at ¶

13). These seven trials included a total of 1,122 participants given various doses of Lipitor and

270 participants given placebo. (Dkt. No. 1063-8 at 119; Dkt. No. 1247-9 at ¶ 13). Dr. Jewell

was provided with the data from these seven clinical trials, including patient-level glucose

readings.

Among other things not relevant here, Table 42 reports how many participants in these

seven clinical trials had an elevated blood glucose reading during the trials, with elevated

glucose defined as > 1.25 ULN [the upper limit of normal]. (Dkt. No. 1063-8 at 119). For most

trials, the upper limit of normal was defined as 100 mg/dL such that 1.25 ULN = 1.25 mg/dL. [3]

However, there was one trial where ULN was defined as 104 mg/dL such that 1.25 ULN was

130 mg/dL. (Dkt. No. 1063-8 at ¶ 18). Table 42 reports that three (3) of the 270 participants on

placebo had an elevated glucose reading "on treatment," i.e. during the trial, and that 37 of the

---

[3] The parties generally agree that fasting blood glucose of $\leq$ 100 mg/dl is normal, multiple
fasting blood glucose levels between 100 mg/dL and 125 md/dL is diagnostic for pre-diabetes,
and multiple fasting blood glucose levels > 125 mg/dL is diagnostic for diabetes.

1,122 participants on Lipitor had an elevated glucose reading on treatment. (Dkt. No. 1063-8 at 119).

Focusing on the data from these seven clinical trials in the NDA data, Dr. Jewell opined that this data "provide less than optimum information about [Lipitor's] effect on glucose metabolism or new-onset diabetes, because of their short duration, relatively small sample sizes, and the unusual imbalance between the number of participants allocated to placebo and atorvastatin treatment." (Dkt. No. 1247-9 at ¶ 6). Yet, Dr. Jewell mined the data and, after some effort, concluded that "[t]he placebo-controlled data [from the NDA trials] show[] a statistically significant three-fold higher incidence of clinically meaningful abnormal increases in blood glucose measurement greater than 1.25 times the upper limit of normal, a level that, if persistent, is diagnostic for diabetes" and that this glucose data "should have alerted Parke-Davis and Defendant to the possibility of increased risk of new-onset diabetes associated with atorvastatin treatment." (Dkt. No. 1247-9 at ¶ 6). To reach this conclusion, Dr. Jewell improperly engaged in a results-driven methodology, performing, in his words, a "whole lot" of analyses of the data, excluding from his report analyses that he "didn't believe . . . supported . . . being the basis of the kinds of opinions I wanted to put in my summary," [4] (Dkt. No. 1247-8 at 230-31), and conducting multiple statistical tests when the first test did not produce the results that he wanted.

B. Relative Risk

Dr. Jewell first looked at adverse event data from these seven clinical trials but could not reach any conclusions from it. Of the 1,392 participants in these trials, only five reported

---

[4] Dr. Jewell did not keep the analyses that were not part of his report. (Dkt. No. 1247-8 at 230).

adverse events of diabetes and five others noted an event such as "elevated glucose."[5] (Dkt. No.

1247-9 at ¶ 14). None of these incidents were reported by participants in the placebo group but

"it's not a shock . . . given th[e] distribution of participants."[6] (Dkt. No. 1247-8 at 206-07).

There are "very few participants" reporting an adverse event of diabetes, so Dr. Jewell "basically

lay[s] to rest the hope that one might glean too much from the adverse event [data]." (Dkt. No.

1247-8 at 206).

Dr. Jewell then turns to glucose measurements. Dr. Jewell "count[ed] the number of

individuals who had elevated glucose measurements at any point *after* baseline" (i.e., during the

trial), and compared this number in the groups on Lipitor and the groups on placebo.[7] (Dkt. No.

1247-9 at ¶ 17 (emphasis in original)). Dr. Jewell took the count from Table 42 of the ISS. (*Id*.).

1. Use of a Single Glucose Measurement

As an initial matter, whether a single elevated glucose measurement can be used as a

proxy for new-onset diabetes (Plaintiffs' alleged injury in this lawsuit) is suspect. Dr. Jewell

himself was unwilling to testify about the role or use of blood glucose as a surrogate marker for

diabetes because he was not a clinician. (Dkt. No. 972-7 at 128). Furthermore, when he used

---

[5] According to Defendant's expert Dr. Wei there were only three adverse events of diabetes in the data, and all three of these participants had baseline glucose levels over 125 mg/dL. (Dkt. No. 1247-12 at ¶ 147). However, whether there were three or five cases of diabetes is not material here.

[6] There were 1,122 participants on Lipitor and 270 participants in placebo groups. (Dkt. No. 1247-9 at ¶ 13).

[7] Dr. Jewell lumped the data from all seven trials together rather than conduct a meta-analysis of the trials. (Dkt. No. 1247-9 at ¶¶ 17-18; Dkt. No. 1247-8 at 218). This can also present methodological concerns. *See In re Zoloft (Sertraline Hydrochloride) Products Liab. Litig.*, 26 F. Supp. 3d 449, 458 n.25 (E.D. Pa. 2014) (contrasting litigation expert's method with that in a published study that "used a well-established method for analyzing data from multiple studies, a meta-analysis following the guidelines for Meta–Analysis of Observational Studies in Epidemiology"), *reconsideration denied*, No. 12-MD-2342, 2015 WL 314149 (E.D. Pa. Jan. 23, 2015).

fasting glucose levels as diagnostic of diabetes in other analyses, he defined diabetes as *two* post-baseline fasting glucose measurements > 125 mg/dL, not one. (*E.g.*, Dkt. No. 1247-9 at ¶ 146). Indeed, Plaintiffs argue in another context that participants with an elevated glucose reading at baseline are not necessarily diabetic "because a diagnosis of diabetes requires more than a single elevated plasma glucose level." (Dkt. No. 1159 at 12).

Presumably, Dr. Jewell used a single glucose measurement in his NDA analysis because, at least for some participants, there is only one post-baseline glucose measurement during the controlled trial. (*See* Dkt. No. 1247-12 at 59-63). However, when Dr. Jewell examined the IDEAL data and realized that a number of participants would have to be excluded from the analysis because they lacked two post-baseline glucose values, Dr. Jewell chose not to look at glucose measurements *at all* and chose not to run *any* statistical analyses using glucose values. (Dkt. No. 1247-9 at ¶¶ 95-110). Instead, he only used adverse event reporting. (*Id.*). Nevertheless, with this NDA data, Dr. Jewell not only performed statistical tests based on a single elevated glucose measurement but opined that data based on a single glucose measurement is sufficient to suggest an increased risk of new-onset diabetes. (*See* Dkt. No. 1247-9 at ¶ 6 (opining that this glucose data "should have alerted Parke-Davis and Defendant to the possibility of increased risk of new-onset diabetes associated with atorvastatin treatment")).

By his deposition testimony, Dr. Jewell lacks the clinical expertise to make any inferences about new-onset diabetes from data regarding a single elevated glucose reading. (Dkt. No. 972-7 at 128). And Plaintiffs state a single elevated glucose measurement is insufficient to infer diabetes. (Dkt. No. 1159 at 12). Therefore, even if the methodological flaws discussed below were not present, the Court would exclude Dr. Jewell's opinion that this data "should have

alerted Parke-Davis and Defendant to the possibility of increased risk of new-onset diabetes

associated with atorvastatin treatment." (Dkt. No. 1247-9 at ¶ 6).

2. Including Participants with Elevated Glucose at Baseline

Dr. Jewell's second methodological flaw in his analysis of the NDA data is that he chose

to include participants that had glucose measures above 125 mg/dL *at baseline* when the study

began. (*See* Dkt. No. 1247-9 at ¶ 19); *see also* Reference Manual on Scientific Evidence 216 (3d

ed. 2011) ("[F]laws in the data can undermine any statistical analysis."). Dr. Jewell's analysis

starts with the count that three (3) of the 270 placebo participants had "abnormal glucose levels"

during the trial and 37 of the 1,122 Lipitor participants had "abnormal glucose levels" during the

trial. (*Id.* at ¶¶ 13, 20). However, 2 of the 3 placebo participants with elevated glucose readings

during the trial had ***baseline*** glucose levels over 125 mg/dL, and 25 of the 37 Lipitor participants

had ***baseline*** glucose levels over 125 mg/dL. (Dkt. No. 1247-12 at ¶ 140; Dkt. No. 974-11 at

131). Seven of the 37 with elevated glucose in the Lipitor group had a documented history of

diabetes before the trial began.[8] (Dkt. No. 974-11 at 131). Thus, in the NDA Medical Review,

the FDA states the following:

> The increased incidence of glucose elevations in the atorvastatin-treated
> participants bears comment. In the placebo-controlled data grouping, of the 37
> atorvastatin participants with glucose elevations, 36/37[9] had elevated glucose at
> baseline, and in 25 of those 36, elevations were > 1.25 X ULN. In addition, 7/37
> had a history of diabetes and 2/37 had a history of glucose intolerance.

---

[8] Plaintiffs note that an elevated glucose reading at baseline does not necessarily mean that
participants have diabetes. (Dkt. No. 1159 at 12). However, Dr. Jewell did not even exclude the
seven (7) participants with a documented history of diabetes. More importantly, as discussed
below, in all other instances where Dr. Jewell performed statistical analyses on data of glucose
values, he did exclude participants with elevated baseline glucose values.

[9] Eleven participants had baseline glucose readings between 100 and 125 mg/dL, the range
considered pre-diabetic. Dr. Jewell was simply looking at people with a reading above 125
mg/dL versus those with a reading below without differentiating those that had elevated readings
in the prediabetic range from those with "normal" glucose readings.

Similarly, in the all-completed studies grouping, of 185 atorvastatin participants with glucose elevations, 115/185 had a history of NIDDM. 174/185 had baseline values > ULN.

In sum, there is little evidence for an effect of atorvastatin on glucose metabolism.

(Dkt. No. 974-11 at 131) (footnote added).[10]

There are two ways that failing to exclude participants with elevated baseline glucose is problematic for Dr. Jewell's analysis. First, there is the potential for confounding.[11] In the NDA trials, 5.3% of all participants on Lipitor had at least one glucose measurement above 125 mg/dL *before the trial*, while only 1.9% of the placebo participants had at least one glucose reading above 125 mg/dL before the trial, and this difference between the two groups was statistically significant. (Dkt. No. 972-36 at ¶ 140). Looking only at the very first pre-trial glucose measurement for each participant, as opposed to all pre-trial glucose measurements, 3.83% of the Lipitor group had an elevated glucose value at this first measurement and only 1.85% of the placebo group had an elevated glucose value at this first measurement, though this difference

---

[10] Parke-Davis' medical monitor reached the same conclusion for the same reasons. (Dkt. No. 1063-8 at 120-121).

[11] Confounding variables are those that correlate with the independent and dependent variable. Reference Manual on Scientific Evidence 285 (3d ed. 2011). Confounding variables cause a correlation to exist between the independent and dependent variables without causation being present. *Id.* at 219, 285. For example, ice creams sales correlate with violent crime; as ice cream sales increase, so does violent crime. However, this does not mean that ice cream sales cause violent crime. There is a confounding variable present: temperature. Both ice cream sales and violent crime rise with temperature. Here, baseline glucose values may be a confounding variable. There may be more participants with elevated on-treatment glucose levels in the Lipitor group, not because the Lipitor caused elevated glucose levels, but because there were more participants with elevated glucose to start with in the Lipitor group.

was not statistically significant. [12] (Dkt. No. 1247-11 at ¶¶ 37, 39). This difference has the potential to confound Dr. Jewell's data.

Second, because Dr. Jewell's analysis includes those with elevated glucose at baseline, it does not compare *new* cases of elevated glucose, i.e., those participants with elevated glucose due to Lipitor. In other words, the analysis does not test for new-onset diabetes, the subject of Dr. Jewell's opinion, or even new-onset elevated glucose. Dr. Jewell claims that he was only concerned with "clinically meaningful abnormal deviations in glucose" or "assessing whether atorvastatin had a potentially deleterious effect on glucose levels" regardless of baseline characteristics. (Dkt. No. 1247-8 at 235.; Dkt. No. 1247-9 at ¶ 19). He argues that "[w]hether the patient had diabetes, was prediabetic, had elevated blood glucose before the trial is irrelevant to me," (Dkt. No. 1247-8 at 235), because all of these groups could still experience significant increases in their glucose levels. (Dkt. No. 1247-9 at ¶ 19).

These claims are belied by Dr. Jewell use of this data to support opinions about "the possibility of increased risk of *new*-onset diabetes," (Dkt. No. 1247-9 at ¶ 6 (emphasis added)), and Plaintiffs argument that this NDA data supports the opinion that 10 mg of Lipitor causes new-onset diabetes. (Dkt. No. 1159 at 11-13). However, even assuming that he did want to assess significant changes in glucose levels, regardless of baseline characteristics, this is not what he assessed. He included 100% of the participants that had an elevated glucose

---

[12] In his supplemental report, Dr. Jewell took issue with how Dr. Wei defined "baseline." (Dkt. No. 1247-11 at ¶ 36). Dr. Jewell would only consider someone to have an elevated glucose measurement at baseline if "the first pre-treatment glucose value" was elevated. (*Id.* at ¶ 37). Thus, for instance, Patient # 47, whose first pre-treatment glucose measurement was 111 but whose glucose increased to 141 *prior to starting Lipitor* was classified as having an elevated glucose at baseline by Dr. Wei but not by Dr. Jewell. (*See* Dkt. No. 1247-12 at 61). The Court notes that the FDA's count of how many participants had elevated glucose at baseline matched Dr. Wei's count. (Dkt. No. 974-11 at 131).

measurement during a trial, regardless of whether they actually experienced any significant

increase in glucose from baseline. For example, Patient #1 had a baseline glucose measurement

of 131 mg/dL and an on-treatment glucose measurement of 133 mg/dL, (Dkt. No. 1247-12 at

59), a change of 2 mg/dL, a change that Plaintiffs' other expert, a clinician, testifies is "minor."

(Dkt. No. 974-1 at 156). However, he included this patient in his analysis, as he included all

participants with an on-treatment elevated glucose measurement.

Dr. Jewell argues that he should be able to assume all of the participants counted in Table

42 had "clinically meaningful abnormal increases in blood glucose," despite data to the contrary,

because of language in Section 5.2 of Defendant's Integrated Summary Safety (ISS).[13] (Dkt.

No. 1247-9 at ¶ 6; Dkt. No. 1247-8 at 167). Whatever Dr. Jewell's reading of Section 5.2, Dr.

Jewell had the data that proved this assumption false and chose to ignore it. He had the

individual glucose readings for all 40 participants at issue. He looked at this data for the purpose

of calculating an average increase in glucose measurements, (Dkt. No. 1247-9 at ¶ 20), and thus

---

[13] ISS Section 5.2 states in full:

5.2 Clinical Laboratory Abnormalities

In the atorvastatin program, clinical laboratory parameters were evaluated for
abnormal values during treatment using normal ranges supplied by the central
laboratory and program-defined criteria. These criteria were established before
studies began and were designed to identify clinically meaningful changes.
Laboratory abnormalities were identified relative to each patient's baseline value.
Thus, during treatment, laboratory parameters with values that met criteria for a
clinically meaningful deviation but were not different from the patient's baseline
value were not identified as abnormalities.

(Dkt. No. 1063-8 at 118). The pre-specified criterion for "clinically meaningful" glucose
elevation was >1.25 ULN (upper limit of normal). (*Id*. at 424). Thus, if a patient had a glucose
reading >1.25 ULN and that reading differed from baseline, even if by 1 mg/dL, then it would
have been reported. Dr. Jewell apparently interpreted the language of Section 5.2 differently,
assuming that all reported cases of elevated glucose were "clinically meaningful abnormal
deviations" *from an individual's baseline*, rather than "clinically meaningful deviations" *from
normal* and also different from baseline.

knew the exact difference between the maximum on-treatment glucose measurement and the

baseline measurement for each patient. He knew, for example, that Patient #1 had a baseline

glucose measurement of 131 mg/dL and an on-treatment glucose measurement of 133 mg/dL.

(Dkt. No. 1247-12 at 59). However, he chose to ignore this data for *this* purpose and simply

assume that every one of these participants had a "clinically meaningful" or "significant" change

from baseline, despite the fact that this assumption was easily verifiable or proven false.

Finally, including participants with elevated baseline glucose is contrary to Dr. Jewell's

methodology in all of his other analyses. In all other instances where Dr. Jewell looked at the

effect of Lipitor on glucose measurements, he excluded participants with elevated baseline

glucose. (*See* Dkt. No. 1247-9 at ¶ 40 (excluding participants with a baseline glucose

measurement > 125 mg/dL), ¶ 146 (excluding participants with a baseline glucose measurement

>125 mg/dL), Dkt. No. 1247-11 at ¶ 19 (same)). Even when looking only at adverse event

reporting of diabetes, he ran the analysis excluding participants with elevated glucose at baseline.

(*See* Dkt. No. 1247-9 at ¶ 35 (excluding participants with baseline glucose values ≤ 125 mg/dL),

¶ 106 (excluding participants with baseline glucose values > 125 mg/dL), ¶ 134 (excluding

participants with baseline glucose values >125 mg/dL)). Indeed in his analysis of the TNT data,

Dr. Jewell states, as if it is obvious, that using the entire population "for analysis seems

inappropriate since these individuals already had Diabetes Mellitus or high glucose levels at

baseline." (Dkt. No. 972-10 at ¶ 145).

In Dr. Jewell's analysis of the TNT data, Dr. Jewell found an increased risk of diabetes

when excluding both those with pre-existing diabetes *and* those with elevated baseline glucose

values. (Dkt. No. 1247-9 at ¶ 145). However, the "increased risk was not . . . present in the

larger population, which included participants with both medical histories of Diabetes Mellitus

and baseline glucose value >125 mg/dL." (*Id.*). Thus, when excluding these participants was the method by which Dr. Jewell could obtain a statistically significant result, he chose to exclude them.

However, he chose not to exclude participants with elevated baseline glucose in his NDA analysis because, as Dr. Jewell states, "exclusion of participants with baseline glucose levels greater than, for example, 125 mg/dL . . . would likely greatly reduce the number of participants in the already limited placebo-controlled data set." (*Id.* at ¶ 19). In other words, doing so would make the already small and limited data set even smaller, such that Dr. Jewell would not be able to obtain a statistically significant result.[14] For the NDA data alone, Dr. Jewell chose to include those with elevated glucose at baseline. This internal inconsistency weighs heavily against reliability. *See, e.g, In re Rezulin Products Liab. Litig.*, 309 F. Supp. 2d 531, 563 (S.D.N.Y. 2004) ("[The expert's] selectivity in defining the universe of relevant evidence thus violated his own standard of proper methodology . . . which suggests that he does not apply the same rigor in the courtroom that he would apply to his medical endeavors.").

3. Turning to a Second Statistical Test

Despite the fact that Dr. Jewell included participants with elevated glucose measurements at baseline in his calculation, he *still* did not get a statistically significant result when he calculated his first set of statistics. Dr. Jewell calculated the estimated Relative Risk of an abnormal glucose measurement for the Lipitor group to be 3.0 with a 95% confidence interval of

---

[14] Defendant's expert, Dr. Wei, verifies this. Dr. Wei states that "in a conventional meta-analysis of the data among those who had no glucose value > 1.25 ULN [upper limit of normal] at baseline, there was no evidence of any statistically significant risk of glucose abnormalities associated with atorvastatin compared to placebo." (Dkt. No. 1247-12 at ¶ 134). In other words, in the participants who did not have a glucose level above 1.25 ULN at baseline, there was no significant effect.

.9 to 9.6. (Dkt. No. 1247-9 at ¶ 17). He made this calculation using Stata, a commonly used

statistical program. (*See* Dkt. No. 1247-8 at 211, 214-17). When Dr. Jewell made these

calculations using Stata he also obtained a p-value of 0.0654 using the Fisher exact test; this p-

value indicates a lack of statistical significance.[15] (*Id*. at 213). Dr. Jewell stated that "I, and

many others, often use the Fisher exact test for computation of the p-value associated with

statistical significance of an estimated Relative Risk." (Dkt. No. 972-10 at 13 n.16); *see also*

Reference Manual on Scientific Evidence 255 n.108 (3d ed. 2011) ("Well-known small-sample

techniques [for testing significance and calculating p-values] include the sign test and Fisher's

exact test."). However, Dr. Jewell did not report this Fisher exact p-value in his report.

Instead, Dr. Jewell turned to the mid-p test, which would "[a]lmost surely" produce a

lower p-value than the Fisher exact test.[16] (Dkt. No. 1247-8 at 213). The software that Dr.

Jewell used to calculate the Relative Risk, confidence interval, and p-value using the Fisher exact

test (Stata) does not calculate the mid-p value. Thus, Dr. Jewell had to use a separate online

piece of software to calculate the mid-p value. (Dkt. No. 1247-8 at 214-15). The p-value, as

calculated using the mid-p test, was .04, under the .05 threshold for statistical significance, and,

thus, Dr. Jewell declared the result statistically significant. (Dkt. No. 1247-9 at ¶ 17).

It is important to note that using the mid-p approach, standing alone, does not render Dr.

Jewell's analysis unreliable. The mid-p approach is used by some statisticians and can be a valid

methodology. (*See* Dkt. No. 972-10 at 13 n.16). For instance, if Dr. Jewell thought the mid-p

approach a better approach than the Fisher exact test, pre-specified the use of the mid-p approach

---

[15] Generally, p-values less than .05 indicate statistical significance. Reference Manual on Scientific Evidence 251 (3d ed. 2011); (*see also* Dkt. No. 972-7 at 240 (stating .05 is the "boundary" of statistical significance)).

[16] This is because the mid-p approach is "more powerful," i.e. less conservative, than the Fisher exact test. (Dkt. No. 1247-8 at 213, Dkt. No. 972-10 at 13 n.16).

from the outset, and consistently used it in all of his analyses, his use of it may be considered

reliable.

The problem with Dr. Jewell's use of the mid-p test is that his use of it was results driven.

He only used this test once the Fisher exact test returned a non-significant result. After he used

the mid-p test to obtain a statistically significant p-value, he did not even bother to determine a

mid-p exact confidence interval but continued to use the prior confidence interval obtained via

Stata and reported with the Fisher exact p-value.[17] (Dkt. No. 1247-8 at 214, 217). This indicates

he was not actually interested in using the mid-p approach but in obtaining a statistically

significant p-value.

Dr. Jewell also did not use the mid-p approach in any of his other analyses, only with this

analysis of the NDA data.[18] (*See* Dkt. No. 1247-9 at ¶ 30 (reporting confidence interval and "p"

value rather than "mid-p" value for Relative Risk ratio), ¶ 31 (same), ¶ 35 (same), ¶ 38 (same), ¶

---

[17] Dr. Jewell testified that he "thinks" the confidence interval was calculated by "flip[ping]" the
Taylor series approximation. (Dkt. No. 1247-8 at 211). Stata's website indicates that, at least
for odds ratios, it calculates the confidence interval by inverting the Fisher exact test. *See*
http://www.stata.com/support/faqs/statistics/fishers-exact-test/, last visited November 16, 2015
("Stata's exact confidence interval for the odds ratio inverts Fisher's exact test.").

[18] The only other use of a mid-p value in Dr. Jewell's report is in the next paragraph where he
calculates the Incidence Rate Ratio based of the same data. (Dkt. No. 1247-9 at ¶ 18). Even
with the less conservative mid-p value, he does not obtain a statistically significant result, though
it is close (mid-p=0.52). Even though this result is not statistically significant, Dr. Jewell states
that "this analysis naturally confirms what we saw with the comparison of proportions of
individuals with abnormal glucose measurements, namely that exposure to atorvastatin increases
the rate of occurrence of such events almost three-fold." (*Id*.). In contrast, when asked about the
hazard ratio being lower for women than men in the ASCOT study (a result unhelpful to
Plaintiffs), Dr. Jewell testifies that the data does not show a trend in that direction "because . .
.the difference . . . is not statistically significant." (Dkt. No. 1247-8 at 131). Thus, his standard
for whether data supports a conclusion changes based on the conclusion at issue. If the data
indicates a lower risk for women (not favorable to Plaintiffs), the data means nothing if it is not
statistically significant. But if the data indicate a higher risk of diabetes (favorable to Plaintiffs),
then it supports this conclusion even if not statistically significant.

45 (same), ¶ 47 (same), ¶ 51 (same), ¶ 58 (same), ¶ 61 (same), ¶ 63 (same), ¶ 73 (same), ¶ 78

(same), ¶ 79 (same), ¶ 108 (same), ¶ 110 (same), ¶ 140 (same), ¶ 144 (same), ¶ 150 (same), ¶ 152

(same), ¶ 157 (same), ¶ 159 (same); *see also* Dkt. No. 1247-8 at 304 (indicating Dr. Jewell used

Stata for his SPARCL analyses); 213 (Stata does not provide mid p value)). He only used the

mid-p test when it was essential for obtaining a statistically significant result. Furthermore, Dr.

Jewell omitted from his report that he obtained the Fisher exact p-value and that this p-value was

not significant.[19] *See Barber v. United Airlines, Inc*., 17 F. App'x 433, 437 (7th Cir. 2001)

("Because in formulating his opinion Dr. Hynes cherry-picked the facts he considered to render

an expert opinion, the district court correctly barred his testimony because such a selective use of

facts fails to satisfy the scientific method and *Daubert*.").

4. Conclusion

 "Coming to a firm conclusion first and then doing research to support it is the antithesis

of [the scientific] method." *Claar v. Burlington N. R. Co*., 29 F.3d 499, 502-03 (9th Cir. 1994).

That is what Dr. Jewell has done here. To reach his conclusion, he included participants with

elevated glucose at baseline, despite the fact that he excludes all such participants in his other

analyses, and reverts to a less conservative test when the first statistical test he used did not

produce the results he wanted. Therefore, the Court excludes this testimony under Rule 702.

B. Average Blood Glucose Increase

Next, Dr. Jewell takes the 40 participants with "abnormal glucose levels" in the seven

NDA trials and calculates the average increase in blood glucose from the first baseline reading to

---

[19] Defendant discovered that Dr. Jewell had calculated this Fisher exact p-value because Dr.
Jewell produced his Stata log showing the 3.0 Relative Risk and accompanying confidence
interval. (*See* Dkt. No. 1247-8 at 215). Dr. Jewell chose not to keep his analyses that were not
part of his report. (*Id*. at 229). Thus, while Dr. Jewell admits that "there is a whole lot I did that
I didn't write in my report," the Court cannot determine if Dr. Jewell conducted other analyses
that reached results contrary to his opinion and he simply chose to ignore them.

the first elevated reading. (*Id.* at ¶ 20). He found the average increase in blood glucose to be

30.8 mg/dL with a standard error of 5.4 mg/dL. (*Id.*). For this calculation, Dr. Jewell lumped all

participants with elevated glucose together, regardless of whether they were in the placebo or

Lipitor group. (*See id.*). However, he attempts to attribute the average increase to Lipitor

stating, "on average, these 40 individuals, ***almost all of them on atorvastatin***, experienced a very

significant increase in blood glucose levels following initiation of treatment." (Dkt. No. 1247-9

at ¶ 20 (emphasis added)).

Plaintiffs, relying on this average increase of the combined placebo and experimental

group, argue that the NDA data show "among those whom it affects, Lipitor can raise blood

glucose sufficiently to take an individual with no prior glucose abnormalities—that, with a

baseline glucose level less than 100 mg/dL—and elevate that individual's glucose beyond the

threshold for new-onset diabetes, at 125 mg/dL."[20] (Dkt. No. 1053 at 34). This is the

convoluted, and flawed, logic: Dr. Jewell determined that all participants with "clinically

meaningful" elevated glucose had an average increase of 30.8 mg/dL, regardless of treatment,

and Dr. Jewell's prior-discussed, flawed analysis of relative risk shows that Lipitor users are

more likely to be in this "elevated glucose" category. Therefore, Plaintiffs' conclude, Lipitor can

raise glucose by 30.8 md/dL.

However, comparing the Lipitor group with the placebo group leads to the opposite

conclusion. The individuals in the placebo group had ***greater*** glucose increases than those in the

Lipitor group–an average increase of 37 mg/dL in the placebo group versus 27.1 mg/dL in the

Lipitor group. (Dkt. No. 1247-12 at ¶ 144).   Thus, to the extent one can infer anything from the

---

[20] Notably, Dr. Jewell does not actually opine that the data make this showing.

"average glucose increases" in this incredibly limited data set of 40, those who experienced "meaningful changes in their glucose level"[21] had lower increases if taking Lipitor.

In briefing, Plaintiffs discuss the average change in glucose levels overall (among all participants) versus the average change in those participants that experienced significant increases in their glucose levels. Plaintiffs state that the average that Lipitor increases blood glucose in general "is not meaningful because Lipitor does not affect all subjects equally." (Dkt. No. 1053 at 33). Plaintiffs argue that one should look only those participants that "experienced meaningful changes in their glucose level," and then find that average. (*Id*. at 34). In other words, only take the participants with significant increases in glucose levels and look at the average of those.

Regardless of the merits of this argument, it misses the point. Dr. Jewell looked at both: all the participants in the NDA trials as well as just the 40 participants reported to have elevated glucose levels during the trials. (*See* Dkt. No. 1247-8 at 228-29). The problem was that, whether he looked at all of the participants or just the 40 participants with elevated glucose levels, a direct comparison of the Lipitor group to the placebo group did not provide results Plaintiffs were seeking. Plaintiffs wish to argue that the NDA show that Lipitor can raise blood glucose sufficiently to take a participant from normal glucose of less than 100 mg/dL to readings over 125 mg/dL. (Dkt. No. 1053 at 34). Thus, they need to establish that Lipitor could raise glucose levels by at least 25 mg/dL.

Considering all participants in the NDA trial, the average increase in glucose was "a small amount," (Dkt. No. 1247-8 at 228-29), and the difference in increases between the Lipitor and placebo groups are very small. The average difference as calculated by Dr. Wei was 0.71

---

[21] This is Plaintiffs' characterization of this group of 40 participants. (Dkt. No. 1053 at 33).

mg/dL (using a fixed-effects model) and 0.12 mg/dL (using a random-effects model) and these differences were not statistically significant.[22] (Dkt. No. 1247-12 at ¶ 145). Dr. Jewell also performed this analysis (comparing the glucose increases between those on Lipitor and those on placebo) and looked at whether there was a statistically significant difference between the two groups. (*Id.* at 227-30). At his deposition, he did not remember if there was a statistically significant difference between the Lipitor and placebo groups but "suspect[ed]" the difference was not significant (*Id.* at 229).

Dr. Jewell performed this analysis and used it "as a basis" for his opinions in Paragraph 21 that "glucose tends to increase more on average for atorvastatin participants that for placebo subjects."[23] (Dkt. No. 1247-8 at 228; Dkt. No. 1247-9 at ¶ 21). However, unlike the other bases of his opinions, he completely excluded this statistical analysis from his report, neglected to state whether the results were statistically significant, which he suspects were not, and did not provide the actual numerical increase for the Lipitor group, which he admits was "small."[24] (*See* Dkt. No. 1247-9 at ¶ 21). When asked why he did not include this analysis in his report, he

---

[22] Dr. Wei conducted a meta-analysis, treating each of the seven trials as separate trials. (Dkt. No. 1247-12 at ¶ 145). Dr. Jewell pooled the data together as if the participants were in a single trial. (*See* Dkt. No. 1247-9 at ¶ 17). This presents additional methodological concerns because the ratios of participants randomized to placebo versus Lipitor vary between trials. (Dkt. No. 1247-12 at ¶ 142).

[23] Again, Dr. Jewell had no problem basing this conclusion on results that were not statistically significant. However, in other contexts, where the data suggests a trend unfavorable to Plaintiffs, Dr. Jewell requires statistical significance to infer any trend from the data. (Dkt. No. 1247-8 at 131).

[24] Without these analyses, his conclusory statement in his report is nothing but the inadmissible "ipse dixit" of Dr. Jewell. *See Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997); *McEwen v. Baltimore Washington Med. Ctr. Inc.*, 404 F. App'x 789, 791-92 (4th Cir. 2010).

responded, "I didn't believe the data . . . supported that being the basis of the kinds of opinions I wanted to put in my summary, and so I did not include it."[25] (Dkt. No. 1247-8 at 231).

Focusing only on these 40 individuals, as Plaintiffs argue should be done, a direct comparison of the Lipitor and placebo groups provides results directly contrary to what Plaintiffs were seeking to argue. The individuals in the placebo group had *greater* glucose increases than those in the Lipitor group–an average increase of 37 mg/dL in the placebo group versus 27.1 mg/dL in the Lipitor group. (Dkt. No. 1247-12 at ¶ 144). Thus, this direct comparison could not be the basis for an argument Lipitor caused significant increases in glucose levels.

The analysis that Dr. Jewell conducted *and* chose to include in his report consisted of lumping the two groups (placebo and experimental) together and attempting to attribute the overall average glucose increase to Lipitor by stating that "these 40 individuals, *almost all of them on atorvastatin*, experienced a very significant increase in blood glucose levels following initiation of treatment." (Dkt. No. 1247-9 at ¶ 20 (emphasis added)). Plaintiffs, relying on this average increase of the combined placebo and experimental group, then argued that the NDA show that Lipitor can raise blood glucose sufficiently to take a patient from normal glucose of less than 100 mg/dL to readings over 125 mg/dL, despite the fact that the average increase in glucose was higher for the placebo group than Lipitor. (Dkt. No. 1053 at 34).

---

[25] Dr. Jewell states that he only considers statistical evidence that will "support a strong opinion one way or the other" and excludes from consideration evidence that would not support a "strong opinion." (Dkt. No. 1247-8 at 231). While Dr. Jewell is not clear about what constitutes a "strong opinion" or what type of statistical analyses and evidence would "support a strong opinion," it appears that he chooses to only consider evidence that shows a strong effect or significant change, whether "negative or positive," and disregard evidence of a small or insignificant effects. For example, once he learned that the average change in glucose overall was "a small amount," he decided the analysis was "not . . . particularly informative" and he "didn't believe that kind of statistical analysis was something I would want to put weight upon in forming my opinions as reflected in the summary." (Dkt. No. 1247-8 at 229; *see also* Dkt. No. 1247-14 at 251 ("I don't find that particularly relevant. I would expect that to be relatively small.")).

Regardless of Plaintiffs' argument, Dr. Jewell's opinion regarding the average increase in glucose is misleading[26] and results driven. It is apparent to the Court that rather than conducting statistical analyses of the data and then drawing a conclusion from these various analyses, Dr. Jewell formed an opinion first, sought statistical evidence that would support his opinion and chose to exclude *his own* contrary analyses from his report. This is unacceptable under *Daubert* and Rule 702. *See, e.g., Claar*, 29 F.3d at 502-03 ("Coming to a firm conclusion first and then doing research to support it is the antithesis of [the scientific] method."); *Barber*, 17 F. App'x at 437 ("Because in formulating his opinion Dr. Hynes cherry-picked the facts he considered to render an expert opinion, the district court correctly barred his testimony because such a selective use of facts fails to satisfy the scientific method and *Daubert*."); *Fail-Safe, L.L.C. v. A.O. Smith Corp.*, 744 F. Supp. 2d 870, 889 (E.D. Wis. 2010) ("[I]t is readily apparent that Dr. Keegan all but 'cherry picked' the data he wanted to use, providing the court with another strong reason to conclude that the witness utilized an unreliable methodology."); *In re Bextra & Celebrex Mktg. Sales Practices & Prod. Liab. Litig.*, 524 F. Supp. 2d 1166, 1176 (N.D. Cal. 2007) (excluding expert testimony where expert "reaches his opinion by first identifying his conclusion . . . and then cherry-picking observational studies that support his conclusion and rejecting or ignoring the great weight of the evidence that contradicts his conclusion").

C. Supplemental Report

Faced with these criticisms of his NDA analysis, Dr. Jewell performed a different analysis in his supplemental report, "adjusting for both protocol and baseline glucose." (Dkt.

---

[26] Even if the opinion were admissible under Rule 702, the Court would exclude it under Rule 403. *See Daubert*, 509 U.S. at 595 ("Expert evidence can be both powerful and quite misleading because of the difficulty in evaluating it. Because of this risk, the judge in weighing possible prejudice against probative force under Rule 403 of the present rules exercises more control over experts than over lay witnesses.").

No. 1247-11 at ¶ 43). As explained above, excluding participants with an elevated glucose at baseline would result in a non-significant result. So, instead, Dr. Jewell attempts to account for the elevated glucose values at baseline by performing a regression analysis that adjusts for baseline glucose (as he defines it)[27] and for variance in the protocols, i.e., trials. He opines that "this analysis demonstrates that the risk of a hyperglycemia lab abnormality (representing a clinically meaningful deviation from the patient's baseline) associated with atorvastatin remains statistically and substantially elevated when adjusted for differences between the protocols and baseline glucose." (*Id.*).

However, to reach this conclusion, Dr. Jewell again had to try multiple statistical models before reaching the result he wanted. He did not specify beforehand which statistical model he would use in his analysis. Instead, he "played around with making sure [he] was getting the right result." (Dkt. No. 1247-14 at 207). His software logs show that he tried at least five different statistical models. (Dkt. No. 1247-13 at ¶ 29). He only reported the results for one model. Even with his cherry-picked model, the model was not a good fit for the data and had extremely high standard errors such as 3.4×10^{29} and 1.58×10^{14}. (Dkt. No. 1247-13 at ¶ 28; *see also id.* ("It is surprising that Dr. Jewell still proceeded to report the treatment difference estimate even though his computer output clearly indicated there is a serious data fitting problem.")).

Dr. Jewell acknowledges that "it's tricky to fit" this data to a model because several of the protocols (i.e., trials) had very few or none of the 40 events of elevated glucose. (Dkt. No.

---

[27] Dr. Wei considered a patient to have an elevated glucose measurement at baseline if any of the pre-treatment measurements were above 125 mg/dL. Dr. Jewell considered a patient to have an elevated glucose measurement at baseline only if the very first pre-treatment measurement was above 125 mg/dL. (Dkt. No. 1247-11 at ¶ 37). Thus, for instance, Patient # 47, had pre-treatment glucose readings of 111, 141, and 174, all taken *prior to starting Lipitor*. (*See* Dkt. No. 1247-12 at 61). This patient was classified as having an elevated glucose at baseline by Dr. Wei but not by Dr. Jewell because the first reading, 111, was below 125 mg/dL.

1247-14 at 207, 208). As he has repeatedly admitted, the characteristics of this data set, such as "small sample sizes and the unusual imbalance between the number of participants allocated to placebo and atorvastatin treatment," make this data "less than optimum" for drawing any conclusions. (Dkt. No. 1247-9 at ¶ 6). But Dr. Jewell shoehorned the new data set[28] into an ill-fitting model anyway. It should also be noted that this model produced a dramatically wide confidence interval, (Dkt. No. 1247-11 at ¶ 43), which in other contexts Dr. Jewell testifies should cause a researcher to "be cautious in [his] interpretation."[29] (Dkt. No. 1247-8 at 131); *see also* Reference Manual on Scientific Evidence 246 (3d ed. 2011) (stating that if standard errors or confidence intervals are large, "the estimate may be seriously wrong"), *id.* at 247 ("[A] broader [confidence] interval indicates less precision."); *id.* at 248 ("A high confidence level with a broad interval means very little."). In sum, Dr. Jewell tried multiple models that did not provide his desired result and kept "playing" with the models until he found, in the words of Dr. Wei, "an ill-fitting model with a large estimated hazard ratio and small p-value." (Dkt. No. 1247-13 at ¶ 29). Under these circumstances, the Court finds his analysis unreliable and excludes it. *See, e.g., Claar*, 29 F.3d at 502-03 ("Coming to a firm conclusion first and then doing research to support it is the antithesis of [the scientific] method.").

---

[28] Dr. Jewell's data set used in his supplemental report differs from that in his original report. (Dkt. No. 1247-13 at ¶ 27). The new data set reduces total patient exposure time disproportionately more in the Lipitor group than the placebo group, resulting in a data set that favors the placebo more than his original data set. (*Id.*). Dr. Jewell's supplemental report does not explain why he used a different data set in his supplemental analysis. (*See* Dkt. No. 1247-11 at ¶ 43).

[29] This model also produced hazard ratios dramatically different from (and larger) than *any* of the hazard ratios found in published, peer-reviewed studies relied on by the parties.

II. ASCOT Analysis in Dr. Jewell's Supplemental Report

A. Background on ASCOT

The Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT) was a randomized placebo-controlled study. The Lipid Lowering Arm (LLA) of ASCOT tested Lipitor's efficacy in primary prevention of coronary heart disease (CHD) in participants with high blood pressure but normal blood lipids ("hypertensive participants who are not conventionally deemed dyslipidaemic"). (Dkt. No. 972-26 at 2). ASCOT-LLA included 10,305 participants, aged 40-79 years with normal cholesterol (6.5 mmol/L or less), with at least three other cardiovascular risk factors, and who had not experienced a cardiovascular event. (*Id.*). Participants were randomly assigned to a 10 mg dose of atorvastatin or placebo. (*Id.*). The study was stopped early because Lipitor "resulted in a highly significant reduction in . . . [coronary heart disease] events compared with placebo and a significant reduction in the incidence of stroke." (*Id.* at 4).

Diabetes was a pre-specified tertiary endpoint in this trial, meaning that the study was designed to contemporaneously collect and adjudicate diabetes data.[30] (Dkt. No. 972-26 at 6, Table 3; *see also* Dkt. No. 972-26 at 3 ("Tertiary objectives were also prespecified. . .."); Dkt. No. 1091-1 at 2). The authors found no statistically significant difference in the rate of new-onset diabetes between those on Lipitor and those in the control group. (*Id.*). In other words, the study did not find an association between Lipitor and new-onset diabetes.

B. Dr. Jewell's Analysis

In his supplemental report, Dr. Jewell conducted his own analysis of the ASCOT-LLA data and opines that there *is* an association between Lipitor and new-onset diabetes, reaching a result contrary to that of the peer-reviewed, published article on ASCOT-LLA:

---

[30] The adjudication process is described in more detail below.

> For ASCOT-LLA participants at risk for the development of new-onset diabetes
> (NOD), atorvastatin use was associated with a significantly increased risk of new-
> onset diabetes compared to placebo when controlling simply for baseline glucose,
> and similarly when also adjusting for three additional significant baseline
> predictors of new-onset diabetes.

(Dkt. No. 1247-11 at 18).

The difference between Dr. Jewell's findings and those of the ASCOT researchers is

largely due to how the researchers versus Dr. Jewell defined, and therefore counted, participants

with new-onset diabetes. The researchers used adjudicated data, and Dr. Jewell used

unadjudicated data.

1. ASCOT Adjudication Process

Randomized controlled trials have pre-specified adjudication processes for determining

whether certain events have taken place. In ASCOT-LLA there was a pre-specified adjudication

process for determining whether particular participants did in fact develop new-onset diabetes

during the trial. The purpose of such an adjudication process is to ensure reliable (i.e.,

consistent) and valid (i.e., accurate) results in the determination of whether individuals have in

fact developed new-onset diabetes. Such adjudication processes include safeguards against bias.

(*See* Dkt. No. 972-40 at 69-72). For example, the adjudication committee in ASCOT was

blinded and did not know whether a particular participant under consideration was in the

experimental or control group and did not even know if the participant was in the lipid-lowering

arm of the trial at all. (Dkt. No. 1091-1 at 3).

In ASCOT, an independent and blinded endpoint committee ("Endpoints Committee")

consisting of clinicians evaluated all potential diabetes events. (Dkt. No. 1091-1 at 2). The

committee used a pre-specified definition of diabetes, namely the definition used by the World

Health Organization at the time. (*Id.* at 3). Specifying the definition to be used *before* the trial is

an important methodology that helps guard against bias.[31]

The WHO definition, as stated in the ASCOT Endpoint Manual, the "Bible" of the

Endpoints Committee, provided that diabetes

> is defined by the World Health Organisation 1999 criteria in one of three ways:
>
> (i) Fasting plasma glucose $\geq$ 7.0 mmol/l on two occasions
> (ii) 2 hour post 75g glucose load plasma glucose $\geq$ 11.1 mmol/l
> (iii) 'Unequivocal hyperglycemia with acute metabolic decompensation or
> obvious symptoms.'

(*Id.*). The manual operationalized this third criterion as "a random plasma glucose $\geq$ 11.1 mmol/l

on two occasions + symptoms consistent with diabetes (e.g. thirst, polyuria, poly dispsia,

excessive weight loss." (*Id.*). If a patient had a fasting glucose of 6-6.9 mmol/l during the study,

she was sent for a glucose tolerance test. (*Id.*).

The committee applied these criteria to determine if a patient had new-onset diabetes.

The committee had access to medical records and case files and "examine[d] and reconcile[d]

data to ensure that the inclusion of a case was accurate." (*Id.* at 3). For instance, they would

review available information to determine whether the blood glucose measurements were in fact

fasting and review medical records and histories to rule out cases where the patient had diabetes

at the beginning of the study. (*Id.* at 3-4).

Two randomly assigned members of the Endpoints Committee looked at each case. (*Id.*

at 3). If they did not reach the same conclusion with respect to a particular individual, all four

---

[31] If the definition is not pre-specified, researchers can look at the data, *post hoc*, observe that if one definition of diabetes is used, certain results are obtained, but if another definition is used, a different result is obtained, and then specify the definition that produced the results expected or desired by the researchers. *See* Reference Manual on Scientific Evidence 2256 (3d ed. 2011) ("If enough comparisons are made, random error almost guarantees that some will yield 'significant' findings, even when there is no real effect.").

members of the committee discussed the case in a face-to-face meeting to reach a conclusion.

(*Id.*). This Endpoints Committee's determination of whether a participant developed new-onset

diabetes during the trial was then coded as a "yes" or "no" for the "diabetes mellitus" event.

(*See* Dkt. No. 1247-14 at 251).   This adjudicated data, even when adjusted for other risk factors

of diabetes, does not show a statistically significant association between Lipitor use and diabetes.

[32] (Dkt. No. 972-26 at 6, Table 3; Dkt. No. 1247-11 at 7-8; Dkt. No. 1247-13 at 2-3).

2. Dr. Jewell's Replacement of Adjudicated Data

Dr. Jewell has not pointed to any methodological flaw on the part of the Endpoints

Committee and does not argue that any of their determinations were incorrect. Although he

initially made an assumption that the endpoint committee used a "non-standard" definition of

diabetes, (Dkt. No. 972-34 at 4), he readily admitted in deposition he simply did not know what

definition of diabetes they used. (Dkt. No. 972-40 at 42). Dr. Jewell testifies that he did not

know whether the Endpoints Committee "got it right or wrong" because he did not know the

definition used by the Endpoints Committee. (Dkt. No. 1247-14 at 250). He "presume[d] that an

explicit definition was provided to study investigators to measure that endpoint"; he just did not

"know exactly what that definition was." (Dkt. No. 972-40 at 42).

It is also important to note that even if Dr. Jewell had asked for and obtained the

definition used by the Endpoints Committee in ASCOT, he lacks the expertise to second guess

---

[32] While Dr. Jewell does not make this argument, Plaintiffs seem to imply in briefing that the published ASCOT results might differ from Dr. Jewell's results because the ASCOT analysis may not have been adjusted for other variables. (Dkt. No. 1053 at 53). However, Dr. Jewell admits that Defendant's 2010 analysis did control for such factors using the adjudicated data and found no association. (Dkt. No. 1247-11 at ¶ 14). Defendant's expert Dr. Wei also "applied Dr. Jewell's adjusted models to [the adjudicated data] and found no statistically significant effects of atorvastin." (Dkt. No. 972-39 at 3-4). In other words, this insinuation in briefing is a red herring. The difference in the authors' conclusion and Dr. Jewell's conclusion stems from their different determinations of whether new-onset diabetes has occurred, not whether researchers adjusted for covariates.

their judgments. Dr. Jewell is a statistician, not a medical doctor or medical professional. (Dkt.

No. 972-7 at 18-19). He has no expertise in diabetes, has never treated participants of any kind,

and is not a clinician. (*Id.*). He even testifies that "I don't quite know what [new-onset diabetes]

means" and that "[d]iabetes diagnosis, that's a clinical question, I'm not prepared to answer

clinical questions." (Dkt. No. 1247-14 at 251; Dkt. No. 972-7 at 65; *see also* Dkt. No. 972-7 at

67). As the Reference Manual points out, although statisticians "are most likely to use

appropriate procedures and correctly interpret the results . . . , the choice of which data to

examine or how best to model a particular process, could require subject matter expertise that a

statistician lacks." Reference Manual on Scientific Evidence 215 (3d ed. 2011); *see also id.*

(providing an example that a subject matter expert may be needed to "supply a definition" of the

relevant data). By contrast, the Endpoints Committee was made up entirely of clinicians who

had to use their clinical judgment in making certain determinations, such as whether, based on

medical records, particular participants had pre-existing diabetes. (Dkt. No. 1091-1 at 4).

In short, ASCOT applied a pre-specified, reliable and valid process to determine whether

participants developed new-onset diabetes, and Dr. Jewell has not pointed to any flaws in this

process. Dr. Jewell does not usually "second-guess the published results in a peer review

literature of any of the authors until a mistake is brought to [his] attention." (Dkt. No. 1245-2 at

4). However, without explanation, Dr. Jewell chose not to run his statistical analysis using this

adjudicated data. Had he done so, he would have reached the same conclusion as the authors of

the published, peer-reviewed article: the data does not show a statistically significant increase in

new-onset diabetes.

Instead, Dr. Jewell, someone with no clinical expertise, chose to replace the adjudication

committee's determination of new-onset diabetes with particular unadjudicated raw data, namely

lab values of his choice: "two or more on-treatment glucose values > 125 mg/dL." (Dkt. No.

972-34 at 11). This is a subset of the definition used by the Endpoints Committee, specifically

the first of the three criteria in the WHO definition. Using these lab values, Dr. Jewell concludes

that, although there is not a statistically significant result using a univariate analysis, there is a

significant association between Lipitor use and new-onset diabetes after controlling for certain

risk factors, with a hazard ratio of 1.31 and a 95% confidence interval of 1.06 to 1.62. (Dkt. No.

1247-11 at 2-3, 12).

Given that Dr. Jewell's lab value criterion for diabetes is a subset of the definition used

by the Endpoints Committee, Dr. Jewell should have found *fewer* total cases of diabetes than the

Endpoints Committee did.   However, he found a *greater* number. Dr. Jewell found 344 cases of

new-onset diabetes (187 Lipitor group and 157 in the placebo group), where the endpoint

committee found 288 cases (154 in the Lipitor group and 134 in the placebo group). (*Compare*

Dkt. No. 1247-11 at 12 *with* Dkt. No. 972-26 at 6, Table 3). This raises serious questions as to

the reliability of Dr. Jewell's determinations. There are two possible explanations for these

different counts: (1) the Endpoints Committee incorrectly counted the number of participants

with two glucose measurements $\geq$ 7.0 mmol/l by at least 56 participants or (2) the Endpoints

Committee determined that some of the participants with these lab values did not in fact have

new-onset diabetes.

Dr. Jewell may have counted the raw data correctly (i.e., who had two glucose readings

over 125 mg/dL), but this is where his analysis ends. The Endpoints Committee started with

these *potential* events of new-onset diabetes and then looked medical records and case files "to

ensure that the inclusion of a case was accurate." (*Id.* at 2-3). For instance, they would review

available information to determine whether the blood glucose measurements were, in fact,

fasting and reviewed medical records and histories rule out cases where the patient had diabetes at the beginning of the study. (*Id*. at 3-4). Dr. Jewell did not have access to these case files and medical records to rule out such cases. He clearly included participants in his count of "new-onset diabetes" that the Endpoints Committee did not.

Dr. Jewell concedes the value in "blinded adjudication of data," (Dkt. No. 972-40 at 70), but rejects the blinded, adjudicated data, ***without any reason to suspect an error in that data***. He then replaced this adjudicated data with particular lab values that he assumed were equivalent with new-onset diabetes and conveniently resulted in a statistically significant finding.[33] In other words, he "cherry-picked data from stud[y] that did not otherwise support his conclusion." *Burst v. Shell Oil Co.*, No. CIV.A. 14-109, 2015 WL 3755953, at *10 (E.D. La. June 16, 2015).

The ASCOT data did not support Dr. Jewell's conclusions, so he first simply "chose" to ignore the study and "not to study the data in ASCOT ." (Dkt. No. 972-7 at 120). When heavily criticized for his cherry picking of studies, Plaintiffs obtained leave from the Court for Dr. Jewell to consider the ASCOT study. Then he, without explanation, chose to ignore and not consider the adjudicated data of new-onset diabetes. Despite the fact that he didn't "quite know what [new-onset diabetes] means," (Dkt. No. 1247-14 at 251), he decided that, instead of using the data adjudicated by a blinded committee of clinicians that did understand the term, he would use unadjudicated raw data (particular lab values) that conveniently resulted in a statistically

---

[33] The finding was only significant in a multivariate analysis. (Dkt. No. 1247-11 at ¶32). In Dr. Jewell's univariate analysis, there was no statistically significant association between Lipitor and diabetes. (*Id.*). When discussing the NDA data, Dr. Jewell argues that "randomization should not allow imbalance" in confounding factors and, thus, he did not look to see if there was an imbalance in the groups. (Dkt. No. 1247-8 at 220). However, here, Dr. Jewell adjusts for confounding factors despite randomization. (Dkt. No. 1247-11 at ¶ 32). While there is not testimony on the point, Plaintiffs argue that Dr. Jewell made these adjustments because they were "pre-specified in the ASCOT protocol and the statistical analysis plan." (Dkt. No. 1159 at 15). Defendant has not contested this fact, and the Court will assume that it is true for the purposes of this motion.

significant finding. This is the very definition of cherry picking data to reach a pre-determined

conclusion and is unacceptable under *Daubert*. *See, e.g., Fail-Safe*, 744 F. Supp. 2d at 889 ("[I]t

is readily apparent that Dr. Keegan all but 'cherry picked' the data he wanted to use, providing

the court with another strong reason to conclude that the witness utilized an unreliable

methodology."); *Burst*, 2015 WL 3755953, at *10 (excluding expert who, among other things,

"cherry-picked data from studies that did not otherwise support his conclusion").

Furthermore, "case law . . . warns against use of medical literature to draw conclusions

not drawn in the literature itself." *Rutigliano v. Valley Bus. Forms*, 929 F. Supp. 779, 785

(D.N.J. 1996) *aff'd sub nom. Valley Bus. Forms v. Graphic Fine Color, Inc*., 118 F.3d 1577 (3d

Cir. 1997). This is not to say that a reanalysis of published data is never admissible, but to be

admissible, the expert must "validate" the reanalysis in some way. *Daubert v. Merrell Dow

Pharm., Inc*., 43 F.3d 1311, 1320 (9th Cir. 1995). An expert could do this by having his

reanalysis published in a peer-reviewed journal or by pointing to methodological flaws in the

published study and explaining how she corrected them. However, an expert cannot simply,

without any explanation for rejecting a published, peer-reviewed analysis, conduct his own

"reanalysis" solely for the purposes of litigation and testify that the data support a conclusion

opposite that of the studies' authors in a peer-reviewed publication. *See Ealy v. Richardson-

Merrell, Inc.*, 897 F.2d 1159, 1162-63 (D.C. Cir. 1990) (rejecting testimony of "the plaintiff's

epidemiology expert . . . [who] tried to refute the validity of the published epidemiological data

through her own unpublished reanalysis"); *Lynch v. Merrell-Nat'l Labs. Div. of Richardson-

Merrell, Inc*., 646 F. Supp. 856, 865 (D. Mass. 1986) ("Even if this Court were to find the

methodology of Dr. Swan's re-analysis credible, this Court still could not accept result-oriented

re-analysis of epidemiological studies . . . , such as that performed here by Dr. Swan, as reliable

data upon which to base an opinion on causation."), *aff'd*, 830 F.2d 1190 (1st Cir. 1987); *see also smith v. Ortho Pharm. Corp.*, 770 F. Supp. 1561, 1579 (N.D. Ga. 1991) ("A scientific study not subject to peer review has little probative value.").

In sum, neither Dr. Jewell's analysis nor any other analysis has called into question the integrity of ASCOT's Endpoints Committee, the methodology it used, or ASCOT-LLA's findings, which were published a peer-reviewed journal. Dr. Jewell's "reanalysis" using cherry-picked data and ignoring adjudicated data without reason to reach the opposite conclusion is results driven and unreliable.

III. Dr. Jewell's "General" Opinions

Defendant argues that Dr. Jewell cannot opine as to causation because, in his initial report, he failed to looked at the one study designed to test for an association between Lipitor and diabetes and based his opinion almost exclusively on his *post hoc* analysis of SPARCL. (Dkt. No. 972 at 14; see also id. at 39-40). Defendant also points out that Dr. Jewell ignored CARDS in his analysis of efficacy and chose to avoid a gender analysis of ASCOT, which would have provided results contrary to those he found in the SPARCL data. (*Id.* at 41, 45-46).

Defendant points out that this is contrary to Dr. Jewell's approach in other litigation where he "looked at all the evidence as of 2014." (Dkt. No. 1245-1 at 7). He testified in the *Zoloft* litigation that a failure to consider published studies was "unacceptable . . . as a thorough investigation of the issue." (*Id.* at 17; *see also id.* at 21 ("We have demonstrated for the record that most, if not all of them, failed to do a comprehensive review of the literature."); *id.* at 22 ("I think it's better to look at . . . the comprehensive review of the literature by a statistician if you are going to write a statistical report.")).

Defendant is correct that general causation opinions cannot be based on cherry-picked studies and the avoidance of all contrary evidence. *In re Zoloft (Sertraline Hydrochloride) Products Liab. Litig.*, 26 F. Supp. 3d 449, 460-61 (E.D. Pa. 2014) ("The Court finds that the expert report prepared by Dr. Bérard does selectively discuss studies most supportive of her conclusions . . . and fails to account adequately for contrary evidence, and that this methodology is not reliable or scientifically sound."), *reconsideration denied*, No. 12-MD-2342, 2015 WL 314149 (E.D. Pa. Jan. 23, 2015). However, the Court finds these arguments mooted by Plaintiffs' representations to the Court that at trial Dr. Jewell will not be offering a causation opinion and Dr. Jewell's testimony will be limited to his opinions in his report. (Dkt. No. 1053 at 8, 23). In his report, Dr. Jewell does not make general opinions based on a comprehensive review of the literature. (*See* Dkt. No. 1247-9). He does not survey the literature and opine that there is an established association between Lipitor and diabetes and does not opine that a survey of the literature shows a greater risk in women than men. (*Id.*).

The opinions in his report are confined to very particular data sets. For example, he opines that "[a]nalyses of the Defendant-sponsored SPARCL trial demonstrate that there was significantly increased risk of new-onset diabetes with 80 mg of atorvastatin compared to placebo" and that the "Relative Risk was greater in women." (Dkt. No. 1247-9 at ¶ 7). Dr. Jewell can be cross-examined on the fact that his gender analyses of TNT and IDEAL did not show statistically significant differences between genders and that he did not conduct a gender analysis of ASCOT.

Because Dr. Jewell chose to only look at particular data sets, his opinions must be limited to these data sets and he will not be allowed to make sweeping opinions that imply he did a
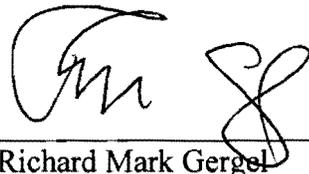
comprehensive literature review. However, Dr. Jewell's opinions in his report are so limited and therefore, not excludable on this basis.

IV. Conclusion

The Court finds that Dr. Jewell's analysis of the NDA data and ASCOT data was results driven, that Dr. Jewell's methodology and selection of relevant evidence changed based on the results they produced, and that Dr. Jewell chose to ignore and exclude from his report his own analyses that did not support his ultimate opinions. It is apparent to the Court that rather than conducting statistical analyses of the data and then drawing a conclusion from these various analyses, Dr. Jewell formed an opinion first, sought statistical evidence that would support his opinion and ignored his own analyses and methods that produced contrary results.

While the particular statistical tests and models used by Dr. Jewell are reliable, Dr. Jewell's methodology for determining the inputs into these statistical models and tests is unreliable and his application of those models (e.g., trying various models until he obtains a particular result) is also unreliable. Therefore, the Court excludes Dr. Jewell's testimony regarding the NDA data and ASCOT data under Rule 702 and *Daubert*.

AND IT IS SO ORDERED.

_____

Richard Mark Gergel
United States District Court Judge

November 20 , 2015
Charleston, South Carolina

35