UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF FLORIDA
PENSACOLA DIVISION

IN RE: ABILIFY (ARIPIPRAZOLE) PRODUCTS LIABILITY LITIGATION

This Document Relates to All Cases

Case No. 3:16-md-2734

Chief Judge M. Casey Rodgers
Magistrate Judge Gary R. Jones

/

AMENDED ORDER[1]

This is a multidistrict product liability action against the manufacturers and marketers of the prescription drug Aripiprazole, more commonly known as Abilify.[2] Plaintiffs allege that, after taking Abilify as prescribed, they developed impulsive and irrepressible urges to engage in certain harmful behaviors, including impulsive gambling, eating, shopping, and sex.[3] Defendants deny the allegations and maintain that Abilify could not, and did not, cause Plaintiffs' impulse control problems.

Defendants have moved for summary judgment on the issue of general causation—that is, whether Abilify is capable of causing uncontrollable impulses to engage in certain harmful behaviors. *See* ECF No. 428. Both the motion, *see id*.,

[1] Portions of the briefing and record in this matter were filed under seal. This redacted version of the Court's Order omits references to information included in sealed materials.

[2] The Court will use these terms interchangeably.

[3] In its pharmacovigilance review of Abilify, the United States Food and Drug Administration ("FDA"), defined "impulse control disorders" to include "pathological gambling (PG; also known as gambling disorder or compulsive gambling), compulsive sexual behavior (*i.e*., hypersexuality or sexual addiction), compulsive buying/shopping (*i.e*., shopping addiction), and compulsive eating (*i.e*., binge eating)." *See* FDA Pharm. Vigil., ECF No. 428-11 at 5. The FDA's pharmacovigilance review is discussed more fully in the body of this Order.

and the response, *see* ECF No. 463, are supported by expert testimony. Each side challenges the other's experts as unreliable and those motions are also pending.[4] A four-day evidentiary hearing was conducted jointly with Magistrate Judge Gary R. Jones of this Court, and Judge James J. Deluca of the New Jersey Superior Court, who presides over multiple similar cases in New Jersey state court. Now, having carefully considered the law, the voluminous record, and the parties' arguments, the Court concludes that Plaintiffs have satisfied their burden to demonstrate that a genuine dispute of material fact exists as to whether Abilify can cause uncontrollable impulsive behaviors in individuals taking the drug.

I. Background

Abilify is an atypical antipsychotic drug developed and manufactured by Defendants Otsuka Pharmaceutical Co., Ltd. and Otsuka America Pharmaceutical, Inc., who jointly market and distribute it in the United States with Defendant Bristol-Myers Squibb Company (collectively, "Defendants"). *See* Master Complaint, ECF

[4] There are ten motions to exclude experts currently pending. Plaintiffs move to exclude the opinions of Defendants' five expert witnesses: Marc N. Potenza, ECF No. 415; Pierre Blier, ECF No. 418; Douglas Weed, ECF No. 419; Deborah B. Leiderman, ECF No. 420; and Catharine Winstanley, ECF No. 422. Defendants move to exclude the general causation opinions of Plaintiffs' five expert witnesses: Antoine Bechara, ECF No. 423, Joseph Glenmullen, ECF No. 424, Eric Hollander, ECF No. 425, Russell Luepker, ECF No. 426, and David Madigan, ECF No. 427. In this Order, "DX-" refers to Defendant's exhibits at the *Daubert* hearing and "PX-" refers Plaintiffs' exhibits at the hearing.

No. 108-1 at 5.[5] In 2002, Abilify was approved by the Food and Drug Administration ("FDA") for the treatment of schizophrenia. Since then, Abilify also has been approved for use in patients with bipolar disorder, irritability associated with autistic disorder, Tourette's Syndrome, and as an add-on treatment for major depressive disorder. *See* Product Label, ECF No. 428-1 at 2. "[T]ens of millions of patients worldwide have used Abilify to help manage the symptoms of these very debilitating mental health conditions." *See* DSJ, ECF No. 428-26 at 9.[6]

In 2010, the first published reports suggesting a possible link between Abilify and pathological gambling began appearing in the medical literature. More published reports followed, as well as hundreds of informal reports from patients and healthcare professionals to Defendants and the FDA, describing the onset of impulsive gambling and other impulse control disorders in patients treated with Abilify. The scientific community, the FDA, Defendants, and public health agencies worldwide took notice and began examining whether Abilify is linked to impulse control disorders. The research findings and conclusions of these bodies are at the heart of the motions currently pending before this Court.

[5] "Master Complaint" refers to Plaintiffs' Master Form Complaint and Jury Demand, ECF No. 108-1.

[6] "DSJ" refers to Defendants' Motion for Summary Judgment on General Causation, ECF No. 28-26.

In 2012, following a safety review of Abilify based on reports of pathological gambling with patients' use of the drug, the European Medicines Agency ("EMA") required Defendants to modify the drug's product label in Europe to include pathological gambling as a possible "post-marketing undesirable effect" of Abilify and to warn of an "increased risk" of pathological gambling in Abilify patients with a prior history of gambling.[7] *See* FDA Pharm. Vigil., ECF No. 428-11 at 5, 12.[8] In November 2015, Health Canada also found an "increased risk" of pathological gambling, as well as hypersexuality, with Abilify use and ordered that the drug's product monograph in Canada be updated to advise of these possible adverse effects.[9] *See id*. at 5, 12. Health Canada's safety review and subsequent product monograph update prompted the FDA to initiate a pharmacovigilance review in the United States to evaluate whether the potential link between Abilify and impulse control disorders presented a "safety issue warrant[ing] any regulatory action." *See id*. at 5. The FDA's review identified an association between Abilify and impulse

[7] The European Medicines Agency is an international public health agency charged with the scientific evaluation, supervision and safety monitoring of medicines for the European Union. *See* http://www.ema.europa.eu/ema/. Abilify's European label history may be found at: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2012/10/WC500134109.pdf.

[8] "FDA Pharm. Vigil." refers to the FDA's Abilify Pharmacovigilance Review dated March 10, 2016, ECF No. 428-11 at 5.

[9] Health Canada, Safety Information for Antipsychotic Drug Abilify and Risk of Certain Impulse-Control Behaviors (Nov. 2, 2015), http://healthycanadians.gc.ca/recall-alert-rappel-avis/hc-sc/2015/55668a-eng.php. Health Canada is Canada's national public health agency. *See* https://www.canada.ca/en/health-canada/corporate/about-health-canada.html.

control disorders, based on an analysis of cases in its adverse event reporting database (FAERS), the published scientific literature, and Defendant's clinical trial and post-marketing patient data.[10] *See id*. at 4, 29. On May 3, 2016, the FDA issued a safety warning that "uncontrollable and excessive urges" to "gamble, binge eat, shop and have sex" had been reported with the use of Abilify, even in patients with no prior history of impulsive behaviors.[11] In August 2016, the FDA required Defendants to modify Abilify's product label in the United States to warn of "post-marketing case reports suggest[ing] that patients can experience intense urges, particularly for gambling, and the inability to control these urges while taking" the drug. *See* Product Label, ECF No. 428-1 at 2, 24. The United States product label was also modified to warn of "[o]ther compulsive urges, reported less frequently, [which] include: sexual urges, shopping, eating or binge eating, and other impulsive or compulsive behaviors." *See id*. At that point, Abilify had been on the market in the United States for almost 14 years.

[10] The FDA's adverse event reporting system, known as FAERS, contains information on adverse event and medication error reports submitted to the FDA by patients, health care professionals, and pharmaceutical companies.

[11] *FDA Drug Safety Communication: FDA Warns About New Impulse-Control Problems Associated with Mental Health Drug Aripiprazole (Abilify, Abilify Maintena, Aristada)*; https://www.fda.gov/Drugs/DrugSafety/ucm498662.htm (last visited Dec. 3, 2017).

A short biochemistry discussion may be helpful at this point.[12] The human brain is a tremendously complex biochemical system. It contains billions of interconnected nerve cells, called neurons, that use chemical and electrical signals to send information throughout the body. The function of a neuron is to process and transmit information—it receives signals from other neurons, integrates and interprets those signals, and transmits signals to other, adjacent neurons. The signals within neurons are carried throughout the brain in the form of electrical impulses. When a signal is sent from one neuron to another, it must cross a microscopic gap between the two communicating neurons. This gap is called a synapse or synaptic cleft. At the synapse, the electrical signal within the neuron is converted to a chemical signal and sent across the synapse towards the receiving neuron.[13] This chemical signal is transported by molecules, called neurotransmitters, that attach to special structures on the outer surface of the receiving neuron, called receptors.[14] There are many different types of receptors, categorized by the type of neurotransmitters with which they interact. The attachment of neurotransmitters to receptors can either stimulate or inhibit electrical activity in the receiving neuron,

[12] This biochemistry discussion is grounded in undisputed expert testimony and reports presented by both sides.

[13] The neuron sending the message is called the presynaptic cell. The neuron receiving the message is called the postsynaptic cell.

[14] The major neurotransmitters include acetylcholine, adrenaline, dopamine, endorphins, GABA, glutamate, norepinephrine, and serotonin.

depending on which neurotransmitter is released and which receptors it activates. In any one synapse, there may be hundreds of neurotransmitters continually moving between, and acting on, neurons, triggering varying physiological effects throughout the brain and the body. Any disruption to the neuronal communication process—whether to the production, release, or attachment of the various neurotransmitters—can alter brain function and, as it relates to this case, human behavior.

Dopamine is a neurotransmitter in the central nervous system that is believed to play an integral role in a number of physiological processes, including movement, cognition, emotional stability, and, relevant to this case, reward-motivated behaviors. It acts on five different receptors—D_1, D_2, D_3, D_4, and D_5—along four major pathways in the brain—the nigrostriatal pathway, the mesocortical pathway, the mesolimbic pathway, the tuberoinfundibular pathway.[15] This case is primarily concerned with the activity of dopamine in the mesolimbic pathway, which regulates pleasure, reward processing, and motivation. Under normal circumstances, the brain responds to rewarding activities or stimuli by releasing dopamine into the mesolimbic pathway, where it binds with dopamine receptors to produce feelings of pleasure. As dopamine levels subside, so do the feelings of pleasure. If the

[15] The Nigrostriatal Pathway covers movement and sensory stimuli. The Mesocortical Pathway covers cognition, memory, attention, emotional behavior, and learning. The Mesolimbic Pathway regulates pleasure, reward processing, and motivation. The Tuberoinfudibular Pathway controls the hypothalamic pituitary endocrine system, and inhibition of prolactin secretions.

rewarding activity is repeated, then dopamine is again released, and more feelings of pleasure are produced. The release of dopamine and the resulting pleasurable feelings serve as positive reinforcements that motivate repetition of the pleasure-inducing activity.

Pharmaceutical companies create drugs that can mimic, duplicate, or block the activity of natural, or "endogenous," dopamine in the brain. The effect of a given drug depends on two pharmacological properties that relate to the manner in which the drug interacts with dopamine receptors: affinity and intrinsic activity. Affinity refers to whether and how tightly the drug binds to dopamine receptors. Intrinsic activity refers to the degree to which the drug, once bound, activates dopamine receptors to produce a measurable physiological effect. Based on these properties, drugs that bind to dopamine receptors can act as agonists or antagonists. A full agonist has both high affinity and 100% intrinsic activity, meaning that it binds tightly to dopamine receptors and mimics the activity of dopamine, producing the same level of physiological response that dopamine naturally produces. Antagonists bind to dopamine receptors, but produce no physiological effects; instead, they simply occupy a receptor site, thereby preventing endogenous dopamine from binding to and activating it. A partial agonist binds to dopamine receptors, but produces less of a response than a full agonist. The functional activity of some partial agonists depends on the presence or absence of endogenous dopamine in the

surrounding area. Where dopamine concentrations are high, the partial agonist functions as an antagonist (*i.e.*, functional antagonist), but where dopamine concentration is low, the partial agonist functions as a full agonist (*i.e.*, functional agonist). In this case, Plaintiffs' position as to how Abilify causes impulse control problems centers on how the drug binds and interacts with two dopamine receptors—D_2 and D_3—to produce physiological effects in the form of impulsive behaviors.

II. Expert Challenges

To establish general causation, Plaintiffs have proffered the testimony of five experts: Dr. Antoine Bechara, Dr. Joseph Glenmullen, Dr. Eric Hollander, Dr. Russell V. Luepker, and Dr. David Madigan. Simply stated, each of Plaintiffs' experts opines that Abilify can cause impulsive behaviors and each presents scientific evidence in support of his conclusion. Defendants challenge the admissibility of Plaintiffs' expert testimony on general causation as unreliable under Federal Rule of Evidence 702 and *Daubert v. Merrell Dow Pharmaceuticals, Inc*., 509 U.S. 579 (1993). More specifically, Defendants contend that Plaintiffs have failed to provide *reliable* scientific evidence demonstrating a statistically significant association between Abilify and impulsive behaviors. According to Defendants, this omission is a fatal flaw in Plaintiffs' case because their remaining evidence is insufficient as a matter of law, even in the aggregate, to establish general causation

by Eleventh Circuit standards. Defendants rely on five experts of their own—Dr. Pierre Blier, Dr. Deborah Leiderman, Dr. Marc Potenza, Dr. Douglas Weed, and Dr. Catherine Winstanley—to support their position that Abilify cannot cause impulse control disorders, each of whom submitted an opinion that Plaintiffs in turn challenge as unreliable.

A. Legal Standard for Expert Testimony

Rule 702, as explained by *Daubert* and its progeny, governs the admissibility of expert testimony. *Rink v. Cheminova, Inc.*, 400 F.3d 1286, 1291 (11th Cir. 2005). Under Rule 702 and *Daubert*, district courts are compelled to act as "gatekeepers" to ensure the reliability and relevancy of expert testimony. *Id*. (quoting *Daubert*, 509 U.S. at 589). Expert testimony is reliable and relevant—and, therefore, admissible—when the following criteria are met: (1) the expert is sufficiently qualified to testify about the matters he intends to address; (2) the methodology used is "sufficiently reliable as determined by the sort of inquiry mandated in *Daubert*; and (3) the testimony assists the trier of fact, through the application of scientific, technical, or specialized expertise, to understand the evidence or to determine a fact in issue." *Id*. The Eleventh Circuit refers to these criteria separately as "qualification, reliability, and helpfulness," *United States v. Frazier*, 387 F.3d 1244, 1260 (11th Cir. 2004), and has emphasized that they are "distinct concepts that courts and litigants must take care not to conflate," *Quiet Tech. DC-8, Inc. v. Hurel-Dubois*

UK Ltd., 326 F.3d 1333, 1341 (11th Cir. 2003). The party offering the expert has the burden of showing, by a preponderance of the evidence, that each of these requirements is met. *Rink*, 400 F.3d at 1292.

To meet the qualification requirement, a party must show that its expert has sufficient "knowledge, skill, experience, training, or education to form a reliable opinion about an issue that is before the court." *Hendrix ex. Rel. G.P. v. Evenflo Co., Inc*., 609 F.3d 1183, 1193 (11th Cir. 2010) (citing Fed. R. Evid. 702) ("*Hendrix II*"), *aff'g* 255 F.R.D. 568 (N.D. Fla. 2009) ("*Hendrix I*"). The qualifications standard for expert testimony is "not stringent" and "[s]o long as the witness is minimally qualified, objections to the level of [his] expertise [go] to credibility and weight, not admissibility." *Hendrix I*, 255 F.R.D. at 585.

To meet the reliability requirement, an expert's opinion must be based on scientifically valid principles, reasoning, and methodology that are properly applied to the facts at issue. *Frazier*, 387 F.3d at 1261-62. The reliability analysis is guided by several factors, including: (1) whether the scientific technique can be or has been tested; (2) whether the theory or technique has been subjected to peer review or publication; (3) whether the technique has a known or knowable rate of error; and (4) whether the technique is generally accepted in the relevant community. *Daubert*, 509 U.S. at 593-94. "[T]hese factors do not exhaust the universe of considerations that may bear on the reliability of a given expert opinion, and a federal court should

consider any additional factors that may advance its Rule 702 analysis." *Quiet Tech.*, 326 F.3d at 1341. The court's focus must be on the expert's principles and methodology, not the conclusions they generate. *Daubert*, 509 U.S. at 595. The test for reliability is "flexible" and courts have "broad latitude" in determining both how and whether this requirement is met. *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 141-42 (1999).

Finally, to satisfy the helpfulness requirement, expert testimony must be relevant to an issue in the case and offer insights "beyond the understanding and experience of the average citizen." *United States v. Rouco*, 765 F.2d 983, 995 (11th Cir. 1985). Relevant expert testimony "logically advances a material aspect of the proposing party's case" and "fits" the disputed facts. *McDowell v. Brown*, 392 F.3d 1283, 1298-99 (11th Cir. 2004). Expert testimony does not "fit" when there is "too great an analytical gap" between the facts and the proffered opinion. *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 147 (1997).

When scrutinizing the reliability and relevance of expert testimony, a court must remain mindful of the delicate balance between its role as a gatekeeper and the jury's role as the ultimate factfinder. *Frazier*, 387 F.3d at 1272. The court's gatekeeping role "is not intended to supplant the adversary system or the role of the jury." *Allison v. McGhan Med. Corp.*, 184 F.3d 1300, 1312 (11th Cir. 1999). Only the jury may determine "where the truth in any case lies" and the court "may not

usurp this function." *Frazier*, 387 F.3d at 1272. Thus, a court may not "evaluate the credibility of opposing experts" or the persuasiveness of their conclusions, *Quiet Tech*, 326 F.3d at 1341; instead, its duty is limited to "ensur[ing] that the fact-finder weighs only sound and reliable evidence, *Frazier*, 387 F.3d at 1272.

B. Reliability of Expert Testimony on General Causation

To prevail in a pharmaceutical products liability case, a plaintiff must establish both general and specific causation through reliable expert testimony.[16] *Chapman v. Procter & Gamble Distributing, LLC*, 766 F.3d 1296, 1303-04 (11th Cir. 2014). General causation is established by demonstrating, often through a review of scientific or medical literature, that a drug or chemical can, in general, cause the type of harm alleged by the plaintiff. *See Hendrix II*, 609 F.3d at 1196. Specific causation is established by showing that exposure to the allegedly toxic drug or chemical actually caused an individual plaintiff's injury. *Id*. Only general causation—whether Abilify is capable of causing impulse control disorders, such as impulsive gambling—is at issue in the motions currently pending before the Court.

The Eleventh Circuit has developed an extensive body of *Daubert* jurisprudence around the reliability of different categories of scientific evidence that

[16] This is true except in the small number of cases where the medical community recognizes and agrees that a particular substance is toxic, in which case, general causation is accepted. *See Chapman*, 766 F.3d at 1303-04. The parties agree there is no consensus in the medical community about whether Abilify causes impulse control disorders.

may support an expert opinion on general causation. The Eleventh Circuit considers three "primary" methodologies "indispensable" for proving that a drug can cause a specific adverse effect: epidemiological studies, dose-response relationship, and background risk of disease. *Chapman*, 766 F.3d at 1308. A general causation opinion that is not supported by at least one of these primary methodologies is unreliable as a matter of law. *See id*. An expert who has reliably applied primary methodologies may bolster his general causation opinion with evidence from "secondary" methodologies, such as: biological plausibility,[17] case studies and adverse event reports, extrapolations from animal and *in vitro* studies, and extrapolations from analogous drugs. *See id*. Importantly, the flaws inherent in the secondary methodologies limit their reliability under *Daubert*. *See id*. For this reason, secondary methodologies alone, even in the aggregate, cannot establish general causation. *See id*.; *see also Hendrix II*, 609 F.3d at 1197.

In this case, the parties' experts offer various combinations of primary and secondary methodologies in support of their general causation opinions. To frame the Court's analysis of the expert opinions, a brief review of the scientific and legal principles governing the reliability of each methodology follows.

[17] Biological plausibility is also referred to as a "plausible biological mechanism of action."

1. Primary Methodologies

a. Epidemiological Studies

The "best evidence of causation in toxic tort actions" is grounded in epidemiology, *Rider v. Sandoz Pharmaceuticals Corp*., 295 F.3d 1194, 1199 (11th Cir. 2002), which is the branch of science that studies the incidence, distribution, and cause of disease in human populations, Reference Manual on Scientific Evidence at 551 ("Ref. Man."). The first step in establishing causation through epidemiology is to demonstrate that exposure to a drug is associated with a particular disease or adverse effect.[18] Ref. Man. at 566. Once an association is identified, scientists next determine whether the association represents "a true cause-effect relationship" between exposure and the disease. Ref. Man. at 597. This is the *sine qua non* of general causation.

This causation inquiry is guided by nine well-established factors, known in the scientific community as the Bradford Hill factors.[19] These include: (1) temporal

[18] Epidemiologists use clinical trials, cohort studies, case-control studies, cross-sectional studies, and/or ecological studies to determine whether exposure increases the risk of developing a particular disease or adverse effect, by comparing individuals exposed to a particular agent with unexposed individuals. Ref. Man. at 555-63.

[19] Sir Austin Bradford Hill was a world-renowned epidemiologist who articulated a nine-factor set of guidelines that is widely accepted in the scientific community for determining whether an observed association between an agent and a disease reflects a true causal relationship. *See* Ref. Man. at 600; *see also* Austin Bradford Hill, *The Environment and Disease: Association or Causation?*, 58 PROCEEDINGS ROYAL SOC'Y MED. 295 (1965) ("Bradford Hill Article"), ECF No. 460-4.

relationship; (2) strength of the association; (3) dose-response relationship; (4) replication of the findings; (5) biological plausibility; (6) consideration of alternative explanations; (7) cessation of exposure; (8) specificity of the association; and (9) consistency with other knowledge. *See* Ref. Man. at 599-600. No one factor is dispositive. *Id*. at 600. Determining whether an association is causal is a matter of scientific judgment, and scientists reliably applying the Bradford Hill factors may reasonably come to different conclusions about whether a causal inference may be drawn. *Milward v. Acuity Specialty Products Group, Inc*., 639 F.3d 11, 18 (1st Cir. 2011); *see also* Ref. Man. at 553, 600.

An epidemiological study identifying a statistically significant association between the use of a drug and a particular adverse effect, accompanied by a reliable expert opinion that the association is causal, is "powerful" evidence of general causation. *See Rider*, 295 F.3d at 1198. The absence of epidemiological evidence, however, does not preclude admission of a general causation opinion in the Eleventh Circuit. *See Kilpatrick v. Breg, Inc.*, 613 F.3d 1329, 1336-37 (11th Cir. 2010); *see also Rider*, 295 F.3d at 1198-99; *Wells v. Ortho Pharm. Corp*., 788 F.2d 741, 745 (11th Cir. 1986). Experts may rely on other, non-epidemiological evidence to prove causation; but where epidemiology is lacking, "the nature of the other evidence . . . becomes that much more important, and [a] court's consideration of such evidence

and the methodologies used must be that much more searching." *See Kilpatrick*, 613 F.3d at 1337 n.9.

b. Dose-Response Relationship

Another primary methodology for establishing causation is through evidence of a dose-response relationship, which is a "relationship in which a change in amount, intensity, or duration of exposure to [a drug] is associated with a change—either an increase or decrease—in risk of" adverse effects from that exposure. *McClain v. Metabolife Int'l, Inc*., 401 F.3d 1233, 1242-43 (11th Cir. 2005). The relationship between dose and response is "the hallmark of basic toxicology" and the "single most important factor to consider" in evaluating the toxicity of a drug. *Id*. at 1242; *see also Chapman*, 766 F.3d at 1307. This is because virtually all substances have the potential to be harmful at high enough doses. *See Chapman*, 766 F.3d at 1307; *see also* Ref. Man. at 636. Inherent in this principle is the fact that, for the vast majority of substances, there are threshold doses below which no individual will respond and doses above which nearly everyone responds. *See McClain*, 401 F.3d at 1241-43. Consequently, a reliable expert opinion on general causation should address what levels of exposure to a drug increase the risk of adverse effects. *See id*. at 1241. Indeed, "[t]he expert who avoids or neglects this principle of toxic torts without justification casts suspicion on the reliability of his methodology." *Kilpatrick*, 613 F.3d at 1339 (quoting *McClain*, 401 F.3d at 1242).

c. Background Risk

A reliable methodology also should take into account the background risk for the disease at issue in the case. *McClain*, 401 F.3d at 1243. Background risk is the risk that members of the general public would have of developing the disease without exposure to the drug. *Id*. It encompasses all causes of the disease, whether known or unknown, except for the drug in question. *Id*. This is important because the aim of the other primary methodologies is to identify "agents that are associated with an increased risk of disease." *See* Ref. Man. at 552. An expert must know the background prevalence of a disease before he can determine whether the risk of that disease is increased as a result of exposure to the agent. *See In re Denture Cream Products Liability Litigation*, 795 F. Supp. 2d 1345, 1355 (S.D. Fla. 2011), *aff'd*, *Chapman*, 766 F.3d 1296. Without background risk to establish a baseline, it is difficult to determine whether any incidence of a disease in individuals exposed to an agent is anything more than a coincidence. *Chapman*, 766 F.3d at 1308. Thus, a failure to identify or describe the background risk of a disease is a "serious methodological deficiency" and "substantial weakness" in an expert's general causation opinion. *See id*.

2. Secondary Methodologies

a. Biological Plausibility

Biological plausibility refers to a credible scientific explanation of the physiological processes or mechanisms by which a drug can cause a particular disease or adverse effect, based on current biological and pharmacological knowledge. *See* Ref. Man. at 604; *see also McClain*, 401 F.3d at 1253. Importantly, biological plausibility is not the same as biological certainty. *See Daubert*, 509 U.S. at 590 ("Of course, it would be unreasonable to conclude that the subject of scientific testimony must be 'known' to a certainty; arguably, there are no certainties in science."); *Jones v. Otis Elevator Co*., 861 F.2d 655, 662 (11th Cir. 1988) (stating that "absolute certainty is not required" from expert testimony). That is, an expert on biological plausibility need not definitively prove the biological means by which a drug acts in the body. *See, e.g.*, *In re Neurontin Mktg. Sales Practices & Prods. Liab. Litig*., 612 F. Supp. 2d 116, 149 (D. Mass 2009) (finding that biological plausibility supported expert's opinion on causation despite the fact that there was "robust debate in the scientific community" on the proposed mechanism); *In re PPA Prods. Liab. Litig*., 289 F. Supp. 2d 1230, 1247 (W.D. Wash. 2003) ("The fact that the mechanism remains unclear does not call the reliability of the opinion into question."). Instead, a biological plausibility opinion is admissible so long as it is derived from and supported by reliable scientific knowledge and reasoning. *See*

Allison, 184 F.3d at 1319, n.23 ("While scientific testimony need not be known to a certainty, *Daubert* does require that assertions be derived from scientific knowledge."); *In re Seroquel Prods. Liab. Litig*., No. 6:06-md-1769, 2009 WL 3817866, *5 (M.D. Fla. Feb. 11, 2009) (finding biological plausibility opinion reliable where each step in expert's methodology had "ample scientific support" and was supported by "sound scientific reasoning"). Although biological plausibility, without more, cannot establish general causation; its existence "lends credence to an inference of causality" drawn from other, more substantial evidence. *See* Ref. Man. at 604; *see also Chapman*, 766 F.3d at 1308; *Rider*, 295 F.3d at 1202; *Milward*, 639 F.3d at 25-26.

b. Case Studies and Adverse Event Reports

Case studies document medical observations occurring coincident with the use of a prescription drug either by a single patient (a case report) or a small number of patients (a case series). *Rider v. Sandoz Pharmaceuticals Corp*., 295 F.3d 1194, 1199 (2002). They tend to be brief recitals of clinical events and do not address prior medical history, use of other medications or drugs, risk factors, or the myriad of other issues necessary to scientifically evaluate whether the drug actually produced the observed adverse effect. *See id*. Moreover, case reports have no controls, are

susceptible to inherent reporting biases,[20] lack statistical context, and are not verifiable through meaningful peer review. *See id*.; *see also Kilpatrick*, 613 F.3d at 1338. The difficulty with case reports is distinguishing between association and causation. For this reason, while case reports may supplement other evidence of causation, they cannot, standing alone, prove causation. *See id.*

One type of case report is more worthy of consideration in the general causation assessment, however. This report documents a patient's dechallenge and rechallenge events while taking a particular drug. A dechallenge event occurs where a patient's adverse side effects partially or completely disappear once the drug is stopped. *Rider*, 295 F.3d at 1199. If the side effects return when the patient resumes taking the drug, that is known as a rechallenge. *See id*. As other courts have noted, dechallenge and rechallenge data is "substantially more valuable than run-of-the-mill case reports because a patient's reactions are measured against his own prior reactions." *Glastetter v. Novartis Pharm. Corp.*, 252 F.3d 986, 990 (8th Cir. 2001); *see also Rider*, 295 F.3d at 1199 ("These reports, which may be analogized to controlled studies with one subject, can be particularly useful in determining whether a causal relationship exists."); *Hollander v. Sandoz Pharmaceuticals Corp*., 289 F.3d 1193, 1212 (10th Cir. 2002); *Giles v. Wyeth, Inc*., 500 F. Supp. 2d 1048, 1051

[20] Inherent biases may include selection bias, conceptual bias, referral bias, or over-reporting of symptoms. *See Barrow v. Bristol-Myers Squibb Co.*, No. 6:96-cv-689, 1998 WL 812318, at *23 n.217 (M.D. Fla. Oct. 29, 1998).

n.7 (S.D. Ill. 2007). Nevertheless, dechallenge and rechallenge events "are still case reports and do not purport to offer definitive conclusions as to causation." *Rider*, 295 F.3d at 1200. Their "value is directly related to the degree of scientific control used in the" dechallenge or rechallenge exercise. *See McClain*, 401 F.3d at 1255. Thus, descriptions of dechallenge and rechallenge events generally are more reliable and probative of causation than a typical case report; however, alone, they cannot establish causation. *See id*.

Adverse event reports describe medical events that occurred during or after an individual's use of a prescription drug, which are submitted directly to the FDA by patients, healthcare professionals, and drug manufacturers.[21] *See* 21 C.F.R. § 314.80(a). Generally speaking, in practice, the FDA adverse events reporting system (FAERS) simply entails "consumers call[ing] in to describe medical problems that they think they are experiencing from taking a product." *McClain*, 401 F.3d at 1250. As a result, the system has several intrinsic limitations, including (1) uncertainty that the drug actually caused the reported event, since the FDA does not require that causation be proven before adverse event data is reported; (2) insufficient detail from which to evaluate causation; (3) information in the reports is unverified and subject

[21] Physicians and patients are encouraged to report such events voluntarily, whereas drug manufacturers are required to do so. *See, e.g.*, 21 C.F.R. § 310.305 (Records and reports concerning adverse drug experiences on marketed prescription drugs for human use without approved new drug applications.); 21 C.F.R. § 312.32 (Investigational new drug application safety reporting.); 21 C.F.R. § 314.80 (Post-marketing reporting of adverse drug experiences.).

to a variety of reporting biases; and (4) the underlying data may be affected by reporting bias stemming from publicity or litigation. *See* Ref. Man. at 731.[22] Consequently, because they contain what amounts to "[u]ncontrolled anecdotal information," adverse event reports are generally considered "one of the least reliable sources" of support for a causation opinion. *McClain*, 401 F.3d at 1250.

c. *In vivo* and *In vitro* Studies

Toxicological knowledge often derives from *in vivo* studies, in which laboratory animals are exposed to a particular drug, with the outcomes monitored and compared to those for an unexposed control group.[23] Ref. Man. at 639. *In vivo* studies offer a number of advantages, including that they can be conducted as true experiments, with exposure controlled and measured, they are replicable, they usually follow a generally accepted methodology, and they present fewer ethical limitations than human experimentation. *See id*. at 563; *see also In re Paoli R.R. Yard PCB Litigation*, 35 F.3d 717, 781 (3d Cir. 1994); *In re Accutane Products Liability*, 511 F. Supp. 2d 1288, 1291 (M.D. Fla. 2007). However, the use of animal studies to prove causation in humans has "two significant disadvantages," which

[22] *See also* United States Food and Drug Administration, *Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment* (March 2005) at 9, https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071696.pdf (last visited Dec. 3, 2017) ("FDA Pharm. Guide").

[23] The term "*in vivo*" encompasses studies conducted on any living organisms, including human subjects. *Kilpatrick*, 613 F.3d at 1340, n.16.

"are almost always fraught with considerable, and currently unresolvable, uncertainty." Ref. Man. at 563. First, extrapolating from animals to humans is difficult because biological "differences in absorption, metabolism, and other factors may result in interspecies variation in responses." *Id*. Second, most animal studies involve significantly higher doses of a drug than would ever be present in humans. *Id*. For these reasons, while animal studies may lend support to a general causation opinion, an expert must explain how and why the studies can be reliably extrapolated to prove comparable effects in humans. *See Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 144-45 (1997); *Kilpatrick*, 613 F.3d at 1338-39; *Rider*, 295 F.3d at 1202.

These limitations apply with equal force to *in vitro* studies, which analyze the effects of drugs on human and animal cells, organs, or tissue cultures in a controlled laboratory setting. *See* Ref. Man. at 639. Observations about a drug's mechanism of action may be more readily gleaned from *in vitro* studies than from other sources, but the chemical reactions that occur in the artificial environment of a test tube or petri dish may differ from how the drug will react in, and impact, the complex biological system that is the human body. Ref. Man. at 564; *Accutane*, 511 F. Supp. 2d. at 1294-95. Thus, *in vitro* evidence alone cannot serve as a basis for a general causation opinion. *See Kilpatrick*, 613 F.3d at 1340-44. However, as with animal studies, *in vitro* data may be used to supplement other types of evidence, provided

the expert explains how the *in vitro* data can be reliably extrapolated to predict a drug's effects in humans. *See id*.

d. Analogous Drugs

In analyzing causation, scientists sometimes draw from existing studies conducted on other drugs in the same class as, or which have a similar chemical structure to, the particular drug at issue in a case. *See McClain*, 401 F.3d at 1244-46; *Rider*, 295 F.3d at 1200-01. This approach is premised on the theory that drugs with similar chemical structures may be expected to have similar properties and produce analogous effects. *See id*.; *see also Richardson v. Richardson-Merrell, Inc.*, 857 F.2d 823, 829 (D.C. Cir. 1988). Although such reasoning by analogy may have valid scientific uses, its value is somewhat limited in the context of establishing legal causation. This is because even within a given class of drugs, there may be "great chemical diversity" and those "minor deviations in chemical structure can radically change a particular substance's properties and propensities." *See Rider*, 295 F.3d at 1201 (quoting *Glastetter*, 252 F.3d at 990). Consequently, extrapolations from drugs within the same class may not support an expert opinion on general causation unless other reliable scientific evidence establishes the validity of the analogy. *See McClain*, 401 F.3d at 1246; *Rider*, 295 F.3d at 1200-01.

3. Weight of the Evidence

The preceding sections addressed the extent to which individual categories of scientific evidence may support an expert opinion on general causation in the Eleventh Circuit. In practice, however, many experts form a general causation opinion by weighing an entire body of scientific evidence. This "weight of the evidence" approach to analyzing causation can be considered reliable, provided the expert considers all available evidence carefully and explains how the relative weight of the various pieces of evidence led to his conclusion. *See Milward*, 639 F.3d at 17; *see also In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 858 F.3d 787, 795-97 (3d Cir. 2017) (*"Zoloft II"*) *aff'g* 26 F. Supp. 3d 449, 464 (E.D. Pa. 2014) ("*Zoloft I*"); *Jones v. Novartis Pharm. Corp*., 235 F. Supp. 3d 1244, 1272-73 (N.D. Ala. 2017); *Magistrini v. One Hour Martinizing Dry Cleaning*, 180 F. Supp. 2d 584, 602 (D.N.J. 2002). Using this methodology, the expert must

> (1) identify an association between an exposure and a disease, (2) consider a range of plausible explanations for the association, (3) rank the rival explanations according to their plausibility, (4) seek additional evidence to separate the more plausible from the less plausible explanations, (5) consider all of the relevant available evidence, and (6) integrate the evidence using professional judgment to come to a conclusion about the best explanation.

Milward, 639 F.3d at 17-18; *Jones*, 235 F. Supp. 3d at 1273. Importantly, because the "weight of the evidence" approach involves substantial judgment on the part of the expert, it is crucial that the expert describe each step in the process by which he

gathered and assessed the relevant scientific evidence. *See Zoloft II*, 858 F.3d at 795-97; *In re Seroquel Products Liability Litigation*, No. 6:06-md-1769, at *4-6 (M.D. Fla. June 23, 2009). To be considered reliable, the expert's weighing process must have been "based on methods and procedures of science, rather than on subjective belief or unsupported speculation." *Zoloft II*, 858 F.3d at 796. Otherwise, the methodology amounts to nothing more than the expert's *ipse dixit*, which the Supreme Court has admonished district courts against admitting into evidence. *See Joiner*, 522 U.S. at 146. Moreover, an expert cannot merely aggregate various categories of otherwise unreliable evidence to form a reliable theory of general causation. *See Siharath v. Sandoz Pharm. Corp.*, 131 F. Supp. 2d 1347, 1371 (N.D. Ga. 2001) (cautioning that expert "cannot lump together lots of hollow evidence" to establish medical causation); *see also Hollander*, 289 F.3d at 1216 ("To suggest that those individual categories of evidence deemed unreliable by the district court may be added to form a reliable theory would be to abandon the level of intellectual rigor of the expert in the field."); *Glastetter*, 252 F.3d at 992 (deciding that neither individual items nor an aggregate of evidence provided a reliable scientific basis for experts' conclusions). Instead, every aspect of the expert's analysis—including his methodology, the combination of facts and scientific evidence on which he relies, and the links between the evidence and his conclusions—must be shown to satisfy Rule 702 and *Daubert*. *See McClain*, 401 F.3d at 1245; *see also Heller v. Shaw*

Industries, Inc., 167 F.3d 146, 155 (3d Cir. 1999). Where an expert is found to have applied the "weight of the evidence" approach with "the same level of intellectual rigor" used by experts in the field, *see Kumho Tire* at 152, his general causation opinion typically will be deemed reliable and admissible. *See Milward*, 639 F.3d at 17; *see also Zoloft II*, 858 F.3d at 795-97; *Jones*, 235 F. Supp. 3d at 1272-73; *Magistrini*, 180 F. Supp. 2d at 602.

C. Reliability of Common Evidence of General Causation

Plaintiffs' experts, each to a greater or lesser extent, rely on much of the same evidence to conclude that Abilify can cause impulsive gambling and other impulse control disorders. The Court addresses the reliability of the common evidence together in this section. In Section II(D), the Court addresses Defendants' expert-specific objections to Drs. Bechara, Glenmullen, Hollander, Luepker, and Madigan.

1. Epidemiological Evidence—The Etminan Study[24]

Three of Plaintiffs' experts—Drs. Glenmullen, Hollander, and Madigan—base their opinions, in part, on an epidemiological study published by Dr. Mahyar Etminan and Dr. Ric M. Procyshyn in February 2017, in which a statistically significant association was found to exist between Abilify and impulse control

[24] The Study's lead author, Dr. Mahyar Etminan, was not called as a witness at the *Daubert* hearing and has not been offered as an expert. However, the Court allowed Defendants to depose Dr. Etminan because of the Study's importance in this case, given that it is the only epidemiological study to date analyzing the connection between Abilify and impulse control disorders, including pathological gambling.

disorder, and also between Abilify and gambling disorder ("Etminan Study"). *See* Mahyar Etminan, *Risk of Gambling Disorder and Impulse Control Disorder with Aripiprazole, Pramipexole, and Ropinirole*, 37 J. CLINICAL PSYCHOPHARMACOLOGY 1 (2017), ECF No. 428-13.[25] The Etminan Study is the only epidemiological study conducted to date that analyzes whether Abilify is associated with an increased risk of gambling and impulse control disorder.[26] Defendants challenge the reliability of the Etminan Study on multiple grounds, arguing that it is "so riddled with flaws as to be inherently unreliable." *See* ECF No. DSJ, 428-26 at 45. The Court disagrees and, for the reasons that follow, finds the Etminan Study sufficiently reliable to support an expert opinion on general causation in this case.

The Etminan Study is an epidemiological case-control study in which the authors analyzed medical and pharmaceutical billing information for over six million individuals, drawn from a large insurance claims database known as LifeLink.[27] *See*

---

[25] The Etminan Study had six co-authors: Dr. Mahyar Etminan, Mohit Sodhi, Dr. Ric M. Procyshyn, Michael Guo, and Dr. Bruce C. Carleton. *See* Etminan Study, ECF No. 428-13 at 1.

[26] Plaintiffs' expert, Dr. Joseph Glenmullen, characterized another study, referred to as the Moore Study, as an epidemiological study. *See* Thomas J. Moore *et al*., *Reports of Pathological Gambling, Hypersexuality, and Compulsive Shopping Associated with Dopamine Receptor Agonist Drugs*, 174 JAMA INTERNAL MED. 1930, 1930-33 (2014) ("Moore Study"), ECF No. 428-10 at 2-5. Despite Plaintiffs' insistence to the contrary, *see* Plaintiffs' Glenmullen Opposition, ECF No. 457-13 at 23, the Moore Study is *not* an epidemiological study. It is a disproportionality analysis. *See* Moore Study, ECF No. 428-10 at 2 (The Moore Study "conducted a *disproportionality analysis* based on . . . adverse drug event reports . . . extracted from the FDA Adverse Event Reporting System.") (emphasis added).

[27] The purpose of an epidemiological case-control study is to determine whether exposure to a drug is associated with a particular outcome (*i.e.*, a disease or adverse effect). *See* Ref. Man. at 559. Researchers identify a group of individuals who have a disease ("cases") and a group of

Etminan Study, ECF No. 428-13 at 1. The database included, *inter alia*, patients' diagnoses, as identified by ICD-9-CM Codes,[28] and all prescriptions they filled between 2006 and 2014. *See id*. Within this data, the authors first identified all individuals whose insurance records reflected a diagnostic code for either pathological gambling or impulse control disorder.[29] These individuals served as the Etminan Study's "case" group. Next, from the same data, the authors drew a random sample of similar individuals whose records contained neither diagnostic code. These individuals served as "controls."[30] The authors then compared the cases (individuals diagnosed with pathological gambling or impulse control disorders) to the controls (individuals with no such diagnoses) based on the prevalence of exposure to Abilify in each group. Exposure to Abilify was defined for the cases as one prescription for Abilify having been filled during the year before the pathological gambling or impulse control disorder diagnosis, and in corresponding

---

similar individuals who do not have the disease ("controls"). *See id*. Then, they compare the two groups in terms of past exposure to the drug. *See id*. If individuals in the case group are found to have a higher proportion of past exposure than the controls, then an association is said to exist between exposure and the disease. *See id*.

[28] "ICD-9-CM Codes" refers to the diagnostic and procedure codes established in the International Classification for Disease, Ninth Edition, Clinical Modification.

[29] There were 355 diagnoses coded as 312.31, which represents pathological gambling, and 4,341 diagnoses coded as 312.3, which represents impulse control disorders. *See* Etminan Study, ECF No. 428-13 at 1-2.

[30] The Etminan Study selected 10 controls for every case. *See* Etminan Study, ECF No. 428-13 at 1. Individuals in the control groups were similar to individuals in the case groups with respect to age, gender, follow-up time, and calendar time. *See id*.

calendar time for the controls. The Study found that individuals exposed to Abilify had a statistically significant higher incidence of pathological gambling and impulse control disorder diagnoses than did unexposed individuals.

The Etminan Study described the existence and strength of the association found between Abilify, pathological gambling, and impulse control disorder in the random sample from the LifeLink database in terms of "rate ratios," also known as relative risk. Relative risk is simply a comparison of the incidence of a disease in exposed individuals with its incidence in unexposed individuals. *See* Ref. Man. at 566. A relative risk of 1.0 means there is no difference in risk between the exposed and unexposed groups; in other words, there is no association between exposure to the drug and the disease. *See* Ref. Man. at 567; *see also Allison*, 184 F.3d at 1315 n.16. A relative risk above 1.0 indicates an increased risk in the exposed group, *see* Ref. Man. at 567, and "[r]isks greater than 2.0 permit an inference that the [disease] was more likely than not caused by the [drug]," *see Allison*, 184 F.3d at 1315 n.16. Relative risk estimates are often accompanied by a "confidence interval," which provides, in essence, a margin of error. *See* Ref. Man. 579-80. Confidence intervals identify the range of likely values, on either side of the relative risk estimate for a population sample, that would be expected to encompass the results a specified percentage of the time (*e.g.*, 95%) if random samples were repeatedly drawn from

the same population as the subject study.[31] *See id*. at 580. Importantly, if the confidence interval contains the value 1.0 or less, then the results of the study are not considered statistically significant. On the other hand, if the lower bound of the confidence interval exceeds 1.0, then the results are considered statistically significant.

In this case, the Etminan Study reported a relative risk of 5.23 for pathological gambling in individuals exposed to Abilify as compared to unexposed individuals, with a 95% confidence interval of 1.78-15.38. *See* Etminan Study, ECF No. 428-13 at 3. This means that the Study predicted that the increased risk of pathological gambling for Abilify patients within any given sample of the entire LifeLink database would likely fall anywhere between 1.78-15.38. Because the lower bound of the confidence interval (1.78) exceeds 1.0, this is statistically significant. The Study also reported a relative risk of 7.71 for impulse control disorder, with a 95% confidence interval of 5.81 and 10.34. This too is statistically significant. Finally, an analysis restricted to patients with bipolar disorder alone yielded a relative risk of

---

[31] As the Fifth Circuit has explained,

> the confidence interval tells one that if repeated samples were drawn from [a population] in the same ways as the instant sample was drawn, the means of the samples drawn would fall within the confidence interval a certain percentage, say 95 percent, of the time. On the basis of this information, researchers customarily conclude that the true means of the [population] falls within the confidence limits.

Univ. Computing Co. v. Mgm't Sci. Am., Inc., 810 F.2d 1395, 1399 n.4 (5th Cir. 1987).

3.38 for pathological gambling in Abilify patients, with a 95% confidence interval of 1.68-8.48, which is also statistically significant. Defendants do not dispute the accuracy of the Etminan Study's relative risk and confidence interval calculations.

Plaintiffs' biostatistician, Dr. David Madigan, analyzed the Etminan Study and found it to be "methodologically sound" with a "highly statistically significant result."[32,33] *See* Madigan Rep., ECF No. 427-1 at 30. This conclusion was based, in part, on the "strong" and "very substantial" relative risk figures reported in the Study; again, numbers that Defendants do not dispute. *See* Madigan Tr., ECF No. 596-4 at 49. Dr. Madigan also calculated a *p*-value for the Study's relative risk finding for pathological gambling. A *p*-value is a separate, widely established indicator of statistical significance, which measures the probability of obtaining the observed results—in this case, the increased risk of developing pathological gambling with exposure to Abilify—if, in reality, there is no true association between the drug and the adverse effect. *See* Ref. Man. at 249-50, 576-77; *see also Matrixx Initiatives, Inc. v. Siracusano*, 563 U.S. 27, 39 n.6 (2011). Stated differently, the *p*-value provides an estimate of the probability that chance alone produced the observed association between the drug and the adverse effect. *See id.*

[32] Biostatistics is a specialty within the field of statistics that involves the application of statistical methods to a wide range of issues in biology, medicine, and public health.

[33] For the reasons discussed more fully in Section II(D)(1)(d) of this Order, the Court finds Dr. Madigan qualified to offer an opinion on the statistical reliability of the Etminan Study.

The lower the *p*-value, the less likely it is that the observed result can be explained by chance alone. *See* Ref. Man. at 250-51; *see also Matrixx*, 563 U.S. at 39 n.6. Generally, *p*-values are considered statistically significant where they are less than or equal to .05 ($p \leq 5\%$). *See* Ref. Man. at 251; *see also Eastland v. Tenn. Valley Auth.*, 704 F.2d 613, 622 (11th Cir. 1983) ("Generally . . . a probability level of .05 is accepted as statistically significant."). By Dr. Madigan's calculation, the *p*-value for pathological gambling in the Etminan Study is .002, which is well below the traditional threshold for statistical significance. *See* Madigan Rep., ECF No. 427-1 at 25. According to Dr. Madigan, this *p*-value indicates that the probability of the Etminan Study producing a 5.23-fold increased risk of pathological gambling by chance alone is one in 500. *See* Madigan Tr., ECF No. 596-4 at 49. Defendants do not dispute the accuracy of this calculation.

Dr. Madigan discussed, at length, the strengths and limitations of case-control studies generally, as well as those of the Etminan Study specifically. *See* Madigan Rep., ECF No. 427-1 at 21-25, Madigan Supp., ECF No. 427-1 at 85-90; Madigan Tr., ECF No. 596-4 at 42-47, 54-58. In particular, with respect to the potential effect of bias on the Study's results, Dr. Madigan explained that the relative risk calculations are simply too "substantial" and "robust" to be explained by investigator bias. *See* Madigan Tr., ECF No. 596-4 at 70. In short, Dr. Madigan opines that a case-control study is "highly unlikely" to yield increased risk estimates like those

found in the Etminan Study "in the absence of a true" association. *See* Madigan Supp., ECF No. 427-1 at 91. In his view, the FDA called for a case-control study "clarify[ing]" the association between Abilify and impulse control disorders, and the Etminan Study reliably did exactly that. *See* Madigan Tr., ECF No. 596-4 at 71.

Defendants argue that numerous methodological flaws render the Etminan Study unreliable under Rule 702 and *Daubert*, including a deficient study design, failure to account for the risk of confounders, and the presence of bias. They also challenge Dr. Madigan's defense of the Etminan Study, which they claim is untenable in light of his prior published research criticizing both healthcare database research and the use of *p*-values as a measure of statistical significance. The Court addresses each category of objections in turn.

a. Study Design

Defendants criticize the Etminan Study's use of the LifeLink database because the database was not designed for research purposes. This criticism has little, if any, merit. The use of health insurance claims databases for epidemiologic research is well-supported by the medical literature, which is an important consideration under *Daubert*.[34] *See Daubert*, 509 U.S. at 589-90. Their

[34] *See, e.g.*, Gianluca Trifirò & Janet Sultana, *The Role of Health Care Databases in Pharmacovigilance of Psychotropic Drugs*, in PHARMACOVIGILANCE IN PSYCHIATRY 73, 90 (Edoardo Spina & Gianluca Trifirò eds., 2016) ("[A]dministrative/claims databases are important data sources to carry out observational studies aimed at quantifying and describing emerging safety issues associated with the use of psychotropic drugs, as shown by the large amount of database safety studies that have been published worldwide in the last decades."), DX-127 at 18; Esther W.

"representativeness [of routine clinical practice], large size, and capacity to contain large quantities of [long-term] clinical data on each patient" can make them a "powerful tool" for studying the use, efficacy, and safety of prescription drugs. Schneeweiss, 48 J. CLIN. EPIDEMIOLOGY at 334, DX-122 at 12. Indeed, large databases are particularly advantageous for studying relatively rare adverse effects of a drug, as in this case, or where multiple possible adverse effects are of interest. *See id*. at 325, DX-122 at 3.

With that said, large database research is not without limitations, one of which is the unavailability of medical records to confirm the accuracy of the data and to provide potentially significant clinical information not reported in the database. Defendants argue that this limitation is fatal to the Etminan Study's reliability under *Daubert*.[35] The Court disagrees. While it is true that the medical literature

---

Chan et al., *Adverse drug reactions – examples of detection of rare events using databases*, 80:4 Brit. J. Clin. Pharmacology 855, 855 (2014) ("Large databases provide an important platform for the undertaking of observational studies to generate clinical data on the effectiveness and safety of drugs."), ECF No. 463-6 at 2; Sebastian Schneeweiss & Jerry Avorn, *A Review of Uses of Health Care Utilization Databases for Epidemiologic Research on Therapeutics*, 58 J. Clin. Epidemiology 323, 323-25 (2005) ("[L]arge health care databases" are "a useful data source for researchers and regulatory agencies to study the safety of drugs."), DX-122 at 1-3.

[35] Defendants raise several scattered arguments arising from this limitation, such as the Etminan Study's inability to: (1) precisely identify the date of onset of pathological gambling or impulse control disorders; (2) verify the medical accuracy of any diagnosis; and (3) ensure against data entry errors, such as inaccurate diagnosis coding. The Court has carefully considered these arguments and finds them lacking in merit, for the reasons more fully stated in the body of this Order. In short, the arguments reflect on the Etminan Study's probativeness, not its admissibility. *See Daubert*, 509 U.S. at 596 ("Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.").

encourages record review, the medical community also recognizes that health information privacy laws have "constrained the availability of" individual medical records "for uses other than the direct care of patients." *See id*. at 327-28, DX-122 at 5-6. Large database research for pharmacovigilance purposes is generally accepted in the scientific community, *see supra* n.34, and has been found to be a reliable methodology by other courts, even where no medical record review occurred. *See Rheinfrank v. Abbott Labs., Inc*., No. 1:13-cv-133, 2015 WL 13022172, at *13 (S.D. Ohio Oct. 2, 2015) (finding "data-mining" of the FDA adverse event reporting system database a reliable methodology for determining whether a signal for developmental delay from in utero exposure to Depakote existed); *In re Fosamax (Alendronate Sodium) Prods. Liab. Litig.*, No. 3:11-cv-05304, 2013 WL 1558690, at *8 (D.N.J. April 10, 2013) ("[D]ata mining in pharmacovigilance[] is generally accepted in the scientific community and has become routine both in the pharmaceutical industry and amongst regulators worldwide.").

The LifeLink database, despite Defendants' criticisms, contains a sufficiently comprehensive dataset of patients, medical diagnoses and prescription claims to reliably serve the epidemiological objectives of the Etminan Study. Indeed, "claims data of this type provide some of the best data on drug exposure in pharmacoepidemiology." Brian L. Strom, *Overview of Automated Databased in*

Pharmacoepidemiology, in PHARMACOEPIDEMIOLOGY 158, 159 (Brian L. Strom et al. eds., 5th ed. 2012), DX-129 at 2. The Etminan Study's statistical analysis of the LifeLink data is capable of being tested, and the Study itself has been subjected to peer review and publication in a reputable medical journal. This is all that Rule 702 and *Daubert* require. *See Chapman*, 766 F.3d at 1305 (citing *Daubert*, 509 U.S. at 593-94). The alleged inadequacies of the LifeLink database may impact the weight afforded to the Etminan Study's conclusions, but not its reliability or admissibility under *Daubert*.

Defendants' next argument, which relates to the database challenges addressed above, is that the Etminan Study is unreliable for its inability to confirm that individual patients in the LifeLink database were ever actually exposed to Abilify; that is, that they actually took the Abilify they were prescribed. The Study did not attempt to validate medication usage, even though its lead author, Dr. Etminan, has done so in other epidemiological studies. This criticism also fails.

All epidemiological studies that make use of large healthcare databases are vulnerable to the risk of drug exposure misclassification, which is the risk of inaccurately measuring actual exposure to a drug. *See* Schneeweiss, 48 J. CLIN. EPIDEMIOLOGY at 328, DX-122 at 6. This is because claims databases only reflect the dispensing of medications and not actual medication use. *See id*. Despite this limitation, the use of pharmacy dispensing data as a proxy for drug usage is seen in

the scientific community as "the gold standard of drug exposure information compared with self-reported information or prescribing records in outpatient medical records." *See id*.; *see also* Strom at 159 ("[C]laims data of this type provide[s] some of the best data on drug exposure in pharmacoepidemiology."), DX-129 at 2. Short of physically monitoring ingestion or requiring study subjects to undergo routine laboratory testing to ascertain medication levels, there appears to be no more reliable means of measuring drug exposure than pharmacy claims data.

The fact that the Etminan Study did not attempt to correct for the risk of drug exposure misclassification does not render it unreliable under *Daubert*. There is no evidence in the record of an established epidemiological protocol for addressing drug exposure misclassification concerns. *See Kumho Tire* at 152 (an expert in the courtroom must use "the same level of intellectual rigor that characterizes the practice of an expert in the relevant field). There also is no evidence that Dr. Etminan violated his own methodological standards with respect to this issue. While Defendants are correct that Dr. Etminan has, in the past, tried to "control" for the risk of drug exposure misclassification in different ways,[36] there are also numerous

[36] At his deposition, Dr. Etminan testified that he has, in the past, used a "buffer" around the prescription date to account for the time it takes patients to fill a prescription, take the medication and develop the disease. *See* Etminan Dep., ECF No. 428-12 at 13. The record includes only one study in which Dr. Etminan employed this "buffer" technique. *See* Mahyar Etminan *et al*., *Oral Contraceptives and the Risk of Gallbladder Disease: A Comparative Safety Study*, 183 Can. Med. Ass'n J. 899 (2011), PX-125. In another published study, Dr. Etminan used two prescriptions as a proxy for exposure on the theory that multiple dispensings increases the likelihood that a patient is actually taking the drug. *See* Mahyar Etminan, *Pharmacoepidemiology*

published, peer-reviewed studies in which Dr. Etminan's treatment of the issue mirrors that used in the Etminan Study.[37] There is no evidence that these studies have been criticized in the scientific community for failing to account for the risk of exposure misclassification in the study design. Moreover, the record in this case suggests that at least one of Dr. Etminan's prior validation techniques—using two prescriptions as a proxy for exposure, instead of a single prescription, *see supra* n.36—might have negatively skewed the Etminan Study results, if used. Both Defendants' expert, Dr. Marc Potenza, and Plaintiffs' expert, Dr. Eric Hollander, cautioned that a study utilizing two Abilify prescriptions as a proxy for exposure could miss individuals who developed symptomology during their first prescription and who subsequently discontinued the drug. *See* Potenza Tr., ECF No. 596-7 at

II: The Nested Case-Control Study—A Novel Approach in Pharmacoepidemiologic Research, 24 Pharmacotherapy 1105 (2004), ECF No. 428-14 at 2.

[37] *See* Mahyar Etminan *et al*., *Risperidone and Risk of Gynecomastia in Young Men*, 25 J. CHILD & ADOLESCENT PSYCHOPHARMACOLOGY 671, 672 (2015) (defining "any use" of the drug as "the use of at least one prescription in the year before the index date"), PX-145; Mahyar Etminan *et al*., *Isotretinoin and Risk for Inflammatory Bowel Disease: A Nested Case-Control Study and Meta-analysis of Published and Unpublished Data*, 149 JAMA DERMATOLOGY 216, 217 (2013) (defining drug "use" as "at least [one] dispensed prescription for [the drug] regardless of dosage during the 365 days before the first IBD claim (and in corresponding calendar time for controls)"), PX-126; Mahyar Etminan *et al*., *Testosterone Therapy and Risk of Myocardial Infarction: A Pharmacoepidemiologic Study*, 35 PHARMACOTHERAPY 72, 73 (2015) (defining testosterone use as having "filled at least one prescription for [testosterone replacement therapy] within [one] year before the index date"), PX-167; Mahyar Etminan *et al*., *Use of Oral Bisphosphonates and the Risk of Aseptic Osteonecrosis: A Nested Case-Control Study*, 35 J. RHEUMATOLOGY 691, 692 (2008) ("Current users were defined as those who received at least one prescription for a bisphosphonate within 90 days of the index date."), PX-124.

84-85;[38] Hollander Tr., ECF No. 596-4 at 116.[39] Under these circumstances, the Etminan Study cannot be considered unreliable for failing to control for the risk of drug exposure misclassification. This objection may be probative on the weight of the Study, but not its admissibility. *See Bazemore v. Friday*, 478 U.S. 385, 400 (1986) ("Normally, failure to include variables will affect the analysis' probativeness, not its admissibility.").

Defendants' next argument with respect to the Etminan Study's design is that the Study cannot reliably measure the incidence of iatrogenic gambling (*i.e.*, medication-induced) in Abilify patients because it identified cases of gambling disorder in the LifeLink database using medical billing codes that are based on the DSM-5 diagnostic criteria for idiopathic gambling (*i.e.*, gambling disorder that occurs spontaneously and with no known cause).[40] The Court disagrees. The LifeLink database classifies diagnoses according to ICD-9-CM codes, not the DSM-5 diagnostic criteria. *See* Etminan Study, ECF No. 428-13 at 1. Although the drafters of these two classification systems have, in recent years, attempted to "harmonize [them] as much as possible," the ICD-9-CM and DSM-5 are not

[38] "Potenza Tr." refers to the redacted version of the official transcript of Dr. Marc Potenza's testimony at the *Daubert* hearing, ECF Nos. 596-6 at 22-72, 596-7 at 3-89.

[39] "Hollander Tr." refers to the redacted version of the official transcript of Dr. Eric Hollander's testimony at the *Daubert* hearing, ECF Nos. 596-4 at 108, 596-5.

[40] "DSM-5" refers to the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (5th ed. 2013).

identical. *See* DSM-5 at 11. Indeed, the two publications serve different purposes and, in some circumstances, diverge or are discordant with one another.[41] *See id.* The Court finds that this case presents one of those circumstances. The DSM-5 expressly recognizes the existence of iatrogenic gambling, but excludes the condition from its diagnostic criteria for "gambling disorder."[42] *See* DSM-5 at 589, ECF No. 428-3 at 68. In other words, the DSM-5 acknowledges the condition of iatrogenic gambling and explains that the diagnostic criteria for idiopathic gambling do not apply to it. *See id*. In contrast, the ICD-9-CM definition of "pathological gambling" does not appear to be limited, either explicitly or implicitly, to idiopathic gambling. *See* ICD-9-CM, PX-063 at 1. By its terms, the ICD-9-CM diagnostic code encompasses *any* "preoccupation with gambling and the excitement that

[41] The International Classification of Disease ("ICD") is the official international classification system, intended to provide a standardized means of documenting, tracking, and billing for medical diagnoses and diseases worldwide. *See* DSM-5 at 10-11. The DSM, which is used primarily in the United States, provides health care professionals with strict criteria and definitions to aid in the clinical diagnosis and treatment of mental disorders. *See* DSM-5 at xii.

[42] It bears repeating that the DSM-5 does not reject the existence or legitimacy of iatrogenic gambling as a "separate and distinct" medical diagnosis, which Defendants seem to suggest. *See* Def. Potenza Opposition, ECF 458-6 at 6-7. In fact, it does exactly the opposite. The DSM-5 provides that "[s]ome patients taking dopaminergic medications (*e.g.*, for Parkinson's disease) may experience urges to gamble. If such symptoms dissipate when dopaminergic medications are reduced in dosage or ceased, then a diagnosis of gambling disorder would not be indicated." *See* DSM-5 at 589, ECF No. 428-3 at 68. The medical term "dopaminergic" means "liberating, activated by, or involving dopamine." *See* Merriam-Webster Online Medical Dictionary, https://www.merriam-webster.com/dictionary/dopaminergic (retrieved Dec. 3, 2017). Because its mechanism of action involves dopamine, Abilify is clearly a dopaminergic medication. The parties' experts agree. *See* Blier Tr., ECF No. 596-8 at 15 (conceding all drugs that work on dopamine "technically" are dopaminergic drugs); Hollander Tr., ECF No. 596-4 at 140 (stating that Abilify is a dopaminergic medication).

gambling with increased risk provides," even where "it may lead [the patient] to lie, steal, or lose a significant relationship, job, or educational opportunity." *See id.* There simply is no carve out for, or exclusion of, iatrogenic gambling (medication-induced) from the ICD-9-CM definition. *See id*. Thus, it would be reasonable for a scientist to conclude that "pathological gambling" under the ICD-9-CM includes iatrogenic gambling. In turn, the Etminan Study's use of that same ICD-9-CM code to identify cases of pathological gambling in the LifeLink database is likewise reasonable. This challenge does not undermine the reliability or admissibility of the Etminan Study.

Defendants' last argument is that the Etminan Study is unreliable because the time between exposure to Abilify and the diagnoses of pathological gambling in the random sample taken from the LifeLink database was too short to be compatible with a cause-effect relationship. This argument is based on the Study's finding that five patients were exposed to Abilify in the year preceding their diagnoses of pathological gambling, with an average, or mean, time to diagnosis of 20 days and a standard deviation of 17.4 days.[43] *See* Etminan Study, ECF No. 428-13 at 3. The

[43] Defendants make much of their claim that five diagnoses of pathological gambling is too small a number to reliably demonstrate the requisite association in this case. Plaintiffs' biostatistician, Dr. Madigan, explains this by reference to the confidence interval. *See* Madigan Tr., ECF No. 596-4 at 53. More specifically, Dr. Madigan testified that the number of adverse medical events is not the relevant inquiry for statistical purposes; rather, the focus is on the lower bound of the confidence interval, which in this case, is substantially greater than 1.0. *See* Madigan Rep., ECF No. 427-1 at 25. This means that the Etminan Study's findings are statistically

standard deviation is a measure of statistical dispersion; that is, the average distance between the five individual time-to-diagnosis data points and the mean.[44] *See* Ref. Man. at 239. None of the five actual time-to-diagnosis periods was individually reported in the Study. *See id.* However, Defendants claim that the standard deviation in the Study indicates that at least one of the five patients was diagnosed with pathological gambling within three days of his exposure to Abilify, which even Dr. Etminan, at his deposition, agreed was "unlikely." *See* Etminan Tr., ECF No. 427-3 at 80; *see also* Def. Madigan Motion, ECF No. 427-20 at 9. Defendants insist that, under the DSM-5 criteria, gambling disorder "takes up to twelve months to develop into a disease," so if the five diagnoses in the Etminan Study followed this criterion, the patients' pathological gambling necessarily would have preceded their exposure to Abilify. DSJ, ECF No. 428-26 at 27.

This criticism is not fatal to the Study's reliability under *Daubert* for several reasons. First, the standard deviation of 17.4 days does not dictate a conclusion that there must have been a three-day period between exposure and diagnosis for at least

significant. In any event, nothing in the published scientific literature criticizes the Study on this basis.

[44] "Technically, a standard deviation is defined as a measure of spread, dispersion, or variability of a group of numbers equal to the square root of the variance of that group of numbers." *Stagi v. National R.R. Passenger Corp*., 391 F. App'x 133, 137 n.6 (3d Cir. 2010) (internal marks omitted). "The variance of the group of numbers is computed by subtracting the mean, or average, of all the numbers, squaring the resulting difference, and computing the mean of these squared differences." *Id.*

one of the five patients in the random sample analyzed in the Study. At his deposition, Dr. Etminan, the Study's lead author, testified that while the time-to-diagnosis for one of the patients could have been "a matter of days," he could "not [be] sure exactly what" the actual times-to-diagnosis were for any of the five patients "without having the data" from the LifeLink database to review.[45] *See Etminan* Dep., ECF No. 457-7 at 42.[46] Dr. Madigan, who is the only statistics expert in this case, calculated at least two possible distributions of the five individual time-to-diagnosis periods, given the mean time of 20 days and the standard deviation of 17.4 days: (1) a distribution of 3, 41, 13, 37, and 8 days, respectively; or (2) a distribution of 50, 16, 8, 8, and 17 days.[47] *See* Madigan PPT, PX-051 at 22;[48] Madigan Rep.,

[45] Defendants mischaracterize Dr. Etminan's testimony as a definitive statement that the standard deviation of 17.4 days means at least one of the five patients was diagnosed three days after his exposure to Abilify. The Court does not read Dr. Etminan's testimony as unequivocal with respect to any minimum number of days. Dr. Etminan initially stated that, given the Study's standard deviation, the least amount of time between exposure and diagnosis was "[p]robably three days." *See* Etminan Dep., ECF No. 427-3 at 42. However, he immediately backpedaled and agreed only that diagnosis could occur "a matter of days" after exposure to Abilify. *See id*. Beyond that, Dr. Etminan testified that he "can't say one way or the other" without reviewing the LifeLink data. *See id*. at 44. Defendants have not suggested that Dr. Etminan should have brought the LifeLink data to his deposition.

[46] "Etminan Dep." refers to the official transcript of Dr. Mahyar Etminan's deposition testimony on May 16, 2017, ECF No. 427-3.

[47] At the *Daubert* hearing, Dr. Madigan also offered a third distribution of time-to-diagnosis dates that he opines are consistent with the mean and standard deviation reported in the Etminan Study. Because this third calculation was not disclosed in Dr. Madigan's expert reports, Plaintiffs agreed to omit it from their argument under *Daubert*. *See* Madigan Tr., ECF No. 596-4 at 62.

[48] "Madigan PPT" refers to the Powerpoint presentation used by Dr. Madigan at the *Daubert* hearing, PX-051.

ECF No. 427-1 at 24. According to Dr. Madigan, there is no way to determine, from the information reported in the Study, which of these two possible distributions represents the actual distribution of time-to-diagnosis periods for the five patients found to have been exposed to Abilify within the year preceding their diagnosis of pathological gambling in the LifeLink database. *See* Madigan Tr., ECF No. 596-4 at 62-63. None of Defendants' experts disputed Dr. Madigan's calculations.[49] Thus, the evidence shows that the minimum time-to-diagnosis in the Etminan Study could have been as few as three, or as many as eight, days.

Second, the possibility of pathological gambling or other impulse control symptoms developing within either three or eight days of exposure to Abilify is consistent with the scientific literature. Multiple published case reports describe a "rapid onset" of such symptoms following exposure to Abilify, with patients' times-to-onset ranging from a few days to a week after starting treatment with the drug.[50]

---

[49] Dr. Weed, who is Defendants' expert epidemiologist, testified that he did not attempt to calculate the possible distribution of time-to-diagnosis periods based on the mean and standard deviation reported in the Etminan Study. *See* Weed Tr., ECF No. 596-8 at 82. "Weed Tr." refers to the official transcript of Dr. Douglas Weed's testimony at the *Daubert* hearing, ECF No. 596-8.

[50] *See, e.g.*, E. Peterson & R. Forlano, *Partial Dopamine Agonist-Induced Pathological Gambling and Impulse-Control Deficit on Low-Dose Aripiprazole*, 25 AUSTRALASIAN PSYCHIATRY 614 (2017) (patient's "strong desire to gamble" developed "after only a few days on" Abilify); L. Gaboriau *et al*., *Aripiprazole: A New Risk Factor for Pathological Gambling? A Report of 8 Case Reports*, 39 ADDICTIVE BEHAVIORS 526, 563 (2014) (one patient's "strong urges to gamble" developed within days of starting Abilify treatment and another patient's "irresistible urge to gamble" developed within days after his Abilify dose was increased from 10 mg to 20 mg per day), ECF No. 425-7 at 2-5; Giles Gavaudan *et al*., *Partial Agonist Therapy in Schizophrenia: Relevance to Diminished Criminal Responsibility*, 55 J. FORENSIC SCI. 1659 (2010) (patient's pathological gambling symptoms began "a few days" after starting Abilify treatment), DX-631 at

See Glenmullen Tr., ECF No. 596-2 at 64. While this anecdotal evidence obviously cannot establish the times-to-onset or diagnosis for the patients in the Etminan Study, it does reliably indicate that a three- or eight-day time-to-onset and diagnosis is not wholly "implausible." *See* Weed Tr., ECF No. 596-8 at 65. Instead, the Court agrees with Dr. Glenmullen that a patient with "a very sudden onset" of "distressing" impulse control problems after starting "a new [and] powerful antipsychotic drug" like Abilify could reasonably be expected to call his doctor's office and get a diagnosis of pathological gambling or impulse control disorder "within days or a week" of beginning treatment. *See* Glenmullen Tr., ECF No. 596-2 at 65.

Finally, as the Court has already found, the DSM-5 diagnostic criteria for gambling disorder do not govern the diagnosis of iatrogenic (medication-induced) gambling. *See* DSM-5, ECF No. 428-3 at 68. Indeed, the DSM-5 clearly contemplates that "patients taking dopaminergic medications" may "experience urges to gamble" that "dissipate when [the] medications are reduced in dosage or ceased." *See id*. In that scenario, the DSM-5 diagnostic criteria—including the requirement that an individual have exhibited at least four "problematic gambling

---

1; M. Kodama & T. Hamamura, *Aripiprazole-Induced Behavioral Disturbance Related to Impulse Control in a Clinical Setting*, 13 INT'L J. NEUROPSYCHOPHARMACOLOGY 549, 550 (2010) (patient's hypersexuality symptoms presented within the first week of taking Abilify), DX-652 at 1; J. Schlachetzki & J. Langosch, *Aripiprazole Induced Hypersexuality in a 24-Year-Old Female Patient with Schizoaffective Disorder?*, 28 J. CLINICAL PSYCHOPHARMACOLOGY 567 (2008) (patient's hypersexuality symptoms developed within "[a] few days" of taking Abilify).

behavior[s]" in a 12-month period—do not apply.[51] *See id*., ECF No. 428-3 at 64, 68. On the other hand, the ICD-9-CM imposes no minimum duration requirements for an individual's "preoccupation with gambling" to qualify as "pathological gambling." *See* ICD-9-CM, PX-063 at 1. Because the LifeLink database classifies medical diagnoses according to ICD-9-CM codes, and not the DSM-5 diagnostic criteria, there is no basis for assuming from the medical billing codes used that the pathological gambling behaviors must, or even may, have preceded the patients' exposure to Abilify.

b. Confounders

Defendants also argue that the Etminan Study is unreliable for its failure to control for potential confounders, *i.e*., known risk factors for pathological gambling, specifically, depressive disorders, anxiety disorder, and personality disorders. *See* DSJ, ECF No. 428-26 at 29. According to Defendants, this flaw is "independently fatal" to the Etminan Study because it means that the pathological gambling diagnoses in the LifeLink database could be attributable to one or more of these underlying conditions rather than Abilify, particularly since they are in many cases

[51] Defendants' notion that idiopathic gambling under the DSM-5 "takes up to *twelve months* to develop into a disease," DSJ, ECF No. 428 at 27, suggests a misunderstanding of the diagnostic criteria for the disorder. On its face, the DSM-5 requires only that an individual exhibit four or more problem gambling symptoms in a 12-month period. *See* DSM-5, ECF No. 428-3 at 64. Under this criterion, the symptoms may develop over 12 months, but they also could manifest within a single month. In either scenario, assuming the other criteria were met, a diagnosis of idiopathic gambling disorder would be appropriate.

the precise medical conditions for which Abilify is prescribed. *See* DSJ Reply, ECF No. 484 at 10-11.[52] Defendants insist that because the Etminan Study failed to take into account the relationship between pathological gambling and these medical conditions, it cannot reliably establish the existence of an association between pathological gambling and Abilify.

When assessing the reliability of an epidemiological study, a court must consider whether the study adequately accounted for confounding factors, or confounders. *See* Ref. Man. at 591; *see also Deutsch v. Novartis Pharm. Corp*., 768 F. Supp. 2d 420, 432 (E.D.N.Y. 2011). Confounding occurs where an extraneous variable, or set of variables, may wholly or partially explain an apparent association between exposure to a drug and a disease, but that variable is not accounted for in the study. *See* Ref. Man. at 591-93. Importantly for purposes of this case, a variable only has the potential to act as a confounder where it is independently related both to exposure and to the disease of interest; that is, where the variable itself is both associated with exposure to the drug and a causal risk factor for the disease. *See id.* at 591;[53] *see also Deutsch*, 768 F. Supp. 2d at 432; Noel S. Weiss & Thomas D.

[52] "DSJ Reply" refers to Defendants' Reply in Support of their Motion for Summary Judgment on General Causation, ECF No. 484.

[53] For example, researchers may conclude from a study that individuals with gray hair have a higher rate of death than those with hair of another color. Ref. Man. at 591. Instead of hair color having an impact on death, the results might be explained by the confounding factor of age. *Id*. If old age is associated differentially with the gray-haired group (those with gray hair tend to be older), old age may be responsible for the association found between hair color and death. *Id*.

Koepsell, EPIDEMIOLOGIC METHODS: STUDYING THE OCCURRENCE OF ILLNESS 216 (2d 2014). The presence of confounders can lead to epidemiological results that do not reflect the true relationship between the variables under study, such as the appearance of an association where none exists or a distortion of the magnitude of an actual association. *See* Ref. Man. at 593, 597. For this reason, when confounders can be identified, they should be accounted for. *Id*. However, failure to control for every conceivable potential confounder does not necessarily render the results of an epidemiological study unreliable. *See id*.; *see also Bazemore*, 478 U.S. at 400 ("Normally, failure to include variables will affect [a statistical analysis'] probativeness, not its admissibility."). Confounders may be addressed through the study design or, after the study is complete, through adjustments and statistical analyses that test the robustness of the methodology and the results.[54]

Thus, researchers would need to separate the relationship between gray hair and risk of death from that of old age and risk of death. *Id*.

[54] These statistical analyses may include sensitivity analyses, stratification, and multivariate analysis. Ref. Man. at 593, 595-97. A sensitivity analysis is used to test whether and how the results of a study change if specific variables or assumptions are changed (*i.e*., indicates whether the results are sensitive to certain variables or assumptions). *See id*. at 595-96. The Etminan Study, for example, included a sensitivity analysis restricted to patients with bipolar disorder. See Etminan Study, ECF No. 428-13 at 3. Stratification involves the use of statistical methods to combine the results of different exposure levels (or strata) to the confounding factor to arrive at one overall estimate of risk. *See id*. at 596-97. Multivariate analysis involves using mathematical modeling to "describe the simultaneous effect of exposure and confounding factors on the increase in risk." *Id*. at 597. The latter two methods modify an observed association to take into account the effect of risk factors are not the subject of the study and that may distort the exposure being studied and disease outcomes." *Id*.

Applying these principles to this case, the Court finds that the failure to control for depressive disorders, anxiety disorders, and personality disorders does not invalidate the results of the Etminan Study. First, Abilify is not indicated for treatment of anxiety or personality disorders. *See* Product Label, ECF No. 428-1 at 2, 4. Nothing in the record as it currently stands suggests that Abilify is even prescribed off-label for these two categories of psychiatric conditions. This is significant because, for epidemiological purposes, no matter how strongly a variable is related to the disease in question, if it is not *also* related in some way to drug exposure, it cannot be a true confounder. *See* Ref. Man. at 591; Weiss & Koepsell at 216. Here, there is no evidence that anxiety and personality disorders are related in any way to Abilify exposure.[55] Therefore, no matter how strongly those disorders may be related to pathological gambling, neither is a true confounder.

Second, the medical literature is inconclusive on the question of whether depressive, anxiety and personality disorders are causal risk factors for pathological gambling. It is true that the DSM-5 provides that "[i]ndividuals with gambling disorder have high rates of comorbidity with" these categories of psychiatric conditions. Am. Psychiatric Ass'n, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 589 (5th ed. 2013) ("DSM-5"), ECF No. 428-3 at 68. It also is

[55] This may be why Defendants direct the substance of their confounders argument at major depressive disorder, for which Abilify *is* indicated as an adjunctive treatment for patients who are also using antidepressants.

true that at least two cross-sectional studies have confirmed the prevalence of problem gambling behavior in the major depression and mood disorder populations.[56] *See* Lena C. Quilty *et al*., *The Prevalence and Course of Pathological Gambling in the Mood Disorders*, 27 J. Gambling Studies 191, 191-92 (2011) (discussing "elevated prevalence" of pathological gambling with mood disorders, depression, and bipolar disorder), ECF No. 427-13; Sidney H. Kennedy *et al*., *Frequency and Correlates of Gambling Problems in Outpatients with Major Depressive Disorder and Bipolar Disorder*, 55 Canadian J. Psychiatry 568, 574 (2010) (discussing "higher prevalence" of "problem gamblers" with major depressive disorder), ECF No. 427-12. But "comorbidity" and "prevalence" are not synonymous with "causative." Notably, both cross-sectional studies explicitly caution against inferring a causal relationship from the presence of these observed comorbidities. *See* Quilty at 198, ECF No. 427-13 at 9; Kennedy at 574, ECF No. 427-12 at 8. The Quilty study, in particular, concluded that there is "no direct association" between pathological gambling and mood disorders, including depression. *See* Quilty at 198, ECF No. 427-13 at 9. In any event, "it is not possible to establish the temporal relation between exposure and disease—that is, that the

[56] A cross-sectional study is an epidemiological study in which exposure and an outcome or disease are measured at the same point in time in a given population. Ref. Man. at 560. Temporal and/or causal relationships between exposure and disease or outcome cannot be established through a cross-sectional study. *Id*. at 560-61.

exposure preceded the disease, which would be necessary for drawing any causal inference"—by reference to a cross-sectional study. Ref. Man. at 560-61. In sum, science has not yet determined with any reliability the precise nature of the relationship, if any, between pathological gambling and depressive, anxiety, and personality disorders. The Court may not simply ignore this gap in scientific understanding by excluding the Etminan Study for failing to consider these psychiatric conditions as causal risk factors for pathological gambling. *See Hendrix II*, 609 F.3d at 1194 ("[C]ourts may only admit the state of science as it is."); *Rider*, 295 F.3d at 1194 ("Law lags science; it does not lead it.").

Finally, other evidence in the record reliably supports the conclusion that depressive, anxiety, and personality disorders did not confound the results of the Etminan Study. For example, as part of a 2016 Pharmacovigilance Review, the FDA performed a disproportionality analysis of its adverse event reporting system database, comparing the relative frequency of pathological gambling reports among 11 different atypical antipsychotics, one of which was Abilify.[57] *See* FDA, Abilify Pharmacovigilance Review 27 (March 10, 2016) ("FDA Pharm. Vigil."), ECF No. 428-11 at 27. The FDA found that only Abilify had a statistically significant percentage of patients reporting pathological gambling. *See id*. In other words,

[57] The 11 atypical antipsychotics that the FDA analyzed were: aripiprazole, olanzapine, quetiapine, risperidone, asenapine, brexpiprazole, clozapine, iloperidone, lurasidone, paliperidone, and ziprasidone. FDA Pharm. Vigil., ECF No. 428-11 at 27.

pathological gambling was "disproportionately reported with [Abilify] relative to all other atypical antipsychotics." *Id*. In fact, seven of the atypical antipsychotics examined reflected no reported cases of pathological gambling.[58] *Id*. The FDA's findings are significant on this issue because all of the atypical antipsychotics it examined treat the same patient population Abilify treats, with the same comorbidity and risk profiles, which include depressive, anxiety, and personality disorders. If the results of the Etminan Study were confounded by these psychiatric conditions, then one would expect comparable percentages of pathological gambling reports for all the atypical antipsychotics.

[* REDACTED ***]**.[59] **[*** REDACTED ***]**. The results of these disproportionality analyses—one by the FDA, **[*** REDACTED ***]**—certainly are not conclusive evidence that Abilify causes pathological gambling. However, they do reliably support a conclusion that controlling for depressive, anxiety, and personality disorders does not decrease the statistical incidence of pathological gambling reports associated with Abilify. Thus, an expert could reasonably conclude that these psychiatric conditions are not confounding factors that must be

[58] No reported cases of pathological gambling occurred with asenapine, brexpiprazole, clozapine, iloperidone, lurasidone, paliperidone, and ziprasidone. *Id*.

[59] **[*** REDACTED ***]**.

controlled for in a reliable epidemiological study of the possible relationship between Abilify and gambling disorder.

In sum, confounding is a "reality" inherent in all epidemiological research. Ref. Man. at 590. As such, confounders "do not reflect an error made by" a particular researcher; rather "they reflect the inherently uncontrolled nature" of observational studies. *See id*. at 593 (internal marks omitted). Identifying and mitigating the effects of confounding is key to ensuring the reliability of an epidemiological study. *See id*. at 591-97; *see also Deutsch v. Novartis Pharm. Corp*., 768 F. Supp. 2d 420, 432 (E.D.N.Y. 2011). In this case, the Etminan Study identified and controlled for the "strong" confounding effects of bipolar disorder, schizophrenia, and substance abuse disorder. *See* Etminan Study, ECF No. 428-13 at 4. It cannot be said that an epidemiological analysis which accounts for these major causal risk factors is unreliable evidence of an association between Abilify, gambling, and impulse control disorders, simply because it did not account for all possible confounders. *See Bazemore*, 478 U.S. at 400. Only when a methodology "is so incomplete as to be inadmissible as irrelevant" should it be excluded. *See id*. at 400 n.10. Such is not the case with the methodology used in the Etminan Study. Therefore, Defendants' objections based on potential confounding variables do not affect the Study's admissibility. "[V]igorous cross-examination" and "presentation

of contrary evidence" are the appropriate means of attacking the Study's limitations. *See Daubert*, 509 U.S. at 595.

c. Bias

Defendants further argue that the Etminan Study is unreliable because its results were compromised by bias. More specifically, they first argue that through his actions, Dr. Etminan created a conflict of interest that affects the integrity of the Study's findings. This argument is based on the fact that Dr. Etminan contacted Plaintiffs' counsel shortly after learning about this litigation on AboutLawsuits.com, before he developed the research protocol for the Study. The implication, of course, is that Dr. Etminan was predisposed towards results that would support Plaintiffs' theory of general causation in this case. The problem with Defendants' position is that the uncontroverted record evidence with respect to this communication does not demonstrate impropriety by either Dr. Etminan or Plaintiffs' counsel. According to Dr. Etminan, the conversation lasted only two minutes, during which he advised Plaintiffs' counsel that he intended to conduct a study on Abilify, gambling, and impulse control disorders, but he did not discuss any specifics about how the study would be designed or what its results might be. *See* Etminan Dep., ECF No. 457-7 at 27-30. Plaintiffs' counsel immediately ended the conversation and refrained from any further discussions about the study with Dr. Etminan until after it was published

in a peer-reviewed journal.[60] *See id*. Moreover, Plaintiffs did not fund the Etminan Study and, although they later retained Dr. Etminan as a consultant in this case, the record is devoid of any evidence that they shaped or directed Dr. Etminan's research.[61] While contacting Plaintiffs' counsel may have been "a strange thing to do," without more, it does not invalidate "the whole exercise, the protocol, and [] the study." *See* Madigan Tr., ECF No. 596-4 at 69.[62]

Importantly, there is no evidence in the record to suggest that any methodological aspect of the Etminan Study or its results was tainted by Dr. Etminan's alleged bias. Although the Etminan Study has its limitations, *see In re Orthopedic Bone Screw Prod. Liab. Litig.*, No. 1014, 1997 WL 230818 (E.D. Pa. May 5, 1997) ("[T]here is no such thing as a perfect epidemiological study."), none of Defendants' experts raised concerns about its impartiality. Not surprisingly, neither did any of Plaintiffs' experts, including Dr. Madigan, who acknowledged the ethical dictate that science be kept separate from economic influence, but found no

[60] Dr. Etminan called Plaintiffs' counsel a second time in October 2016, after the Etminan Study had been selected for publication by the Journal of Clinical Psychopharmacology. *See* Etminan Dep., ECF No. 457-7 at 28. Plaintiffs' counsel again ended the conversation after telling Dr. Etminan that they could speak once the study was actually published. *See id*.

[61] The Etminan Study was funded by the British Columbia Health Services Authority, a third-party Canadian health agency that is not involved in this litigation. *See* Etminan Study, ECF No. 428-13 at 1.

[62] "Madigan Tr." refers to the redacted version of the official transcript of Dr. David Madigan's testimony at the *Daubert* hearing, ECF No. 596-4 at 4-107.

indication of bias in the Etminan Study. *See* Madigan Tr., ECF No. 596-4 at 69. Finally, as aptly observed by Dr. Madigan, to the extent that Dr. Etminan may have been biased, that bias was mitigated by the participation of his co-author, Dr. Ric M. Procyshyn, who at the time had an ongoing relationship with Defendant Otsuka Pharmaceutical Company.[63] In sum, Defendants' argument that Dr. Etminan's alleged bias in favor of Plaintiffs renders the Etminan Study unreliable is not supported by the record. Concerns about litigation bias go to the Study's weight, not its admissibility. *See Adams v. Lab. Corp. of Am.*, 760 F.3d 1322, 1334 (11th Cir. 2014) (holding that bias in an expert witness's testimony generally goes to its weight, not its admissibility).

Defendants next argue that the Etminan Study is unreliable due to its failure to control for detection bias, also called reporting bias. In this context, detection bias refers to the possibility that pathological gambling and impulse control disorders were more likely to be detected and diagnosed in the exposed group (individuals taking Abilify) than in the unexposed group, due to increased medical awareness of the particular problems allegedly associated with Abilify. *See Zoloft I*, 26 F. Supp. at 464 ("Detection bias means that an abnormality . . . is more likely to be detected in the exposure group, often due to increased medical vigilance of exposed

[63] Dr. Procyshyn has served on advisory boards and speaker's bureaus for Otsuka. Etminan Study, ECF No. 428-13 at 4.

[individuals].”). Heightened public awareness, the argument necessarily goes, would have led doctors to probe for signs of pathological gambling and impulse control disorders in individuals from the exposed group, but to search less vigilantly for those signs in unexposed individuals, thereby increasing the likelihood of diagnoses in the former and decreasing the relative proportion of diagnoses in the latter.

The only evidence offered by Defendants to demonstrate the potential for detection bias in this case is a single “suspect adverse reaction report” by a patient who only reported having experienced an “urge to gamble” when he took Abilify after seeing an advertisement in 2014, which said that Abilify “causes compulsive gambling and was a bad drug.” *See* ECF No. 460-13 at 1. A single adverse reaction report is insufficient to discredit the Etminan Study as a whole. As the Court has already noted, adverse event reports are “one of the least reliable sources” of scientific information. *See McClain*, 401 F.3d at 1250. Moreover, this particular adverse reaction report is distinguishable from the data analyzed in the Etminan Study in that it reflects an apparent layperson’s self-assessment of his own symptoms, rather than a physician’s diagnosis of pathological gambling.[64] The fact

[64] The suspect adverse reaction report indicates that the consumer “stopped taking his [Abilify] after seeing the advertisement two or three days ago and said that his urge to gamble had stopped. At the time of the report, the consumer had not talked with his [health care provider] about stopping [Abilify] treatment.” ECF No. 460-13 at 2.

that a single patient may have been prompted, by an advertisement, to attribute his gambling "urge" to Abilify is not evidence that the gambling disorder diagnoses in the Etminan Study were prompted by publicity or are medically incorrect. In fact, during the time period analyzed by the Etminan Study, there was very little public information available about a possible link between Abilify and pathological gambling. Even according to Defendants' epidemiology expert, Dr. Douglas Weed, the first case reports raising the issue were not published in the medical literature until 2010. *See* Weed Supp., 419-23 at 13-14.[65] Although several more case reports were published in 2011 and 2014, *see id*., and at least one patient saw an advertisement in 2014, *see* ECF No. 460-13, most of the publicity around the issue did not begin until well after the time period covered by the Etminan Study.[66] Under these circumstances, the Court agrees with Dr. Madigan that it is "hard . . . to imagine" how publicity-related detection bias "could exert a lot of influence" on the diagnoses compiled by the Etminan Study, *see* Madigan Tr. at 66, especially to a degree that necessitates exclusion of the Study under *Daubert*. As with Defendants' litigation bias concerns, this publicity-related detection bias challenge implicates the weight, not admissibility, of the Etminan Study. *See Adams*, 760 F.3d at 1334.

[65] "Weed Supp." refers to Dr. Douglas L. Weed's Supplemental Expert Report, ECF No. 419-23 at 2-20.

[66] Dr. Weed identified three case reports published in the medical literature in 2010, five published in 2011, and eight published in 2014. *See* Weed Supp., ECF No. 419-23 at 14.

d. Dr. Madigan's Etminan Analysis[67]

Defendants raise two objections to Dr. Madigan's statistical analysis of the Etminan Study.[68] First, Defendants argue that his defense of the Study, an observational healthcare database study, is untenable in light of his published research criticizing observational healthcare database studies.[69] According to Defendants, Dr. Madigan's opinion in this case "violates his own standard of proper methodology" and suggests that he "does not apply the same rigor in the courtroom that he would apply to his scientific endeavors." *See* Def. Madigan Motion, ECF No. 427-20 at 15 (quoting *In re Rezulin Products Liability Litigation (MDL No. 1348)*, 309 F. Supp. 2d 531, 563 (S.D.N.Y. 2004)). The Court disagrees because Dr. Madigan's opinion in this case is not inconsistent with his prior publications.

---

[67] Again, the Court separately analyzes Dr. Madigan's qualifications to offer expert opinions in this case, and also the reliability of his opinions with respect to other scientific evidence, in Section II(D)(1)(d).

[68] Defendants also challenge Dr. Madigan's analysis of the Etminan Study on the grounds that he failed to properly address: (1) any potential confounders that were "not controlled for or otherwise addressed" in the Study, such as major depressive disorder; (2) possible investigator bias; (3) potential misclassification bias, and (4) the rapid onset time for symptoms of pathological gambling in patients from the LifeLink database. The Court has carefully considered these arguments and finds them lacking in merit, for the reasons more fully stated in the body of this Order. In short, the arguments reflect only on the probative value of Dr. Madigan's opinion, not its admissibility.

[69] *See, e.g.*, David Madigan *et al.*, *A Systematic Statistical Approach to Evaluating Evidence from Observational Studies*, 1 ANNUAL REVIEW OF STATISTICS AND ITS APPLICATION 11 (2014) ("Madigan 2014"), DX-117; David Madigan *et al.*, *Empirical Performance of the Case-Control Method: Lessons for Developing a Risk Identification and Analysis System*, 36 (Supp. 1) DRUG SAFETY S73 (2013) ("Madigan 2013a), ECF No. 427-6 at 2; David Madigan *et al.*, *Evaluating the Impact of Database Heterogeneity on Observational Study Results*, 178 AM. J. EPIDEMIOLOGY 645 (2013) ("Madigan 2013b"), ECF No. 427-8 at 2.

From 2009 to 2013, Dr. Madigan served as principal investigator for the Observational Medical Outcome Partnership ("OMOP"), a public-private partnership between the FDA, academia, and the pharmaceutical industry that was established, in part, to empirically evaluate the strengths and weaknesses of observational healthcare database studies of the effects of medical products.[70] Dr. Madigan testified that the focus of his research with the OMOP was large-scale automated observational health database research, which involves the use of high volume, generic data mining techniques to uncover hidden relationships of potential clinical significance to drug safety. Madigan Dep., ECF No. 427-1 at 106, 133-35; Madigan Tr., ECF No. 596-4 at 43-44. His work with the OMOP culminated in the publication of a series of peer-reviewed, scientific articles criticizing generic automated database analysis for its potential for bias, confounding, and inaccurate results.

Dr. Madigan's published criticisms of generic automated database research do not contradict his endorsement of the Etminan Study because, in short, the Etminan Study is not the product of generic automated database research. Rather, it was custom-designed to analyze a very specific clinical question—whether Abilify

[70] The OMOP was originally managed by the Foundation for the National Institutes of Health ("FNIH"). On its completion in June 2013, the OMOP was transferred from FNIH to the Reagan-Udall Foundation for the FDA and became known as the Innovation in Medical Evidence Development and Surveillance program. *See* https://fnih.org/what-we-do/major-completed-programs/omop (last visited Dec. 3, 2017).

is associated with pathological gambling and other impulse control disorders—and narrowly tailored to account for a number of factors unique to the LifeLink database and to the Abilify patient population. *See* Madigan Tr., ECF No. 596-4 at 44-45, 88-89; Madigan Dep., ECF No. 427-21 at 40. This customized design and implementation distinguishes the Etminan Study from the generic automated studies criticized by Dr. Madigan and the OMOP. Indeed, several of Dr. Madigan's publications actually recommend such "customizing [of] analyses to databases" and "thoughtful and careful study design" as means of improving the accuracy and performance of generic automated database research. *See* Madigan 2014, DX-117 at 27, 35; *see also* Madigan 2013b, ECF No. 427-8 ("It is conceivable that customizing the analytical approach to [drug-outcome] pairs could lead to greater consistency across databases."). The Etminan Study also implements many of the specific design-level and analytical strategies that Dr. Madigan and the OMOP suggest for reducing potential errors in generic automated database research, including careful matching of cases and controls, adjusting and controlling for potential confounders, and the use of sensitivity analyses to assess the potential consequences of unknown confounders. *See* Madigan 2014, DX-117 at 15-17. None of Defendants' experts disputed Dr. Madigan's explanation of the distinctions between customized database research, like the Etminan Study, and generic database research, like the studies criticized by Dr. Madigan. To the contrary, on this record,

Dr. Madigan's opinion with respect to the Etminan Study is consistent with his prior published literature and does not "violate[] his own standard of proper methodology." *See Rezulin*, 309 F. Supp. 2d at 563. Defendants' objection based on Dr. Madigan's published literature goes to the weight of his opinion in this case, not its admissibility.

Defendants also argue that Dr. Madigan's reliance on *p*-values to demonstrate the validity of the statistical evidence in this case is untenable in light of his published research criticizing traditional *p*-values for their vulnerability to systematic error, such as bias. *See* Martijn J. Schuemie *et al.*, *Interpreting Observational Studies: Why Empirical Calibration is Needed to Correct P-Values*, 33 STATISTICS MED. 209 (2014) ("Schuemie 2014"), ECF No. 427-5 at 2, 3. The Court disagrees. The *p*-value is a generally accepted statistical technique for evaluating the significance of the results of a statistical analysis. *See* Ref. Man. at 249-56, 258; *see also Jones v. City of Boston*, 752 F.3d 38, 43 (1st Cir. 2014) (describing "customar[y]" use of *p*-values in statistical analysis); *In re Lipitor (Atorvastatin Calcium) Mktg., Sales Practices and Prods. Liab. Litig.*, 174 F. Supp. 3d 911, 914-15 (D.S.C. 2016) (stating that the calculation of a *p*-value is a "common" way to evaluate whether an observed association is due to chance). Even Defendant's expert epidemiologist, Dr. Douglas Weed, agrees that *p*-values are proper for evaluating the statistical significance of an observed association. *See*

Weed Rep., ECF No. 419-3 at 21.[71,72] Thus, the Court finds Dr. Madigan's use of *p*-values to evaluate the significance of the statistical evidence in this case a reliable part of his methodology.

While it is true that Dr. Madigan and his colleagues have proposed "a new empirical framework" for evaluating statistical significance, called a calibrated *p*-value, which they assert will "minimiz[e] the potential effects of bias when interpreting observational study results," *see* Schuemie 2014, ECF No. 427-5 at 3, this calibrated *p*-value framework is "controversial" and has not yet gained general acceptance or approval in the scientific community, Madigan Tr., ECF No. 596-4 at 50. Indeed, Dr. Madigan testified at his deposition that, at least as of that time, "the only people to have ever calculated calibrated *p*-values [were him]self and [his] coworkers." *See* Madigan Dep., ECF No. 427-1 at 148.[73] Perhaps one day Dr. Madigan's calibrated *p*-value framework will become the gold standard for evaluating statistical significance, but it simply is not there yet and, for now at least, it is not altogether clear whether the framework is sufficiently well-established and

[71] "Weed Rep." refers to Dr. Douglas L. Weed's Expert Report, ECF No. 419-3 at 2-78.

[72] In his initial expert report, Dr. Weed stated that the Bradford Hill factors "are only applied . . . once a statistical association has been established, *i.e.* once at least one (and typically several) epidemiological studies has revealed increased risks of a disease (or condition) that could be called "statistically significant," most often at the level of ($p < 0.05$), whether evaluated using *p*-values or confidence intervals." *See* Weed Rep., ECF No. 419-3 at 21.

[73] "Madigan Dep." refers to the official transcript of Dr. David Madigan's deposition testimony on June 28, 2017, ECF No. 427-1 at 94-238.

reliable to satisfy Rule 702 and *Daubert*. *See Rider*, 295 F.3d at 1202 ("Law lags science; it does not lead it."). In any event, Dr. Madigan did not use calibrated *p*-values, so the reliability of that statistical technique is not before the Court.[74] Dr. Madigan explained that he "was not in a position to calculate a calibrated *p*-value" for the Etminan Study because he did not have access to the information (*i.e.*, the underlying data from the LifeLink database) needed to perform the calculation. *See* Madigan Dep., ECF No. 427-1 at 148; Madigan Tr., ECF No. 596-4 at 50-52. Dr. Madigan explained that this did not affect his ability to opine on the statistical significance of the Etminan Study because he was able to calculate a *p*-value for pathological gambling and, in any event, the Study's relative risk findings were too "robust" to be explained by bias. *See* Madigan Tr., ECF No. 596-4 at 70. Defendants have offered no evidence to the contrary. Under these circumstances, the Court finds that Dr. Madigan's failure to calculate a calibrated *p*-value was not, as Defendants suggest, a failure to "employ in the courtroom the same level of intellectual rigor that characterizes the practice of" an expert statistician in the field. *See Kumho Tire*, 526 U.S. at 152. Thus, Defendants' objection on that basis goes to the weight of Dr. Madigan's opinion, not its admissibility.

[74] The Court notes that if Dr. Madigan had offered calibrated *p*-values in support of his opinion instead of traditional *p*-values, there undoubtedly would have been reliability challenges to that methodology.

Based on the foregoing, the Court finds the Etminan Study is a scientifically sound epidemiological study and, therefore, reliable evidence of general causation in this case.

2. Dose-Response Relationship

In addition to epidemiology, Plaintiffs' experts offer a series of case studies and adverse event reports as evidence of a dose-response relationship between Abilify, impulsive gambling, and other impulse control disorders.[75] These materials describe the onset of new impulse control problems in individual patients after their doses of Abilify were increased, problems which disappeared when the Abilify doses were reduced or discontinued.[76] While the Court finds this evidence suggestive of a dose-response relationship, it ultimately lacks the intrinsic reliability that is the hallmark of a primary methodology under Eleventh Circuit *Daubert* jurisprudence.

In the Eleventh Circuit, the use of dose-response evidence as a "primary" means of establishing causation generally requires a scientifically reliable showing of a correlation between dosage and disease, the minimum dose at which adverse effects are seen and the dose at which a substance is lethal. *See McClain*, 401 F.3d

[75] Dr. Joseph Glenmullen's expert report ("Glenmullen Rep."), ECF No. 424-1 at 1-137, 88-90, 103-04, 134 (compiling case and adverse events reports involving a "dose-dependent" relationship between Abilify, pathological gambling, and other impulse control disorders); Dr. Eric Hollander's Initial General Causation Report ("Hollander Rep."), ECF No. 459-1 at 30-31 (same); Dr. David Madigan's Rebuttal Report ("Madigan Rep."), ECF No. 427-1 at 79-92, 83 (referencing FDA discussion of case and adverse event reports involving dose response relationship).

[76] *See* Glenmullen Rep., ECF No. 424-1 at 88-90, 103-04 (collecting case studies).

at 1241-43. Scientists may reliably establish the existence of a dose-response relationship in several ways. One approach is through a well-controlled clinical dose-response study, which may be designed to measure dose-response relationships in humans at both the individual and population levels.[77] A dose-response study "allow[s] observations of benefits and risks at different doses" and "can help ensure that excessive doses [] are not used, offering some protection against unexpected and unrecognized dose-related toxicity." FDA Dose Response Guide, *supra* note 77 at 4. Because of the ethical limitations inherent in human experimentation, scientists must usually rely on *in vitro* and *in vivo* animal studies to examine the relationship between dose and response. *See* Ref. Man. at 563-64, 639. Animal toxicological research "often provides the best scientific information about" dose-response relationships and other toxicity risks associated with exposure to a drug. *See id*. at 639, 641.

In this case, Plaintiffs' experts have not presented any controlled, experimentally derived evidence of a dose-response relationship between Abilify and impulse control disorders. While the absence of such evidence is not fatal to the experts' general causation opinions, it does weaken the force and reliability of their

[77] U.S. Dep't of Health and Human Serv., FDA, *Guidance for Industry: Exposure-Response Relationships—Study Design, Data Analysis, and Regulatory Applications* (April 2003), https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm072109.pdf (last visited Dec. 3, 2017) ("FDA Dose-Response Guide").

conclusions as to dose-response. Nonetheless, the Court agrees with the FDA that a number of the published case studies—those describing positive dechallenge and rechallenge events, in particular—indicate a "temporal relationship between the initiation of [Abilify] treatment and the onset of" impulse control problems. *See* FDA Pharm. Vigil., ECF No. 428-11 at 21-25. Some of these case studies and adverse event reports also strongly suggest that an increase in a patient's dose of Abilify may increase that patient's risk of impulse control problems, while a decrease in dose may correspondingly decrease the risk.[78] This evidence is "substantially more valuable than run-of-the-mill case reports." *See Glastetter*, 252 F.3d at 990. However, these materials do not provide a sufficient evidentiary basis from which to delineate the threshold exposure, or even a potential range of threshold exposures, at which Abilify causes adverse effects. *See McClain*, 401 F.3d at 1241 (a reliable dose-response opinion addresses "the dose or level of exposure at which [a drug] causes harm"). More significantly, the lack of meaningful scientific controls limits the weight that these case studies and adverse event reports may reliably bear on an expert's general causation opinion under Eleventh Circuit

[78] *See, e.g.*, L. Gaboriau *et al*., *Aripiprazole: A New Risk Factor for Pathological Gambling? A Report of 8 Case Reports*, 39 ADDICTIVE BEHAVIORS 526, 563 (2014) ("irresistible urge to gamble" developed "immediately" after Abilify daily dose increased from 10 mg to 20 mg per day; urges stopped after dose decreased to 15 mg); Milton G. Roxanas, *Pathological Gambling and Compulsive Eating Associated with Aripiprazole*, 44 AUSTRALIAN & NEW ZEALAND J. PSYCHIATRY 291 (2010) (urges to gamble and eat developed six months after daily dose of Abilify was increased from 10 mg to 15 mg; urges stopped one month after discontinuing Abilify).

standards. *See Chapman*, 766 F.3d at 1308 (stating that "generalized case reports" are "secondary methodologies," not recognized as "primary" or "indispensable" means of proving the effect of a drug). Therefore, the Court finds that Plaintiffs' anecdotal evidence of a dose-response relationship between Abilify and impulse control disorders is relevant and admissible, but only as a supplement to the other, more substantial evidence of general causation (*i.e.*, the Etminan Study).[79]

3. Background Risk

Two of Plaintiffs' experts, Drs. Glenmullen and Hollander, provide the background risk or prevalence of various impulse control disorders, including compulsive gambling, in the general population, as reflected in the scientific literature. More specifically, Dr. Glenmullen, relying on the DSM-5, stated that the past-year prevalence rate of gambling disorder is approximately 0.2%-0.3% in the general population. *See* Glenmullen Rep., ECF No. 424-1 at 59.[80] According to the DSM-5, the lifetime prevalence rate of gambling disorder in the general population

[79] The Court is aware that at least one district court in this Circuit appears to have considered "dechallenge-rechallenge reports" to be reliable, primary evidence of a dose-response relationship under *Daubert*. *See In re Chantix (Varenicline) Prods. Liab. Litig.*, 889 F. Supp. 2d 1272 (N.D. Ala. 2012). Given the Eleventh Circuit's clearly drawn distinction between "primary" and "secondary" methods of proving general causation in pharmaceutical product liability cases, *see Chapman*, 766 F.3d at 1308, the Court finds that the case studies and adverse event reports describing dechallenge and rechallenge events in this case, while certainly constituting strong secondary evidence, are not sufficiently reliable to warrant elevation to the status of a primary methodology.

[80] *See* DSM-5 at 587, ECF No. 428-3 at 66.

is about 0.4%-1.0%. *See* DSM-5, ECF No. 428-3 at 66. Dr. Hollander referenced several scientific studies identifying the background risks for compulsive shopping (6%-7%) and hypersexuality (3%-6%) in the general population, and for compulsive eating (1.6%) in the general adolescent population. *See* Hollander Rep., ECF No. 459-1 at 8-9. The Court finds that these figures constitute reliable evidence of background risk.

The fact that Plaintiffs' experts do not offer a more expansive analysis of background risk in this case does not present a "serious methodological deficiency" or "substantial weakness" in their general causation opinions. *See Chapman*, 766 F.3d at 1308. According to the FDA, at least as recently as March 2016, there had been "no large studies of the life-time prevalence of most [impulse control disorders] in the general population, except pathological gambling." *See* FDA Pharm. Vigil., ECF No. 428-11 at 7. Thus, this is not a case where Plaintiffs' experts simply ignored the available evidence regarding the background risk of various impulse control disorders. *See Kilpatrick*, 613 F.3d at 1342 (stating that an expert who "ignore[s]" available evidence about background risk "place[s] the reliability of [his] conclusions in . . . doubt"). Rather, the uncontroverted record evidence indicates that Plaintiffs' experts identified and accounted for the limited body of scientific literature currently available on background risk for impulse control disorders. In so doing, they satisfied Rule 702 and *Daubert*.

4. Biological Plausibility

Plaintiffs' experts share the opinion that the biological mechanism by which Abilify can cause pathological gambling and impulse control disorders is its effect on dopamine neurotransmission in the brain.[81] Dr. Antoine Bechara provides the most thorough analysis of the medical literature offered in support of this position. *See* Bechara Rep., ECF No. 423-1.[82] The Court first briefly summarizes Dr. Bechara's biological plausibility opinion and then considers Defendants' reliability challenges.[83]

According to Dr. Bechara, Abilify binds to over 90% of postsynaptic D_2 receptors in the nucleus accumbens and acts as a functional antagonist, occupying the D_2 receptors and preventing dopamine molecules from attaching and activating them, thereby blocking dopamine neurotransmission at those sites. The brain compensates for the resultant decrease in dopaminergic activity by increasing, or upregulating, the number of dopamine receptors in the nucleus accumbens; and also by increasing the sensitivity of those receptors, so that when activated by dopamine they produce greater, more "potentiated" physiological responses than would occur

[81] *See* Dr. Antoine Bechara's Expert Report, ECF No. 423-1; Dr. Joseph Glenmullen's Expert Report, ECF No. 424-1 at 38-56; Dr. Eric Hollander's Expert Report, ECF No. 425-1 at 9-11, 13-19; Dr. Russell V. Luepker's Expert Report, ECF No. 462-1 at 6-7.

[82] "Bechara Rep." refers to Dr. Antoine Bechara's Expert Report, ECF No. 423-1.

[83] The Court separately analyzes Dr. Bechara's qualifications to offer expert opinions in this case in Section II(D)(1)(a).

naturally. *See* Bechara Tr., ECF No. 596-3 at 100-101.[84] At the same time, the "displaced" dopamine molecules, unable to bind to D_2 receptors, diffuse towards the other available receptors in the nucleus accumbens, most of which are D_3 receptors. Abilify occupies and "strongly stimulates" about 30% of these D_3 receptors, producing in them between 50% and 100% of the physiological response that dopamine naturally produces. *See* Bechara Rep., ECF No. 423-1 at 7. The remaining D_3 receptors, which are upregulating and hypersensitive, bind with the endogenous dopamine to produce "supercharge[d]" reward-seeking behavior. *See* Bechara Tr., ECF No. 596-3 at 98. Finally, Dr. Bechara also opines that Abilify's functional antagonism at D_2 receptors may disrupt the brain's ability to process the consequences of negative behavior.

Defendants challenge the reliability of Plaintiffs' experts' proposed mechanism of action on the ground that it lacks evidentiary support and instead is premised on "pure speculation." Def. Bechara Motion, ECF No. 423-10 at 17.[85]

a. Displacement

Defendants argue that there is no scientific support for the proposition that endogenous dopamine is "displaced" when Abilify occupies a majority of D_2

[84] "Bechara Tr." refers to the redacted version of the official transcript of Dr. Antoine Bechara's testimony at the *Daubert* hearing, ECF No. 596-1 at 74-168.

[85] "Def. Bechara Motion" refers to Defendants' Motion to Exclude the General Causation Opinion on Antoine Bechara, ECF No. 423-10.

receptors in the nucleus accumbens. *See id*. at 18. The Court disagrees. Even Defendants' own psychopharmacology expert, Dr. Pierre Blier, agrees that Abilify displaces dopamine at D_2 receptors. *See* Blier Dep., ECF No. 455-2 at 241, 248, 258-59.[86] Moreover, **[*** REDACTED ***]**.[87] **[*** REDACTED ***]**. The fact of displacement also appears to be supported by the scientific literature, which indicates that, at therapeutic doses, Abilify occupies and blocks over 90% of postsynaptic D_2 receptors in the nucleus accumbens, leaving only "10% or fewer D_2 receptors [] available for endogenous dopamine to bind" with. *See* Takashi Hamamura & Toshiki Harada, *Unique Pharmacological Profile of Aripiprazole as the Phasic Component Buster*, 191 PSYCHOPHARMACOLOGY 741, 742 (2007) ("Hamamura 2007"), PX-020 at 2.[88] Given this evidence, the "displacement" premise of Plaintiffs' experts' opinions cannot be considered "pure speculation." *See* Def. Bechara Motion, ECF No. 423-10 at 21-22.

---

[86] "Blier Dep." refers to the official transcript of Dr. Pierre Blier's deposition testimony on June 19, 2017, ECF No. 455-2 at 175-374.

[87] **[*** REDACTED ***]**.

[88] *See also* K. Maeda *et al*., *Brexpiprazole I: In Vitro and In Vivo Characterization of a Novel Serotonin-Dopamine Activity Modulator*, 350 J. PHARMACOLOGY & EXPERIMENTAL THERAPEUTICS 589 (2014) (clinically effective dose range of Abilify leads to "80% to 90% D_2 receptor occupancy") ("Maeda 2014"), DX-062; Takashi Hamamura *et al*., *Intrinsic Activity of Aripiprazole is Not 30% of Dopamine, But Only About 6% Under Ideal Antipsychotic Therapy*, 69 J. CLINICAL PSYCHIATRY 863, 843 (2008) ("Hamamura 2008") (same), DX-057 at 3; M. Kodama & T. Hamamura, *Aripiprazole-Induced Behavioral Disturbance Related to Impulse Control in a Clinical Setting*, 13 INT'L J. NEUROPSYCHOPHARMACOLOGY 549, 550 (2010) ("Kodama 2010") ("Because [Abilify] has a high affinity to D_2 receptors, about 90% of D_2 receptors are occupied [and blocked] by it."), ECF No. 453-17 at 3.

Defendants also argue that there is no evidentiary support for a "key assumption" of the "displacement theory," namely, that Abilify "occup[ies] relatively more D_2 receptors than D_3 receptors." *See id.*, ECF No. 423-10 at 22 (citing Bechara Rep., ECF No. 423-1 at 13). Again, the Court disagrees. As explained in Maeda 2014, which is a peer-reviewed, published article relied on by Dr. Blier, *see* Blier Rep., ECF No. 455-1 at 15-16, the only available method for directly measuring human receptor occupancy is PET imaging, *see* Maeda 2014 at 600, DX-062 at 12. Where PET imaging is unavailable, *in vitro* studies of receptor affinity are used to predict receptor occupancy.[89] *See id*. In this case, there are peer-reviewed, published *in vitro* studies in the scientific literature describing Abilify's higher affinity for D_2 receptors than for D_3 receptors, at least one of which found the drug's affinity to be over three-fold higher for D_2 receptors than for D_3 receptors.[90]

[89] As the Court discussed in Section I of this Order, affinity is a measure of whether and how strongly a drug binds, or attaches, to a particular receptor. Affinity is distinct from intrinsic activity, which measures the degree to which the drug, once bound, activates dopamine receptors to produce a physiological effect.

[90] Yoshihiro Tadori *et al*., *Functional Potencies of Dopamine Agonists and Antagonists at Human Dopamine D_2 and D_3 Receptors*, 666 EUROPEAN J. PHARMACOLOGY 43, 45 (2011) ("Tadori 2011a") (Abilify exhibiting more than three times higher affinity for D_2 than for D_3), PX-021 at 3; *see also* Béla Kiss *et al*., *Cariprazine (RGH-188), a Dopamine D3 Receptor-Preferring, D_3/D_2 Dopamine Receptor Antagonist-Partial Agonist Antipsychotic Candidate: In Vitro and Neurochemical Profile*, 333 J. PHARMACOLOGY & EXPERIMENTAL THERAPEUTICS 328, 332 (2010) (Kiss 2010) (describing Abilify's higher affinity for D_2 than D_3 receptors); T. Hirose & T. Kikuchi, *Aripiprazole, A Novel Antipsychotic Agent: Dopamine D2 Receptor Partial Agonist*, 52 J. MEDICAL INVESTIGATION SUPPL. 284, 288 (2005) ("Hirose 2005") (describing Abilify's higher affinity for D_2 than for D_3), DX-281 at 5.

Moreover, **[*** REDACTED ***]**.[91] Finally, Defendants' psychopharmacology expert, Dr. Blier, agrees that Abilify has "a presence three to ten times lower at D_3 [receptors] than at D_2 receptors." *See* Blier Rep., ECF No. 455-1 at 16. This evidence supports a plausible conclusion that Abilify binds to D_2 receptors (affinity), and thereby occupies them, with greater frequency and strength than it does to D_3 receptors. For this reason, the Court finds Dr. Bechara's opinion that "Abilify would occupy relatively more D_2 receptors" than D_3 receptors supported by reliable scientific evidence in this case. *See* Bechara Rep., ECF No. 423-1 at 13.

Dr. Bechara's opinion that "displaced" endogenous dopamine diffuses towards, and binds with, other available receptors in the nucleus accumbens, most of which are D_3 receptors, is also biologically plausible based on the scientific evidence in this case. It appears to be well-established in the scientific literature that, in humans, D_3 receptors predominate in the mesolimbic regions of the brain, including the nucleus accumbens.[92] In fact, only one expert, Dr. Blier, has offered a

---

[91] **[*** REDACTED ***]**.

[92] *See, e.g.*, Tadori 2011a at 43 ("Dopamine D_3 receptors are predominantly expressed in the nucleus accumbens"), PX-021 at 1; Kodama 2010 ("[T]he dopamine D3 receptor [ ] is highly enriched in the nucleus accumbens and plays an important role in reward."); Kiss 2010, 333 J. PHARMACOLOGY & EXPERIMENTAL THERAPEUTICS at 328 ("Dopamine D_3 receptors . . . are most abundant in the mesolimbic regions (*i.e.*, nucleus accumbens, island of Calleja")); Tadori 2008 at 30 ("[I]n humans, dopamine D_3 receptors are predominantly expressed in regions of the limbic striatum, including the nucleus accumbens."), DX-058 at 1; Gurevich & Joyce, *Distribution of Dopamine D_3 Receptor Expressing Neurons in the Human Forebrain: Comparison with D_2 Receptor Expressing Neurons*, 20 NEUROPSYCHOPHARMACOLOGY 60, 64, 74 (1999) "Gurevich 1999") (finding the nucleus accumbens and ventral striatum "enriched" with "more abundant" concentrations of D_3 receptors than D_2 receptors); Murray *et al.*, *Localization of Dopamine D3*

different opinion. In his expert report, Dr. Blier stated that D_1, D_2, D_3, and D_4 receptors "are all expressed in the nucleus accumbens, the first three in high density." *See* Blier Rep., ECF No. 455-1 at 16. Dr. Blier offers a single citation in support of this opinion, a 2003 book by Jack R. Cooper. *See id*. (citing Cooper *et al*., THE BIOCHEMICAL BASIS OF NEUROPHARMACOLOGY (Oxford Univ. Press 2003)). However, there is no corresponding exhibit in the record (*i.e*., a copy of the relevant chapter or pages) for the Court to compare with Dr. Blier's statement. Given that virtually all of the other scientific literature in this case indicates that the mesolimbic reward pathway is a predominantly D_3-rich environment, Dr. Blier's representation to the contrary must be supported by more than just his *ipse dixit*. *See Joiner*, 522 U.S. at 146 ("[N]othing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert."). Moreover, the record reflects **[*** REDACTED ***]**. To the extent Dr. Blier now opines otherwise, the Court rejects that opinion as unsupported by the scientific evidence in this case. Having said that, even assuming Dr. Blier's opinion on this issue was supported by the evidence, the opinion might

Receptors to Mesolimbic and D2 Receptors to Mesostriatal Regions of Human Forebrain, 91 PROCEEDINGS NAT'L ACAD. SCI. U.S.A. 11271, 11274 (1994) (finding "high concentration" of D_3 receptors in the nucleus accumbens); *see also* FDA Pharm. Vigil., ECF No. 428-11 at 8 "[T]he D_3 dopamine receptor subtype is found predominantly in the limbic regions of the brain.").

create an area of reasonable debate for the experts, but it would not detract from the plausibility of Dr. Bechara's proposed mechanism of action.

Relatedly, the record evidence also reflects, and Defendants' experts have not disputed, that endogenous dopamine has "higher affinity for" D_3 receptors than D_2 receptors. *See* Tadori 2011a at 51, PX-021 at 9; *see also* Gurevich 1999 at 78 (stating that dopamine has "significantly higher affinity" for D_3 receptors than for D_2 receptors). Dr. Blier testified that displaced endogenous dopamine can diffuse "everywhere," *see* Blier Tr., ECF No. 596-8 at 27, including postsynaptic dopamine receptors "not located in precise apposition to" the presynaptic, sending neuron, *see* Blier Rep., ECF No. 455-1 at 15; *see also* Blier Dep., ECF No. 455-2 at 56-58. According to Dr. Blier, endogenous dopamine "diffuses all around the neuron" and may act on any of the "different types of receptors . . . in that region." *See* Blier Tr., ECF No. 596-8 at 27. The scientific literature confirms Dr. Blier's opinion on this issue.[93]

Taken together, this evidence reasonably and reliably supports the plausibility of Dr. Bechara's displacement opinion. In short, the scientific evidence reflects that (1) endogenous dopamine has a high affinity for D_3 receptors; (2) D_3 receptors are

[93] *See* L. Descarries, *et al*., *Dual Character, Asynaptic and Synaptic of the Dopamine Innervation in Adult Rat Neostriatum: A Quantitative Autoradiographic and Immunocytochemical Analysis*, 375 J. COMPARATIVE NEUROLOGY 167, 183 (1996), DX-251 at 17 (stating that dopamine may spill or diffuse away from a synapse and activate dopamine receptors not located in precise apposition to the presynaptic neuron).

the predominant dopamine receptors in the mesolimbic pathway, including the nucleus accumbens; and (3) displaced endogenous dopamine may diffuse throughout that region of the brain. Under these conditions, it is certainly plausible that displaced endogenous dopamine would diffuse to, and activate, at least some of the D_3 receptors so abundant in the mesolimbic pathway.

At this point, the Court finds it important to emphasize that determining whether an expert's opinion is "biologically plausible" is a far different inquiry than determining whether an opinion is "biologically certain." *See Daubert*, 509 U.S. at 590 ("[I]t would be unreasonable to conclude that the subject of scientific testimony must be 'known' to a certainty; arguably, there are no certainties in science."); *Jones v. Otis Elevator Co*., 861 F.2d 655, 662 (11th Cir. 1988) (stating that "absolute certainty is not required" from expert testimony). In other words, for *Daubert* purposes, the proponent of an expert opinion "does not have the burden of proving that [the opinion] is scientifically correct, but that by a preponderance of the evidence, it is reliable." *Allison*, 184 F.3d at 1312. This means that the expert must "know[] of facts which enable him to express a reasonably accurate conclusion as opposed to conjecture or speculation." *Jones*, 861 F.2d at 662. In this case, the Court finds Dr. Bechara's conclusions about displacement to be reasonably reliable, as opposed to merely speculative, in light of the known science about Abilify and the biochemistry of the brain. To the extent Defendants' experts draw a different

conclusion from those same facts, this presents a proverbial "battle of the experts," which appropriately should be decided by a jury. *See Allapattah Servs., Inc. v. Exxon Corp.*, 61 F. Supp. 2d 1335, 1341 (S.D. Fla. 1999), *aff'd*, 333 F.3d 1248 (11th Cir. 2003), *aff'd sub nom*., *Exxon Mobil Corp. v. Allapattah Servs., Inc.*, 545 U.S. 546 (2005) ("Merely because two qualified experts reach directly opposite conclusions using similar, if not identical data bases, or disagree over which data to use or the manner in which the data should be evaluated, does not necessarily mean that, under *Daubert*, one opinion is *per se* unreliable."); *see also Quiet Tech*, 326 F.3d 1341 (stating that *Daubert* does not permit district courts to "evaluate the credibility of opposing experts and the persuasiveness of competing scientific studies"). Defendants' challenges to Dr. Bechara's displacement opinion go to its weight, not its admissibility.

b. Upregulation and Sensitization

Defendants also argue there is no medical support for Plaintiffs' experts' opinions that Abilify's functional antagonism of D_2 receptors triggers upregulation and sensitization of dopamine receptors in the brain. This is incorrect. Both phenomena have been described in published case reports[94] and, more significantly,

[94] L. Gaboriau *et al*., *Aripiprazole: A New Risk Factor for Pathological Gambling? A Report of 8 Case Reports*, 39 ADDICTIVE BEHAVIORS 562 (2014) (observing that upregulation and sensitization may "hyperstimulat[e]" D_3 receptors in Abilify patients, particularly those treated with antipsychotic dopamine antagonists in the past), ECF No. 428-8; Julien Cohen *et al*., *Aripiprazole-Induced Pathological Gambling: A Report of 3 Cases*, 6 CURRENT DRUG SAFETY 51 (2011) ("The appearance of [pathological gambling] in these [patients] could have been caused by

demonstrated via peer-reviewed, published *in vivo* studies.[95] **[*** REDACTED ***]**. Dr. Blier testified that he too was "aware of other studies that have shown the sensitization" and upregulation effects of Abilify, though "not necessarily to any greater extent than" a pure D_2 receptor antagonist. *See* Blier Tr., ECF No. 596-8 at 30. Finally, the FDA also has acknowledged, based on the medical literature, that D_3 partial agonism, upregulation, and sensitization are plausible biological mechanisms that "could" explain the onset of impulse control disorders after Abilify exposure, although it cautioned that more research was necessary before any final conclusion can be reached. *See* FDA Pharm. Vigil., ECF No. 428-11 at 29. Given this evidence, Dr. Bechara's opinion as to upregulation and sensitization cannot be considered "improper *ipse dixit*." *See* Def. Bechara Motion, ECF No. 423-10 at 26.

The fact that none of the other atypical antipsychotics is associated with a higher incidence of impulse control disorders, even though those drugs may "cause more upregulation than Abilify," *see id*. at 25, does not detract from the reliability

---

the aberrant stimulation of the [meso-cortico-limbic] pathway by [Abilify]."), ECF No. 428-6 at 4.

[95] Yong Kee Choi *et al*., *Long-term Effects of Aripiprazole Exposure on Monoaminergic and Glutamatergic Receptor Subtypes: Comparison with Cariprazine*, 22 CNS SPECTRUMS 484 (2017) (*in vivo* study finding chronic treatment with Abilify upregulated D_2 and D_3 receptor levels in rats, but did not increase D_3 receptors in shell of the nucleus accumbens); Jun Gao *et al*., *Repeated Administration of Aripiprazole Produces a Sensitization Effect in the Suppression of Avoidance Responding and Phencyclidine-Induced Hyperlocomotion and Increases D_2 Receptor-Mediated Behavioral Function*, 29 J. PSYCHOPHARMACOLOGY 390 (2015), DX-067 (*in vivo* study finding sensitization and upregulation of D_2 and D_3 receptors in rats treated with therapeutic dose of Abilify).

of Dr. Bechara's opinion that Abilify can cause upregulation and sensitization. None of the scientific evidence in this case indicates that earlier atypical antipsychotics are agonists at D_3 receptors.[96] In any event**,** the record makes clear that Abilify has a "unique" pharmacological profile; therefore, analogies between Abilify and other atypical antipsychotics are not permissible unless reliable scientific evidence establishes the validity of the analogy. *See McClain*, 401 F.3d at 1246; *Rider*, 295 F.3d at 1200-01. Defendants have not offered any such evidence or even identified the specific atypical antipsychotics with which they are asking the Court to draw a comparison. Defendants' challenges to Dr. Bechara's upregulation and sensitization opinion go to its weight, not its admissibility.

c. Direct Agonism

Defendants argue that Plaintiffs' experts do not have "any methodologically reliable basis" for concluding that Abilify acts differently at D_3 receptors than it does at D_2 receptors. *See* Def. Bechara Motion, ECF No. 423-10 at 12. More specifically, Defendants argue that (1) the *in vitro* studies offered in support of this conclusion may not reliably support an expert opinion on general causation; and (2) the only way Dr. Bechara can show any difference between Abilify's effects at D_2 and D_3

[96] To the contrary, the record indicates that many atypical antipsychotics are in fact antagonists at D_3 receptors. *See* Yoshiro Tadori *et al*., *Characterization of Aripiprazole Partial Agonist Activity at Human Dopamine D3 Receptors*, 597 EUROPEAN J. PHARMACOLOGY 27, 31 (2008) (*in vitro* study) ("Tadori 2008"), DX-58 at 5.

receptors is by "cherry-picking" data from the scientific landscape to give the illusion of clarity in favor of his biological plausibility opinion. Again, the Court disagrees.

Before addressing the merits of Defendants' objection, it is necessary to clarify what is *not* in dispute. Defendants do not dispute the apparent scientific consensus that Abilify is considered a partial agonist that acts as a functional antagonist at postsynaptic D_2 receptors.[97] *See id.*, ECF No. 423-10 at 11-12; *see also* Blier Tr., ECF No. 596-8 at 47; Def. Bechara Reply, ECF No. 479-8 at 3; Bechara Tr., ECF No. 596-3 at 113-14.[98] Indeed, **[*** REDACTED ***]**. **[*** REDACTED ***]**; *see also* Tadori 2011a at 51 ("All antipsychotics are anti-'psychotic' due to their ability to block phasic postsynaptic D_2 receptor signals."), PX-021 at 9.[99] There also is no dispute that Abilify acts as a partial agonist at *presynaptic* D_2 receptors. *See* Pl. Bechara Opposition, ECF No. 453-32 at 33;[100] **[*** REDACTED ***]**. Even Defendants' expert, Dr. Blier, recognizes that Abilify's partial agonism at presynaptic D_2 receptors is thought to slow the firing

---

[97] **[*** REDACTED ***]**.

[98] "Def. Bechara Reply" refers to Defendants' Reply to Exclude the General Causation of Antoine Bechara, ECF No. 479-8.

[99] Yoshihiro Tadori *et al*., *Functional Potencies of Dopamine Agonists and Antagonists at Human Dopamine D2 and D3 Receptors*, 666 EUROPEAN J. PHARMACOLOGY 43, 48 (2011) ("Tadori 2011a"), PX-021.

[100] "Pl. Bechara Opposition" refers to Plaintiffs' Memorandum of Law in Opposition to Defendants' Motion to Exclude the General Causation Opinion of Antoine Bechara, ECF No. 453-32.

rate and release of dopamine into the synapses.[101] *See* Blier Rep., ECF No. 455-1 at 6, 13-14. Thus, for purposes of this challenge at least, the *only* dispute is whether Dr. Bechara has a sufficient scientific basis to opine that Abilify acts with greater intrinsic activity at postsynaptic D_3 receptors than at *postsynaptic* (as opposed to presynaptic) D_2 receptors. The Court finds that he does.

First, there is support in the scientific literature for the fact that "the actions of [Abilify] differ markedly across [dopamine] receptor systems." David A. Shapiro *et al*., *Aripiprazole, A Novel Atypical Antipsychotic Drug with a Unique and Robust Pharmacology*, 28 NEUROPSYCHOPHARMACOLOGY 1400, 1407-08 (2003) ("Shapiro 2003"), DX-45 at 8-9. For example, one *in vitro* study using animal cells transfected with human dopamine receptors found that, at least in a controlled laboratory environment, Abilify was "sometimes an antagonist (*e.g.*, D_2), sometimes a partial agonist (*e.g.*, D_2), and sometimes a full agonist (D_3, D_4)." *See id*. at 1408, DX-45 at 9. This finding is consistent with subsequent scientific literature, which has characterized Abilify as "functionally selective" for its "markedly different" effects at individual D_2 and D_2-like receptors in the various dopamine pathways. *See* Richard B. Mailman and Vishakantha Murthy, *Third Generation Antipsychotic Drugs: Partial Agonist or Receptor Functional Selectivity*, 16 CURRENT PHARM.

[101] Presynaptic receptors, also called autoreceptors, provide a mechanism by which dopamine neurons regulate functions such the release and synthesis of dopamine.

DESIGN 488, 492-93 (2011), DX-069 at 5-6. This evidence reliably supports a fundamental premise of Plaintiffs' experts' proposed mechanism of action, that Abilify *can* act differently at postsynaptic D_3 receptors than it acts at postsynaptic D_2 receptors.

Second, there is scientific literature reflecting that Abilify is *not* an antagonist at D_3 receptors. Yoshiro Tadori *et al*., *Characterization of Aripiprazole Partial Agonist Activity at Human Dopamine D3 Receptors*, 597 EUROPEAN J. PHARMACOLOGY 27, 31 (2008) (*in vitro* study) ("Tadori 2008"), DX-58 at 5. This fact alone supports a distinction between Abilify's effects at postsynaptic D_2 receptors (functional antagonism) and its effects at D_3 receptors (no antagonism).[102] **[*** REDACTED ***]**.[103] This evidence reliably supports Plaintiffs' experts' opinions that Abilify *does* act differently at D_3 receptors than it acts at postsynaptic D_2 receptors.

Finally, multiple *in vitro* studies show that Abilify exhibits strong partial to full agonist activity at D_3 receptors,[104] while exhibiting very low partial agonist to

[102] Again, antagonists bind to dopamine receptors and produce no physiological effects. Partial agonists bind to dopamine receptors, but produce less of a physiological effect than endogenous dopamine would produce. Full agonists bind to dopamine receptors and mimic the activity of dopamine, producing the same level of physiological response that dopamine naturally produces.

[103] **[*** REDACTED ***]**.

[104] Yoshihiro Tadori *et al*., *In Vitro Pharmacology of Aripiprazole, its Metabolite and Experimental Dopamine Partial Agonists at Human Dopamine D2 and D3 Receptors*, 668 EUROPEAN J. PHARMACOLOGY 355, 357 (2011) ("Tadori 2011b") (exhibiting 54.9% and 43.9%

antagonistic activities at postsynaptic D_2 receptors.[105] As the Court has already observed, **[*** REDACTED ***]**.[106] Thus, the medical science and record in this case reliably support Plaintiffs' experts' opinions that Abilify acts with greater intrinsic activity at postsynaptic D_3 receptors than at postsynaptic D_2 receptors.

The fact that Plaintiffs' experts rely primarily on *in vitro* data does not invalidate their conclusions as to this proposed mechanism of action. *In vitro* studies are often "the only or best available evidence" of a drug's effects at the cellular level. *See* Ref. Man. at 564. However, because of its limitations (*e.g.*, ethical concerns, problems extrapolating from laboratory experimental findings to humans), *in vitro* evidence cannot be the sole basis for a general causation opinion. *See Kilpatrick*, 613 F.3d at 1340-44. It only may be used to supplement other, more substantial evidence of general causation, provided the expert explains how the *in vitro* data can be reliably extrapolated to predict the drug's effects in humans. *See id.*

---

intrinsic activity at human D_3 receptors), DX-061 at 3; Tadori 2011a at 48 (exhibiting 50.5% and 49.9% intrinsic activity at human D_3 receptors), PX-021 at 6; Tadori 2008 at 30 (exhibiting 51.2% and 47.8% intrinsic activity at human D_3 receptors), DX-058 at 4; Liesbeth A. Bruins Slot *et al.*, Action of Novel Antipsychotics at Human Dopamine D3 Receptors Coupled to G Protein and ERK1/2 Activation, 53 Neuropharmacology 232, 235 (2007) ("Bruins Slot 2007") (exhibiting 55% intrinsic activity at human D_3 receptors); Shapiro 2003 at 1401, 1408 (exhibiting both partial and full agonist actions at D_3 receptors), DX-045 at 2, 9.

[105] Tadori 2011a at 48 (exhibiting no intrinsic activity at human postsynaptic D_2 receptors), PX-021 at 6; Yoshihiro Tadori *et al.*, *Differences in Agonist/Antagonist Properties at Human Dopamine D2 Receptors Between Aripiprazole, Bifeprunox and SDZ 208-912*, 574 EUROPEAN J. PHARMACOLOGY 103, 105 (2007) ("Tadori 2007") (exhibiting -2.8% intrinsic activity at human postsynaptic D_2 receptors), DX-056 at 3.

[106] **[*** REDACTED ***]**.

In this case, Dr. Bechara acknowledged the limitations of *in vitro* research and explained which of the *in vitro* findings about Abilify reliably "reflect what happens in a normal human brain." *See* Bechara Tr., ECF No. 596-3 at 91. This is all that Rule 702 and *Daubert* require. Moreover, Defendants have not contradicted Dr. Bechara's explanation, nor have they argued that any specific *in vitro* study he presented is methodologically flawed. Finally, Defendants **[*** REDACTED ***]**.[107] In any event, here, the *in vitro* evidence supporting biological plausibility is not the sole basis for Plaintiffs' experts' general causation opinions. To the contrary, the *in vitro* data is offered in support of more "powerful" evidence of general causation, namely, the Etminan Study. *See Rider*, 295 F.3d at 1198. Thus, the Court finds that Plaintiffs' *in vitro* evidence of biological plausibility may supplement the more substantial evidence of general causation in this case (*i.e.*, the Etminan Study). *See Kilpatrick*, 613 F.3d at 1340-44; Ref. Man. at 604.

Defendants argue that Dr. Bechara "cherry picked" two data points that support his biological plausibility opinion and "ignore[d] the numerous studies that are inconsistent with" it. Def. Bechara Motion, ECF No. 423-10 at 18-19. The first data point was reported in Shapiro 2003, which stated that Abilify exhibited both partial and full agonist actions at D_3 receptors. *See* Shapiro 2003, DX-045 at 2, 9. The second data point was drawn from Hamamura 2008, which calculated Abilify's

[107] **[*** REDACTED ***]**.

intrinsic activity at D_2 receptors to be approximately 6%. *See* Hamamura 2008 at 864, DX-057 at 3. Defendants claim Dr. Bechara deliberately presented only these two data points to the Court, and no others, in order to accentuate the alleged disparity between Abilify's effects at D_2 and D_3 receptors. The Court is not persuaded. The uncontroverted evidence—including Dr. Bechara's expert reports and his testimony at the *Daubert* hearing—shows that Dr. Bechara performed an extensive and systematic review of the scientific literature on Abilify's intrinsic activity at D_2 and D_3 receptors. *See, e.g.*, Bechara Rep., ECF No. 423-1 at 21-25; Bechara Supp., ECF No. 423-1 at 448-454; Bechara Tr., ECF No. 596-3 at 82-83.[108] Dr. Bechara testified that he reported the Shapiro 2003 and Hamamura 2008 data points, as well as the other data points he found, for purposes of completeness, because they are part of the body of scientific evidence on this issue. *See* Bechara Tr., ECF No. 596-3 at 86, 120. The fact that Dr. Bechara ultimately concluded that Abilify is a partial agonist at D_3 receptors is evidence that he did not put undue weight on Shapiro 2003's description of Abilify as a full agonist at D_3 receptors. Instead, his partial agonism conclusion is based on "the majority of the articles" in the scientific literature, which report Abilify's partial agonism at D_3 receptors to be in the 50% range. *See id.* at 86. This was a reliable approach to analyzing and

[108] "Bechara Supp." refers to Dr. Antoine Bechara's Supplemental Report, ECF No. 423-1 at 442-54.

reporting the scientific evidence on Abilify's activity at D_3 receptors. Defendants' arguments with respect to Dr. Bechara's inclusion of the Shapiro 2003 and Hamamura 2008 data points go to the weight of Dr. Bechara's biological plausibility opinion, not its admissibility.

Finally, Defendants' argument that Dr. Bechara "ignore[d] the numerous studies that are inconsistent with his opinion," *see* Def. Bechara Motion, ECF No. 423-10 at 18-19, fails because it misrepresents the findings of the "numerous" *in vitro* studies that Defendants claim Dr. Bechara ignored. While it is true that there are studies in the medical literature evidencing Abilify's partial agonism at D_2 receptors, at least six of those studies report findings related to Abilify's effects at *presynaptic* D_2 receptors.[109] *See id*. at 19. In that respect, those six studies do not "logically advance" the resolution of questions about Abilify's effects at

[109] *See* Cindy P. Lawler *et al*., *Interactions of the Novel Antipsychotic Aripiprazole (OPC-14597) with Dopamine and Serotonin Receptor Subtypes*, 20 NEUROPSYCHOPHARMACOLOGY 612, 613, 622 (1999), DX-051 at 7 (stating that D_{2S} receptors are preferentially expressed presynaptically); *see also* Tadori 2011b (showing 26.3% intrinsic activity at presynaptic D_{2S} receptors), DX-061 at 3; Tadori 2011a (showing 25.5% intrinsic activity at presynaptic D_{2S} receptors), PX-021 at 6; Yoshihiro Tadori, *et al.*, *Receptor Reserve-Dependent Properties of Antipsychotics at Human Dopamine D2 Receptors*, 607 EUROPEAN J. PHARM. 35, 37 (2009) ("Tadori 2009") (exerting 31.9% intrinsic activity at presynaptic D_{2S} receptors), DX-059 at 3; Tadori 2007 (exerting 17.2% intrinsic activity at presynaptic D_{2S} receptors), DX-056 at 3; Lisbeth A. Bruins Slot, *et al*., *Differential Profile of Antipsychotics at Serotonin 5-HT$_{1A}$ and Dopamine D2S Receptors Coupled to Extracellular Signal-Regulated Kinase*, 534 EUROPEAN J. PHARM. 534, 66 (2006) (exerting 58.7% intrinsic activity at presynaptic D_{2S} receptors), DX-054 at 4; Yoshihiro Tadori, *et al*., *Aripiprazole's Low Intrinsic Activities at Human Dopamine D2L and D2S Receptors Render It a Unique Antipsychotic*, 515 EUROPEAN J. PHARM. 10, 14 (2005) (exerting 20% intrinsic activity at presynaptic D_{2S} receptors) ("Tadori 2005"), DX-053 at 5.

postsynaptic D_2 and D_3 receptors. It was not unreliable for Dr. Bechara to discount those six studies in reaching his opinion.

Moreover, according to Dr. Bechara, five studies, including all but one of the studies discussed above involving presynaptic D_2 receptors, also report *in vitro* findings based on tissue cultures that were "artificially manipulated" by researchers to increase the number of D_2 receptors, called a receptor reserve, far beyond the amount present in the human brain.[110] *See* Bechara Tr., ECF No. 596-3 at 90-92; Bechara Dep., ECF No. 423-1 at 290-96.[111] Dr. Bechara testified that such artificially manipulated tissues "do not reflect what actually happens in a normal brain," therefore the findings about a drug's effects on those tissues cannot be reliably extrapolated to predict the drug's effect in humans. *See* Bechara Tr., 596-3 at 90-92; Bechara Dep., ECF No. 423-1 at 290-96. Defendants essentially concede this point, responding only that "most" of the *in vitro* studies of postsynaptic D_3 receptors also involve artificially high receptor density. This response is

[110] Tadori 2011b (exerting 90.6% intrinsic activity at artificially high-density postsynaptic D_{2L} receptors and 95.9% intrinsic activity at artificially high-density presynaptic D_{2S} receptors), DX-061 at 3; Tadori 2011a (exerting 96.8% intrinsic activity at artificially high-density presynaptic D_{2S} receptors), PX-021 at 6; Tadori 2009 (exerting 95.5% intrinsic activity at artificially high-density postsynaptic D_{2S} receptors); Tadori 2007 (exerting 86% intrinsic activity at artificially high-density postsynaptic D_{2L} receptors), DX-056 at 3; Tadori 2005 (exerting 76% intrinsic activity at postsynaptic D_{2L} artificially high-density postsynaptic D_{2L} receptors), DX-053 at 5.

[111] "Bechara Dep." refers to the official transcript of Dr. Antoine Bechara's deposition testimony on June 7, 2017, ECF No. 423-1 at 28.

unpersuasive. None of Defendants' experts criticized the *in vitro* data involving D_3 receptors on this basis. *See Marmo v. Tyson Fresh Meats, Inc*., 457 F.3d 748, 759 (8th Cir. 2006) ("The function of rebuttal testimony is to explain, repel, counteract or disprove evidence of the adverse party."). Moreover, Defendants do not identify which of the *in vitro* studies of D_3 receptors is allegedly tainted by artificial manipulation. On this record, the Court finds that it was entirely reliable for Dr. Bechara to rely on *in vitro* studies involving D_3 receptors and to exclude from consideration the studies involving artificially high D_2 receptor reserves.

There are two other scientific articles containing data points that Defendants claim Dr. Bechara ignored. The first, Lawler 1999, DX-051, Dr. Bechara clearly cites in his expert report, *see* Bechara Rep., ECF No. 423-1 at 12. Moreover, at the *Daubert* hearing, Dr. Bechara referenced the Lawler article as part of the basis for his opinion that Abilify exerts very low partial agonist to antagonistic action at postsynaptic D_2 receptors. *See* Bechara Tr., ECF No. 596-3 at 92. The second article, Maeda 2014, DX-062, described an *in vitro* finding that Abilify exhibited 61% intrinsic activity at postsynaptic D_2 receptors. At the *Daubert* hearing, Dr. Bechara acknowledged that he did not rely on or cite the Maeda article as part of his opinion in this case. *See* Bechara Tr., ECF No. 596-3 at 126. However, there is no evidence that Dr. Bechara knew of and willfully excluded the Maeda article from his analysis. There also is no evidence that his search of the scientific literature was

in any other way infirm. Under these circumstances, the Court finds that Dr. Bechara's failure to cite a single article out of the vast body of scientific literature connected with this case cannot render his entire analysis and opinion unreliable. This issue may be fodder for vigorous cross-examination, but it is not grounds for exclusion of Dr. Bechara's testimony.

In sum, the Court finds there is a methodologically sound basis for Dr. Bechara's conclusion that Abilify acts with greater intrinsic activity at D_3 receptors than it does at postsynaptic D_2 receptors. Dr. Bechara did not simply ignore the medical science that did not support his opinion. Instead, he analyzed all of the available medical literature and explained how and why certain studies did not alter or undermine his opinion regarding Abilify's functional antagonism at postsynaptic D_2 receptors and partial agonism at D_3 receptors. Defendants' criticisms go to the credibility, and thus the weight, of Dr. Bechara's opinion, not its admissibility.

d. Negative Reward Prediction Error

Defendants also argue there is insufficient evidentiary support for Dr. Bechara's opinion that Abilify impairs negative reward prediction error, also called reversal learning. On this issue, the Court agrees. Reward prediction error learning refers to the process by which the brain learns from associations between actions and consequences. The mesolimbic dopamine system plays a central role in such reward-motivated behavior. *See* Roy A. Wise, *Brain Reward Circuitry: Insights*

from Unsensed Incentives, 36 NEURON 229, 234 (2002) ("Wise 2002").[112] In brief, mesolimbic dopamine neurons have been shown to fire in two modes, tonic and phasic, which are thought to modulate two distinct aspects of human behavior. The tonic firing mode involves a slow and steady release of dopamine that maintains the baseline levels necessary for proper brain function. Tonically fired dopamine binds with and activates autoreceptors on presynaptic neurons, which regulate dopamine synthesis and release; it cannot activate or produce a physiological response in postsynaptic receptors. In phasic mode, dopamine neurons sharply increase or decrease their firing rate based on events or stimuli with motivational significance. For example, where an action yields more reward than predicted (positive prediction error), the neurons fire in short, high frequency bursts, rapidly increasing the concentration of synaptic dopamine. The release of dopamine and accompanying positive feelings motivate repetition of the rewarding activity. In contrast, where a reward is worse than predicted (negative prediction error), phasic firing activity drops, sharply decreasing synaptic dopamine concentrations. This "dip" in dopamine levels and the accompanying negative feelings motivate avoidance of the aversive activity. In either scenario, learning occurs and behavior changes.

[112] *See also* Kent C. Berridge & Terry E. Robinson, *What is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience?*, 28 BRAIN RESEARCH REVIEWS 309 (1998).

Dr. Bechara opines that Abilify may cause pathological gambling and impulse control disorders by preventing the dopamine "dip" that is critical for "teach[ing]" an individual to avoid activities with negative consequences. *See* Bechara Rep., ECF No. 423-1 at 11. According to Dr. Bechara, Abilify's functional antagonism at 90% of D_2 receptors signals a drop in baseline dopamine concentration, which triggers increased tonic transmission of dopamine. The excess tonically fired dopamine accumulates in the synapses and offsets, or blocks the effects of, the phasic dopamine dips associated with negative prediction errors. Without the phasic dips, no negative dopaminergic signal is sent when an activity should be stopped and avoided. The behavioral effect is an increase in risky, reward-directed activities, such as pathological gambling and other impulse control disorders, despite the potential for and occurrence of negative consequences.

As support for his negative reward prediction error opinion, Dr. Bechara relies heavily on the findings of a 2006 study investigating reward and punishment processing in Parkinson's disease patients taking one or more dopamine replacement medications, none of which was Abilify.[113] *See* Roshan Cools *et al.*, *Reversal Learning in Parkinson's Disease Depends on Medication Status and Outcome Valence*, 44 NEUROPSYCHOLOGIA 1663 (2006) ("Cools Study"), DX-143. The Cools

[113] The dopaminergic medications included sinemet (*i.e.*, L-DOPA), pramipexole (*i.e.*, Mirapex), pergolide, amantadine, comtan, methylphenidate, and modafinil.

Study found that patients on dopamine replacement medications exhibited "significantly impaired" capacity to process and learn from unexpected negative outcomes (*i.e.*, punishment) relative to healthy controls. *See id*. at 1670, DX-143 at 8. Consistent with theoretical models proposed in earlier medical literature, the Cools Study attributed the patients' negative reward prediction errors to artificially high tonic dopamine levels induced by the dopamine replacement medications, which the authors hypothesized as having functionally eliminated the effectiveness of phasic dopamine dips. *See id*. at 1669, DX-143 at 7.

The Court finds that the Cools Study cannot reliably establish biological plausibility in this case because it involved dopamine replacement medications that either directly increase dopamine levels in the brain (*e.g.*, Levodopa) or directly stimulate dopamine receptors, including D_2 receptors, by mimicking the activity of endogenous dopamine (*e.g.*, full agonists like Mirapex). None of the drugs in the Cools Study were functional antagonists at D_2 receptors like Abilify, blocking dopamine neurotransmission at those sites. Importantly, none of Plaintiffs' experts explained why the Cools Study's findings as to the behavioral effects of stimulating D_2 receptors through dopamine replacement therapy can be extrapolated to reliably predict the behavioral effects of a D_2 receptor antagonist. *See McClain*, 401 F.3d at 1246 (extrapolation between drugs permissible only with reliable evidence establishing validity of the analogy). The Court recognizes that both categories of

drugs appear to enhance tonic transmission of dopamine to some degree. But the extent to which Abilify actually does so—that is, whether its antagonism at D_2 receptors increases tonic dopamine concentrations to a level that could trigger negative reward prediction errors—is far from clear based on the record as it currently stands. At most, the record reflects that, with Abilify, phasic dopaminergic transmission (*i.e.*, bursts and dips associated with motivational stimuli) is "relatively more suppressed than" tonic transmission. *See* Hamamura 2007, PX-020 at 2. Without more, this evidence cannot bridge the gap between D_2 antagonism and the "excessive [tonic dopamine] levels" thought to have been produced by the dopamine replacement therapies in the Cools Study. *See* Cools 2006 at 1669, DX-143 at 7. Therefore, the Cools Study must be excluded.

Additionally, a 2015 study of reversal learning in individuals taking the D_2 receptor antagonist Sulpiride, which was cited by Dr. Hollander, also presents an extrapolation problem, although it is a closer call. *See* L. Janssen *et al.*, *Abnormal Modulation of Reward Versus Punishment Learning by a Dopamine D2-Receptor Antagonist in Pathological Gamblers*, 232 PSYCHOPHARMACOLOGY 3345 (2015) ("Janssen Study"), DX-189. The Janssen Study found that administration of the drug Sulpiride impaired reversal learning in healthy controls, but did not appear to alter reversal learning in pathological gamblers. The Study's authors also described the "seemingly paradoxical" state of the scientific literature with respect to a

relationship between D_2 receptor antagonism and reward prediction error learning: some studies report that D_2 receptor antagonists impaired reward prediction error learning, while other studies report that they improved it. *See id*. at 3350-51, DX-189 at 6-7. None of the studies investigating D_2 receptor antagonism and reward prediction errors involved Abilify. Moreover, none of Plaintiffs' experts even attempted to explain how or why Abilify is sufficiently similar to Sulpiride in its mechanism of action to warrant an extrapolation. This is significant because within a given class of drugs—such as D_2 receptor antagonists—there may be "great chemical diversity" and those "minor deviations in chemical structure can radically change a particular substance's properties and propensities." *See Rider*, 295 F.3d at 1201 (quoting *Glastetter*, 252 F.3d at 990). Without evidence establishing the validity of the analogy, an extrapolation between Abilify and Sulpiride is impermissible and the Janssen Study must be excluded. *See McClain*, 401 F.3d at 1246. Accordingly, because the scientific literature offered in support of Plaintiffs' experts' negative reward prediction error opinion is inadmissible under *Daubert*, Defendants' motions to exclude this aspect of their biological plausibility opinions is due to be granted.

e. Conclusion

In sum, the Court finds Plaintiffs' experts' biological plausibility opinions that Abilify can cause impulse control problems through its effects on dopamine

neurotransmission in the brain to be scientifically reliable, based on current biochemistry and pharmacological knowledge. Each element of this proposed mechanism of action is adequately supported by peer-reviewed, published scientific literature and sound scientific reasoning. Moreover, Plaintiffs' experts' opinions are consistent with the FDA's assessment, based on the scientific literature, that Abilify's partial agonism, upregulation, and sensitization "could theoretically stimulate dopamine transmission in the mesolimbic pathway, a core component of the brain reward circuitry, providing biological plausibility for treatment-emergent [impulse control disorders]." *See* FDA Pharm. Vigil., ECF No. 428-11 at 4, 29. The opinions are also consistent with **[*** REDACTED ***]**. Finally, Plaintiffs' experts' proposed mechanism of action is consistent with **[*** REDACTED ***]**.[114] **[*** REDACTED ***]**. Although this biological plausibility evidence, standing alone, cannot establish general causation, it may "lend[] credence to an inference of causality" drawn from other, more substantial evidence. *See* Ref. Man. at 604; *see also Chapman*, 766 F.3d at 1308; *Rider*, 295 F.3d at 1202; *Milward*, 639 F.3d at 25-26. Again, the relevant standard is "biological plausibility," not

[114] **[*** REDACTED ***]**. Defendants sought to exclude this internal communication, and others like it, because, in their view, while it may implicate issues of notice, it is not relevant to general causation. Plaintiffs disagreed, arguing that these materials do bear on general causation because many of them are authored by Dr. McQuade, who essentially led the research team that developed Abilify and whose observations about the drug thus are an integral, reliable part of the body of scientific evidence on its mechanism of action. The Court finds it unnecessary to consider Defendants' internal communications, including emails and meeting minutes, for purposes of this Order; therefore, Defendants' challenges to the admissibility of these materials are moot.

"biological certainty," *Jones*, 861 F.2d at 662 (stating that "absolute certainty is not required" from expert testimony), and the Court finds Plaintiffs' experts' proposed mechanism of action to be biologically plausible. Defendants' objections to the experts' biological plausibility opinions are granted with respect to negative reward prediction error, but denied in all other respects.

5. Case Studies and Adverse Event Reports

Since 2010, there have been hundreds of reports of gambling and/or impulse control disorders in patients treated with Abilify. A number of the reports are published case studies that contain details about dosage, duration of use, concomitant medications, comorbid conditions, and other pertinent clinical information.[115] Many more involve adverse event reports submitted to Defendants or the FDA, with varying levels of narrative detail about the patient and the relevant medical circumstances.[116] Defendants argue, broadly, that case studies and adverse

[115] *See, e.g.*, L. Gaboriau *et al.*, *Aripiprazole: A New Risk Factor for Pathological Gambling? A Report of 8 Case Reports*, 39 ADDICTIVE BEHAVIORS 562 (2014) (eight case reports), ECF No. 428-8; EunJin Cheon *et al.*, *Two Cases of Hypersexuality Probably Associated with Aripiprazole*, 10 PSYCHIATRY INVESTIGATION 200 (2013) (two case reports), PX-07; Neil Smith *et al.*, *Pathological Gambling and the Treatment of Psychosis with Aripiprazole: Case Reports*, 199 BRIT. J. PSYCHIATRY 158 (2011) (three case reports), ECF No. 428-7; Julien Cohen *et al.*, *Aripiprazole-Induced Pathological Gambling: A Report of 3 Cases*, 6 CURRENT DRUG SAFETY 51 (2011) (three case reports), ECF No. 428-6; Giles Gavaudan *et al.*, *Partial Agonist Therapy in Schizophrenia: Relevance to Diminished Criminal Responsibility*, 55 J. FORENSIC SCI. 1659 (2010) (two case reports), ECF No. 428-5; Milton G. Roxanas, *Pathological Gambling and Compulsive Eating Associated with Aripiprazole*, 44 AUSTL. & N.Z. J. PSYCHIATRY 291 (2010) (one case report), ECF No. 428-4.

[116] *See, e.g.*, Otsuka DA, PX-44 (discussing 236 post marketing spontaneous reports made to Bristol-Myers Squibb regarding pathological gambling in patients using Abilify); FDA Pharm.

event reports cannot reliably support a general causation opinion in the Eleventh Circuit. This is misplaced.

Although it is true that case studies and other anecdotal evidence may not, standing alone, support a general causation opinion, *Rider*, 295 F.3d at 1199, these materials may reliably bolster other, more substantial evidence of general causation. *See id*. In this case, none of Plaintiffs' experts relied solely on case studies and adverse event reports in forming his general causation opinion. Instead, each used those materials to supplement his review of the epidemiological evidence (*i.e*., the Etminan Study), medical literature evidencing a plausible biological mechanism of action, and clinical trial data. This was an entirely reliable use of case studies and adverse event reports under Eleventh Circuit precedent, particularly given the substantial volume of case studies and adverse event reports associating Abilify with pathological gambling and impulse control disorders. *See Rider*, 295 F.3d at 1199, 1202 (stating in dicta that reliable evidence on causality includes, *inter alia*, "a very large number of case reports"). As with Defendants' other evidentiary objections, concerns about Plaintiffs' experts' reliance on case studies and adverse event reports affect only the weight to be given to their general causation opinions, not the admissibility.

Vigil., ECF No. 428-11 (discussing 167 FAERS case reports of impulse control related diagnoses, including pathological gambling, with Abilify use).

6. Disproportionality Analyses

Several statistical analyses, called disproportionality analyses, have been conducted by the FDA and Defendant Otsuka on the FAERS and VigiBase adverse event reporting databases, comparing the relative frequency of pathological gambling reports among various patient populations.[117] Disproportionality analysis is an industry standard pharmacovigilance technique used to detect and evaluate safety signals—that is, the existence of an excess of reported adverse medical events—associated with the use of FDA-regulated medical products. *See* FDA Pharm. Guide, DX-299 at 7, 11; *see also Fosamax*, 2013 WL 1558690, at *8. The statistic generated by a disproportionality analysis, called an EB05, quantifies the strength of an association between a drug and reports of a particular adverse effect. *See* FDA Pharm. Guide, DX-299 at 11. An EB05 of 1.0 means there is no association between the drug and reports of the adverse effect. An EB05 greater than 1.0 is indicative of elevated adverse event reporting with the subject drug relative to the reporting rate with comparator drugs. An EB05 greater than 2.0 is the "widely-accepted threshold" indicator of a statistically significant association between adverse event reports and a drug, signaling potential safety issues that

[117] As discussed more fully in Section II(D)(1)(e), Dr. Madigan also performed disproportionality analyses of the FAERS database as of four different dates he considered significant in the life of this case. All four of his disproportionality analyses reflected statistically significant percentages of pathological gambling reports, which is consistent with the findings of the disproportionality analyses performed by both Otsuka and the FDA. *See* Section II(D)(1)(e).

require further investigation. *See* Madigan Tr., 596-4 at 20; *see also* FDA Pharm Vigil., ECF No. 428-11 at 28. In layman's terms, an EB05 of 2.0 means that a particular adverse effect has been reported twice as often with the subject drug, as compared to the background reporting rate of the adverse effect. *See* Madigan Tr., 596-4 at 20.

As the Court already discussed, the FDA's disproportionality analysis of the FAERS database found a statistically significant, disproportionately higher proportion of patients reporting pathological gambling with Abilify relative to all other atypical antipsychotics. *See* FDA Pharm. Vigil., ECF No. 428-11 at 28. The FDA calculated an EB05 score of 6.304 for this finding, which represents a statistically significant result. Defendant Otsuka **[*** REDACTED ***]**.

Defendants argue, again broadly, that disproportionality analyses cannot establish causation. The Court agrees. The safety signals identified through disproportionality analyses in this case "do not, by themselves, demonstrate [a] causal association." *See* FDA Pharm. Vigil., ECF No. 428-11 at 28. However, they do reliably support a conclusion that pathological gambling is disproportionately reported with Abilify relative to all other antipsychotics. *See id.* The disproportionality analyses, in combination with the adverse event reports they represent, also provide evidence of the "frequency, character, or severity" of pathological gambling reports in the respective adverse event reporting databases.

See Fosamax, 2013 WL 1558690, at *8; *see also* FDA Pharm. Guide, DX-299 at 11 (data mining "is especially useful for assessing patterns" and "time trends"). In short, the disproportionality analyses provide statistical context for the substantial number of adverse event reports in the databases. *See Kilpatrick*, 613 F.3d at 1338 (noting that anecdotal evidence is less reliable because it lacks statistical context); *Rider*, 295 F.3d at 1199, 1202 (stating *in dicta* that reliable evidence of causation includes, *inter alia*, "a very large number of case reports"). While such evidence may not, standing alone, support a general causation opinion, it may reliably supplement other, more substantive evidence of general causation, *see Rider*, 295 F.3d at 1199, which is exactly how it was used by Plaintiffs' experts in this case. Defendants' challenge to the disproportionality analyses affect only the weight of these materials, not their admissibility.

D. Expert-Specific Challenges

1. Plaintiffs' Experts

a. Antoine Bechara, Ph.D.

Dr. Antoine Bechara is a professor of neuroscience and psychology at the University of Southern California with extensive professional experience in neurobiology and, in particular, the anatomical and neurotransmitter systems involved in human decision-making, behavioral addictions, and gambling. ECF No. 423-1 at 2. He is offered primarily for the purpose of explaining how Abilify can

cause Plaintiffs' impulse control disorders. More specifically, Dr. Bechara has offered his opinion that there is a biologically plausible mechanism by which Abilify causes impulsive gambling and other impulse control behaviors, namely, its effect on dopamine neurotransmission in the brain. Defendants challenge Dr. Bechara's testimony on qualification and reliability grounds. The Court has already found that the science on which Dr. Bechara based his biological plausibility opinion is reliable. *See supra* Section II(C). Therefore, the only question that remains to be resolved with respect to Dr. Bechara is whether he is qualified to render an expert opinion in this case.

Defendants argue that Dr. Bechara is not qualified to offer opinions on biological plausibility or general causation because he does not have a medical degree or a degree in pharmacology, has never diagnosed or treated patients with impulse control disorders, and has conducted no independent studies into how Abilify affects brain chemistry. As to biological plausibility, the Court disagrees. First, the record evidence reflects that Dr. Bechara does, in fact, have "a university degree in pharmacology from the University of Toronto," as well as a Ph.D. in neuroscience. ECF No. 423-1 at 80-82. More importantly, Dr. Bechara has over 25 years of clinical experience studying, publishing, and teaching courses on brain function and the effects of drugs—both prescription and street drugs—on human behavior. Of particular relevance to this case is the extensive research Dr. Bechara

has conducted on dopamine systems and the neurobiological mechanisms of human decision-making, substance use and abuse, and behavioral and psychiatric disorders, including impulsive gambling and other impulse control disorders. He has written and collaborated on hundreds of peer-reviewed articles, papers, and book chapters on these subjects. Finally, Dr. Bechara developed the Iowa Gambling Task, which is currently used worldwide to detect and measure brain dysfunction and decision-making deficits in numerous clinical populations. As Dr. Bechara has not been offered as a medical doctor, it is irrelevant that he lacks a medical degree. Given the breadth of Dr. Bechara's knowledge and clinical experience in the field of neurobiology, the fact that he had not studied Abilify until he became involved in this case does not disqualify him from reviewing the scientific literature and offering an expert opinion on Abilify's mechanism of action.[118]

As to a more comprehensive general causation opinion—that is, an opinion beyond the neurobiological mechanisms by which Abilify can cause pathological gambling or impulse control disorders—the Court agrees with Defendants. Although Dr. Bechara's expert reports frequently frame his conclusions in very broad and definitive language (*i.e.*, explaining that Abilify "causes" impulsive

[118] *Trilink Saw Chain, LLC v. Blount, Inc.*, 583 F. Supp. 2d 1293, 1304 (N.D. Ga. 2008) ("[A]n expert with the education or background to permit him to analyze a given set of circumstances . . . can through reading, calculations, and reasoning from known scientific principles make himself very much an expert [regarding a] particular product even though he has not had actual experience with the product.").

behaviors, rather than how it "can cause" the behaviors), most of the scientific support for his positions relates only to biological plausibility. His opinion does not meaningfully depend on any of the three categories of primary evidence considered indispensable in the Eleventh Circuit. He did not perform a Bradford Hill or weight-of-the-evidence analysis of general causation.[119] These facts do not disqualify him from testifying as an expert regarding Abilify's biological mechanism of action, but they do preclude him from offering a comprehensive general causation opinion. Accordingly, Defendants' Motion to Exclude the General Causation Opinion of Antoine Bechara, ECF No. 423, is granted with respect to medical causation and denied with respect to biological plausibility.

b. Joseph Glenmullen, M.D.

Joseph Glenmullen, M.D. is a board-certified psychiatrist and lecturer in psychiatry at Harvard Medical School, with more than 30 years of clinical experience treating psychiatric patients in private practice. Dr. Glenmullen offers a general causation opinion that Abilify is capable of causing pathological gambling and impulse control disorders. Dr. Glenmullen supports his opinion with epidemiological evidence (*i.e.*, the Etminan Study), medical literature evidencing a plausible mechanism of action, case and adverse event reports, disproportionality analyses and clinical trial data, all of which he analyzed under the Bradford Hill

[119] *See supra* Section II(B).

factors.[120] Defendants challenge Dr. Glenmullen's testimony on qualification and reliability grounds.

i. Qualification

Defendants argue that Dr. Glenmullen lacks the requisite expertise to offer opinions related to general causation, such as biological plausibility, epidemiology, toxicology, biostatistics, FDA regulations, and pathological gambling. In essence, Defendants contend that Dr. Glenmullen's medical education, post-graduate training, and professional experience in the field of psychiatry do not translate into qualifications that enable him to testify competently based on the scientific evidence in this case. Defendants read the "qualification" prong of Rule 702 too stringently. "An expert is not necessarily unqualified simply because [his] experience does not precisely match the matter at hand." *Furmanite Am., Inc. v. T.D. Williamson, Inc.*, 506 F. Supp. 2d 1126, 1129 (M.D. Fla. 2007) (citing *Maiz v. Virani*, 253 F.3d 641, 665 (11h Cir. 2001)).[121] Again, "so long as the [expert] is minimally qualified, objections to the level of [his] expertise go to credibility and weight, not

[120] *See supra* n.19. Dr. Glenmullen also offers what he characterizes as dose-response evidence, but the Court has already found this evidence insufficient to establish the existence of a dose-response relationship. *See supra* Section II(C).

[121] *See also Kipperman v. Onex Corp.*, 411 B.R. 805, 843 (N.D. Ga. 2009) ("[A]n expert's training does not always need to be narrowly tailored to match the exact point of dispute in a case."); *Trilink*, 583 F. Supp. 2d at 1304 ("[A]n expert with the education or background to permit him to analyze a given set of circumstances . . . can through reading, calculations, and reasoning from known scientific principles make himself very much an expert [regarding a] particular product even though he has not had actual experience with the product.").

admissibility." *Hendrix I*, 255 F.R.D. at 585. The critical question for qualification purposes is whether the proffered expert has such "knowledge, skill, experience, training, or education" that his opinion will aid the trier of fact in understanding the evidence or resolving a factual issue. *See* Fed. R. Evid. 702.

In this case, the Court finds Dr. Glenmullen at least minimally qualified to offer expert opinions that will assist the trier of fact in understanding and resolving general causation issues related to biological plausibility, epidemiology, toxicology, and pathological gambling. Again, Dr. Glenmullen is a medical doctor and board-certified psychiatrist who has spent most of his career training psychiatric residents at Harvard Medical School and treating psychiatric patients in private practice, including "plenty of" patients with pathological gambling or impulse control disorders, albeit none with a diagnosis of drug-induced pathological gambling.[122] *See* Glenmullen Tr., ECF No. 596-2 at 135-36. Notably, he has authored two books on the side effects of psychiatric medications and co-authored five peer-reviewed published studies related to the neuropsychopharmacology of psychiatric or dopamine agonist medications. *See* Glenmullen Curriculum Vitae, ECF No. 457-1 at 140-41; *see also* Glenmullen Tr., ECF No. 596-2 at 13. In recent years, much of his time has been devoted to forensic or expert consulting work on legal cases

[122] At the *Daubert* hearing, Dr. Glenmullen testified that he has never diagnosed a patient with drug-induced pathological gambling. *See* Glenmullen Tr., ECF No. 596-2 at 136.

involving adverse side effects of psychiatric medications. Indeed, Dr. Glenmullen has been qualified as an expert on general causation by numerous federal courts in pharmaceutical products liability cases.[123]

The fact that Dr. Glenmullen is not an epidemiologist does not disqualify him from testifying about epidemiological studies. *See United States v. Thorn*, 317 F.3d 107, 114-15 (2d Cir. 2003) (medical doctor specializing in asbestos-related disease permitted to testify about various epidemiological studies of asbestos exposure); *DeLuca v. Merrell Dow Pharm*., 911 F.2d 941, 953 (3d Cir. 1990) (pharmacologist qualified to testify about his interpretation of epidemiological evidence), *abrogated on other grounds by Paoli*, 35 F.3d at 748; *In re Mirena IUD Products Liability Litigation*, 169 F. Supp. 3d 396, 426 (S.D.N.Y. 2016) (stating that "medical doctors do not need to be epidemiologists in order to testify regarding epidemiological studies"). He has formal training in epidemiology and has practical experience evaluating epidemiological evidence as part of his research and forensic consulting

[123] *See, e.g., Chantix,* 889 F. Supp. 2d 1272 (N.D. Ala. 2012) (rejecting defense arguments that Dr. Glenmullen was unqualified to offer opinions on general causation and biological plausibility because he is not an epidemiologist, is not formally trained in pharmacovigilance, and has no degree in pharmacology, chemistry, or neuroscience); *Tucker v. SmithKline Beecham Corp*., 701 F. Supp. 2d 1040, 1062-63 (S.D. Ind. 2010) (no challenge to Dr. Glenmullen's qualifications to offer general causation opinion that Paxil can cause suicidality); *Giles v. Wyeth, Inc*., 500 F. Supp. 2d 1048 (S.D. Ill. 2007) (noting absence of dispute that Dr. Glenmullen was qualified to offer general causation opinion that Effexor can cause suicidality); *Laisure-Radke v. Par Pharm., Inc*., No. 2:03-cv-03654, 2006 WL 829102, at *1-3 (W.D. Wash. 2006) (finding Dr. Glenmullen qualified to offer general causation opinion that selective serotonin reuptake inhibitors can cause suicidality).

work. His academic background and professional experience with epidemiology, though limited, is sufficient to support his proposed testimony in this case. A witness need not be the best or most qualified authority in a field to be admitted as an expert. *See, e.g.*, *Burgett v. Troy-Bilt LLC*, 579 F. App'x 372, 378 (6th Cir. 2014) ("[I]t is an abuse of discretion to exclude testimony simply because the trial court does not deem the proposed expert to be the best qualified or because the proposed expert does not have the specialization that the court considers most appropriate.") quoting *Pineda v. Ford Motor Co*., 520 F.3d 237, 244 (3d Cir. 2008); *Robinson v. GEICO Ins. Co*., 447 F.3d 1096, 1101 (8th Cir. 2016) (same); *Bracey v. Jolley*, No. 1:10-cv-4064, 2012 WL 12870257, at *3 (N.D. Ga. 2012) ("Rule 702 does not require a party to produce the 'most qualified' expert."). Dr. Glenmullen need only possess enough general knowledge of epidemiology that his testimony would likely assist the trier of fact. *See, e.g.*, *Maiz*, 253 F.3d at 665 (economist was properly qualified to estimate damages resulting from real estate investment scheme even though he had no experience in real estate development). Considering Dr. Glenmullen's extensive experience in the field of psychiatry, his knowledge of the relevant scientific principles and methods, and the liberal standard for admission of expert testimony under Rule 702, *see Frazier*, 387 F.3d at 1294, the Court finds him qualified to offer his general causation opinion (encompassing biological plausibility,

epidemiology/biostatistics, toxicology, and pathological gambling) in this case.[124] Objections to the level of his expertise go to the credibility and weight of his opinion, not its admissibility.

ii. Reliability and Helpfulness

Defendants challenge the reliability of Dr. Glenmullen's methodology on two primary grounds. First, they argue that the evidence on which Dr. Glenmullen bases his opinion is unreliable and, thus, insufficient to support his opinion. Second, Defendants maintain that Dr. Glenmullen did not reliably apply the Bradford Hill factors in reaching his conclusion on general causation. With one exception, the Court disagrees.

With respect to the evidence, the Court has already found that most of the scientific literature on which Dr. Glenmullen relied, including the Etminan Study, is sufficiently reliable to support or bolster his general causation opinion. *See supra* Section II(C). However, one of the studies cited by Dr. Glenmullen, referred to by the parties as the Moore Study, must be excluded as unhelpful in this case. *See* Thomas J. Moore *et al*., *Reports of Pathological Gambling, Hypersexuality, and*

[124] Defendants also object to Dr. Glenmullen's qualifications to offer an expert opinion on "FDA regulations" and "FDA regulatory compliance." *See* Def. Glenmullen Opposition, ECF No. 424-15 at 13-14. Dr. Glenmullen has not been offered as an expert on FDA regulations or regulatory compliance, therefore this challenge is denied as moot. To the extent Defendant intended this objection to exclude Dr. Glenmullen's testimony regarding the FDA's disproportionality analysis of its adverse event reporting database, the challenge is denied for the same reasons he has been found qualified to testify on epidemiology.

Compulsive Shopping Associated with Dopamine Receptor Agonist Drugs, 174 JAMA INTERNAL MED. 1930, 1930-33 (2014) ("Moore Study"), ECF No. 428-10 at 2-5. The Moore Study is based on a disproportionality analysis of the FAERS database examining the association between six dopamine receptor agonist drugs and "unusual but severe" impulsive behaviors.[125] *See id*. at 1931, ECF No. 428-10 at 3. The Moore Study found that those six drugs had a statistically significant higher proportion of patients reporting impulse control disorders from 2003 to 2012, when compared to "all other drugs" in the FAERS database. *See id*. at 1931, ECF No. 428-10 at 3.

The Court has no reason to doubt the reliability of the Moore Study's methodology or findings as to those six drugs. But as to Abilify, "there is simply too great an analytical gap" between the Moore Study and any conclusion about a possible association between Abilify and reports of severe impulsive behaviors. *See Joiner*, 522 U.S. at 147. This is because Abilify was not a focus of the Moore Study; it was mentioned only once, more or less peripherally, in the article. *See* Moore Study at 1932, ECF No. 429-10 at 4 ("We also found a weaker signal for [Abilify],

[125] Again, FAERS refers to the FDA's Adverse Event Reporting System. The Moore Study analyzed FAERS data, for the time period covering 2003 and 2012, for the following six dopamine receptor agonist drugs: Apomorphine, Bromocriptine, Cabergoline, Pramipexole, Ropinirole, and Rotigotine. *See* Moore Study at 1933, ECF No. 428-10 at 2. The 10 "impulsive" behaviors examined by the Moore Study were pathological gambling, hypersexuality, compulsive shopping, gambling, poriomania, binge eating, excessive masturbation, compulsive sexual behavior, kleptomania, and excessive sexual fantasies. *See id*.

an antipsychotic classified as partial agonist at the D_3 receptor."). The Moore Study does not describe any of the underlying FAERS data related to Abilify, such as the number and type of reported adverse events or individual characteristics of the reporting population, although it does so for the six dopamine agonist receptor drugs featured in the study. More importantly, the Moore Study explicitly describes its findings, conclusions, and implications only in terms of those six specific drugs.[126] *See id*. Dr. Glenmullen may not extrapolate from this information a finding, conclusion, or implication about Abilify that the Moore Study authors themselves did not make.[127] *See McClain*, 401 F.3d at 1248 (affirming exclusion of expert who showed "lack of scientific rigor" by "draw[ing] unauthorized conclusions from limited data—conclusions the authors of the study do not make"); *Happel v. Walmart Stores, Inc.*, 602 F.3d 820, 825-26 (7th Cir. 2010) (affirming exclusion of expert opinion based on medical literature that "stops short of" supporting the expert's conclusion); *Anderson v. Bristol Myers Squibb Co.*, No. 4:95-cv-00003, 1998 WL 35178199, at *9-11 (S.D. Tex. April 20, 1998) (finding that an expert may not use studies purporting to prove one fact in order to infer that the same studies

[126] For example, the Moore Study states that its "findings confirm and extend the evidence that dopamine receptor agonist drugs are associated with serious impulse control disorders; the associations were significant, the magnitude of the effects was large, and the effects were seen for *all 6 dopamine receptor agonist drugs*." *See* Moore Study at 1932, ECF No. 428-10 at 4 (emphasis added).

[127] Dr. Glenmullen is one of the Moore Study's authors. *See* Moore Study at 1930, ECF No. 428-10 at 2.

prove a different fact). Notably, Dr. Glenmullen did not even attempt to explain how or why the six dopamine receptor agonists were sufficiently similar to Abilify to warrant an extrapolation. *See McClain*, 401 F.3d at 1246 (extrapolation between drugs permissible only with reliable evidence establishing validity of the analogy). In short, the Moore Study does not "fit" the disputed facts in this case or "logically advance" resolution of the only material question at this stage, namely, whether Abilify is capable of causing uncontrollable compulsions to engage in certain behaviors. *See McDowell*, 392 F.3d at 1299 (stating that relevant expert testimony "logically advances a material aspect of the proposing party's case" and "fits" the disputed facts). Therefore, it is inadmissible and may not support Dr. Glenmullen's general causation opinion.

Defendants final argument for excluding Dr. Glenmullen is that he did not reliably apply the Bradford Hill factors in reaching his general causation opinion. More specifically, Defendants contend that Dr. Glenmullen erred by applying the Bradford Hill factors at all because there is no reliable epidemiological study in existence finding a statistically significant association between Abilify and impulsive behaviors. Since the Court has already found that the Etminan Study reliably establishes the requisite association, this challenge is moot.

Defendants also argue that even if Dr. Glenmullen had statistically significant epidemiological evidence of an association between Abilify and compulsive

behaviors, he misapplied the Bradford Hill factors by "giv[ing] all of the criteria equal weight" and "discussing each criteria in check-the-box fashion." *See* Def. Glenmullen Motion, ECF No. 424-15 at 37.[128] The Court disagrees. As discussed in Section II(B) above, the following nine Bradford Hill factors guide scientists in making judgments about causation: (1) temporal relationship; (2) strength of the association; (3) dose-response relationship; (4) consistency or replication of the findings; (5) biological plausibility; (6) consideration of alternative explanations; (7) cessation of exposure; (8) specificity of the association; and (9) consistency with other knowledge. *See* Ref. Man. at 599-600. Importantly, "[t]here is no formula or algorithm that can be used to assess whether a causal inference is appropriate based on the" Bradford Hill factors. *See* Ref. Man. at 600. The drawing of causal inferences "requires judgment and searching analysis . . . informed by scientific expertise" and reasonable scientists reliably applying the Bradford Hill factors may come to different conclusions about whether a causal inference is appropriate. *See id*.; *see also Milward*, 639 F.3d at 18. Thus, the fact that Dr. Glenmullen found that all of the Bradford Hill factors supported a causal inference does not, standing alone, render his methodology unreliable.

[128] "Def. Glenmullen Motion" refers to Defendants' Motion to Exclude the General Causation Opinion of Joseph Glenmullen, ECF No. 424-15.

Dr. Glenmullen began his Bradford Hill analysis by assessing the "experiment" and "strength of association" factors, which address whether a body of experimental findings exists showing a statistically significant association between a drug and disease of interest. *See* Ref. Man. at 602. Dr. Glenmullen identified several experimental studies showing a statistically significant association between either (1) Abilify and diagnoses of gambling and other impulse control disorders[129] or (2) Abilify and FDA adverse event reports of pathological gambling, hypersexuality, impulsive behavior, and other impulse control disorders.[130] *See* Glenmullen Rep., ECF No. 424-1 at 132-33.[131] It is worth noting that Defendants have not challenged the accuracy of these statistical calculations. Dr. Glenmullen also addressed the consistency factor, *see id*. at 133, which is met where an association is consistently observed by different researchers using different methodologies on different population samples, *see* Ref. Man. at 604. Dr. Glenmullen found that the various types of research evidence in this case—including

[129] This refers to the Etminan Study, which found a 5.23- and 7.71-fold increase in risk of gambling and impulsive control disorder diagnoses, respectively, for patients on Abilify, when compared to individuals not taking Abilify. *See* Etminan Study, ECF No. 428-13 at 2. Dr. Glenmullen also references the Moore Study, which the Court has found inadmissible.

[130] This refers to several disproportionality analyses of the FAERS database related to Abilify, including: (1) the FDA Pharmacovigilance Report, which found a 6.3-fold increase in risk of pathological gambling reports associated with Abilify, as compared to all other atypical antipsychotics, *see* FDA Pharm. Vigil., ECF No. 428-11 at 27; (2) **[*** REDACTED ***]** and (3) **[*** REDACTED ***]**.

[131] "Glenmullen Rep." refers to Dr. Joseph Glenmullen's Expert Report, ECF No. 424-1.

epidemiological evidence, analyses of clinical trial data, and disproportionality analyses of the FDA adverse event reporting database—consistently demonstrate the existence of an association between Abilify and impulsive behaviors. *See* Glenmullen Rep., ECF No. 424-1 at 133. Therefore, this is not a case in which there is a completely novel theory of causation.

Regarding temporality and cessation of exposure, Defendants concede that the numerous case studies cited by Dr. Glenmullen "show . . . a temporal relationship" between the use of Abilify and "a change in the presence or severity of symptoms" of gambling and other impulse control disorders. *See* Def. Glenmullen Motion, ECF No. 424-15 at 24. These case studies, as well as the many adverse event reports Dr. Glenmullen describes, are strongly suggestive of a dose-response relationship between Abilify, gambling and other impulse control disorders. Dr. Glenmullen also references **[*** REDACTED ***]**. *See* Glenmullen Rep., ECF No. 424-1 at 134. Defendants do not question Dr. Glenmullen's characterization of the clinical trial data.

Dr. Glenmullen also considered alternative explanations for the association between Abilify, impulsive gambling, and other impulse control disorders by reference to the Etminan Study, FDA's disproportionality analysis, Bristol-Myers Squibb disproportionality analyses and clinical trial data, and other medical literature. *See id*. at 108-09. According to Dr. Glenmullen, these materials, to

varying degrees, controlled for possible confounding factors such as age, gender, substance abuse disorders, bipolar disorder, schizophrenia, concomitant use of other dopamine agonists or recreational drugs, and history of impulse control disorder of interest. *See id*. at 108. Dr. Glenmullen found this evidence supportive of an inference that none of those possible confounders explains the association between Abilify, pathological gambling and impulse control disorders. *See id*. at 108-09.

As to biological plausibility, Dr. Glenmullen concluded from the medical literature, as did Plaintiffs' other experts, that the biological mechanism by which Abilify can cause impulsive behaviors is its effect on dopamine neurotransmission in the brain. *See id*. at 135. The Court has already found this reasoning on biological plausibility sufficiently reliable to support a general causation opinion. *See supra* Section II(C). Dr. Glenmullen also stated that the association in this case is specific in that Abilify has been shown to be strongly associated with a single, very specific adverse effect, which manifests as uncontrollable impulsivity to engage in certain, harmful behaviors. *See* Ref. Man. at 605-06; Glenmullen Rep., ECF No. 424-1 at 134-35.

On balance, the Court finds that Dr. Glenmullen's application of the Bradford Hill factors is sufficiently reliable to support a conclusion that the observed association between Abilify and impulsive behaviors, such as pathological gambling, reflects a "true cause-effect relationship." *See* Ref. Man. at 597.

Accordingly, Defendants' Motion to Exclude the General Causation Opinion of Joseph Glenmullen, ECF No. 424-15, is due to be denied.[132]

c. Eric Hollander, M.D.

Dr. Eric Hollander is a board-certified psychiatrist and clinical professor at Albert Einstein College of Medicine, with specialized training and experience in the fields of psychopharmacology and neuropsychopharmacology,[133] as well as over thirty years of clinical practice experience treating individuals with impulsive and compulsive behaviors, including pathological gambling, depression, schizophrenia, autism spectrum disorder, and bipolar disorder. *See* Hollander Rep., ECF No. 459-1 at 3-4. Dr. Hollander offers a general causation opinion that Abilify can cause impulsive behaviors, including drug-induced gambling. *See id*. at 8-9, 12.[134] Dr. Hollander supports his opinion with epidemiological evidence (*i.e*., the Etminan Study), scientific literature supporting a plausible biological mechanism of action, animal and *in vitro* data, case and adverse event reports, including dechallenge and

[132] At the *Daubert* hearing, Defendants objected to Dr. Glenmullen's testimony regarding whether case reports exhibiting dechallenge and rechallenge events are caused by a placebo effect, arguing that the testimony was outside the scope of his expert report. *See* Glenmullen Tr., ECF No. 596-2 at 72. Defendants' objection on this basis is overruled, as the issue was sufficiently raised at Dr. Glenmullen's deposition. *See* Glenmullen Dep., ECF No. 424-1 at 206.

[133] Psychopharmacology and neuropsychopharmacology are interrelated fields of study. Psychopharmacology is the study of how drugs affect mood, perception, thinking, and behavior. Neuropsychopharmacology is the study of how drugs affect the nervous system and how those nervous system changes alter behavior.

[134] "Hollander Rep." refers to Dr. Eric Hollander's Expert Report, ECF No. 459-1.

rechallenge events, disproportionality analyses, and Defendants' clinical trial data, all of which he analyzed using the Bradford Hill factors and a weight-of-the-evidence methodology. *See* Hollander Tr., ECF No. 596-4 at 117-18. Defendants challenge Dr. Hollander's testimony on qualification and reliability grounds.

i. Qualification

Defendants argue that Dr. Hollander is not qualified to offer a general causation opinion because he lacks sufficient training and experience in the fields of epidemiology and toxicology. The Court disagrees. As already explained, a witness may be qualified "by knowledge, skill, experience, training, or education" to offer an expert opinion that will help the trier of fact understand the evidence or resolve a factual issue. *See* Fed. R. Evid. 702. The qualification standard is "not stringent" and "so long as the witness is minimally qualified, objections to the level of [his] expertise go to credibility and weight, not admissibility." *Hendrix I*, 255 F.R.D. at 585.

In this case, the Court finds that Dr. Hollander is amply qualified to offer an expert opinion on whether Abilify can cause impulse control disorders. Dr. Hollander is a medical doctor and board-certified psychiatrist with over thirty years of experience researching, publishing, and teaching in the fields of psychopharmacology and neuropsychopharmacology. *See* Hollander Curriculum Vitae, ECF No. 459-1 at 38-96. This background, with its emphasis on the study of

how psychiatric drugs affect brain chemistry and behavior, well equips Dr. Hollander to assist the trier of fact in understanding the biological mechanisms by which Abilify can cause impulsive behaviors. Dr. Hollander also has formal training in epidemiology, has "worked closely with epidemiologists to publish epidemiologic papers in peer review[ed] journals," and as an academic, has taught "epidemiologic principles as it relates to psychiatry and psychopharmacology" to "medical students, residents, and fellows." *See* Hollander Tr., 596-4 at 114-15. This specialized education and experience with epidemiology qualifies Dr. Hollander to give an expert opinion about the epidemiological study in this case (*i.e.*, the Etminan Study). *See Thorn*, 317 F.3d at 114-15 (medical doctor specializing in asbestos-related disease permitted to testify about various epidemiological studies of asbestos exposure).

Notably, Defendants do not dispute that Dr. Hollander is a leading expert on the etiology and treatment of impulse control disorders, including impulsive gambling. Indeed, Dr. Hollander was a member of the research agenda workgroup that, quite literally, wrote the DSM-IV diagnostic criteria for pathological gambling and, several years later, he oversaw the reorganization of the DSM-5 diagnostic criteria for gambling disorder and impulse control disorders.[135] *See* Hollander Tr.,

[135] For this reason, the Court overrules Defendants' *Daubert* hearing objections as to Dr. Hollander's testimony on both the history of the DSM-5 and the distinction made in the publication between idiopathic and iatrogenic gambling. Given his professional experience working on the

596-4 at 113-14. Dr. Hollander also has written and collaborated on hundreds of peer-reviewed articles related to uncontrollable impulsive behaviors, such as pathological gambling, and the brain circuitry underlying these psychiatric problems. *See* Hollander Curriculum Vitae, ECF No. 459-1 at 57-96. Finally, Dr. Hollander has treated thousands of patients with impulse control disorders, including hundreds with gambling disorders. *See* Hollander Tr., ECF No. 596-4 at 111. Given the breadth of Dr. Hollander's academic and clinical experience in the fields of psychiatry, psychopharmacology, and neuropsychopharmacology, the Court finds him qualified to offer an expert opinion on general causation in this case. Objections to the level of Dr. Hollander's expertise go to the credibility and weight of his opinion, not its admissibility. *See Hendrix I*, 255 F.R.D. at 585.

ii. Reliability

Defendants challenge the reliability of Dr. Hollander's general causation opinion on two primary grounds. First, they argue that the evidence on which Dr. Hollander bases his opinion is unreliable and, thus, insufficient to support his opinion. Second, Defendants maintain that Dr. Hollander did not reliably analyze the evidence in reaching his conclusion on general causation.

DSM-5, and also his deposition testimony, these opinions are clearly within the scope of his expertise and his reports.

Regarding the evidence, the Court has already found that most of the scientific literature on which Dr. Hollander relied, including the Etminan Study, is sufficiently reliable to support or bolster his general causation opinion. *See supra* Section II(C). Only one of Defendants' specific evidentiary challenges warrants additional comment. As part of the support for his general causation opinion, Dr. Hollander cites a 2016 article from the scientific literature comparing the characteristics of "possibly iatrogenic" problem gambling in patients taking Abilify with the characteristics of such gambling in patients taking a full dopamine replacement therapy.[136,137] *See* Marie Grall-Bronnec *et al*., *Pathological Gambling Associated with Aripiprazole or Dopamine Replacement Therapy: Do Patients Share the Same Features? A Review*, 36 J. CLINICAL PSYCHOPHARMACOLOGY 63, 64 (2016) ("Grall-Bronnec Article"), ECF No. 425-4 at 2, 3. Defendants argue that the Grall-Bronnec

[136] A "problem gambler" was defined as a patient who has exhibited three or more of the DSM-IV diagnostic criteria for "pathological gambling" in the preceding 12 months. *See* Grall-Bronnec Article, ECF No. 425-4 at 3. The Article's authors stated that, although the presence of at least five of the DSM-IV diagnostic criteria are required for a formal diagnosis of pathological gambling, the presence of at least three criteria is enough to suggest "at risk gambling" or "problem gambling." *See id*. The authors selected a threshold of three criteria because, in their view, both pathological and problem gamblers require care. *See id*.

[137] The dopamine replacement medications were cabergoline, pergolide, piribedil, pramipexole, ropinirole, levodopa, and carbidopa. *See* Grall-Bronnec Article, ECF No. 425-4 at 3.

Article is not a valid epidemiological study and, therefore, cannot support a general causation opinion. *See* Def. Hollander Motion, ECF No. 425-14 at 13.[138]

The problem with Defendants' challenge to the Grall-Bronnec Article is that neither Dr. Hollander nor the Article's authors offered the Article as epidemiological evidence of causation. *See* Hollander Rep., ECF No. 459-1 at 22; Grall-Bronnec Article, ECF No. 425-4 at 7. Defendants' argument appears to be based on a misinterpretation of the authors' use of the term "cohort" to describe the "problem gamblers" who were interviewed as part of their research. A "cohort," as defined by Merriam-Webster, is "a group of individuals having a statistical factor (such as age or class membership) in common in a demographic study." *See* Merriam-Webster Online Dictionary, https://www.merriam-webster.com/dictionary/cohort (retrieved Dec. 3, 2017). In this case, the Grall-Bronnec Article's authors evaluated a group of individuals with the statistical factors of problem gambling and Abilify use in common. Thus, the term "cohort" aptly describes the subjects of their work. A "cohort," as used in this context, differs from a "cohort study," which is a type of epidemiological study used to "measure and compare the incidence of disease" in certain populations. *See* Ref. Man. at 557. Neither Dr. Hollander nor the Article's authors characterize the Grall-Bronnec Article as a cohort study, and neither treats

[138] "Def. Hollander Motion" refers to Defendants' Motion to Exclude the General Causation Opinion of Eric Hollander, ECF No. 425-14.

the Article's findings as carrying the same weight as epidemiology. Indeed, the Grall-Bronnec Article's authors actually *recommend* a cohort study as "a promising way to obtain further evidence" on causation. *See* Grall-Bronnec Article, ECF No. 425-4 at 7.

While the Court agrees that the Grall-Bronnec Article is not based on epidemiology, that fact does not preclude Dr. Hollander from relying on the Article as support for his general causation opinion. The Grall-Bronnec Article's authors' objective was to further scientific understanding of the nature of iatrogenic gambling by analyzing the sociodemographic profiles, gambling characteristics, comorbidities, and personality traits of patients whose "problem gambling" behaviors "could possibly result from an adverse drug reaction after the administration of a dopamine medication." *See id*., ECF No. 425-4 at 3-4. The authors identified nine published case reports and conducted in-person clinical evaluations of eight individual patients in treatment for problem gambling, resulting in 17 discrete cases involving the use of Abilify.[139] *See id*. From this anecdotal data, the Article's authors concluded it was "possible" that the gambling behavior in 16

[139] The Grall-Bronnec Article's authors identified 17 published case reports describing gambling behaviors in Abilify patients, eight of which involved individual patients who also participated in the in-person clinical evaluations. *See* Grall-Bronnec Article, ECF No. 425-4 at 4. The authors identified 42 published case reports describing gambling behaviors in patients taking dopamine replacement therapy, in addition to six such patients who participated in the in-person clinical evaluations. *See id*.

of the 17 cases was "actually due to" Abilify, but cautioned that more research would be necessary before the relationship could be characterized as causal.[140] *See id*. at 4, 7. Notably, Defendants do not challenge the reliability of the authors' methodology, findings, or substantive conclusions.

The Court finds that the Grall-Bronnec Article presents a reliable and probative analysis of 17 patients' personal experiences while taking Abilify. In particular, the authors' in-depth clinical assessments and comparisons of the eight individual patients facilitate understanding and evaluation of the characteristics associated with "possibly" iatrogenic (medication-induced) gambling.[141] Nevertheless, the Grall-Bronnec Article is not quantitative research and, as acknowledged by the authors, its results cannot be generalized to definitively establish that Abilify causes compulsive gambling. *See id*. at 7. The Article essentially is a compilation of thoroughly examined case studies and, therefore, it cannot, standing alone, prove general causation. *See Rider*, 295 F.3d at 1199. But it may supplement other, more substantive evidence of causation, *see id*., which is exactly how it was used by Dr. Hollander in this case, *see* Hollander Rep., ECF No.

[140] The authors also concluded it was "possible" that the gambling behaviors in 46 of the 48 dopamine replacement therapy cases "was actually due to" the dopamine replacement therapy. *See id*. at 4.

[141] The same is true of the authors' in-person clinical assessments of the six patients taking dopamine replacement therapy.

459-1 at 22.[142] Defendants' challenge to the Grall-Bronnec Article affects only its weight, not its admissibility.

Defendants' last argument for excluding Dr. Hollander is that he did not reliably analyze the scientific evidence in this case. More specifically, Defendants contend that Dr. Hollander improperly relied on the Naranjo Scale and WHO-UMC criteria, which are specific causation methodologies, to reach his general causation opinion.[143] Defendants also argue, in the alternative, that Dr. Hollander did not reliably apply the Bradford Hill factors. The Court disagrees.

With respect to the Naranjo Scale and WHO-UMC criteria, Defendants' argument is misplaced because Dr. Hollander did not employ either technique during his analysis of the scientific evidence in this case. Dr. Hollander testified that he used a weight-of-the-evidence methodology and fully considered all of the Bradford Hill factors. *See* Hollander Tr., ECF No. 596-4 at 117-19. This is evident from Dr. Hollander's initial expert report, which, in addition to explicitly citing his reliance

[142] Neither Dr. Hollander nor the Grall-Bronnec Article's authors characterized, or treated, this evidence as epidemiology.

[143] The Naranjo Adverse Drug Reactions Probability Scale, also referred to as the Naranjo Scale or Naranjo Algorithm, is a scientific tool designed to assess the probability that a discrete adverse medical event was drug-induced, rather than the result of other factors. Christopher R.J. Pace, *Admitting and Excluding General Causation Expert Testimony: The Eleventh Circuit Construct*, 37 AM. J. TRIAL ADVOCACY 47, 61 n.60 (2013). Similarly, the Causality Assessment System of the World Health Organization-Uppsala Monitoring Centre ("WHO-UMC criteria") is a tool for assessing causality with respect to individual suspected adverse drug reactions. *See* WHO-UMC, *The Use of the WHO-UMC System for Standardised Case Causality Assessment*, at 1, https://www.who-umc.org/media/2768/standardised-case-causality-assessment.pdf (last visited Dec. 3, 2017) ("WHO-UMC Reference").

on the Bradford Hill factors, demonstrates that he applied those factors, in substance, in reaching his general causation opinion. *See* Hollander Rep., ECF No. 459-1. Dr. Hollander only ever mentions the Naranjo Scale or WHO-UMC criteria twice—once to identify them as "additional method[s]" for determining whether a drug caused an isolated adverse medical event, *see id*. at 12, and once to disclose that the authors of the Grall-Bronnec Article used the Naranjo Scale to assess whether Abilify caused the "problem gambling" behaviors exhibited by individual patients in their cohort, *see id*. at 22.[144] Defendant has not referenced, and the Court has not found, any other instance in which Dr. Hollander further discussed either the Naranjo Scale or the WHO-UMC criteria in his report or in his testimony, much less relied on them in forming his causation opinion. Consequently, this challenge fails.

Defendants' challenge to the reliability of Dr. Hollander's application of the Bradford Hill factors also fails. As an initial matter, Defendants' argument that Dr. Hollander erred by considering the Bradford Hill factors at all is moot because the prerequisite for applying those factors—that is, an epidemiological study reliably establishing a statistically significant association between the use of a drug and an adverse medical effect—is satisfied by the Etminan Study. *See* Section II(C). Thus,

[144] The Naranjo Scale is used in determining the causal link between a drug and an individual clinical event (*i.e.*, specific causation), which is what the Grall-Bronnec authors were assessing. As noted, the Naranjo Scale is not a general causation tool.

the only remaining question is whether Dr. Hollander reliably weighed the Bradford Hill factors as part of his general causation analysis.[145] The Court finds that he did.

Dr. Hollander found a "very strong" association between Abilify and impulse control disorders based on his assessments of the Etminan Study, the FDA's 2016 Pharmacovigilance Review and **[*** REDACTED ***]**. *See* Hollander Tr., 596-5 at 37. As the Court has already observed, this evidence reliably establishes the existence of a statistically significant association between both (1) Abilify and medical diagnoses of pathological gambling and other impulse control disorders; and (2) Abilify and adverse event reports of pathological gambling and other impulse control disorders. The Bradford Hill factor of specificity is also met, as the association in this case involves only the very narrow and specific adverse effect of impulse control problems. Dr. Hollander found that the evidence also reliably demonstrates the Bradford Hill factor of consistency, in that the association has been shown to be consistently present in a number of different analyses using different criteria, populations and methods. *See* Hollander Rep., 459-1 at 27; Hollander Tr., ECF No. 596-5 at 36. As Dr. Hollander noted, there does not appear to be "any

[145] Again, the nine Bradford Hill factors are: (1) temporal relationship; (2) strength of the association; (3) dose-response relationship; (4) consistency or replication of the findings; (5) biological plausibility; (6) consideration of alternative explanations; (7) cessation of exposure; (8) specificity of the association; and (9) coherence with other knowledge. *See* Ref. Man. at 600.

evidence at all that suggests" there is *no* association between Abilify and impulsive behaviors. *See* Hollander Tr., ECF No. 596-5 at 38.

As to temporality, Dr. Hollander cites numerous case studies and adverse event reports in which impulsive behaviors emerged only after a patient's exposure to Abilify. *See* Hollander Rep., ECF No. 459-1 at 23. These case studies and adverse event reports, although not dispositive of the issue, also are strongly suggestive of a dose-response relationship between Abilify and impulse control disorders. *See* Section II(C); Hollander Rep., ECF No. 459-1 at 30-31. Moreover, the reports of dechallenge events, in particular, satisfy the Bradford Hill factor of cessation of exposure, as they reliably demonstrate that, with many individual patients, impulse control problems disappeared once Abilify was decreased or discontinued.[146]

Much like Dr. Glenmullen, Dr. Hollander ruled out alternative explanations for the association between Abilify and impulsive behaviors by reference to the Etminan Study and the FDA's 2016 Pharmacovigilance Review, as well as to case studies describing dechallenge events. According to Dr. Hollander, the FDA's disproportionality analysis of 11 different atypical antipsychotics is particularly significant because, although those medications all treat the same patient population with the same underlying conditions, only Abilify showed a statistically significant

[146] Again, a dechallenge event occurs where a patient's adverse side effects partially or completely disappear once the drug is stopped. *Rider*, 295 F.3d at 1199.

incidence of adverse event reports involving uncontrollable impulsive behaviors. *See* Hollander Tr., ECF No. 596-1 at 25-27. Dr. Hollander concluded that if the patients' underlying psychiatric conditions caused the uncontrollable impulses, then all of the atypical antipsychotics should have had statistically significant reporting of impulse control problems. *See id*. Dr. Hollander observed that the Etminan Study specifically controlled for bipolar disorder, schizophrenia, and substance abuse disorder, which ruled out those conditions as possible explanations for the association between Abilify and impulsive behaviors. *See* Hollander Rep., ECF No. 459-1 at 33; Hollander Tr., 596-5 at 71.

Regarding the Bradford Hill factor of biological plausibility, Dr. Hollander's opinion as to Abilify's mechanism of action "mirror[s]" that of Plaintiffs' neurobiology expert, Dr. Antoine Bechara, *see* Def. Hollander Motion, ECF No. 425-14 at 24, and is reliable for the same reasons, *see* Section II(C). This proposed biological mechanism of action—Abilify's effect on dopamine neurotransmission in the brain—is coherent with existing scientific knowledge about psychopharmacology, neuropsychopharmacology, and the biochemistry of the brain. *See* Section II(C). Importantly, Dr. Hollander demonstrated a comprehensive understanding of the scientific evidence in support of his opinion. In sum, the Court finds that Dr. Hollander reliably considered all of the Bradford Hill factors in reaching his opinion that Abilify can cause impulse control problems. Dr. Hollander

also reliably explained his methodology, reasoning and conclusions at length, both in his expert reports and at the *Daubert* hearing. *See* Hollander Rep., ECF No. 459-1 at 1-36; Hollander Supp., DX-641;[147] Hollander Tr., ECF Nos. 596-4 at 108-45, 596-5. For these reasons, Dr. Hollander's expert opinion is sufficiently reliable under Rule 702 and *Daubert*. Accordingly, Defendants' Motion to Exclude the General Causation Opinion of Eric Hollander, ECF No. 425-14, is due to be denied.

d. Russell V. Luepker, M.D.

Dr. Russell V. Luepker is a board-certified cardiologist, and a professor of public health and medicine at the University of Minnesota. He holds a master's degree in epidemiology from Harvard University, is certified in epidemiology by the American College of Epidemiology, and served as head of the Division of Epidemiology at the University of Minnesota for over 13 years. Dr. Luepker also has over 40 years of experience researching, publishing, and teaching on the "design, implementation and interpretation of clinical research" methods. *See* Luepker Rep., ECF No. 462-1 at 3.[148] There is no dispute that Dr. Luepker possesses the formal credentials necessary to offer an expert opinion on medical causation. The Court's concern with respect to Dr. Luepker, based on his expert reports and deposition

[147] "Hollander Supp." refers to Dr. Eric Hollander's Rebuttal/Supplemental Report, DX-641.

[148] "Luepker Rep." refers to Dr. Russell V. Luepker's Expert Report, ECF No. 462-1 at 2-15.

testimony, is that he does not appear to have brought his considerable expertise to bear in his analysis of the evidence in this case. *See* Luepker Rep., ECF No. 462-1 at 2-15; Luepker Supp., ECF No. 426-5 at 2-10; Luepker Dep., ECF No. 462-1 at 97-181.[149] This presents a reliability problem under *Daubert* that Plaintiffs have not overcome.

The "primary focus" of Dr. Luepker's "teaching, research, and clinical career" has been epidemiology and other types of clinical research in humans. *See* Luepker Rep., ECF No. 462-1 at 2. The Court is of the view that this background well equips him to offer unique insights into the methodological soundness of the only epidemiological evidence in this case, the Etminan Study. Yet, Dr. Luepker devotes just a single paragraph of his initial expert report to the Etminan Study, in which he provides only a cursory statement of the Study's findings and nothing more. *See id.* at 10. Almost two pages of his rebuttal report discuss the Etminan Study further, but this too lacks any meaningful analysis beyond general assertions about health insurance claims database research becoming a "major trend[] in epidemiology over the past 10 years." *See* Luepker Supp., ECF No. 426-5 at 3. Dr. Luepker also failed to meaningfully examine the background risk of pathological gambling and other impulse control problems in either the general or psychiatric patient populations.

[149] "Luepker Supp." refers to Dr. Russell V. Luepker's Rebuttal of Defendants' Expert Reports, ECF No. 426-5 at 2-10. "Luepker Dep." refers to the official transcript of Dr. Luepker's deposition testimony on June 16, 2017, ECF No. 462-1 at 97-181.

Although Dr. Luepker expressed "some hesitation" and "worry" about Dr. Potenza's background risk estimates, he did not attempt to independently verify the accuracy of those figures. *See* Luepker Dep., ECF No. 462-1 at 118-19. Since the aim of epidemiology is to "identif[y] agents that are associated with an increased risk of disease," the Court would expect a more robust background risk analysis from an expert epidemiologist. *See* Ref. Man. at 552. Equally, if not more troubling is Dr. Luepker's opinion that a published case series, and even a single case report, are "definitely" types of epidemiological studies. *See* Luepker Dep., ECF No. 462-1 at 128. Both from a scientific perspective and for legal causation purposes, the distinction between epidemiological evidence and anecdotal evidence (*i.e.*, case series and case reports) is substantial and consequential. Dr. Luepker, apparently, disagrees.

There are also *Daubert* reliability problems with Dr. Luepker's general causation analysis. First, he employed the WHO-UMC causality criteria, which, as the Court has already discussed, *see* Section (II)(D)(1)(c)(ii), is a scientific tool designed to assess specific causation, *see* WHO-UMC Reference at 1. It "cannot" be used to prove general causation. *See id*. Second, Dr. Luepker's explanation of the biological mechanism by which Abilify can cause impulse control problems is inadequate, likely because much of the subject matter is, by his own admission,

beyond the scope of his expertise.[150] Indeed, he does not appear to have much more than a superficial understanding of how dopamine functions in the brain. He readily conceded as much at his deposition, testifying that he has only "some ancillary knowledge" about dopamine binding, intrinsic activity, *in vivo* and *in vitro* toxicological studies of activity at dopamine receptors, and dopaminergic drugs, including Abilify. *See* Luepker Dep., ECF No. 462-1 at 114-16. As a result, Dr. Luepker's initial expert report speaks in overly broad and general terms about the complex and nuanced proposed mechanism of action in this case. While the report contains a lengthy appendix of scientific literature that he "reviewed" or "relied upon" as part of his general causation analysis, at no point does he directly connect these publications to his own biological plausibility analysis.[151] In other words, there is very little evidence, based on Dr. Luepker's written submissions and deposition

[150] During his deposition, Dr. Luepker initially testified that he "certainly [has] some knowledge and understanding of dopamine and dopamine receptors because of an interest in that and some family health issues." *See* Luepker Dep., ECF No. 462-1 at 110. However, when questioned about the various aspects of Plaintiffs' proposed biological mechanism of action, he repeatedly stated that he "did not know" much about, and "would not hold [himself] up as an expert" on, virtually every issue. *See id.*, ECF No. 462-1 at 114 (mechanism of action of dopaminergic drugs or animal studies of such drugs), 115 (intrinsic activity, dopamine binding sites, or *in vivo* and *in vitro* studies of those sites)

[151] This is true except with respect to an analogy Dr. Luepker draws between Abilify's mechanism of action and that of dopamine replacement therapies used to treat Parkinson's Disease. *See* Luepker Rep., ECF No. 462-1 at 6, 16 (citing a peer-reviewed scientific article on impulse control disorders in Parkinson's patients taking dopamine replacement therapies). However, he never offers any evidence, or even an explanation, establishing the reliability of the analogy. *See Rider*, 295 F.3d at 1200-01 (stating that extrapolations between drugs are impermissible unless reliable scientific evidence establishes the validity of the analogy). Thus, Dr. Luepker's drug analogy, and the scientific article he offered in support of it, are inadmissible.

testimony, that these materials meaningfully informed his biological plausibility opinion. This is unacceptable, given Dr. Luepker's unfamiliarity with the biochemistry of the brain.

Taken together, these are not insignificant failings and they cannot be cured by the fact that Dr. Luepker's conclusions are consistent with those of Plaintiffs' other experts. *See In re Polypropylene Carpet Antitrust Litig.*, 93 F. Supp. 2d 1348, 1357 (N.D. Ga. 2000) (expert "may not simply repeat or adopt the findings of another expert without attempting to assess the validity of the opinions relied upon"). The focus of the *Daubert* reliability inquiry "must be solely on [an expert's] principles and methodology, not on the conclusions that they generate." *Daubert*, 509 U.S. at 595; *see also McDowell*, 392 F.3d at 1298 (same). This requires a district court to "undertake an independent analysis of each step in the logic leading to the expert's conclusions; if the analysis is deemed unreliable at any step, the expert's entire opinion must be excluded." *See Hendrix I*, 255 F.R.D. at 578. For the reasons discussed above, the Court has deemed several critical steps in Dr. Luepker's analysis to be unreliable; therefore, his entire opinion must be excluded. This decision was not made lightly. It is obvious from Dr. Luepker's curriculum vitae that he is a prominent, highly respected cardiologist and research scientist, and deservedly so. However, it appears that the general causation questions presented in this case are beyond the ken of his expertise, which hindered his ability to provide

a reliable expert opinion. Accordingly, Defendants' Motion to Exclude the General Causation Opinion of Russell Luepker, ECF No. 426-16, is due to be granted.

e. David Madigan, Ph.D.

Dr. David Madigan is a biostatistician with over thirty years of experience researching, publishing, teaching, and consulting in the fields of statistics, biostatistics, epidemiology, and pharmacovigilance. He is offered for the purposes of providing: (1) a biostatistical analysis of the scientific evidence in this case; (2) background and contextual information about pharmacovigilance practices, as well as the design and analysis of clinical trials; and (3) a medical causation opinion that Abilify is capable of causing the specific adverse effects of pathological gambling and impulse control disorder. Defendants challenge Dr. Madigan on qualification and reliability grounds.

i. Qualification

Defendants argue that Dr. Madigan lacks the medical knowledge and experience to offer a general causation opinion. The Court agrees. "Dr. Madigan is a man of statistics, not medicine." *See* Def. Madigan Motion, ECF No. 427-20 at 9. He is not a medical doctor, toxicologist, pharmacologist, or psychologist. He also has no specialized knowledge of, or clinical experience with, pathological gambling or impulse control disorders. The Court finds that Dr. Madigan's admitted lack of expertise in the aforementioned fields precludes him from offering a medical or

scientific opinion that Abilify is capable of causing pathological gambling and impulse control disorder.

Nevertheless, the Court finds Dr. Madigan amply qualified to offer a biostatistical analysis of the evidence in this case, as well as opinions related to pharmacovigilance and clinical trials generally, as his credentials in these fields are well beyond reasonable challenge. Briefly, Dr. Madigan is a Professor of Statistics at Columbia University, where he is also Dean of the Faculty and Executive Vice-President for Arts and Sciences. He holds a bachelor's degree in Mathematical Sciences and a doctorate in Statistics. He is an elected Fellow of both the Institute of Mathematical Sciences and the American Statistical Association. He has published more than 160 peer-reviewed academic articles in the areas of statistics, biostatistics, epidemiology, and pharmacovigilance. Drug safety, with a focus on the development and application of statistical methods for pharmacovigilance, is a "significant research interest" of Dr. Madigan's. *See* Madigan Rep., ECF No. 427-1 at 2. Over the years, he has served the FDA in a number of different capacities related to the identification and evaluation of safety risks of medical products, and he currently serves the FDA as a consultant.[152] He has also consulted for various

[152] For example, Dr. Madigan served as an investigator for the FDA's Mini-Sentinel project, the goal of which was "to inform and facilitate development of a fully operational active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products." *See* Madigan Rep., ECF No. 427-1 at 2. From 2009-2013, Dr. Madigan served as principal investigator for the OMOP, which studied the strengths and weaknesses of healthcare database research for identifying and evaluating safety and benefit issues of FDA-regulated drugs.

pharmaceutical companies on issues related to statistics, drug safety, and clinical trials. Finally, Dr. Madigan has "extensive experience" designing and analyzing clinical trials, both as a consultant and in the academic context. *See* Madigan Tr., ECF No. 596-4 at 7.

Defendants do not dispute that Dr. Madigan is a leading expert on biostatistics, pharmacovigilance, and clinical trials. Indeed, Dr. Madigan has been qualified to offer expert opinions in these areas in numerous federal and state courts.[153] Likewise here, the Court finds Dr. Madigan qualified to offer expert opinions on the design

---

See id. at 2-3. From 2010-2011, Dr. Madigan was a member of a subcommittee of the FDA Science Board charged with reviewing the Center for Drug Evaluation and Research's pharmacovigilance program. *See id*. at 3. From 2011-2014, Dr. Madigan was a member of the FDA's Drug Safety and Risk Management Advisory Committee, which advises the FDA Commissioner of risk management, risk communication, and quantitative evaluation of adverse events reports for FDA-regulated medical products. *See id*. at 3.

[153] *See, e.g.*, *Rheinfrank*, 2015 WL 13022172, at *11-13 (finding Dr. Madigan qualified to offer expert opinion about the presence and timeline of safety signal in FAERS database for developmental delay from *in utero* exposure to Depakote, based on statistical analysis of FAERS database); *Fosamax*, 2013 WL 1558690, at *7-8 (finding Dr. Madigan qualified to offer expert opinion on the existence and strength of a safety signal indicating an association between Fosamax and reports of bone turnover and atypical femur fractures, based on "industry standard pharmacovigilance techniques and data sources"); *In re Yasmin and YAZ (Drospirenone) Marketing, Sales Practices and Products Liability Litigation*, No. 3:09-md-2100, 2011 WL 6302573, at *16-17 (S.D. Ill. Dec. 16, 2011) (finding Dr. Madigan qualified to offer expert opinion as to the detection and assessment of "a pharmacovigilance safety signal" in FAERS database concerning increased rate of venous thromboembolic disease with YAZ and Yasmin); *In re Pfizer Inc. Securities Litigation*, No. 05-md-1688, 2010 WL 1047618, *4 (S.D.N.Y. 2010) (finding Dr. Madigan qualified to offer expert opinion as to drug safety and the import of a statistical meta-analysis he performed on data that was in existence during the relevant period to determine its significance with respect to the cardiovascular safety of Celebrex); *see also In re Accutane Litigation*, 451 N.J. Super. 153 (2017) (reversing trial court's exclusion of Dr. Madigan's expert opinion based on statistical analyses of epidemiological studies and the FAERS database), *cert. granted*, No. 079958, 2017 WL 6728709 (N.J. Dec. 8, 2017).

and analysis of clinical trials, pharmacovigilance, and to provide expert biostatistical analyses of the scientific evidence in this case, including, but not limited to, the Etminan Study, adverse event reports in the FAERS database, various disproportionality analyses by Defendants and the FDA, and Defendants' clinical trial data.

ii. Dr. Madigan's Opinion

Dr. Madigan used a series of different statistical analyses to assess whether and to what extent the evidence in this case indicates the presence of an association between Abilify, pathological gambling, and impulse control disorders. The Court has already discussed Dr. Madigan's statistical analysis of the Etminan Study and his opinion that the Study evidences a "strong" association between Abilify and these two adverse effects. *See supra* Section II(C).

Dr. Madigan also conducted several disproportionality analyses of the FDA's adverse event reporting database (FAERS), comparing the relative frequency of pathological gambling reports among Abilify and 10 other atypical antipsychotics.[154] To address concerns about potential litigation-driven reporting (*i.e.*, reporting bias), Dr. Madigan excluded all adverse event reports from lawyers and analyzed the data

[154] Dr. Madigan analyzed the same 10 other atypical antipsychotics that the FDA analyzed as part of its 2016 Pharmacovigilance Review, which were: olanzapine, quetiapine, risperidone, asenapine, brexpiprazole, clozapine, iloperidone, lurasidone, paliperidone, and ziprasidone. *See* Madigan Rep., ECF No. 427-1 at 16; FDA Pharm. Vigil., ECF No. 428-11 at 27.

as of four different dates he considered significant in the life of this case.[155] He also controlled for potential confounders by performing separate disproportionality analyses restricted to pathological gambling reports involving patients with bipolar disorder or schizophrenia. On each of the four dates, and in each of the separate analyses, Dr. Madigan found that Abilify had statistically significant percentages of pathological gambling reports.[156] In other words, according to Dr. Madigan, the association between Abilify and pathological gambling reports was "very strong" at all times.[157] Madigan Tr., ECF No. 596-4 at 23-24; Madigan Rep., ECF No. 427-1 at 19. None of the other atypical antipsychotics exhibited an association with pathological gambling reports. *See* Madigan Tr., ECF No. 596-4 at 23; Madigan Rep., ECF No. 427-1 at 19. Importantly, Defendants do not dispute Dr. Madigan's

[155] Those four dates were: (1) **[*** REDACTED ***]**; (2) Quarter 4, 2015 – prior to first lawsuits being filed in early 2016; (3) Quarter 2, 2016 – prior to warning added in August 2016; and (4) the most recent FAERS data available as of the date of Dr. Madigan's analysis. *See* Madigan Rep., ECF No. 427-1 at 17.

[156] Dr. Madigan measured disproportionality in adverse event reporting by calculating an EB05 score for each drug as of each date. Again, the EB05 is a statistic used by the FDA to estimate the strength of an observed association between a drug and reports of a particular adverse effect. *See* Section II(C). All of Abilify's EB05 scores, as calculated by Dr. Madigan, were greater than 2.0, which is the "widely accepted threshold" indicator of a statistically significant association between adverse event reports and a drug, signaling potential safety issues that require further investigation. *See* Madigan Tr., 596-4 at 20.

[157] More specifically, the EB05 score for Abilify in the unrestricted analyses as of the four different dates was: (1) Quarter 3, 2014 – 2.90; (2) Quarter 4, 2015 – 13.16; (3) Quarter 2, 2016 – 18.55; and (4) Quarter 3, 2016 – 20.79. *See* Madigan Rep., ECF No. 427-1 at 17. The EBO5 score for Abilify in the analyses restricted to bipolar and schizophrenia patients was: (1) Quarter 3, 2014 – 2.41; (2) Quarter 4, 2015 – 3.10; (3) Quarter 2, 2016 – 3.55; and (4) Quarter 3, 2016 – 3.90. *See id*. at 18.

statistical calculations or findings. Indeed, Dr. Madigan's findings are consistent with the findings of **[*** REDACTED ***]**, and of the FDA's disproportionality analysis in 2016, *see* FDA Pharm. Vigil., ECF No. 428-11 at 27.

Next, Dr. Madigan analyzed Defendants' clinical trial materials related to pathological gambling. He found that the clinical trials results reflected no patient reports of pathological gambling, so he analyzed the underlying trial data to determine whether the trial was sufficiently powered—essentially, whether it was large enough—to detect a statistically significant association between Abilify and pathological gambling, if in fact such an association existed.[158] He found that it was not. More specifically, Dr. Madigan's analysis showed that Defendants' clinical trials were not large enough to detect a statistically significant increased risk of pathological gambling.[159] Madigan Rep., ECF No. 427-1 at 30. Dr. Madigan further

[158] "Power" is a statistical concept that quantifies the ability of a study to detect an association that truly exists. *See ASARCO, Inc. v. Occupational Safety and Health Admin*., 746 F.2d 483, 493 n.19 (9th Cir. 1984); *see also Kuhn v. Wyeth, Inc*., 686 F.3d 618, 622 n.5 (8th Cir. 2012) ("Power analysis can be used to calculate the likelihood of accurately measuring a risk that manifests itself at a given frequency in the general population based on the sample size used in a particular study."). Larger associations are more readily detectable than associations of a small magnitude. *See ASARCO*, 746 F.2d at 493 n.19. An increase in the amount of data increases the chance or "power" of observing a given association. *See id*. Dr. Madigan testified that a power calculation only is relevant where a study finds no statistically significant results. *See* Madigan Tr., ECF No. 596-4 at 46. According to Dr. Madigan, "[o]nce you have a statistically significant finding, power is moot" because the study's ability to detect an association (*i.e*., its power) is established by the fact that it actually detected an association. *See id*.

[159] By Dr. Madigan's calculations, the clinical trials, given their small size, had a 6% probability or chance of detecting a 25% increased risk of pathological gambling with Abilify, a 9% chance of detecting a 50% increased risk, a 14% chance of detecting a 75% increased risk, and a 22% chance of detecting a 100% increased risk. *See* Madigan Rep., ECF No. 427-1 at 25-26.

determined that, to have been adequately powered to detect a statistically significant association, the clinical trials would have needed to have been approximately five times larger than they were.[160] *See* Madigan PPT, PX-051 at 10; Madigan Tr., ECF No. 596-4 at 33. Again, Defendants do not dispute Dr. Madigan's calculations.

Dr. Madigan also analyzed **[*** REDACTED ***]**. Dr. Madigan first calculated the frequency of these three categories of adverse effects in patients exposed to Abilify as compared with the frequency of those effects with comparator drugs and a placebo. This analysis yielded, in statistical terms, an estimate of the relative risk of developing the adverse effect among the three groups.[161] *See* Ref. Man. at 566; *see also Allison*, 184 F.3d at 1315 n.16. Dr. Madigan then calculated *p*-values for the relative risk findings to determine whether the increased risk of hypersexuality, impulsive behavior, and increased libido with Abilify was statistically significant.[162]

---

[160] Dr. Madigan calculated **[*** REDACTED ***]**. By Dr. Madigan's calculation, **[*** REDACTED ***]**.

[161] Again, a relative risk of 1.0 means there is no difference in risk between the exposed, comparator, and placebo groups; in other words, there is no association between exposure to the drug and the adverse effect. *See* Ref. Man. at 567; *see also Allison*, 184 F.3d at 1315 n.16. A relative risk above 1.0 indicates an increased risk in the exposed group, *see* Ref. Man. at 567, and "[r]isks greater than 2.0 permit an inference that the [adverse effect] was more likely than not caused by the [drug]," *see Allison*, 184 F.3d at 1315 n.16.

[162] As discussed in Section II(C), the *p*-value is an indicator of statistical significance, which, in this context, provides an estimate of the probability that chance alone produced the association between Abilify, hypersexuality, impulsive behavior, and increased libido. Generally, *p*-values are considered "statistically significant" where they are less than or equal to 5% ($p \leq 0.05$). *See* Ref. Man. at 251; *see also Eastland v. Tenn. Valley Auth.*, 704 F.2d 613, 622 (11th Cir. 1983).

Based on Dr. Madigan's calculations, **[*** REDACTED ***]**. Dr. Madigan testified that the other five *p*-values, although not statistically significant, still indicate that the probability that chance alone explains the other relative risk ratios is "very slim." *See* Madigan Tr., ECF No. 596-4 at 37. **[*** REDACTED ***]**. According to Dr. Madigan, although the corresponding *p*-value, 0.08, is not statistically significant, it is still relevant because it indicates that the likelihood of the 2.7 relative risk being explained by chance is only 8%.[163] *See* Madigan Tr., ECF No. 596-4 at 38-39. Dr. Madigan characterizes his findings with respect to the hypersexuality, impulsive behavior, and increased libido events reported during the clinical trials as evidence of a "trend or concern." *See* Madigan Tr., ECF No. 596-4 at 39; *see also* Madigan Rep., ECF No. 427-1 at 29 ("This shows a concerning trend against [Abilify]."), **[*** REDACTED ***]**.

iii. Reliability

Defendants challenge the reliability of Dr. Madigan's methodology on multiple grounds.[164] Since the Court has already found that Dr. Madigan is not qualified to offer expert opinions on medical causation, several of Defendants'

[163] Dr. Madigan testified that **[*** REDACTED ***]**.

[164] The Court separately analyzed the reliability of Dr. Madigan's opinions regarding the Etminan Study in Section II(C).

reliability objections are now moot. The Court addresses the remaining objections in turn.

Defendants first challenge Dr. Madigan's expert opinions on the grounds that his initial expert report offers "no discernible methodology" as to how he reached his conclusions and that, at any rate, his opinions are not supported by any of the methodologies—epidemiology, dose-response, and background risk—that the Eleventh Circuit considers "indispensable" for proving that a drug can cause an adverse effect.[165] *See Chapman*, 766 F.3d at 1308. This is incorrect.

Dr. Madigan's opinions are very clearly based on the application of widely accepted statistical methods to data drawn from the "indispensable" field of epidemiology (*i.e.*, the Etminan Study), the FAERS database, and Defendants' clinical trials. *See* Madigan Rep., ECF No. 427-1 at 2-30; Madigan Supp., ECF No. 427-1 at 79-92.[166] Dr. Madigan's reports and testimony describe, in great detail, the precise steps he took to (1) evaluate the strengths and limitations of the Etminan Study; (2) to determine whether and to what extent a safety signal for pathological gambling existed for Abilify in the FAERS database; and (3) to assess whether Defendants' clinical trials were capable of detecting an association between

[165] To the extent this objection is directed, more broadly, at Dr. Madigan's opinion on general causation, the objection is denied as moot because Dr. Madigan's general causation opinion has been excluded.

[166] "Madigan Supp." refers to Dr. David Madigan's Rebuttal Report, ECF No. 427-1 at 79-92.

pathological gambling and Abilify. Notably, Defendants do not challenge the accuracy and reliability of Dr. Madigan's calculations or offer an expert statistician to contradict or refute his mathematical conclusions. Given the scientific support for Dr. Madigan's methodologies and the lack of rebuttal evidence to discredit any step in his analysis, the Court finds no basis for Defendants' argument that his methodology is not discernible or reliable enough to be admissible. To the contrary, in the Eleventh Circuit, mathematical and statistical analyses are well-recognized as reliable and acceptable means of supporting an expert opinion. *See City of Tuscaloosa v. Harcros Chemicals, Inc.*, 158 F.3d 548, 565-66 (11th Cir. 1998) (statistician's expert opinion reliable where based on "well-established" mathematical and statistical methods); *State of Ga., Dept. of Human Res. v. Califano*, 446 F. Supp. 404, 409 (N.D. Ga. 1977) (same). Because the Court has already found Dr. Madigan qualified to testify on the basis of statistics in this case, and because Dr. Madigan thoroughly explained the steps in his analysis, the Court finds his methodology sufficiently transparent and reliable to support his expert opinion on the strength of the statistical evidence in this case.

Defendants also raise arguments regarding Dr. Madigan's analysis of their clinical trial data. First, they challenge as "pure speculation" Dr. Madigan's opinion that the clinical trials were not powered to detect a statistically significant increased risk of pathological gambling, which Defendants appear to claim is based solely on

his observation that placebo subjects in the trials had no reports of pathological gambling. The premise of this argument is incorrect. When a study, such as the randomized clinical trials in this case, fails to find a statistically significant association between a drug and an adverse effect, an "important question is whether [that] result tends to exonerate the [drug's] toxicity or is essentially inconclusive with regard to toxicity." *See* Ref. Man. at 582. A statistical power analysis is a well-established scientific means of evaluating whether the study's outcome is exonerative or inconclusive. *See id*.[167] Again, "power" quantifies the ability of a study to detect a statistically significant association of a given magnitude, if it exists, in light of the sample sizes used in the study. *See id*.; *ASARCO*, 746 F.2d at 493 n.19. A power analysis depends on several factors, including the sample size, the level of statistical significance specified, the background incidence of the disease at issue, and the level of increased risk that scientist is testing whether the study would detect. *See id*. Using these factors, Dr. Madigan calculated the power of Defendants' clinical trial data and found that they simply were not large enough (*i.e*.,

[167] *See Kuhn*, 686 F.3d at 622 n.5 ("Power analysis can be used to calculate the likelihood of accurately measuring a risk that manifests itself at a given frequency in the general population based on the sample size used in a particular study."); *ASARCO*, 746 F.2d at 493 n.19 ("'Power' is a statistical concept which quantifies the ability of a study to detect an excess risk that truly exists."); *Cooley v. Lincoln Elec. Co.*, 693 F. Supp. 2d 767, 773 (N.D. Ohio 2010) ("[W]hen a study fails to find a statistically significant association, an important question is whether the result tends to exonerate the agent's toxicity or is essentially inconclusive with regard to toxicity, due to lack of sufficient statistical power."); *Smith v. Wyeth-Ayerst Laboratories Co.*, 278 F. Supp. 2d 684, 693 (W.D.N.C. 2003) ("[T]he concept of power is key because it's helpful in evaluating whether the study's outcome . . . is exonerative or inconclusive.).

sufficiently powered) to detect a statistically significant increased risk of pathological gambling with Abilify. *See* Madigan Rep., ECF No. 427-1 at 26-28. Accordingly, Dr. Madigan determined that that the clinical trials were essentially inconclusive with respect to the question of whether Abilify is associated with an increased risk of pathological gambling. *See id*.; *see also* Madigan Supp., ECF No. 427-1 at 85. Defendants do not dispute the accuracy of Dr. Madigan's statistical power calculation. Given Dr. Madigan's clear explanations of his source of data, statistical method, and his conclusions, his opinion as to the power of Defendants' clinical trial data cannot be considered "pure speculation."

Finally, Defendants challenge as unreliable Dr. Madigan's use of five statistically insignificant data points as the basis for his opinion that the clinical trials show a "concerning trend" of increased risk of impulsive behaviors with Ability patients.[168] *See* Madigan Rep., ECF No. 427-1 at 29. On this issue, the Court agrees. Statistical significance, by itself, does not mechanically control whether a statistical analysis is sufficiently reliable under *Daubert*. *See In re Viagra Products Liability Litigation*, 572 F. Supp. 2d 1071, 1081 (D. Minn. 2008) ("There is persuasive authority stating that on a *Daubert* motion involving general-causation evidence in an MDL matter, lack of statistical significance under some circumstances does not detract from the reliability of the study."). However, federal courts have routinely

[168] **[*** REDACTED ***]**. *See* Madigan Rep., ECF No. 427-1 at 29.

required that statistically insignificant evidence bear other indicia of scientific reliability to be admissible. *See Lipitor*, 174 F. Supp. 3d at 926 (excluding expert opinion where plaintiffs failed to show that use of non-statistically significant "trend" data was generally accepted in the relevant field, supported by peer-reviewed literature, or governed by statistical standards); *Zoloft I*, 26 F. Supp. 3d at 456-57 (excluding expert opinion where expert failed to show that reliance on "trends" in statistically insignificant findings is accepted within her scientific community).[169]

In this case, the strength of Dr. Madigan's opinions lies in the statistical context they provide, which is based primarily on his assessments of the statistical significance of various categories of scientific evidence. For example, Dr. Madigan's defense of the Etminan Study is premised, in large part, on its "highly statistically significant results." *See* Madigan Rep., ECF No. 427-1 at 30. Similarly, according to Dr. Madigan, the statistically significant percentages of pathological gambling reports in the FAERS database evidence their "very strong" association with Abilify. *See* Madigan Tr., ECF No. 596-4 at 23-24. Dr. Madigan cannot now,

[169] *See also Joiner*, 522 U.S. at 145-47 (holding that study showing a statistically insignificant increase in disease incidence following exposure to the alleged causal chemical could properly be excluded as a foundation for an expert's opinion); *Pluck v. BP Oil Pipeline Co*., 640 F.3d 671, 680 (6th Cir. 2011) (affirming district court's exclusion of expert evidence because, *inter alia*, the expert relied on studies with statistically insignificant results); *United States v. Morrow*, 374 F. Supp. 2d 51, 68 (D.C. Cir. 2005) (holding that "even DNA evidence with relatively low statistical significance may be admitted as probative evidence, provided that certain safeguards are afforded"); *Kadas v. MCI Systemhouse Corp*., 255 F.3d 359, 362-63 (7th Cir. 2001) (stating that the level of statistical significance required to present a particular study to the factfinder depends on the context of the study and the case).

with respect to the statistically insignificant "trends" in the clinical trial data, abandon statistical significance as a measure of reliability without a thorough explanation of why doing so is sound and supportable scientific practice. *See Lipitor*, 174 F. Supp. 3d at 926; *Zoloft I*, 26 F. Supp. 3d at 456-57. Without more, the statistically insignificant data undoubtedly will tend to confuse the issues and mislead the jury. *See* Fed. R. Evid. 403, 702. The Court thus finds that Dr. Madigan's five statistically insignificant findings from the clinical trials, and also his characterization of those findings as a trend, must be excluded as unreliable.[170]

In sum, Defendants' Motion to Exclude the General Causation Opinion of David Madigan, ECF No. 427-20, is due to be granted in part and denied in part. Dr. Madigan may not offer an expert opinion on medical causation and also may not testify about the five statistically insignificant *p*-values he calculated from the clinical trial data. In all other respects, his expert opinion is admissible.

2. Defendants' Experts

Plaintiffs have moved to exclude the opinions of Defendants' five proposed experts—Drs. Blier, Leiderman, Potenza, Weed, and Winstanley—on multiple grounds. With respect to Drs. Blier, Potenza, and Weed, the Court has carefully

[170] The Court notes that Defendants have not objected to Dr. Madigan's reliance on the one statistically significant *p*-value (.03) he calculated from the clinical trials, based on the 6.16-fold increase in the risk of impulsive behavior with Abilify relative to comparator drugs. This *p*-value is reliable and admissible.

considered Plaintiffs' arguments and finds them to be lacking in merit.[171] As an initial matter, Plaintiffs have not challenged these three experts' qualifications to testify in this case and the Court finds, from their reports and testimony at the *Daubert* hearing, that each is amply qualified to offer the expert opinions provided. The Court also finds the three expert opinions reliable, except to the extent otherwise indicated in this Order.[172] Briefly, each of the three experts prepared a standard report of the type the Court would expect to see in response to Plaintiffs' experts' reports. *See* Blier Rep., ECF No. 455-1; Potenza Rep., ECF No. 458-1; Weed Rep., ECF No. 419-3. Their opinions were, essentially, critiques of Plaintiffs' experts' evidence, methodologies, and conclusions. *See id*. This was entirely appropriate. There is no requirement that a defense expert offer a competing general causation opinion, for example; his opinions properly may be limited to criticizing the analysis

[171] With two exceptions, Plaintiffs' criticisms of Dr. Weed affect only the weight to be afforded his opinion, not its admissibility. First, Dr. Weed will not be permitted to testify that the Bradford Hill factor of consistency can only be satisfied by the existence of multiple epidemiological studies, *see* Weed Rep., ECF No. 419-3 at 45, because this opinion is not supported by the scientific literature. *See, e.g.*, Ref. Man. at 604; Bradford Hill Article, ECF No. 460-4. Second, Dr. Weed will not be permitted to represent the AMSTAR criteria as the "minimum requirements for a scientifically rigorous systematic review," such as the literature reviews performed by Plaintiffs' experts. See id. at 33. The record reflects that the developers of AMSTAR recommend "further testing" of their system before "strong recommendations can be made on its use." *See Shea, et al., Development of AMSTAR: A Measurement Tool to Assess the Methodological Quality of Systematic Reviews*, 7 BMC MED. RESEARCH METHODOLOGY 10 (2007), ECF No. 419-35 at 4,7.

[172] The Court finds Dr. Blier's opinion reliable except with respect to his opinion as to the high concentrations of D_1 and D_2 receptors in the nucleus accumbens, which the Court has already excluded as inconsistent with the record. *See* Section II(C)(4)(a).

and conclusions presented by another party. *See In re Zyprexa Products Liability Litigation*, 489 F. Supp. 2d 230, 285 (E.D.N.Y. 2007) ("[D]efendant's experts have a less demanding task, since they have no burden to produce models or methods of their own; they need only attack those of plaintiff's experts."). In this case, Drs. Blier, Potenza, and Weed identified alleged shortcomings in Plaintiffs' experts' opinions, provided a reasoned basis for each criticism, and furnished reference materials in support of their positions. The Court is satisfied that their opinions are sufficiently grounded in science to render them reliable, and thus admissible, under *Daubert*.

The Court next addresses, in turn, the admissibility of the proposed expert testimony of Drs. Leiderman and Winstanley.

a. Deborah B. Leiderman, M.D., M.A., FAAN

Dr. Deborah B. Leiderman is a licensed physician and board-certified neurologist. She has extensive experience in clinical research, as well as drug development, regulation and policy, including just over seven years at the FDA as the Director of the Controlled Substances Staff, during which time she served as the agency's "lead physician and official on issues related to the Controlled Substances Act, abuse liability assessment, and domestic international drug scheduling and

prescription drug abuse." Leiderman Rep., ECF No. 420-3 at 2-3.[173] In this role, she "consulted on proposed and draft language for many drug labels across multiple therapeutic areas." *Id.* She also spent ten years at the National Institutes of Health, where she was responsible for "various aspects of new drug clinical development including clinical trial design, clinical trial conduct and oversight, safety, and drug labeling." *Id.* Through her own business, CNS Consulting, LLC, she provides consulting services related to the clinical and regulatory aspects of drug development. *See* Leiderman Dep., ECF No. 456-2 at 9.[174] Based on this experience, Dr. Leiderman offers the following opinions to rebut Plaintiffs' experts' assertions that Abilify can cause pathological gambling and other impulse control disorders: (1) the approved FDA warning label does not support the conclusion that Abilify causes compulsive behaviors; (2) when the FDA determines that a drug product causes a particular adverse effect, FDA safety communications will so indicate and the FDA will require clear language in the revised approved product label; and (3) the FDA's comprehensive Pharmacovigilance Review of March 2016 and its May 2016 Drug Safety Communication do not support the conclusion that Abilify causes compulsive behaviors or impulse control disorders. *See* Leiderman

[173] "Leiderman Rep." refers to Dr. Deborah B. Leiderman's Expert Report, ECF No. 420-3.

[174] "Leiderman Dep." refers to the official transcript of Dr. Deborah B. Leiderman's deposition testimony on June 21, 2017, ECF No. 456-2.

Rep., ECF No. 420-3 at 4-11; Leiderman Dep., ECF No. 456-2 at 19-20. Plaintiffs attack Dr. Leiderman's opinions on numerous reliability grounds, namely: (1) that she fails to use any intelligible methodology to reach her conclusions, (2) her opinions are not supported by the evidence she relies on, and (3) she demonstrates a lack of understanding of FDA regulations.[175]

The main thrust of Dr. Leiderman's opinions is that the FDA warning label and Pharmacovigilance Review concerning Abilify "do not support the conclusion that Abilify causes compulsive behaviors or impulse control disorders."[176] *See* Leiderman Rep., ECF No. 420-3 at 5, 7. As a threshold matter, Plaintiffs' experts will not be permitted to testify at trial that the FDA warning label and Pharmacovigilance Review, standing alone or together, are definitive proof of causation. *See Rider*, 295 F.3d at 1201 (affirming district court's determination that FDA statement withdrawing approval of drug's indication for the prevention of lactation could not, itself, prove causation). However, Plaintiffs' experts may rely on these materials as part of their Bradford Hill and/or weight-of-the-evidence

[175] Although Plaintiffs argue that Dr. Leiderman failed to properly explain how her experience led her to the conclusions she reached in this case, they have not specifically argued that Dr. Leiderman is not qualified to opine on matters pertaining to FDA regulations.

[176] Dr. Leiderman testified that she is not a causation expert and has not offered a general causation opinion that Abilify can cause pathological gambling or other impulse control disorders. *See* Leiderman Dep., ECF No. 456-2 at 19-20. She stated that her opinion only "address[ed] FDA's interpretations of the data and their findings of an association and not causality." *See id.* Dr. Leiderman's position is that the "FDA has not found evidence and data to be sufficient to reach" a finding of causation and that "that's what [her opinion is] describing." *See id.*

causation analyses.[177] The FDA warning label and Pharmacovigilance Review are both a part of the body of scientific evidence on Abilify. To suggest they cannot support a general causation analysis in any way is incorrect. *See Neurontin*, 612 F. Supp. 2d at 137 (concluding that FDA findings and decision to require warning label are not "definitive proof" of causation, but nonetheless may support a causation opinion).

With respect to Dr. Leiderman's opinions, the Court finds that she may testify to the purpose behind the FDA's pharmacovigilance process and how it is conducted; however, she may not testify about language in a warning label or pharmacovigilance review unless her testimony is supported by a specific FDA regulation, rule, policy, or official agency guidance (*e.g.*, FDA Pharm. Guide, DX-15). In other words, Dr. Leiderman may not simply read FDA materials to the jury and then testify to what the FDA meant or intended by including or excluding certain language.[178] The FDA warning label and Pharmacovigilance Review for Abilify

[177] *See, e.g.*, Glenmullen Rep., ECF No. 424-1 at 132; Hollander Rep., ECF No. 459-1 at 36; Luepker Rep., ECF No. 462-1 at 7.

[178] *See, e.g., Jones v. Novartis Pharmaceuticals Corp.*, 235 F. Supp. 3d 1244, 1254 (N.D. Ala. 2017) (allowing former FDA official to offer expert testimony about the FDA regulatory process, but excluding her proposed testimony as to the existence of either causation or a causal association, or as to the drug company's knowledge, state of mind, intent, or motive); *In re Fosamax Products Liab. Litig.*, 645 F. Supp. 2d 164, 192 (S.D.N.Y. 2009) (excluding expert testimony "as to the knowledge, motivations, intent, state of mind, or purposes of [the defendant drug company], its employees, the FDA or FDA officials"); *Rezulin*, 309 F. Supp. 2d at 546 (excluding proposed expert testimony regarding the intent, motives, or states of mind of corporations or regulatory agencies); *In re Diet Drugs*, No. MDL 1203, 2001 WL 454586, at *2

speak for themselves. Dr. Leiderman's opinions will be limited to the FDA regulatory process and what can or cannot be concluded about Abilify from that process based on specific, established FDA regulations, rules, policies, or official guidance. She may draw inferences based on her experience, but only if the inferences are reasonably supported by FDA regulations, rules, policies, or guidance. Plaintiffs' *Daubert* Motion to Exclude the Testimony of Defendants' Expert Deborah B. Leiderman, M.D., M.A., FAAN, ECF No. 420, is granted in part and denied in part, as discussed above.

b. Catharine A. Winstanley, Ph.D.

Dr. Catharine Winstanley is a professor in the psychology department at the University of British Columbia. She holds an undergraduate degree in psychology and physiology, as well as a doctorate in behavioral neuroscience. She has extensive professional experience in the design and use of *in vivo* studies of "impulsivity and risky decision-making" in rodents to investigate and understand these conditions in humans. *See* Winstanley Rep., ECF No. 461-1 at 2.[179] In brief, Dr. Winstanley offers an opinion that none of the available *in vivo* data supports Plaintiffs' experts' conclusions that Abilify can cause gambling disorder or impulsivity. Plaintiffs

(E.D. Pa. Feb. 1, 2001) (excluding all proposed expert testimony concerning the intent of the defendant drug company or any other entity, such as the FDA).

[179] "Winstanley Rep." refers to Dr. Catharine A. Winstanley's Expert Report, ECF No. 461-1.

challenge Dr. Winstanley's testimony on a number of grounds, most notably, her lack of qualifications to testify on general causation and her failure to adequately explain how and why *in vivo* findings about the effects of various drugs on impulsivity in rodents may be reliably extrapolated to prove that Abilify would have comparable effects on gambling and impulsivity in humans.

As to Dr. Winstanley's qualifications, the Court agrees that she is not qualified to offer a comprehensive general causation opinion. She is not a medical doctor, toxicologist, pharmacologist, or epidemiologist, and she has no specialized knowledge of, or clinical experience with, Abilify, pathological gambling, or impulse control disorders in humans. Dr. Winstanley's admitted lack of expertise in the aforementioned areas precludes her from offering a medical or scientific opinion that Abilify cannot cause impulse control problems. However, given Dr. Winstanley's knowledge and experience with animal models of impulsivity, the Court finds her qualified to offer an expert opinion regarding any such evidence that is ultimately deemed admissible in this case.

Dr. Winstanley's opinion is based on her review of *in vivo* studies investigating impulse control and gambling-related behavior in rodents, studies that she considers contradictory to Plaintiffs' experts' opinions regarding the biological mechanism by which Abilify can cause gambling and impulsivity problems in

humans. *See* Winstanley Tr., ECF No. 596-8 at 107.[180] Dr. Winstanley refers to these *in vivo* studies as "translationally valid," by which she means that the behavioral test each study used to assess impulsivity in rodents is a scientifically accepted, valid model of impulsivity in humans. *See* Winstanley Tr., ECF No. 596-8 at 110-11. Although Dr. Winstanley defines, in general terms, the three components of a translationally valid animal model, it is not entirely clear from her expert report or her testimony how or why the animal models she cites in this case meet that definition.[181] This is problematic because, as gatekeeper for the expert evidence presented to the jury, the Court must ensure that any extrapolations from rodents to humans are based on more than just the *ipse dixit* of an expert. *See McDowell*, 392 F.3d at 1299 (noting "there is no fit where a large analytical leap must be made between the facts and the [proposed expert's] opinion," such as the proffer of animal studies concerning one type of cancer in mice to establish a different type of cancer in humans).

[180] "Winstanley Tr." refers to the official transcript of Dr. Catharine A. Winstanley's testimony at the *Daubert* hearing, ECF No. 596-8 at 93-11.

[181] According to Dr. Winstanley, the three components of a translationally valid animal model are: (1) face validity, which refers to how much the behavioral task designed for use in rodents resembles, on its face, a behavioral task that a human subject might perform; (2) construct validity, which refers to whether the same brain regions and neurotransmitters that regulate a particular behavior in humans regulate that same behavior in rodents; and (3) predictive validity, which refers to whether rodent models of a particular behavioral task have been shown to accurately predict human performance on the same task. *See* Winstanley Rep., ECF No. 461-1 at 7-8; Winstanley Tr., ECF No. 596-8 at 96-97.

Dr. Winstanley testified that it is "impossible" to translationally validate any model of gambling disorder in rodents and that, to her knowledge, Abilify has never been tested in any rodent models of gambling-related behavior. *See* Winstanley Tr., ECF No. 596-8 at 99, 108. Therefore, to the extent she has extrapolated from the *in vivo* studies in this case any conclusions about gambling disorder or gambling-related behavior in humans, the Court finds those conclusions inadmissible at trial. Dr. Winstanley also offers opinions based on the findings of *in vivo* studies involving drugs that are pharmacologically different from Abilify.[182] However, she failed to offer any evidence establishing the reliability of an analogy between those other drugs and Abilify. *See Rider*, 295 F.3d at 1200-01 (stating that extrapolations between drugs are impermissible unless reliable scientific evidence establishes the validity of the analogy). Therefore, the *in vivo* studies concerning other drugs (*i.e.*, not Abilify) are inadmissible and may not support Dr. Winstanley's opinion. Finally, with respect to the remaining *in vivo* studies (*i.e.*, those involving Abilify), the Court has lingering questions about how and why the behavioral tests used on

[182] *See, e.g.*, P. Cocker *et al.*, *Chronic Administration of the Dopamine D2/3 Agonist Ropinirole Invigorates Performance of a Rodent Slot Machine Task, Potentially Indicative of Less Distractible or Compulsive-Like Gambling Behavior*, 234 PSYCHOPHARMACOLOGY 137 (2017) (ropinirole); Rokosik & Napier, *Pramipexole-Induced Probabilistic Discounting: Comparison Between a Rodent Model of Parkinson's Disease and Controls*, 37 NEUROPSYCHOPHARMACOLOGY 1397 (2012) (pramipexole); L. Cervo *et al.*, *Selective Antagonism at Dopamine D3 Receptors Attenuates Cocaine-Seeking Behaviour in the Rat*, 10 INT'L J. NEUROPSYCHOPHARMACOLOGY (2007) (a D_3 receptor antagonist); P. Flores & R. Pellón, *Antipunishment Effects of Diazepam on Two Levels of Suppression of Schedule-Induced Drinking in Rats*, 67 PHARMACOLOGY, BIOCHEMISTRY & BEHAVIOR 207 (2000) (diazepam).

rodents are translationally valid, such that it is reliable to extrapolate from them a conclusion about Abilify's effects in humans. At the *Daubert* hearing, Dr. Winstanley spent a total of 21 minutes on the witness stand.[183] *See* Winstanley Tr., ECF No. 596-8 at 93-11. This was neither party's fault; rather, time constraints and the late hour at which her testimony began limited each side's ability to thoroughly flesh out Dr. Winstanley's positions. Under these circumstances, the Court finds it appropriate to defer ruling on the reliability of this single aspect of Dr. Winstanley's opinion and give Defendants an opportunity to recall her at the next *Daubert* hearing, at which time Plaintiffs may also cross-examine her on the issue. Accordingly, Plaintiffs' Daubert Motion to Exclude the Testimony of Defendants' Expert Catharine Winstanley, Ph.D., ECF No. 422, is due to be granted in part, denied in part, and deferred in part, as discussed above.

E. Conclusion

When ruling on challenges to expert testimony under Rule 702 and *Daubert*, the Court is charged with the responsibility of acting as a gatekeeper, excluding "junk science" and other unreliable information, *see Joiner*, 522 U.S. at 153, but allowing—for a jury's consideration—testimony that derives from sound scientific knowledge, methodologies, and reasoning, *see Daubert*, 509 U.S. at 590. While the

[183] For frame of reference, Plaintiffs' expert, Dr. Bechara, spent over two hours on the witness stand. *See* Bechara Tr., ECF No. 596-3.

Daubert inquiry must be "rigorous," it "is not intended to supplant the adversary system or the role of the jury." *Rink*, 400 F.3d at 1291, 1293 n.7. Again, only the jury may determine "where the truth in any case lies" and the court "may not usurp this function." *See Frazier*, 387 F.3d at 1272. Accordingly, a district court may not "evaluate the credibility of opposing experts" or "the persuasiveness of competing scientific studies." *Quiet Tech.*, 326 F.3d at 1341. For *Daubert* purposes, the Court's duty is limited to ensuring that a proposed expert's testimony is based on "sound and reliable evidence." *Frazier*, 387 F.3d at 1272.

In this case, Plaintiffs have shown, by a preponderance of the evidence, that their general causation evidence is sound and reliable. For starters, there is reliable evidence of a broad scientific consensus regarding the existence of an association between Abilify and increased risk of impulse control problems.[184] The FDA, EMA, and Health Canada have all concluded as much, based on their reviews of largely the same scientific literature and statistical analyses discussed in this Order, and, as a result, have required that safety warnings be added to the Abilify product labels. In 2015, **[*** REDACTED ***]**. Several months later, **[*** REDACTED ***]**. The FDA's 2016 pharmacovigilance review "confirm[ed]" Defendants' conclusions as to a "possible causal association" between Abilify use and impulse control

[184] For the Court's discussion of the distinction between an association and causation, see Section II(B)(1)(a).

disorders, and it also called for a case-control study to "help clarify" the nature of the association. *See* FDA Pharm. Vigil., ECF No. 428-11 at 29-30. The Etminan Study—an epidemiological case-control study that is peer-reviewed, published, and unchallenged in the scientific literature to date—reliably confirms the association as causal. Moreover, the biological mechanism by which Abilify can cause impulse control problems has been reliably established by peer-reviewed, published scientific literature, and notably, Plaintiffs' experts' biological plausibility opinions are consistent with both the FDA and Defendants' assessments of Abilify's mechanism of action. *See* Section (II)(C)(4)(e). Defendants' experts hotly dispute Plaintiffs' general causation evidence, but a hot dispute is not a basis for excluding Plaintiffs' experts' opinions. "[T]he subject of scientific testimony [need not be] 'known' to a certainty; arguably, there are no certainties in science." *Daubert*, 509 U.S. at 590. Instead, the scientific evidence must only be sound and reliable. Disputes over the relative persuasiveness of either sides' reliable evidence "should be tested by the adversary process—competing expert testimony and active cross-examination—rather than excluded from a jury's scrutiny." *Neurontin*, 612 F. Supp. 2d at 159 (quoting *Ruiz-Troche v. Pepsi Cola of Puerto Rico Bottling Co*., 161 F.3d 77, 85 (1st Cir. 2009)). In sum, Plaintiffs have demonstrated their experts' opinions as to general causation are admissible under Rule 702 and *Daubert*.

III. Summary Judgment

Defendants have moved for summary judgment on general causation under *Daubert* based on Plaintiffs' lack of admissible expert testimony. Because the Court has found that most of Plaintiffs' evidence on general causation—including epidemiology (*i.e.*, Etminan Study), background risk, biological plausibility, disproportionality analyses, *in vivo* and *in vitro* studies, voluminous case and adverse event reports (including dose-response, dechallenge, and rechallenge events), FDA materials, Defendants' investigative findings, and Plaintiffs' experts' Bradford Hill and weight-of-the-evidence analyses—satisfies Rule 702 and *Daubert*, there exists a genuine dispute of material fact on the issue of whether Abilify can cause uncontrollable impulses in individuals taking the drug. Therefore, Defendants' Motion for Summary Judgment on General Causation, ECF No. 428, is due to be denied.

Accordingly, it is **ORDERED**:

1. Plaintiffs' *Daubert* Motion to Exclude the Testimony of Defendants' Expert Marc N. Potenza, ECF No. 415, is **DENIED**.

2. Plaintiffs' *Daubert* Motion to Exclude the Testimony of Pierre Blier, M.D., Ph.D., ECF No. 418, is **GRANTED** in part and **DENIED** in part, as discussed in the body of this Order.

3. Plaintiffs' *Daubert* Motion to Exclude the Testimony of Defendants' Expert Douglas Weed, M.D., ECF No. 419, is **GRANTED** in part and **DENIED** in part, as discussed in the body of this Order.

4. Plaintiffs' *Daubert* Motion to Exclude the Testimony of Defendants' Expert Deborah B. Leiderman, M.D., M.A., FAAN, ECF No. 420, is **GRANTED** in part and **DENIED** in part, as discussed in the body of this Order.

5. Plaintiffs' *Daubert* Motion to Exclude the Testimony of Defendants' Expert Catharine Winstanley, Ph.D., ECF No. 422, is **GRANTED** in part, **DENIED** in part, and **DEFERRED** in part, as discussed in the body of this Order.

6. Defendants' Motion to Exclude the General Causation Opinion of Antoine Bechara, ECF No. 423, is **GRANTED** in part and **DENIED** in part, as discussed in the body of this Order.

7. Defendants' Motion to Exclude the General Causation Opinion of Joseph Glenmullen, ECF No. 424, is **GRANTED** in part and **DENIED** in part, as discussed in the body of this Order.

8. Defendants' Motion to Exclude the General Causation Opinion of Eric Hollander, ECF No. 425, is **DENIED**.

9. Defendants' Motion to Exclude the General Causation Opinion of Russell Luepker, ECF No. 426, is **GRANTED**.

10. Defendants' Motion to Exclude the General Causation Opinion of David Madigan, ECF No. 427, is **GRANTED** in part and **DENIED** in part, as discussed in the body of this Order.

11. Defendants' Motion for Summary Judgment on General Causation Based on Plaintiffs' Lack of Admissible Expert Testimony Under Daubert, ECF No. 428, is **DENIED**.

SO ORDERED, on this 15th day of March, 2018.

M. Casey Rodgers

M. CASEY RODGERS
CHIEF UNITED STATES DISTRICT JUDGE