UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE: Acetaminophen – ASD-ADHD Products Liability Litigation **This Document Related To: All Cases**	**22md3043 (DLC)**

MEMORANDUM OF LAW IN SUPPORT OF PLAINTIFFS' RULE 702 MOTION TO EXCLUDE DR. CRAIG M. POWELL

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Cases: **Page**

Other Authorities:

INTRODUCTION

Plaintiffs' experts' opinion that prenatal exposure to acetaminophen ("APAP") is capable of causing autism spectrum disorder ("ASD") and attention-deficit/hyperactivity disorder ("ADHD") in offspring is based on robust evidence, including the 26 out of 27 peer-reviewed preclinical in vivo experiments finding measurable neurodevelopmental changes in animals exposed to APAP at periods corresponding to human gestational neurodevelopment.[1] Plaintiffs' experts did not evaluate these findings in isolation. They viewed them in the context of a large body of epidemiological evidence consistently showing an association between prenatal APAP exposure and ASD/ADHD in humans.

Rather than engage with this data, Dr. Powell waves it off, decrying "speculative theories like those in this litigation," going so far as to accuse Plaintiffs' experts of "the misuse of science." Ex. 1, Amended Expert Report of Craig M. Powell ("Powell Report"), ¶ 228. But science is an invaluable tool to uncover the truth. The only misuse of that discipline is to refuse to engage with substantial and compelling data that contradicts ill-informed conventional wisdom. That is precisely what Dr. Powell did. Of the 99 preclinical studies related to prenatal APAP exposure that he identified, Dr. Powell deems 97 of them irrelevant. That glaring act of cherry-picking cannot defeat Plaintiffs' position.

The first study Dr. Powell deigned to consider, Saad 2016, finds that prenatal APAP exposure affects neurodevelopment and is capable of causing ASD and ADHD. This despite attempting to negate the effect through an inexplicable statistical correction that Dr. Powell does not even use in his own work. The second study, Baker 2023, is one of the stronger animal studies

[1] Preclinical research is done in a laboratory and includes animal studies (in vivo), cell-culture studies (in vitro), studies of organs and tissue outside the human body (ex vivo), and computational studies (in silico).

to show the biological association between prenatal APAP exposure and ASD/ADHD. Even if Dr. Powell had a reliable basis for disregarding the vast majority of relevant preclinical evidence (and he does not), he should have "grave concerns" not about Plaintiffs' experts' opinions, but about the neurodevelopmental impact of the product he is being paid to defend. Ex. 1, Powell Report ¶ 228. Dr. Powell's report thus fails to support its own rhetoric, at least as applied to Plaintiffs' experts rather than to the report itself.

All agree that the relevant questions regarding the preclinical evidence presented in this case are whether that evidence supports the neurodevelopmental effect of APAP exposure observed in humans, specifically in the development of ASD and ADHD in offspring and locates plausible biological mechanisms for that effect. Indeed, that is the analysis that Dr. Powell insisted at deposition that he conducted in his report. But the report speaks for itself and contains no such analysis. No weighing of any preclinical evidence—including most of the preclinical data analyzed by Plaintiffs' experts Dr. Brandon Pearson and Dr. Robert Cabrera—is to be found in Dr. Powell's review. His project was, instead, to exclude as much of that evidence as he could.

Dr. Powell's opinions fail the *Daubert* standard. His exclusionary approach does not reliably account for the vast majority of preclinical evidence that conflicts with his conclusions about APAP's neurodevelopmental effects and mechanisms of action. Indeed, his methodology is not clear from his own report, as evidenced by his belated efforts at deposition to recharacterize his methodology as a weight-of-evidence methodology, which it is not. Even were his methodology clear, it suffers from fundamental flaws: Dr. Powell deems irrelevant several animal studies supporting the observed association between prenatal APAP exposure and ASD/ADHD on the ground that the test animals received doses of APAP that are higher than the therapeutic dose recommended for humans. But his own dosing threshold is simply miscalculated and, therefore,

arbitrarily low. This is compounded by another foundational error: Powell's original report excluded numerous relevant studies for failing to apply a statistical correction to their data that was, in fact, applied. He attempted to rectify this in his Amended Report by simply removing the errant endnotes from the string citation without describing how this remarkably different understanding of the data affected his opinions. *See* Ex. 3, Powell Dep. Tr. at 99:20–100:8; Ex. 2, Redline of Amended Powell Report ¶ 72, 91d. Reliable expert testimony cannot start from a flawed premise.

Shorn of a methodology, Dr. Powell can draw no reliable conclusions about what the preclinical evidence shows. Nor can he reliably criticize Plaintiffs' experts' opinions about the weight of that evidence, which, in contrast to his opinions, are based on objective and transparent—that is, reliable and scientific—methodologies. The only way Dr. Powell can reach the contrary conclusions he does is by disregarding the evidence Defendants do not like. Dr. Powell's report and testimony should be excluded.

LEGAL STANDARD

Plaintiffs refer the Court to the Rule 702 legal standard set forth in Plaintiffs' Memorandum in Support of their Rule 702 Motion to Exclude Dr. Wendy Chung, Dkt. 1138 at 3–5.

ARGUMENT

Dr. Powell's report offers two related opinions: that the published animal studies do not support either the association between prenatal APAP exposure and ASD/ADHD or a biological mechanism underlying that association, and consequently, that this body of evidence does not support Plaintiffs' experts' opinions to the contrary. Reasonable scientists can, of course, disagree. What matters under Rule 702 is whether their opinions are based on reliable scientific methodologies. Neither Dr. Powell's opinions about the association between prenatal APAP

exposure and ASD/ADHD or a biological mechanism underlying that association nor his conclusory assertions to Plaintiffs' experts' opinions pass this test.

I. Dr. Powell's "Systematic Review" Is Unreliable.

For his assessment of the animal studies, Dr. Powell opted to use a "systematic review." In contrast to the weight-of-evidence methodologies used by Plaintiffs' experts (which are themselves systematic), Dr. Powell's version of "systematic review" and the sources upon which he relied to create this "methodology" are aimed at a specific context, namely the study of therapeutics in human trials. Dr. Powell describes his methodology as "an amalgam of Gurusamy, et al., and the ARRIVE guidelines, and my own list of what I think was most important and critical in terms of scientific experimentation." *See* Ex. 3, Powell Dep. Tr. at 131:4–7. The first of those sources, Gurusamy, describes a "tool" that is "***currently being piloted***" and "intended for all preclinical researchers and clinical researchers considering translation of preclinical findings to first-in-human clinical trials." *See* Ex. 4, Gurusamy et al. (2021) at 2–3 (emphasis added). The scope of the tool is limited to assessing "the likelihood that therapeutic preclinical findings can be translated into improvement in the management of human diseases" and is explicitly "*not* [intended] for assessment of the quality of the study." *Id*. at 10 (emphasis added). Other preclinical systematic review guidance documents Dr. Powell relies upon are similarly limited in scope. *See* Ex. 1, Powell Report ¶ 58 (citing nn.39 & 41, additional sources describing systematic review as intended for determining whether a therapeutic tested in animals should move on to human testing). The ARRIVE 2.0 guidelines, for their part, are simply "a checklist of information to include in a manuscript to ensure that publications contain enough information to add to the knowledge base." Ex. 5, Percie du Sert et al. (2020), at 2; *see also* Ex. 1, Powell Report ¶ 58 (citing nn.17 (Percie du Sert 2020), 37, 38, 42 (further sources with similar descriptions)). In other

words, the ARRIVE guidelines are a checklist of information that should be included in the "methods" section of a study manuscript – not a methodology for systematically evaluating data from a variety of studies to determine a compound's potential neurotoxic hazard. Dr. Powell admits as much. *See* Ex. 3, Powell Dep. Tr. at 127:19–24.

When assessing whether a therapeutic tested in animals is safe to test in humans, it makes perfect sense to be highly cautious—or, in Dr. Powell's description of his "systematic review," to "narrow the evidence" and ensure that drugs are tested in humans based only on the most definitive preclinical evidence. *See id.* ¶ 58. But when the question is not one of therapeutics and instead one of toxicology—*i.e.*, whether a drug that already is commonly taken by humans is causing adverse health outcomes—the approach is necessarily different. As established in the preclinical systematic-review guidance literature, such reviews should take in "the totality of evidence from animal studies, just as is done for studies on humans, rather than giving excessive weight to any one or two animal test results in one species." Ex. 6, Hoojimans & Ristkes-Hoitinga (2013) at 3.

Dr. Powell recognized as much at deposition, insisting that, after narrowing the animal evidence of the association between prenatal APAP exposure and ASD/ADHD, he nevertheless assessed the association based on the totality of that evidence. *See* Ex. 3, Powell Dep. Tr. at 132:12–25. Not only is this supposed second step of his analysis reflected nowhere in his report, nowhere does he identify what standards he used at this step. At the same time, he acknowledges that the ARRIVE guidelines are not meant to evaluate causation and that he did not use them to do so. *See* Ex. 3, Powell Dep. Tr. at 127:18–24. Perhaps an expert could fashion a relevant and reliable methodology to assess causation in this case out of the remaining systematic-review guidance literature. But that remains hypothetical because Dr. Powell did not do so. Systematic review requires "systematic and explicit methods to identify, select, and critically appraise relevant

research, and to collect and analyze data from the studies that are included in the review." Ex. 7, Avey et al. (2015) at 154. Above all, the review requires "transparency and objectivity of the techniques used." Ex. 8, Sena, et al. (2014), at 738. Rather than review the totality of the animal evidence in this manner, Dr. Powell gave any real consideration to only two studies (namely, the Saad 2016 study and the Baker 2023 study, which he misinterprets) and dismissed all other relevant animal studies on opaque or clearly erroneous grounds.

A. Dr. Powell's Report Relies on Basic Errors.

To begin with, Dr. Powell's analysis depends in significant part on basic errors. The first is a basic math error. Dr. Powell deemed several studies irrelevant because test animals received doses of APAP that were purportedly in excess of the equivalent therapeutic dose for humans, *see* Ex. 1, Powell Report ¶¶ 74, 91a, which he places at 200 milligrams (of APAP) per kilogram (of body weight) per day in rats and mice, *see id*. ¶ 56. This threshold, as stated, is artificially low.

Dr. Powell calculated this daily dose threshold based on the maximum *single* therapeutic APAP dose for humans (1,000 mg). *See* Ex. 3, Powell Dep. Tr. at 88:3–10, 207:2–7, 221:20–222:10. The maximum *daily* human therapeutic dose is 4,000 mg (four single doses every six hours). By Dr. Powell's own math, the maximum single APAP dose for a human, based on a body-surface calculation, would equate to approximately 90 milligrams per kilograms in a rat. *See id*. at 207:2–7. Multiplying by four to reach an equivalent *daily* dose for the rat yields a dose of approximately 360 milligrams per kilograms per day. The threshold would be even higher for mice. *See* Ex. 9, FDA, *Guidance for Industry*, at 7 tbl.1 ("Conversion of Animal Doses to Human Equivalent Doses Based on Body Surface Area").

Thus, Dr. Powell's 200 mg/kg/day threshold is in no way "conservative." *See* Ex. 1, Powell Report ¶ 56. It is arbitrary to exclude studies without explanation for using a translationally

relevant *daily* dose simply because they exceed Dr. Powell's artificially low threshold which is based on a *single* dose. And it is doubly arbitrary to apply the same maximum dose threshold to both mice and rats given that, as Dr. Powell acknowledges, rats are not in danger of hepatotoxicity (liver toxicity) at the translationally relevant dose. *See id*. ¶ 55 (explaining that hepatotoxicity is observed in rats at APAP doses of 500+ mg/kg/day). Dr. Powell also disputes that translationally relevant doses should be calculated by body-surface-area conversions (also known as allometric scaling) in this context, noting that the above-cited FDA guidance document is specifically focused on first-in-human toxicity studies of new drugs, and the liver toxicity of APAP is already well-understood. *See id*. ¶ 54–55. But that directly conflicts with his reliance on such first-in human sources as a basis of his own methodology, as noted above. In any event, the liver is not the issue here; neurodevelopment is. And nothing in the FDA's guidance or other guidance literature supports Dr. Powell's otherwise bald assertion that this established method of dose conversion cannot be used to study APAP's effect on neurodevelopment. *See* Ex. 10, Rebuttal Expert Report of Brandon Pearson ("Pearson Rebuttal Report") at 6.

By arbitrarily limiting his scope of "relevant" animal studies to those that administered only 200 mg/kg/day of APAP—little more than half of the 360 mg/kg *daily* threshold that he should have applied according to his own numbers—Dr. Powell improperly diminishes studies that would otherwise have merited more serious review even on his own terms. *See In re Mirena IUS Levonorgestrel-Related Prod. Liab. Litig. (No. II)*, 341 F. Supp. 3d 213, 242 (S.D.N.Y. 2018), *aff'd sub nom* 982 F.3d 113 (2d Cir. 2020) ("Where an expert ignores evidence that is highly relevant to his conclusion, contrary to his own stated methodology, exclusion of the expert's testimony is warranted."). For example, out of the four studies addressed in the short section purportedly comprising the second, "weighting" step of his analysis, he criticizes two rat studies

that support Plaintiffs' causal claims because they administered APAP at 350 mg/kg/day. *See* Ex. 1, Powell Report ¶ 92a. Those doses are within the correctly calculated (360 mg/kg) daily dose threshold.[2]

Applying the correct daily threshold, therefore, Dr. Powell would have needed to actually account for this evidence that contradicts his conclusions. But, because he operates on a false premise, he unduly disregards this and other significant evidence of APAP's effect on neurodevelopment. And because he unduly disregards that data, he cannot offer reliable conclusions about what it shows. *See Gopalratnam v. Hewlett-Packard Co.*, 877 F.3d 771, 784 (7th Cir. 2017) (affirming exclusion of testimony by expert whose "central underlying premise . . . was not only unsupported, but in fact contrary to generally accepted . . . science"); *Amorgianos v. Nat'l R.R. Passenger Corp.*, 303 F.3d 256, 267 (2d Cir. 2002) ("The reliability analysis applies to all aspects of an expert's testimony: the methodology, the facts underlying the expert's opinion, the link between the facts and the conclusion, *et alia*." (quoting *Heller v. Shaw Indus., Inc.*, 167 F.3d 146, 155 (3d Cir. 1999))) (cleaned up).

Second, Dr. Powell deemed approximately 30 studies irrelevant due to so-called "statistical flaws"—specifically, failing to correct for multiple comparisons—before realizing that appropriate statistical corrections had, in fact, been applied. *See* Ex. 2, Redline of Amended Powell Report ¶¶ 72, 91d. Indeed, this non-existent statistical error is the only *methodological* flaw Dr. Powell found in at least five studies he excluded: Philippot 2018 (n.60), Philippot 2017 (n.61), Viberg 2014 (n.62), Herrington 2022 (n.64), and Harshaw 2022 (n.65).[3] *See* Ex. 1, Powell Report, ¶¶ 71-

[2] Dr. Powell's string cite regarding excessive dose is also problematic in that it includes at least three studies that actually meet his arbitrary dosing cutoff. Lalert 2020 (n.80), Wanasuntronwong 2017 (n.84), and Yisarakum 2014 (n.99) are all listed in paragraph 74, and all three of these rat studies administered an APAP dose of 200 mg/kg/day. *See* Ex. 1, Powell Report ¶ 74 (citing nn.80, 84, 99); *see also* Lalert 2020 at 456; Wanasuntronwong 2017 at 2; Yisarakun 2014 at 37).

[3] Dr. Powell also asserts that these studies have not been replicated. *See* Ex. 1, Powell Report ¶¶ 77-80.

76 (discussing methodological "flaws" in relevant studies and providing string citations to same). Nevertheless, Dr. Powell claims that correcting his misunderstanding of these studies did not change his opinion (*see* Ex. 3, Powell Dep. Tr. 99:20-100:8), which raises the question: If finding that approximately 30 studies do not suffer from this alleged statistical error has no impact Dr. Powell's opinion, why does he consider failing to correct for multiple comparisons a "critical flaw"?"

B. Dr. Powell's Report Has No Transparent, Objective Methodology.

Those errors aside, Dr. Powell's report provides no way to assess—or, in important respects, even understand—how he reached his conclusions. In contrast to Plaintiffs' experts, *see infra* Part II, Dr. Powell did not apply the objective, transparent methodology called for by his own cited guidance literature. *See* Ex. 8, Sena, et al. (2014), at 738; Ex. 1, Powell Report, at 104 n.41. That much is clear just from the presentation. Dr. Powell's "systematic review" amounts to several long, bullet-point lists strewn with footnotes indicating the preclinical studies he deemed irrelevant to whether prenatal APAP exposure contributes to ASD or ADHD. *See* Ex. 1, Powell Report ¶¶ 72–91. Consider the first five (of twenty-four) entries in the list of studies measuring biochemical effects of APAP exposure that he deemed irrelevant because they have different findings:

a. Serotonin (no change[48,50,51,53-55,93], decrease[55], increase[50,51,93,112,115]);

b. Serotonin in striatum (no change[50,51], increase[51]);

c. Noradrenaline (no change[48,51,53,54,93], decrease[54,98], increase[51,93,112]);

d. Dopamine (no change[50,51,53,55,93,112], decrease[54,55,98], increase[48,51,98]);

e. Glutamate (no change[48,51-55], decrease[48,52], increase[51]).

Ex. 1, Powell Report ¶ 83. An expert may not rely on string-cites in place of analysis. *See Daniels-Feasel v. Forest Pharms., Inc.*, No. 17CV4188, 2021 WL 4037820, at *17 (S.D.N.Y. Sept. 3, 2021) (excluding expert whose references to evidence were "bare-bones, often in the form of string cites, and entirely omit[ted] any analysis, much less a thorough one, of the strengths and weaknesses of the underlying conclusions"). Yet to find, for example, which of the above serotonin studies Dr. Powell deemed irrelevant, readers must turn to entries 48, 50, 51, 53, 54, 55, 93, 112, and 115 of the List of Materials Considered at the end of this report. They must do so again and again to understand these lists. *See* Ex. 3, Powell Dep. Tr. at 86:11–16 ("Q. Okay. And for someone to figure out why you found a study non-relevant or what flaws that study had, they need to look at the string of citations at the end of some of these sentences; is that fair? A. Among other things in the list format in subsequent pages, yes."). And nowhere will they find an explanation as to *why* these studies should be deemed irrelevant simply for reaching different results rather than be considered and weighed accordingly. Worse, Dr. Powell's reliance on string cites deprives the reader of any meaningful analysis of the relevant data, including a discussion of the endpoints evaluated in individual studies and the extent to which outcomes are consistent across the data set.

That points to the fundamental difference between Dr. Powell's approach and a true weight-of-evidence approach. The "'weight of the evidence" approach to making causal determinations involves a mode of logical reasoning often described as 'inference to the best explanation,'" and its steps include, among other things, "consider[ing] *all of the relevant available evidence*" and "integrat[ing] the evidence using professional judgment to come to a conclusion about the best explanation." *Milward v. Acuity Specialty Prod. Grp., Inc.*, 639 F.3d 11, 17–18 (1st Cir. 2011) (emphasis added; citations omitted); *accord Daniels-Feasel*, 2021 WL

4037820, at *6. Those steps are demonstrated in Plaintiffs' expert reports. By contrast, Dr. Powell claims to have conducted a "systematic review," which he describes this way: "Performing a systematic review of preclinical studies allows scientists to *narrow the evidence* to only those studies that are performed in a rigorous, reproducible, and robust manner." Ex. 1, Powell Report ¶ 58 (emphasis added). And in effect, Dr. Powell used this approach to *eliminate* most of the available evidence—deeming only two of ninety-nine studies capable of having any relevance to whether prenatal APAP exposure is capable of causing ASD or ADHD. That is called cherry-picking, which an expert "must not" do. *Daniels-Feasel*, 2021 WL 4037820, at *5.

To be sure, particular preclinical studies have particular limitations, as Plaintiffs' experts fully recognize. But dismissing studies is no way to weigh them, and, consequently, it is no way to assess causation. *See Reference Manual on Scientific Evidence: Third Edition* (2011), at xiv ("Fundamentally, the task [of establishing causation] is an inferential process of weighing evidence and using judgment to conclude whether or not an effect is the result of some stimulus."). Study design might make a study more or less capable of *showing* causation, but it does not necessarily make the study *irrelevant* to the causation questions it addresses.

Seeming to recognize as much, at deposition Dr. Powell contended that his review actually has *two* parts. *First*, he said, he narrowed the evidence as described above. But *then*, he said, he "ignored all the potential scientific flaws and took all the results at face value to determine, in a most conservative manner, what's left that replicates and that's believable." Ex. 3, Powell Dep. Tr. at 83:9–20. This consideration of "all the results at face value" apparently occurs in the section of his report titled "Evaluation of Experimental Findings or Endpoints that Are Replicated Consistently," which explicitly "focuses *only* on those publications that are replicated." Ex. 1, Powell Report ¶ 92 (emphasis added); *see* Ex. 3, Powell Dep. Tr. at 83:15–17. And Dr. Powell's

analysis of this artificially limited range comprises two subparagraphs about two findings across four of ninety-nine studies, still focusing on study design, clearly not taking results at face value. *See* Ex. 1, Powell Report, ¶ 92a–b.

The sum total of Dr. Powell's "systematic review," then, is this: some discussion of two studies (the Saad and Baker studies, *see id*. ¶¶ 67–70); a brief discussion of the designs of four studies he deemed irrelevant, *see id*. ¶ 92; and footnotes for ninety-three others. Despite Dr. Powell's insistence that his review does not exclude the vast majority of preclinical literature relating to the neurodevelopmental effects of APAP exposure, *see* Ex. 3, Powell Dep. Tr. at 84:6–17, his report cannot reasonably be read any other way.[4] His opinions thus cannot "help the trier of fact" at all, let alone reliably, to determine whether prenatal APAP exposure can cause ASD and ADHD in offspring, which is the point of expert testimony. Fed. R. Evid. 702(a). Indeed, as Dr. Powell admits, the ARRIVE 2.0 guidelines that form part of the basis of his systematic review are not meant to be—and he did not apply them as—a methodology for evaluating causation. *See* Ex. 1, Powell Report ¶¶ 35–37; Ex. 3, Powell Dep. Tr. at 127:19–24 ("Q. Okay. And is it your understanding that the ARRIVE 2.0 guidelines are meant to be a methodology to evaluate causation in the context of the safety of a compound? A. That's not how I've applied those guidelines, and, no.").

[4] At deposition, Dr. Powell seemed to realize that the title of the section of his report that lists all the excluded studies—"Explanation of Critical Flaws in and/or Lack of Relevance of *Remaining* Publications Systematically Reviewed"—indicates that those studies were in fact excluded from his ultimate conclusions. Ex. 1, Powell Report ¶ 33 (emphasis added); *see* Ex. 3, Powell Dep. Tr. at 85:21–86:3 ("Q. And when you say 'remaining publications,' are you talking about all the papers besides Baker and Saad? A. When I talk about the remaining publications, explanation of critical flaws in and/or lack of relevance of remaining publications systematically reviewed. Oh, yeah, that's a good point that you point out. Remaining may—is probably the incorrect term that I would have used.").

What is more, Dr. Powell did not even reliably apply his own exclusionary criteria. There are inconsistencies across the lists of excluded studies that not even Dr. Powell can explain. *See* Ex. 3, Powell Dep. Tr. at 90:6–91:6 (ultimately agreeing that the studies listed in paragraphs 74 and 91a of his report as excluded for purportedly excessive dosing "aren't perfectly in alignment"); *see also* Ex. 12, Demonstrative Chart (showing studies that Dr. Powell's report inconsistently identifies in paragraphs 74, 91a, and 116 as excluded for administering purportedly excessive doses). As mentioned, Dr. Powell was also forced to amend his report to remove 30 studies from the list of those faulted for failure to correct for multiple comparisons, among several other substantive changes. *See* Ex. 2, Redline of Amended Powell Report ¶¶ 72, 91d. The reason he gave at deposition was that, when initially conducting his review, he searched for words like "multiple comparisons." Ex. 3, Powell Dep. Tr. 95:11–22. Belatedly considering the studies more closely, indeed *after already serving his expert report and receiving Dr. Pearson's rebuttal report*, he realized that most did include this statistical correction where appropriate and could not be faulted on this basis.[5]

Because he did not apply his own criteria accurately, he improperly deemed various studies irrelevant. Powell states that Saad 2016 and Baker 2023 are the only two papers that passed criteria for relevance and appropriateness of experimental design. He describes this as: whether studies used animals (not humans or cells in a dish), a relevant dose (≤200 mg/kg/day, a conservatively high number, as discussed above and below) (which, again, is inexplicably limited), relevant timing of drug administration (≤ postnatal day 14), appropriate statistical methods, and a relevant

[5] *See* Ex. 3, Powell Dep. Tr. at 96:20–97:3 ("Q. Okay. And if I'm understanding it right, when you did your initial review, you looked for words like multiple comparisons, comparisons, false discovery rate, in those studies; is that fair? A. Yes. And my expectation was that if you did the corrections for multiple corrections, you would actually say that in your methods, and I realized at some point that that wasn't always the case.").

outcome measure (e.g., one publication excluded because it studied acetaminophen's ability to protect against another toxin but not the direct effects of acetaminophen). *See* Ex. 1, Powell Report ¶ 66. However, as provided above, with removal of his criticism of failing to correct for multiple comparisons from Philippot 2017, Philippot 2018, Viberg 2014, Herrington 2022, and Harshaw 2022, they similarly pass the criteria for relevance and appropriateness of experimental design along with Saad 2016 and Baker 2023 and should be considered along with them. This is just one example of Dr. Powell's misapplied criteria resulting in a fundamentally flawed analysis. His methodology is neither reliable nor reliably applied. *See Amorgianos*, 303 F.3d at 267 ("[A]ny step that renders the analysis unreliable under the *Daubert* factors renders the expert's testimony inadmissible.").

Regardless, these elements of study design and execution are no basis to exclude studies *en masse* rather than consider the data in the context of its strengths and limitations. *See* Ex. 10, Pearson Rebuttal Report at 5-6 (explaining that "uncertainties, confounders, variabilities, etc." are inescapable in scientific experimentation but diversity of study designs lends strength to the ultimate conclusion); *see also infra* Part II.B. And though Dr. Powell briefly states his general views about the importance of these criteria as study-design features, *see* Ex. 1, Powell Report ¶¶ 72–77, he nowhere explains how he thinks they might affect a study's reliability or results, nor does he use them to assess the relative consistency or significance of outcomes. Reviewing the evidence is, needless to say, important to systematically reviewing the evidence. Dr. Powell has no reliable methodology for doing so.

C. Dr. Powell's Conclusions Are Unfounded.

In the end, Dr. Powell, having blinded himself to the overwhelming majority of the evidence, finds no evidence of an association between prenatal APAP exposure and ASD/ADHD

or of a biologically plausible mechanism for that association. Yet he acknowledges that many mechanisms, including many of the mechanisms identified in Plaintiffs' expert reports, can be mechanisms for neurodevelopmental effects—specifically, that oxidative stress can adversely affect neurodevelopment, *see* Ex. 3, Powell Dep. Tr. at 61:2–25, that endocrine disruption adversely affects neurodevelopment, *see id*. at 71:4–22, that APAP can affect serotonergic signaling in the brain, *see id*. at 72:8–14, that changes in gene expression can lead to pathologies, *see id*. at 74:2–5, and that the developing brain is vulnerable to abnormal energetic demands, *see id*. at 193:8–19.

Having not weighed the evidence, Dr. Powell cannot reliably reject any of these mechanisms based on the weight of the evidence. Indeed, even if his report were intended to opine on the weight of the evidence, as he suggested at deposition, his opinions would necessarily fail to follow from his methodology and thus fail the *Daubert* test. *See Amorgianos*, 303 F.3d at 267 ("[I]t is critical that an expert's analysis be reliable at every step."). He assigns no comparative weights to any evidence. In contrast to Plaintiffs' experts' conclusions, therefore, any weighting he might theoretically have conducted is entirely untestable—and hence unscientific. *See, e.g.*, *United States v. Gissantaner*, 990 F.3d 457, 463–64 (6th Cir. 2021) ("An untestable scientific theory is all theory and no science. In the absence of proof that a technology can be tested, there is no way to show whether it works (its 'refutability' or 'falsifiability,' a scientist would say) and no way to give it 'scientific status.'") (cleaned up).

As for the two studies that Dr. Powell deemed strong enough to warrant any meaningful space in his "systematic review," the senior author of one study (Baker 2023) is none other than Dr. Pearson. There is thus consensus that Dr. Pearson is an authority in this space. His own report reviews this study objectively (not even assigning it the highest weight of all the studies he

reviewed) and thoroughly details how, contrary to Dr. Powell's misreading, the study supports a causal association between prenatal APAP exposure and ASD/ADHD. *See* Ex. 13, Pearson Report at 112–14; Ex. 10, Pearson Rebuttal Report at 4–5.

The other study (Saad 2016) actually found behavioral effects consistent with APAP's causal association in neurodevelopmental disorders, including a significant effect of APAP on hyperactivity. *See* Ex. 13, Pearson Report at 102–05; Ex. 14, Cabrera Report at 83–84. Inexplicably, however, the authors negated these findings in their own conclusions by applying an extreme threshold for statistical significance that, as Dr. Pearson shows, they do not apply in their other work. *See* Ex. 13, Pearson Report at 102–05. Dr. Powell's report does not acknowledge these elements of this study. Disregarding relevant aspects of an assertedly relevant study is not the mark of a scientific methodology. *See In re Rezulin Prods. Liab. Litig.*, 309 F. Supp. 2d 531, 564, n.146 (S.D.N.Y. 2004) ("[T]he dispositive fact here is that [the expert] pointedly ignored directly relevant scientific data in violation of his own standards.").

Dr. Powell also makes no attempt to evaluate any of the other preclinical evidence—in vitro/ex utero or in silico data—weighed in Plaintiffs' holistic analyses. After all, the systemic-review literature from which he amalgamated his methodology is geared toward preclinical in vivo (animal) evidence. *See* Ex. 4, Gurusamy et al. (2021) at 10. And he offers no reason, much less a reliable one, why such evidence should not be included in a weight-of-evidence review. He no doubt disagrees that this evidence, like the rest of the preclinical evidence, supports the observed causal association between prenatal APAP exposure and ASD/ADHD. And perhaps he could have presented a reliable opinion along those lines. Instead, he opted to "'pick and cho[o]se' from the scientific landscape and present the Court with what he believes the final picture looks like." *In*

re Rezulin Prods. Liab. Litig., 309 F. Supp. 2d at 563 (quotation marks omitted). The picture is woefully inadequate.

II. Dr. Powell Has No Reliable Criticisms of Plaintiffs' Experts' Methodologies.

A significant portion of Dr. Powell's report is dedicated to "opinions" that are in fact merely criticisms of the analyses of Plaintiffs' experts. *See* Ex. 1, Powell Report ¶¶ 49–99. Indeed Section IX. A., for example, is entitled "Dr. Pearson and Dr. Cabrera's Weight-of-Evidence Reviews in Support of Their Conclusion that Behavioral Animal Studies Support a Causal Connection Between In Utero Acetaminophen Exposure and ASD or ADHD Are Deeply Flawed." *Id.* at 51. Not only is this not an appropriate expert opinion (the Court is the gatekeeper), these "opinions" lack support. Fed. R. Evid. 702.

A. Animal Studies Are a Reliable and Generally Accepted Source of Causation Evidence.

Dr. Powell frames his response to Plaintiffs' experts by undermining his own discipline. In a particularly egregious illustration of cherry-picking, Dr. Powell disputes that any animal experiments can be used to study the causes of ASD or ADHD in humans. *See* Ex. 1, Powell Report ¶¶ 3, 28–29. This assertion is puzzling from someone whose work purportedly "focuse[s] on the study of genetic animal models of relevance to ASD," *id.* ¶ 17, who acknowledges that "[m]odel species are biologically similar to humans" and can "allow us to study the brain in a more invasive and detailed manner," *id.* ¶ 25, and who otherwise emphasizes that "animal models are helpful tools in ongoing research to help improve the lives of children and adults diagnosed with ASD and ADHD," *id.* ¶ 228. He further admitted at deposition that animal studies allow for tighter controls and thus less confounding than human observational studies, *see* Ex. 3, Powell Dep. Tr. at 32:4–10, 32:18–22, 32:24–33:6, and that rodents and humans share many biological systems, including the endocannabinoid system and endocrine system, and both experience oxidative stress,

id. at 30:17–31:22. Plaintiffs' experts opine that these are biologically plausible pathways through which prenatal use of APAP can cause ASD and ADHD. Dr. Powell's wholesale disregard of the animal literature is made for litigation, not the approach he would undertake in his daily work as a scientist. *See In re Terrorist Attacks on Sept. 11, 2001*, No. 03MD01570, 2023 WL 3116763, at *2 (S.D.N.Y. Apr. 27, 2023) (indicia of unreliability include whether an opinion "was developed expressly for purposes of testifying"); *Restivo v. Hessemann*, 846 F.3d 547, 577 (2d Cir. 2017).

The reason animal studies remain such an important part of studying endpoints in humans is supplied by Dr. Powell himself: Randomized controlled trials cannot ethically be conducted in humans to test biological mechanisms. *See id*. at 42:21–43:3 ("[Dr. Powell:] If you're speaking about randomizing pregnant women to receive doses of Tylenol without a clinical indication, I would say, no, that study would not be possible. . . . Q. Right. It wouldn't be ethical, right? A. I would agree with that."). That leaves observational studies in humans and animal studies. Dr. Powell attacks a strawman when he opines that "[n]o scientific research can begin and end with the animal model," Ex. 1, Powell Report ¶ 28, because none of Plaintiffs' experts assert that it can. The preclinical animal studies must be evaluated and weighed in the context of a wide body of human epidemiological evidence. In his work outside this case, Dr. Powell claims at least to "care about the human context to some degree when [he is] working with the mice." *Id*. at 41:15–25. But not in this litigation, where the human epidemiological studies are absent from his report's "systematic review" and where he fails to bring the same rigor that he applies in his daily work. *See In re Fosamax Prods. Liab. Litig.*, 645 F. Supp. 2d 164, 188 (S.D.N.Y. 2009).

Researchers must certainly be careful in drawing conclusions about human disorders from animal models. Plaintiffs' experts Dr. Pearson and Dr. Cabrera certainly agree with this proposition. Dr. Powell, by contrast, only agrees selectively. He is perfectly comfortable opining

that animal models can be useful tools to study potential *genetic* contributors to ASD and ADHD, *see* Ex. 1, Powell Report ¶ 17, but he cannot explain why, with appropriate caution, they are nevertheless unavailable to study potential *environmental* contributors to neurodevelopmental disorders.[6] Animal-models-for-me-but-not-for-thee is not an accepted methodology or the product of a reasonable debate between good-faith scientists. It is outcome-oriented testimony that Rule 702 forbids. *See Faulkner v. Arista Records LLC*, 46 F. Supp. 3d 365, 381 (S.D.N.Y. 2014) ("[M]ethodology . . . aimed at achieving one result . . . is unreliable, and . . . must be excluded."); *In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, 858 F.3d 787, 796–800 (3d Cir. 2017) (affirming exclusion of "conclusion-driven" analysis).

B. Plaintiffs' Experts Reliably Weigh the Causation Evidence.

Dr. Powell does not object to the weight-of-evidence methodology. *See* Ex. 3, Powell Dep. Tr. at 50:12–16 ("I think that the weight of the evidence and/or a systematic review is a reasonable way to analyze the body of literature, and if it's applied consistently and in a reproducible manner, I think that it's a reasonable way to go."). Nor could he. The Federal Reference Manual on Scientific Evidence notes that, "[f]undamentally, the task [of establishing causation] is an inferential process of *weighing evidence* and using judgment to conclude whether or not an effect is the result of some stimulus." *Reference Manual on Scientific Evidence: Third Edition* (2011), at xiv (emphasis added). Indeed, a weight-of-evidence review is what Dr. Powell claims he did for his own report, though, as seen, he fails to appreciate the distinctions between weight-of-

[6] Indeed, Dr. Powell only demonstrates his misunderstanding when he asserts that Dr. Pearson "wrongly equates any change in *behavior* with a biologically plausible mechanism, when behavioral differences are not mechanisms at all, but merely outcomes (that might or might not imply an underlying change in the brain)." Ex. 1, Powell Report ¶ 111. The only such conflation occurs in that sentence of Dr. Powell's. As Dr. Pearson's report thoroughly explains, the behavioral manifestations of a toxicant working through a biological mechanism is not the mechanism itself. Rather, they can be used to assess whether a studied mechanism has an effect that is translationally relevant to humans.

evidence methodologies and the "systematic review" he conducted. *See* Ex. 3, Powell Dep. Tr. at 49:13–16 ("Q. So why did you frame it as a systematic review rather than a weight of evidence analysis? A. Because they're essentially, in my mind, very similar.").

Dr. Powell's only available ground for disagreement, therefore, lies in how Plaintiffs' experts weighed the evidence in this case. For example, he criticizes Dr. Pearson's framing, homing in on Dr. Pearson's observation that, "[o]ver the last four decades, the overwhelming majority of preclinical studies investigating the effect of acetaminophen on neurodevelopment show that acetaminophen causes *neurodevelopmental disruption*." Ex. 13, Pearson Report at 4 (emphasis added). Dr. Powell does not disagree with that description of the studies Dr. Pearson reviewed. Rather, he construes it to suggest that Dr. Pearson "equates *any* change in the brain as a plausible biological mechanism whereby acetaminophen could cause ASD or ADHD." Ex. 1, Powell Report ¶ 111. Dr. Powell calls this suggestion "unscientific." *Id*. And it may be, but it is not in Dr. Pearson's report. ASD and ADHD have shared symptoms and etiologies, but they are also heterogenous, meaning that they manifest in a variety of outcomes. *See* Ex. 13, Pearson Report at 22–32. As Dr. Powell provides in his report: "If you've seen one child with autism, you've seen one child with autism." Ex. 1, Powell Report ¶ 50; *see also* Ex. 3, Powell Dep. Tr. at 33:21–25 ("Q. Would you agree with me that in humans autism has a pretty heterogenous presentation? A. Autism spectrum disorder is a spectrum and it has a wide—a fairly broad range of, I guess, severity of general symptoms and comorbidities.").

Accordingly, a thorough weight-of-evidence review of the causes of ASD and ADHD must include a range of studies. Dr. Pearson was careful, however, to include only studies implicating neurological systems involved in the development of ASD and ADHD, and he explicitly excluded those that did not. *See* Ex. 13, Pearson Report at 69 ("Because the focus of this report is the

developmental neurotoxicity of APAP, studies from other disciplines and studies investigating other, unrelated parameters were screened out at the evidence collection stage.") (emphasis added); Ex. 10, Pearson Rebuttal Report at 4 ("I did not include published studies on developmental APAP in rodents that measure pain or analgesia, nor mood, anxiety, aggression, sexual behavior, nor that evaluate stereological differences in sex-specific cellular composition in the brain.").

Dr. Powell also criticizes Dr. Pearson for using the scoring system detailed in his report to assess and compare study designs rather than some different and supposedly "established scoring systems." Ex. 1, Powell Report ¶ 118. This criticism is not credible. Dr. Pearson's weight-of-evidence methodology, including his system for assessing data quality, is "drawn from the broadly applicable and systematic approach described in *Guiding Principles and Key Elements for Establishing a Weight of Evidence for Chemical Assessment* (OECD, 2019)" and incorporates "[r]elevant concepts from other guidance documents." Ex. 13, Pearson Report at 6; *see also id*. at 73–74. He also used the scoring system to transparently convey the analysis he typically performs in his peer review capacity. *Id.* at 6; *see also* Ex. 3, Pearson Dep. Tr. At 147:13–149:25. For his part, Dr. Powell describes his own approach as an "amalgam," indeed, an amalgam based on his "own list of criteria that [he] would use" and that he then cross-checked with inapposite guidance literature. Ex. 3, Powell Dep. Tr. at 79:22 –80:9, 131:4–7. That is not a reliable standard. It is tantamount to Dr. Powell saying "trust me, I have criteria." Nothing in *Daubert* or the Federal Rules of Evidence requires the Court to accept that. *See Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 157 (1999).

Nor is Dr. Pearson's scoring system in any way the sort of homespun method Dr. Powell makes it out to be. Although not required to do so under the guidance literature, Dr. Pearson "used a coarse quantitative scoring system to enable objective comparisons between the studies

reviewed." Ex. 13, Pearson Report at 73. As he explained at his deposition, the scores themselves did not drive his weight-of-evidence conclusions. The point of scoring a given study was simply "to understand the characteristics of the study and give a transparency into my work into understanding the parameters of controls and those sorts of characteristics of the study." Ex. 3, Pearson Dep. Tr. at 131:5–11. The scores thus do not replace the qualitative analysis that is ultimately required when weighing evidence; rather, they make Dr. Pearson's qualitative findings easily testable. Objectivity, transparency, testability—all hallmarks of reliability in scientific analysis, as Dr. Powell would agree,[7] despite failing to achieve them in his own report.

The scoring system could not include entries for every study-design parameter without turning Dr. Pearson's report into a textbook. *See id.* at 131:18–22. But Dr. Powell voices no disagreement with the fundamental parameters (*e.g.*, quality of controls, sample size) that it does include. And Dr. Pearson provides a narrative discussion of those and other parameters for each study reviewed. Dr. Powell fails to explain why either of the two "established" scoring systems that he located, Ex. 1, Powell Report ¶ 118, would have been a better fit for this analysis. His own report does not use any objective, transparent scoring system, let alone one of these two. He likewise fails to explain how either of these systems would have changed the analysis.[8]

Dr. Powell lodges a medley of other criticisms at the studies included in Dr. Pearson's weight-of-evidence analysis. But his criticisms of how Dr. Pearson weighed the evidence stem

[7] *See* Ex. 3, Powell Dep. Tr. at 82:17–83:2 ("Q. Okay. But do you agree that, in general, when you're reviewing—excuse me, a body of literature, it's good to be objective? A. It's always good to be objective, in my opinion. Q. Okay. And it's good to be transparent, right? A. I would agree with that. Q. And it's good to be transparent so someone can go behind you and reproduce your analysis if they need to, right? A. Well, insofar that there is an analysis done, I think that would be important, yes.").

[8] These systems, both developed for regulatory reviews, have scoring criteria similar to Dr. Pearson's. *See* Ex. 1, Powell Report ¶ 118 n.11. They also have many others geared toward their regulatory context. Including these on top of Dr. Pearson's existing scores would, again, have simply made his report unnecessarily long.

from the criteria that Dr. Powell used to *exclude* the evidence, which are themselves irrelevant or misapplied. He argues that Dr. Pearson fails to address the coherence of the studies' findings. *See id*. ¶ 117. Yet Dr. Powell's apparent definition of coherence—that studies can be considered consistent only if they show the exact same behavioral marker of ASD or ADHD resulting from the exact same biological mechanism—is both unsupported in his report and insupportable in this context. As explained, a thorough weight-of-evidence analysis of the causes of ASD and ADHD requires a range of studies, and the consistency among those studies is accounted for by weighing them, as Dr. Pearson (but not Dr. Powell) did.

The same applies to Dr. Powell's mistaken focus on whether certain findings have been replicated. The neurodevelopmental effect of in utero APAP exposure has been seen across a variety of studies. Dr. Powell, who has never worked on an animal model of ADHD, *see* Ex. 3, Powell Dep. Tr. at 47:21–48:2, or any developmental-neurotoxicity study, *see id*. at 46:11–13, would require that the same, resource-intensive study be run at least twice before its findings could be considered as potential causation evidence. But evidence can be weighed even if not so strictly replicated. *See e.g.* Ex. 11, Tyl et al. at 373 (2008). Dr. Powell also criticizes Dr. Pearson for weighing some studies that administered purportedly excessive doses of APAP to the test animals. *See* Ex. 1, Powell Report ¶ 116. As seen, however, Dr. Powell's own views on dosage are mistaken, and this criticism in particular is based on cherry-picked data. *See* Ex. 10, Pearson Rebuttal Report at 6–8.

In all events, Dr. Pearson discusses dosage, other relevant study-design features, and their impact on a study's weight throughout his report, and Dr. Powell fails to identify any that Dr. Pearson disregards. Disagreements with the ultimate results of Dr. Pearson's weighting analysis, which Dr. Powell advances without attempting any sort of weighting analysis himself, are out of

place at this stage. Dr. Powell's argument that Dr. Robert Cabrera's "approach is flawed for similar reasons," Ex. 1, Powell Report ¶ 126, fails for similar reasons.

CONCLUSION

Dr. Powell's report and testimony should be excluded.

Dated: September 19, 2023

Respectfully submitted,

KELLER POSTMAN LLC

/s/Ashley C. Keller
Ashley C. Keller – *Pro Hac Vice*
KELLER POSTMAN LLC
150 N. Riverside Plaza Suite 4100
Chicago, IL 60606
Phone: (312) 741-5220
Fax: (312) 971-3502
ack@kellerpostman.com

WATTS GUERRA LLC
Mikal C. Watts (*Pro Hac Vice*)
Millennium Park Plaza RFO
Ste. 410, C112
Guaynabo, Puerto Rico 00966
(210) 447-0500
mcwatts@wattsguerra.com

THE LANIER LAW FIRM
W. Mark Lanier (*Pro Hac Vice*)
Tower 56
126 East 56th St., 6th Floor
New York, New York 10022
(212) 421-2800
mark.lanier@lanierlawfirm.com

Counsel for Plaintiffs