Primate Model of Effects of Vaccination on Autism-Like Behaviors and Neuropathology: A Review of the Data from Phases I & II

Sallie Bernard, Lyn Redwood, Laura Bono
SafeMinds Board Members and Co-Founders
SAFEMINDS. October 19, 2015

EXECUTIVE SUMMARY

Since 2003, a team of investigators have been using a primate model to study biological and behavioral response to vaccine schedules and components in relation to autism. This effort was implemented in two sequential phases, Phase I and Phase II, each leading to several peer reviewed publications as well as conference abstracts and presentations.

The conclusions from the most recent Phase II publications by Curtis et al and Gadad et al contradict those arising from Phase I. Conclusions of these Phase II papers also contradict those from earlier Phase II conference abstracts and materials as well as internal study reports. Phase I papers observed consistent significant differences in brain, behavior, learning and gastrointestinal outcomes similar to those found in autism, from vaccine exposed compared to controls. Early Phase II documents reported significant differences in brain anatomy of those exposed. Later Phase II papers and abstracts, however, concluded no effects from vaccinations. Phase II papers from Curtis and Gadad report differences on measures of learning and behavior but ultimately concluded that vaccine exposure "does not result in autism-like behavior or neuropathology".

While complex research can find contradictory results in an ongoing search for answers, we have identified potential problems with the findings and conclusions from Phase II Curtis and Gadad papers.

FINDINGS

(a) The investigators have provided a sole argument for the contradiction between Phase I and II: that a larger sample size in two Phase II gave more accurate results. Though the control sample size increased modestly in Phase II compared to Phase I, the size for the key exposure group (1990s Primate) was the same (N=12) in both Phases. Further, statistically speaking, we should have more confidence in a positive result in a smaller sample as reported in Phase I, because it suggests the exposure causes clear adverse effects, than the Phase II negative result in a slightly larger but still small and thus underpowered sample.

(b) The authors explained the discrepancy in neuropathology findings between early Phase II reports, submitted as either a conference poster or abstract, and the later Gadad and Curtis papers, as being due to sample size changes. We have found evidence of no differences in sample sizes used between the earlier and later Phase II efforts, contradicting the authors' explanation.

(c) Significant differences in behavior and learning reported in Phase II Gadad and Curtis papers were determined to be irrelevant by the authors without substantiation.

(d) Outcomes in Gadad and Curtis papers were reported as mean values without showing individual level data, and some data was log transformed, which may have obscured idiosyncratic or outlier effects relevant to an animal model of autism.

(e) The Ns are inconsistent across various tables and figures from Phase II. If the larger N argument is made, then at minimum the study needs to explain why the reported sample sizes varied repeatedly.

(f) Explanations other than sample size which could address the differences between Phase I and II, including plausible confounding and biases, were never addressed as relevant discussion points in either Gadad or Curtis, constituting a form of reporting bias.

CONCLUSIONS

Given the Phase II issues around sample size, inconsistent findings over time, and data reporting choices described in this review, the Gadad paper overreaches in its conclusion regarding vaccine safety: *"Our data strongly support the conclusion that childhood TCVs do not produce ASD-like neuropathology or behavioral changes in the nonhuman primate."*

 (a)   Even if there were no biases and confounders in the study, it is not powered to show a lack of an effect for a condition like autism, which occurs in 1-3% of children.

(b)   Phase 1 showed large, significant results. No suggestions have been made by the authors that Phase I was deficient, and Phase II followed the same methodology and protocols.

(c)   There are significant differences between the Phase I and Phase II results, and between earlier and later versions of Phase II findings, and those discrepancies need to be explained. They cannot be explained by sample sizes. Misclassification, sample bias or other confounders should be explored.

(d)   There are discrepancies in the samples reported for the various versions of the papers, and these need to be explained.

(e)   The choice of data reporting approaches may be obscuring outlier effects. Such effects are likely in an animal model of autism and exposures, as the non-genetic risk for autism is hypothesized by the authors to be an uncommon or idiosyncratic response to environmental agents.

(f)   Across the known study documents – papers, abstracts, posters and reports for both phases - significant differences between exposed and unexposed have been observed, including brain pathology, learning and behavior problems, and gastrointestinal illness. These findings should be recognized in drawing conclusions on vaccine side effects.

(g)   Reporting bias including publication bias may be a factor in publication of Phase II and the strong conclusions of the Gadad paper, as well as the lack of publication of additional Phase I papers.

The autism-vaccine primate research spanned over 10 years, comprised two phases, many arms, and many outcome measures. The complexity of the effort cannot be adequately contained within a few journal articles. Our review, and interpretation by others wishing to evaluate this research, is limited by lack of access to the raw data (both published and unpublished), all study protocols, and all study reports. To resolve the concerns raised here and allow verification, these materials should be posted publicly and immediately, even if the authors intend to conduct future additional analyses. Further, the unpublished papers from Phase I should be published, to correct reporting biases, and the authors should refrain from diminishing the importance of Phase I by claiming superiority of Phase II based on weak and statistically inaccurate arguments. Since Phase II of this research is being messaged as absolving vaccine regimens from having a role in autism risk, these recommendations have important implications for public health, vaccine safety, and autism prevention.

BACKGROUND

The prevalence of autism has increased from 1 in 2000 before 1980 to 1 in 68 children today. Autism is estimated to cost the country $268 billion annually and reach $1 trillion by 2025 if prevalence trends continue upward. (1) Finding the causes of autism is vital for prevention as well as for development of rational treatment approaches. In a recent survey, 42%t of parents of a child with autism agreed or strongly agreed that vaccines played a part in the development of their child's autism.(2) Studies over the past two decades have both found and not found associations between vaccines or vaccine components and autism or autism-like outcomes. A primate model of autism and vaccine safety has been developed over the past decade by scientists led by Laura Hewitson and partially funded by SafeMinds. This model

was designed to examine vaccines as a potential contributor to some cases of autism.

The vaccine-autism primate study investigated the vaccine components thimerosal and several US infant vaccine schedules as a whole. It consisted of two phases. Phase I began in 2003 and analysis and manuscript preparation extended to 2010. Phase II began with animal breeding from 2008-2012 and subsequent analysis and publication to the present (2015). Phase I found a series of negative effects in infant reflexes, learning, brain growth and gastrointestinal function among those exposed to vaccines. Phase II found inconsistent effects, but due to a preponderance of no effects, concluded that vaccines and the vaccine component thimerosal do not lead to autism-like behaviors or neuropathology. These conclusions were described in the two published papers from Phase II: one, by Curtis and colleagues, appearing in 2014 in Environmental Health Perspectives (EHP) (3); the second, by Gadad and colleagues, appearing in the Proceedings of the National Academy of Sciences (PNAS) in 2015. (4)

This review examines the reasons that Phase II is at odds with Phase I. The inconsistency is problematic because the conclusions being drawn from Phase II – which ignore Phase I - have lead to several media reports characterizing the study as disproving an autism-vaccine link, and the publications in which the Phase II papers appeared – PNAS and EHP – are widely read in the medical field. This review represents our investigation into the details as currently available of the entire vaccine-autism primate research effort. Our goals are to foster transparency and full reporting, potentially leading to a full reanalysis of the data and further publication of Phase I "invisible" studies, so that the most reliable research is available on the safety of health related medical interventions (vaccines) and for understanding autism etiology.

REVIEW

STUDY HISTORY & FINDINGS OVER TIME

Genesis of the Primate Model. In early 2000 concerns arose that American infants had been exposed to mercury through the preservative thimerosal in excess of Federal safety guidelines based on daily exposure. (5) As part of the FDA Modernization Act, the FDA found that vaccine manufacturers were not required to evaluate thimerosal's safety in animal studies prior to its introduction as a vaccine preservative (6), although modern federal regulations require this. (7) Thimerosal was developed by Eli Lilly and Company in the 1920s and grandfathered into use as generally accepted as safe.

One mechanism for investigating the safety of multiple vaccine schedules and vaccine components such as thimerosal is to use animal models for research. Vaccine safety studies in humans have examined individual vaccines or vaccine components in isolation (8); none have looked at the potential for synergy and interaction in the context of adverse events from multiple vaccinations. Macaques are commonly used in pre-clinical vaccine safety testing. Primates may exhibit biological differences and behaviors that have been associated with autism in humans. Thus, this research model was capable of providing insights into mechanisms and causes of autism, as well as having utility in translational studies concerning treatment and prevention.

Phase I: Model Development, Findings, & Publications. In 2003, as part of a research initiative on the effects of mercury exposure and autism, SafeMinds awarded a grant for development of an autism-vaccine primate model to a team of researchers at the University of Pittsburgh headed by Laura Hewitson, PhD (subsequently of the Johnson Center for Child Health and Development). The first phase was a multidisciplinary, observer-blinded, placebo-controlled study which examined the hypothesis that the vaccine exposures during the 1990s consisting of thimerosal-containing vaccines (TCVs) and the MMR

vaccine were associated with neurodevelopmental and gastrointestinal pathology.

Two peer-reviewed papers were published from this phase. In one paper (9), the results on neuroimaging were given. Based on 12 exposed primates and 3 saline-injected controls, the results showed that:

> "exposed animals did not undergo the maturational changes over time in amygdala volume that was observed in unexposed animals. After controlling for left amygdala volume, the binding of the opioid antagonist [11C]diprenorphine (DPN) in exposed animals remained relatively constant over time, compared with unexposed animals, in which a significant decrease in [11C]DPN binding occurred. These results suggest that maturational changes in amygdala volume and the binding capacity of [11C]DPN in the amygdala was significantly altered in infant macaques receiving the vaccine schedule."

The other publication from Phase I reported the results on neonatal reflexes of the birth dose of the thimerosal-containing Hepatitis B vaccine (HBV) (10). In exposed animals (n=13) "there was a significant delay in the acquisition of root, snout, and suck reflexes, compared with unexposed animals" (n=7: 3 saline-injected controls and 4 non-injected controls). Interaction models indicated that lower birth weight (BW) and/or lower gestational age (GA) "exacerbated the adverse effects following vaccine exposure." Low BW and GA are factors that increase autism risk.

According to the investigators in a 2008 report to funders, analysis of the toxicology, immunology, and neuropathology data was underway. Three abstracts were given at the annual IMFAR conference in 2008, including one reporting altered gene expression profiles indicative of chronic inflammation from GI biopsy of the vaccinated group. (Link to IMFAR abstracts) The authors expected to derive a number of additional papers from this phase, referring to the HBV-Neonatal Reflexes paper as "the first of at least 10 papers" and "I will continue with publishing data from Phase I and have several manuscripts close to submission." A third and fourth paper (11,12) were submitted for publication but never published. One manuscript reported that "a majority of rhesus macaques exposed to a childhood vaccine schedule developed a progressive colonic immunopathology culminating in a Crohn's-like phenotype." This study was based on 9 controls (both saline injected and no injection) and 9-12 (depending on the analysis) exposed with TCVs and MMR. The other unpublished manuscript concerned cognitive development. An abstract (13), submitted to the Neurobehavioral Teratology Society (NBTS) annual conference in 2011, reported on behavior and learning (n=16). It found that:

> "Infant macaques exposed to thimerosal via vaccination scored significantly lower in the reversal phase of discrimination learning tests compared with control animals. The poor performance by some exposed animals indicates that they were particularly prone to failure on the reversal phase suggesting that they were resistant to extinction, i.e. unable to adjust their behavior when the consequences of previous actions had been changed. This behavior and the impaired ability to organize behavior temporally are salient features of ADHD and suggest that environmental neurotoxicants may contribute to the behavioral manifestations of this disorder."

In addition to mean group differences, the NBTS abstract noted that some exposed animals were particularly impaired, suggesting variable response to vaccination:

> "For the reversal phase, exposed animals demonstrated longer mean latency, increased number of balks and longer time-to-criterion compared with unexposed animals. Some exposed animals scored particularly poorly on this test, achieving only one or two correct responses per day, registering more than twenty balks per day, having abnormally long latencies, and failing to reach criterion at 120 days of testing."

Phase II: Methodology & Reporting. Due to the strong statistically and clinically significant results of

the primate model from Phase I, in 2008 the investigators requested funding to move into Phase II. Phase II was proposed to utilize the same methodology as Phase I to examine which vaccine or vaccine combinations may have contributed to the observed effects on behavioral, brain and gastrointestinal differences in the infant macaques.

To meet these expanded objectives, five groups of infants were to be studied (Table 1) using 60 animals, with the larger N to address the additional exposure arms, not to increase the N per exposure group. In the first group, 10 unexposed infants would receive only placebo saline injections and be used as controls. In the second group, 10 infants were to receive TCVs only with the MMR replaced with a saline injection. In the third group, 10 infants were to receive the MMR only, with all other vaccines replaced with saline injections. In the fourth group, the complete childhood vaccine regimen corresponding to the US practice in 1994 to 1999 would be given, which was the same as the exposed group from Phase I. For this group, 10 additional animals were proposed for entry into a treatment arm as part of a Phase III. In the final group, the complete vaccine regimen currently recommended (proposed as 2005 and later reported as "2008") was to be given. This group would include a single influenza vaccine administered to pregnant animals, as recommended at the time by the CDC for pregnant women. All other pregnant animals would receive a placebo saline injection.

Table 1. Vaccine Groups – Treatment Condition and Sample Size. As proposed by Hewitson et. al. in 2008 and contained in request for funding to SafeMinds.

Group	Code	N	Vaccines administered
Group 1	Th^- MMR^-	10	None, all saline placebos
Group 2	Th^+ MMR^-	10	All thimerosal-containing vaccines and saline placebo for MMR
Group 3	Th^- MMR^+	10	MMR only, all others replaced with saline placebo
Group 4*	Th^+ MMR^+	20	Previous vaccine regimen as recommended through the 1990's
Group 5	$Th^{+/-}$ MMR^+	10	Current vaccine regimen recommended in 2005, including the MMR (both thimerosal-containing and thimerosal-free vaccines given under the current regimen)

Th - thimerosal; MMR - Mumps, measles, rubella vaccine; N – number of infants; $Th^{+/-}$ - with or without thimerosal; $MMR^{+/-}$ - with or without MMR vaccine; *10 infants in Group 4 will enter a treatment study (Phase III).
_____

By 2009, Phase II of the study was transferred from the Department of Obstetrics and Gynecology at the University of Pittsburgh School of Medicine to the Department of Psychiatry at the University of Texas Southwestern.

In 2012 one of the investigators, Dwight German from the University of Texas Southwestern, requested additional funding for a sixth exposure arm. The new arm (termed the 1990s Pediatric group in the Gadad and Curtis papers) would expose the primates to the 1990s vaccine schedule that was not accelerated based on the more rapid development of the infant macaques compared to human infants (4 human years to 1 primate year). At the time of the funding request, the researchers were reporting significant changes in neuropathology between the 1990s Primate schedule and controls consistent with autism brain findings:

> "based on our preliminary studies, we have found a significant decrease in Purkinje cell number and CA1 hippocampal cell size in monkeys given the 1990s vaccination schedule (which includes a cumulative total of 41.58 ug of EtHg). The thimerosal-containing vaccines (TCVs) were administered at 6 different time points (from birth to 12 months of age, and animals sacrificed at ~18 months). The animals received

injections of TCVs at a 4:1 schedule compared to the pediatric schedule. For example, humans receive the vaccines at birth, 2 month, 4 months, 6 months, 12-18 months and 48 months of age. The macaques received injections at birth, 2 weeks, 4 weeks, 6 weeks, 12-18 weeks and 52 weeks. The rationale for the 4:1 schedule is based upon much more rapid brain development of the macaque vs. human. However, in order to be certain that the observed neuronal changes in the macaque, that parallel those observed in post-mortem autism brains, are not due to the accelerated schedule of vaccine administration (and to satisfy any reviewer concerns at publication), we feel that it is very important to add another group of infant macaques to our study that will receive injections on the same schedule as given to human infants (i.e. birth, 2 month, 4 months, 6 months, 12-18 months and 48 months of age). The final injection at 48 months, will not be given to macaques in this group so that we are able to perform stereological analyses for all animals at ~18 months of age".

In November 2013, findings from Phase II were first publicly presented in an Abstract (14) and Poster (link to poster) at the annual Society For Neuroscience (SFN) conference. Like the findings update in the 2012 request for funding to SafeMinds above, the SFN Abstract and SFN Poster reported significant differences in brain pathology between exposed and control groups, as discussed below (Phase II Results).

Subsequently, in 2014 and 2015, study presentations were reporting _no_ effects from vaccine exposures. Five abstracts were submitted to the NBTS conference, one of which was on neuropathology and which reported that "no neuronal cellular or protein changes were observed in vaccinated animals compared to controls." (15) The paper by Curtis et. al. on learning and behavior appearing in 2014 and the paper by Gadad et. al. on neuropathology and behavior published in 2015 declared finding no or no consistent adverse effects consistent with those observed in autism from any of the vaccine exposure regimens.

PHASE II RESULTS

Same Data/Opposite Results on Hippocampus CA1 Size

Here we compare the discrepant findings between Phase II neuropathology results of the SFN poster and abstract and the Gadad paper. The Gadad Results section on brain pathology reported no significant findings in CA1 cell size between the groups (Figure A). This finding is based on 16 animals in the Control group, 12 in the 1990s Primate, and 8 in the 2008 groups, as shown in the legend of Gadad Fig 4 (Figure B).

The poster from the 2013 SFN conference is based on the same sample of 16 in the Control group, 12 in the 1990s Primate, and 8 in the 2008 groups reported in the Gadad paper. That the sample size is the same can be ascertained by a visual count of the data points given for the CA1 results from the poster (Figure C). This sample size is corroborated by the legend to the SFN poster Fig. 1 which references group cells sizes of 8-16 (Figure D). The Poster results section reports a significant reduction of 11% in CA1 size for the 1990s Primate vs. Control group (Figure D), in direct contradiction of the later Gadad paper.

Figure A. Results section from Gadad on Hippocampus CA1 cell size, reporting no difference between controls and exposed.

> *Hippocampus.* The CA1 neurons in the hippocampus have been reported to be reduced in size in postmortem brains from children with autism (18).
>
> *CA1 cell size.* Cell size (area) was measured in Nissl-stained sections at a rostral (section 100), middle (section 200), and a caudal (section 300) level of the CA1 region (Fig. 4). Approximately 250–450 cells were measured per animal, each with a clear nucleolus at the three levels of the nucleus. There was no significant reduction in cell area for the 1990s Primate group vs. Control group or for the 2008 group vs. Control group.

Figure B. Gadad Fig 4 showing CA1 results, and text reporting the sample size for the exposure groups. The sample size for the controls is 16, for 1990s Primate is 12, and for 2008 is 8.

Fig. 4. CA1 cells in the hippocampus. (A) Location of the CA1 region. (B) Neurons at a higher magnification. Red arrows point to cells with a visible nucleolus. (Inset, Right) High magnification view of two neurons one with a visible nucleolus. (C) Cell size data for Control (n = 16), 1990s Primate (n = 12), and 2008 (n = 8) groups. [Scale bars, (A) 1 mm and (B) 25 µm.]

Gadad et al.
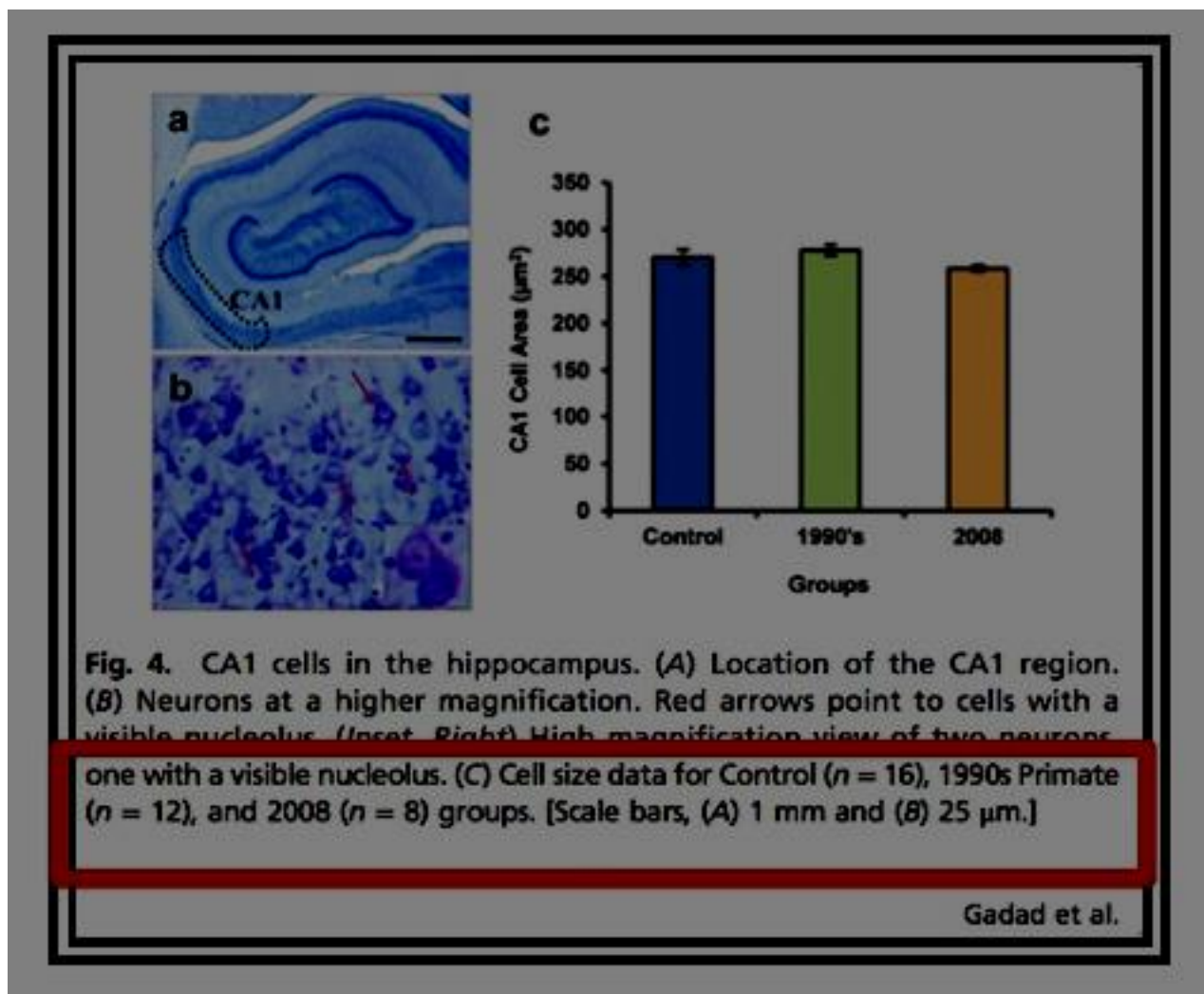
Figure C. SFN Poster section of the results of CA1 cell size for the controls (left box), 1990s Primate (middle

box) and 2008 (right box), providing individual level data. The number of dots, that is the sample size, for the controls is 16, for 1990s Primate is 12, and for 2008 is 8.
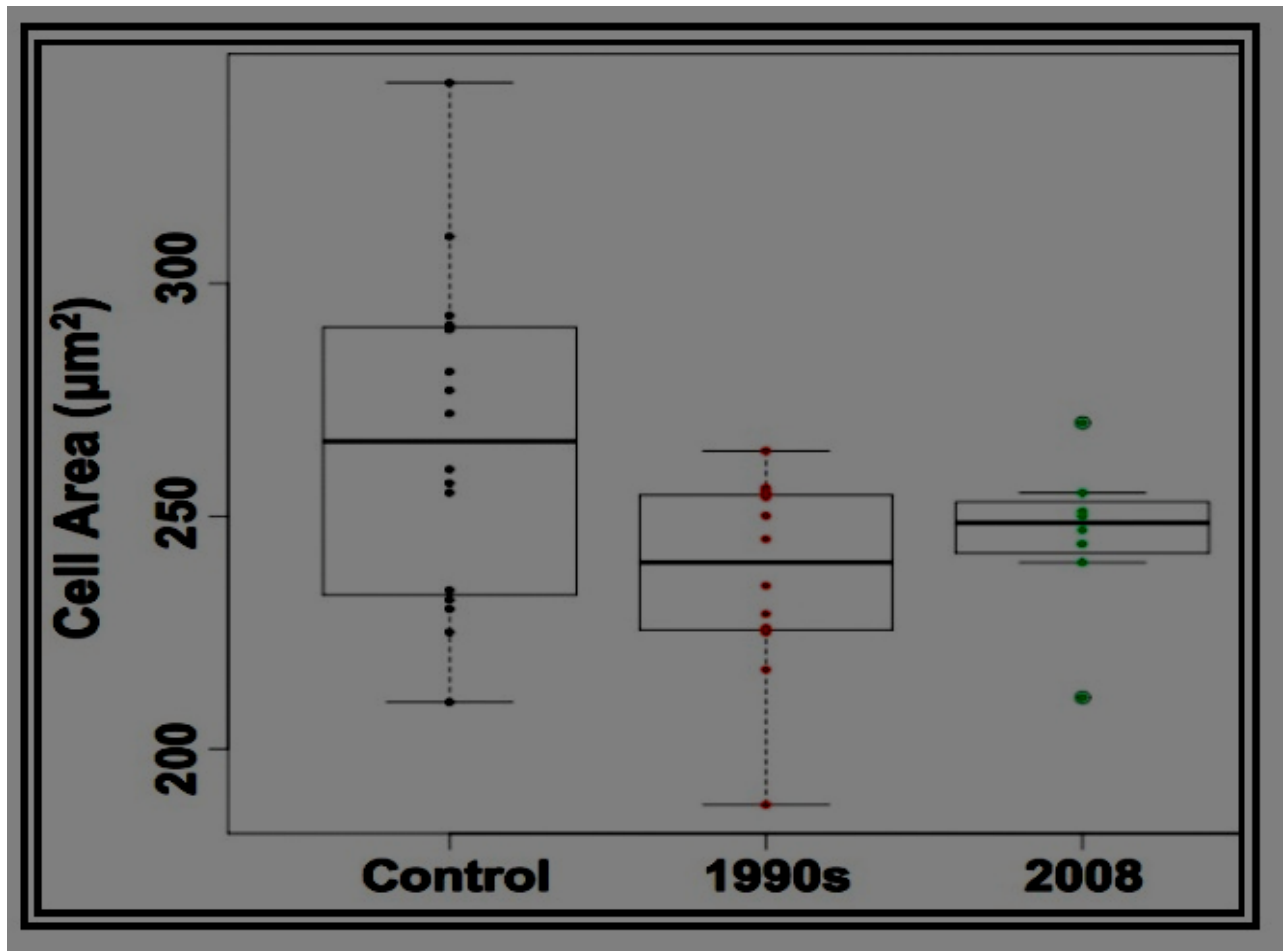
Figure D. SFN poster section of CA1 results with legend below figure which state that the difference in CA1 cell size is significantly different between controls and 1990s Primate, and the group cell size ranged from 8 to 16.

Figure 1: CA1 soma size was measured in the 5 groups. (A) CA1 cells measured contained a clear nucleolus. (B) There was an 11% reduction in soma area in the 1990s group vs. the saline control (p<0.05). N=8-16/group.
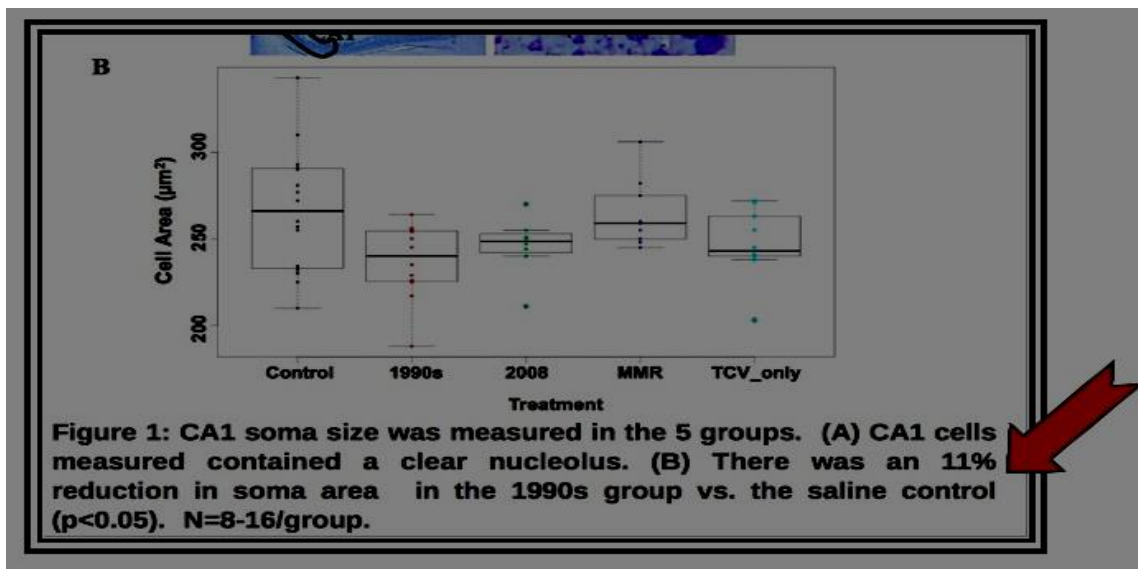
Disappearance of Purkinje Cell Number Significance

Besides the SFN poster, an abstract was submitted to the 2013 SFN conference. Abstracts are generally due several months in advance of a scientific conference, thus the abstract would pre-date the 2013 SFN poster by several months. The abstract reported a significant reduction of 7% in Purkinje cell count (Figure E). This Abstract also stated the same CA1 findings as the Poster, that is, an 11% reduction in CA1 cell area, as above (Figures C & D). Since the neuropathology examinations would have been undertaken at the same time, it is likely that any Purkinje findings would be based on the same sample as CA1 findings if they are reported in the same communication. Yet by the time the SFN conference poster was created (ie, after the abstract), the previously reported significant difference in Purkinje cell count had disappeared. (Figure F)

Figure E. Except from 2013 SFN Abstract reporting reduction in Purkinje cell count and hippocampus CA1 size.

Significant reductions in Purkinje cell number (7%) and CA1 cell size (11%) were observed in animals in the 1990's vaccination group, but no difference in amygdala cell number or cell size. CA1 cell size was not reduced in the 2008 vaccination group. These data from 1990's vaccine/Hg treated macaques are consistent with Purkinje cell and limbic abnormalities reported in the brains of people with ASDs.
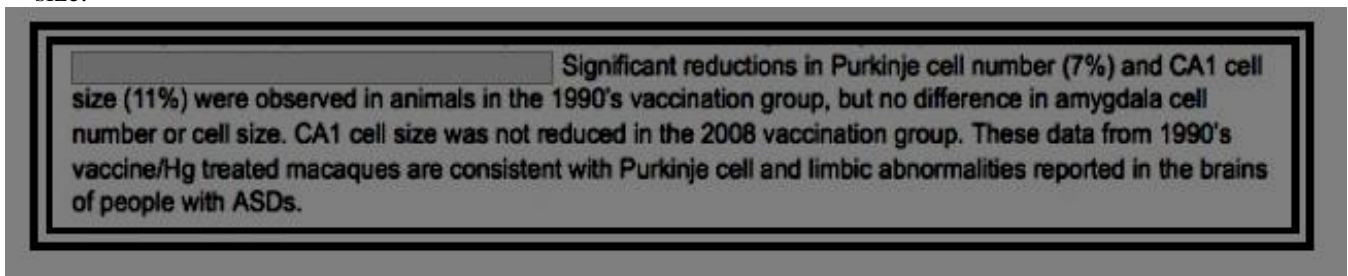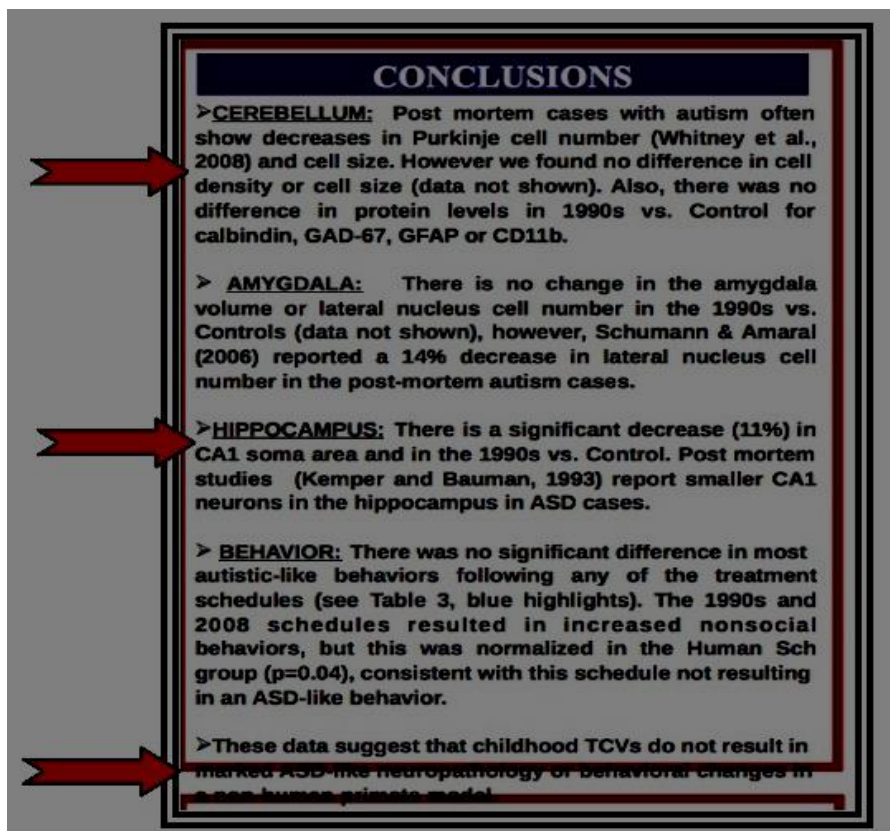
Figure F. Conclusions section from the 2013 SFN Poster showing no difference in Purkinje cell count, a decrease in CA1 area, and increase in non-social behaviors, between exposed and control groups, and conclusions of no effect overall.

CONCLUSIONS

➤CEREBELLUM: Post mortem cases with autism often show decreases in Purkinje cell number (Whitney et al., 2008) and cell size. However we found no difference in cell density or cell size (data not shown). Also, there was no difference in protein levels in 1990s vs. Control for calbindin, GAD-67, GFAP or CD11b.

➤ AMYGDALA: There is no change in the amygdala volume or lateral nucleus cell number in the 1990s vs. Controls (data not shown), however, Schumann & Amaral (2006) reported a 14% decrease in lateral nucleus cell number in the post-mortem autism cases.

➤HIPPOCAMPUS: There is a significant decrease (11%) in CA1 soma area and in the 1990s vs. Control. Post mortem studies (Kemper and Bauman, 1993) report smaller CA1 neurons in the hippocampus in ASD cases.

➤ BEHAVIOR: There was no significant difference in most autistic-like behaviors following any of the treatment schedules (see Table 3, blue highlights). The 1990s and 2008 schedules resulted in increased nonsocial behaviors, but this was normalized in the Human Sch group (p=0.04), consistent with this schedule not resulting in an ASD-like behavior.

➤These data suggest that childhood TCVs do not result in marked ASD-like neuropathology or behavioral changes in a non-human primate model.

Likewise, in the Gadad paper, no significant finding of Purkinje cell reduction was reported (Figure G),

with the same sample size of 16 for controls, 12 for 1990s Primate, and 8 for 2008 described for the SFN poster. Rather, the Gadad Supplement Table S2 showed a reduction of 2.3%, not significant (Figure H).

Thus, what has happened is that for the neuropathology for which we have data reported over a period of 4 years (2012-2015), that is, for CA1 and Purkinje cells, we see a gradual elimination of reported significance over time, even when the same sample size is reported. A critical question is, why did the results differ so significantly between the abstract, poster and final paper.

Figure G. Gadad Fig2 showing results for Purkinje cell number and reporting no significant difference. The sample size is 16 Controls, 12 1990s Primate, and 8 2008.
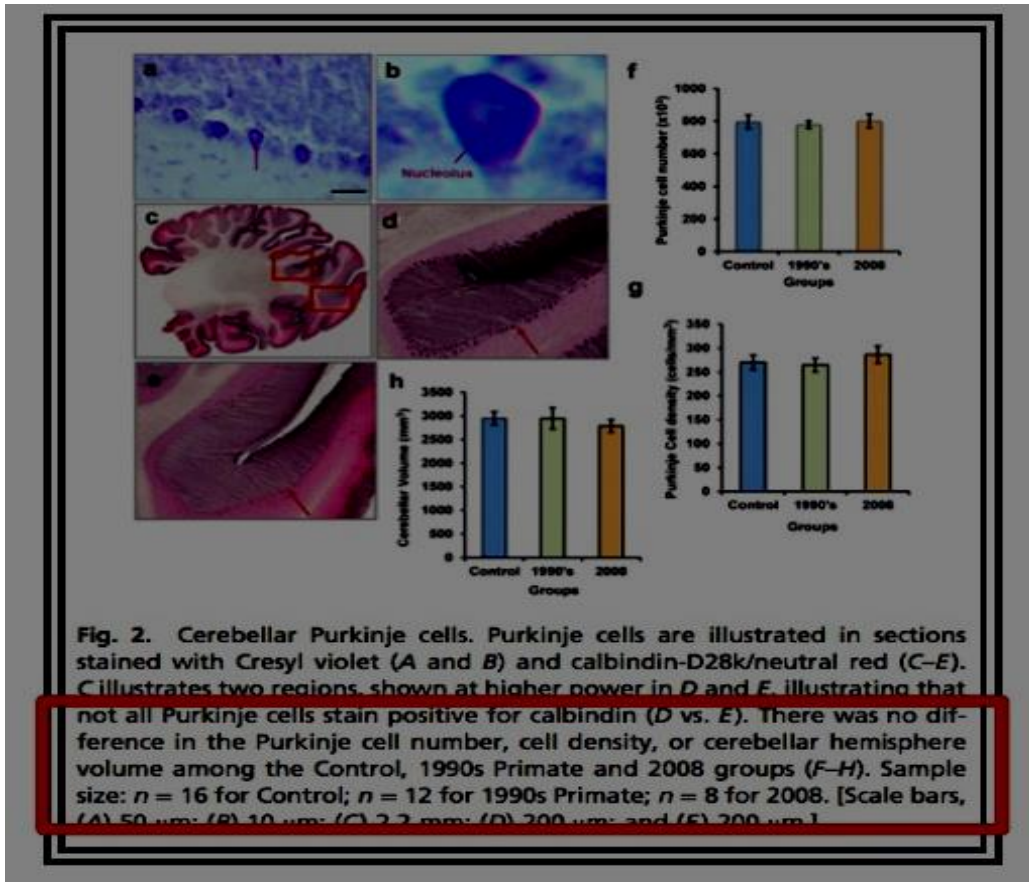

Fig. 2. Cerebellar Purkinje cells. Purkinje cells are illustrated in sections stained with Cresyl violet (A and B) and calbindin-D28k/neutral red (C–E). C illustrates two regions, shown at higher power in D and E, illustrating that not all Purkinje cells stain positive for calbindin (D vs. E). There was no difference in the Purkinje cell number, cell density, or cerebellar hemisphere volume among the Control, 1990s Primate and 2008 groups (F–H). Sample size: $n = 16$ for Control; $n = 12$ for 1990s Primate; $n = 8$ for 2008. [Scale bars, (A) 50 μm; (B) 10 μm; (C) 2.2 mm; (D) 200 μm; and (E) 200 μm.]

Figure H. Gadad Supplement Table 2 showing a lower Purkinje cell count number for the 1990s group which is not significant.

Table S2.   Purkinje cell number (mean ± SEM) for five groups of animals

Group	Mean ± SEM
Control ($n = 16$)	795,754 ± 10,544
1990s Primate ($n = 12$)	777,423 ± 6,560
2008 ($n = 8$)	804,689 ± 31,610
TCV ($n = 5$)	824,977 ± 18,129
MMR ($n = 5$)	787,866 ± 8,422

ANOVA indicated that there was no difference among the groups ($F = 1.55$, $P = 0.207$).

Sample Increases and Disappearance of Significance

We have shown in the previous section that for CA1 neuron area, the earlier SFN poster and abstract and the later Gadad paper findings are based on the same sample, and that it is likely that the Purkinje cell number data is also based on the same sample of animals. The Gadad et. al. authors have argued that the sample is different, despite what was written in the SFN abstract and poster, and that the published Gadad paper is really based on a larger sample and thus its finding of no significance (in terms of p value) is somehow more accurate than a finding of significance from a smaller sample. As Dr. Hewitson said, the data in the earlier reports "should be treated as preliminary until all of the animals had completed the study." (16) In an email to SafeMinds, Dr. Hewitson stated:

> "As you well know, when you increase a sample size, the effect size will go down and so it is not surprising that p values may change as we add more animals to a study group."

This statement is technically incorrect: changing N does not change the effect size, and increasing N decreases the standard error, which decreases the p value making the results *more* significant, not *less* significant. As explained in more detail below in the section on comparing Phase I and Phase II opposing findings, significant findings are more likely to be found with increased N - not less likely. The contention is wrong on grounds of basic biostatistics.

Reporting of Mean Values vs. Individual Level Data

An alternative explanation to sample size for the Phase II neuropathology discrepancies may be bias. Animals in Phase II were added to the study over five breeding seasons (2008-2012). The animals added later may be different in key ways from those included earlier. Misclassification or other forms of bias are common and often hard to deal with, creating misleading results regardless of sample size. One way to examine whether the new animals differed in key ways is to provide the data on an individual level in addition to group average of values. The individual data points could be identified by when they were included in the study analyses, to see any differences in results between the earlier time when neuropathology findings were significant and later times when they were not.

A further benefit of reporting individual level data as well as mean values is the ability to detect rare outcomes by identifying subjects with abnormal values in each group. Autism occurs in 1 in 68 children, and if a vaccine association is real, it would likely arise as an unusual or idiosyncratic response to the vaccination in only a few individuals. This point in fact was made by the investigators: (9)

> "We purposefully assigned a larger number of animals to the exposed group in order to optimize the chances of observing what we anticipated to be an uncommon or idiosyncratic effect."

Yet reporting only means may obscure this effect if it occurs in a few animals per group. The SFN Poster does provide a look at the individual level data (Figure D above), and variability is evident.

Another example is from Gadad Table S2 which displays the Purkinje cell number for the various groups and the standard error of the mean (SEM), from which the standard deviation (SD) can be calculated. The Purkinje cell number data for the 2008 group has a high standard deviation indicating that at least one animal had an unusual number of Purkinje cells. The control group SD is also large. Such idiosyncratic or outlier effects, if only occurring in 1-2% of a population (like autism) would not be detected with a mean value reporting method even with a relatively large sample size.

Behavior Findings

A number of observed differences between placebo and exposed groups on learning and behavior were reported in the text and figures of Gadad and Curtis, listed below. Yet both papers ignore these findings in their conclusions:

> Curtis: "This comprehensive 5-year case–control study, which closely examined the effects of pediatric vaccines on early primate development, provided no consistent evidence of neuro-developmental deficits or aberrant behavior in vaccinated animals."

> Gadad: "These data indicate that administration of TCVs and/or the MMR vaccine to rhesus macaques does not result in...aberrant behaviors, like those observed in ASD."

There was, in fact, a significant difference in acquisition of one of the 19 neonatal reflexes measured in Curtis: "There were no significant differences between groups in days to criterion for the acquisition of neonatal reflexes except for hand top of counter.... This effect was driven by the 1990s Pediatric group." This finding may or may not have occurred by chance due to multiple measures and may or may not be important.

In Supplemental Figure 5 of Curtis, Group C - TCV has a marked deficit in learning as compared to the control Group A, but the authors dismiss this in the main paper because it was not recapitulated in the other exposure groups. However, if outlier or idiosyncratic responses are expected in only a few animals (as discussed previously), one might only find a response in one exposure group and not others. The authors also try to dismiss the Group C - TCV deficit by saying there was no strategy deficit, but in Group E – 1990s Pediatric, there is a slope difference in learning deficit, which would indicate a strategy problem in learning. The authors also try to dismiss this association by saying that they performed well in other tasks, hence, the evidence is inconsistent. This line of reasoning is incorrect on two points. First, as the authors point out, mercury exposure in macaques has been shown in other studies to facilitate learning for some tasks even as it leads to deficits in others. Second, children on the autism spectrum tend to show inconsistent strengths and weaknesses in different skills, so if a study is using monkeys as a model for autism, it should employ logic consistent with autism.

Gadad reported abnormal findings in non-social explore behavior, which they brushed aside. This behavior had the highest duration and frequency of all measured behaviors and hence the most data. Non-social explore was categorized as an "autistic behavior" in the SFN Poster (Figure I). The Gadad paper says that "non-social explore behavior was significant in all the vaccine exposed groups compared to the controls, with the control animals exhibiting significantly more non-social explore behavior at the beginning of social living". They go on to say this difference in behavior disappeared after 6 months of social living, and conclude that "low dose TCVs via vaccination in our study did not significantly impact behavior." Similarly, the Curtis paper examination of behaviors in the earlier age group (2-12 months) reported significant differences between Controls and the 2008 and 1990s Primate groups at 2 months of age in negative behaviors, consisting of withdrawal, fear/disturbance, rock-huddle-self-clasp, and stereotypy. These are also categorized as "autistic behavior" (Figure I). Yet Curtis concludes no adverse effects because the significance disappeared by age 12 months (the equivalent of age 4 years in human development). However, the non-social explore and the negative behaviors may have normalized for a number of reasons. For example, the control animals may have habituated to their environment after 6 months or the exposed primates were able to ultimately model and learn this behavior over time from their control peers. They do not provide evidence against a developmental effect.

Figure I. SFN Poster results of behavior data for age 2-12 months. "Autistic" behaviors are highlighted in blue per the legend, and include non-social explore and the negative behaviors of withdraw, fear-disturbance, rock-huddle-self-clasp and stereotypy.

BEHAVIOR

Table 3. Social and Nonsocial behavioral activity   of Groups 1-6. N=12/group.

Dependent Variable	Duration F	p	Frequency F	p	Significant Effects
OVERALL					
Total Nonsocial	1.63	0.16	2.48	.04	Human Sch. < 2008, 1990's
Total Social	0.74	0.80	0.64	0.87	
Self-Directed	0.90	0.49	1.13	0.35	
Total Behavior Ch			1.73	0.14	
Scratch	1.23	0.30	0.66	0.65	
Groom	---	---	---	---	Too few occurrences to analyze
NONSOCIAL					
Passive	0.66	0-.85	1.86	0.11	
Explore	1.19	0.32	2.18	.065	
Withdraw	---	---	---	---	Did not occur
Fear-Disturb	---	---	---	---	Too few occurrences to analyze
Rock-Huddle-SC	---	---	---	---	Too few occurrences to analyze
Stereotypy	---	---	---	---	Too few occurrences to analyze
Play	1.82	0.12	1.96	0.10	
Sex	---	---	---	---	Too few occurrences to analyze
Aggress	---	---	---	---	Too few occurrences to analyze
SOCIAL					
Passive	3.23	.011	1.99	0.09	No significant multiple comparisons
Explore	1.08	0.38	3.24	.011	No significant multiple comparisons
Withdraw	0.64	0.67	0.67	0.64	
Fear-Disturb	0.65	0.66	0.73	0.60	
Rock-Huddle-SC	1.02	0.42	0.88	0.52	
Stereotypy	---	---	---	---	Did not occur
Play	2.28	0.02	1.26	0.29	No significant multiple comparisons
Sex	2.80	0.02	3.63	.006	2008 > Saline
Aggress	---	---	---	---	Too few occurrences to analyze

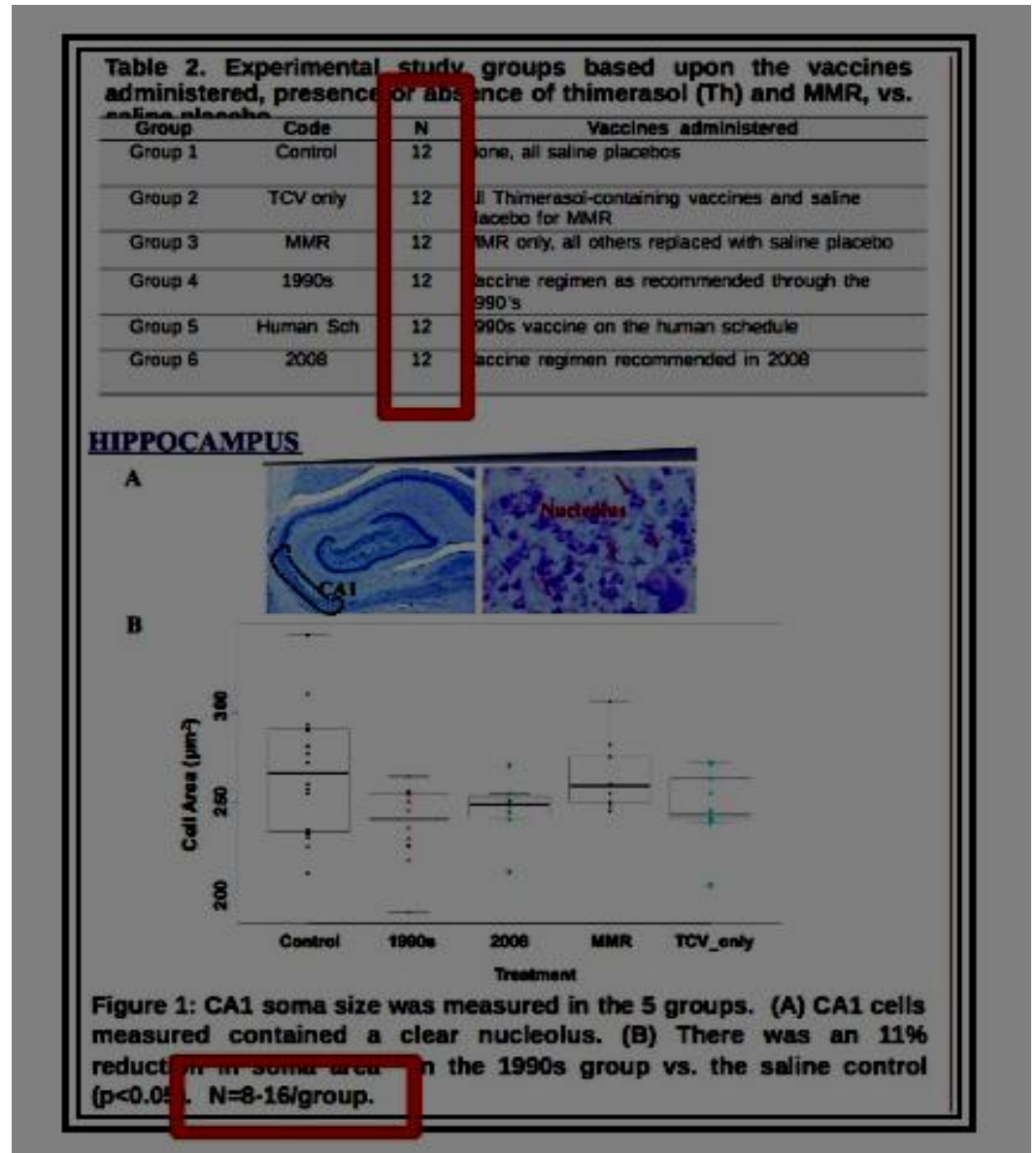*Autistic behaviors highlighted in blue.

Idiosyncratic or outlier responses on the behavioral measures might be obscured through statistical reporting of means only, as opposed to also showing individual level data and calculation of percent falling below the norm, as elaborated on previously for the neuropathology sections. Further, the Curtis paper notes: "Duration values were natural log-transformed to reduce the possibility of disproportionate influence from extreme values" in their analysis of social and non-social behaviors. Likewise, Fig 1 in Gadad refers to the behavioral data as "back-transformed with antilog". Applying log transformation, when extreme observations are present, will have the effect of lessening their influence on the means. The justification for using log transformation was not given, and it is not an appropriate way to deal with extreme values /outliers when such observations are expected as part of the model used.

Sample Size Inconsistencies

The reported sample size per exposure group (Ns) was inconsistent between and within the SFN Poster, the Gadad paper and the Curtis paper, and results for some groups for some analyses are not reported at all, as summarized in Table 2. At the least, the paper should make it clear in a table how the numbers fit together, provide justification for changing or not reporting the numbers, and state what the implications might be. Examples of inconsistent reporting of samples sizes follow.

(a) The Poster experimental design description has all groups being of equal size, N=12, while the Figure just below it shows data points for 16 in the Control, 12 in the 1990s Primate and 8-9 in the remaining 3 exposure groups, and the text for the figure state a sample size of 8-16. (Figure J)

——————

Figure J. Sample sizes reported from the 2013 SFN Poster. The Table 2 at the top gives group sizes of 12 each, which do not match the plots and figure text below it, which range from 8-16.

Table 2. Experimental study groups based upon the vaccines administered, presence or absence of thimerasol (Th) and MMR, vs. saline placebo.

Group	Code	N	Vaccines administered
Group 1	Control	12	None, all saline placebos
Group 2	TCV only	12	All Thimerasol-containing vaccines and saline placebo for MMR
Group 3	MMR	12	MMR only, all others replaced with saline placebo
Group 4	1990s	12	Vaccine regimen as recommended through the 1990's
Group 5	Human Sch	12	1990s vaccine on the human schedule
Group 6	2008	12	Vaccine regimen recommended in 2008

HIPPOCAMPUS

Figure 1: CA1 soma size was measured in the 5 groups. (A) CA1 cells measured contained a clear nucleolus. (B) There was an 11% reduction in soma area in the 1990s group vs. the saline control (p<0.05. N=8-16/group.

(b) Curtis Table 1 gives different group sample sizes than the SFN Poster Table 2 or Figure 1, some higher and some lower. (Figure K) The 1990s Primate group now shows 16 animals and the Controls show 12 animals, the reverse of the plot data points of the Poster. The MMR group shows 15, rather than the 12, which was reported in the Poster's Table 2. (Figure J)

_____

Figure K. Table 1 taken from Curtis. Group sample sizes show 12 for Controls, 15 for MMR, and 16 for 1990s Primate.

_____

(c) The Curtis Supplement Table S1 gives different group sizes than the main Curtis paper Table 1, with 16 in the Control (the "20" in the Total N column is an obvious typo) and 12 in the 1990s Primate. (Figure L)

Table 1. Study group sample sizes (*n*), and schedules for vaccine administration.

Group	n	Birth	2 weeks	4 weeks	6 weeks	15 weeks	52 weeks
Control	12		Saline	Saline	Saline	Saline	Saline
		Saline	Saline	Saline	Saline	Saline	Saline
			Saline	Saline	Saline	Saline	
MMR	15		Saline	Saline	Saline	MMR	MMR
		Saline	Saline	Saline	Saline	Saline	Saline
			Saline	Saline	Saline	Saline	
TCV	12		Hep B	Hep B	Hep B	Saline	Saline
		Hep B	DTaP	DTaP	DTaP	DTaP	DTaP
			Hib	Hib	Hib	Hib	
1990s Primate	16	Hep B	Hep B	Hep B	Hep B	MMR	MMR
			DTaP	DTaP	DTaP	DTaP	DTaP
			Hib	Hib	Hib	Hib	
1990s Pediatric[a]	12	Hep B	Hep B	Hep B	Hep B	MMR	None
			DTaP	DTaP	DTaP	DTaP	
			Hib	Hib	Hib	Hib	
2008	12	See Supplemental Material, Table S3, for details					

Abbreviations: Hep [...] vaccine; DTaP, diphtheria, tetanus, acellular pertussis vaccine; Hib, Haemophilus influenza B vaccine; MMR, measles, mumps, rubella vaccine; TCV, thimerosal-containing vaccines.
[a]For the 1990s Pediatric group, vaccines were administered at birth, 2 months, 4 months, 6 months and 15 months; the MMR and DTaP boosters were not administered at 52 months because animals were sacrificed at approximately 18 months.

VOLUME 123 | NUMBER 6 | June 2015 · Environmental Health Perspectives

Figure L. Table S1 taken from Curtis Supplement. Group sample sizes show 16 for Controls, 15 for MMR, and 12 for 1990s Primate.

Table S1. Assignment of animals to each study group across the 5 breeding seasons.

Study Group[a]	2008 N	2009 N	2010 N	2011 N	2012 N	Total N
Control	4	4	0	8	0	20
MMR	3	0	4	8	0	15
TCV	0	4	4	4	0	12
1990s Primate	0	4	8	0	0	12
1990s Pediatric	0	0	0	0	12	12
2008	0	4	4	4	0	12
Total	7	16	20	20	12	79

[a]Animals in each study group were derived from pregnancies from multiple breeding seasons. For example, animals in the control group were included in years 2008, 2009 and 2011, whereas animals in the TCV group were included in 2009, 2010 and 2011. The only exception to this was made for animals in the 1990s Pediatric group. This group was added to the study protocol in 2011 as a protocol modification, so all pregnancies for this group were derived in the last year of the study (2012).

(d) Consistent with Curtis S1 but not Curtis Table 1, the Gadad paper has 16 in the Control group, 15 for MMR and 12 for the 1990s Primate as well as the remaining exposure groups. (Figure M)

Figure M. Table 1 taken from Gadad. Group sample sizes show 16 for Controls, 15 for MMR, and 12 for 1990s Primate.

Table 1. Vaccination schedules used for the six groups of animals

Group	N	Vaccines administered
Control	16	None, all saline placebos
1990s Pediatric	12	Vaccine regimen as recommended in the 1990s
1990s Primate	12	Vaccine regimen as recommended in the 1990s accelerated fourfold
TCV's	12	All TCVs and saline placebo for MMR
MMR	15	MMR only, all others replaced with saline placebo
2008	12	Vaccine regimen recommended in 2008

(e) Figures 2 & 4 in Gadad showing the findings on brain pathology describe an N of 16 for Controls and N of 12 for 1990s Primate, but only 8 for 2008. Western Blots of cerebellar proteins in Gadad Figure 3 report values of just 8 for each of the controls, 1990s Primate and 2008 groups. Figures 5 & 6 from Gadad of the dentate gyrus and amygdala staining respectively report on only 12 Controls, with 12 in the 1990s Primate and 8 in the 2008. Table S2, while having N=16 for Controls and N=12 for 1990s Primate, only has data for 8, 5 and 5 for the 2008, TCV and MMR groups respectively.

Table 2. Summary of sample sizes reported across Phase II iterations. N/A means no data was reported for this group.

Group	SFN Poster Table 2	SFN Poster Figure 1	Curtis Table 1	Curtis Table S1	Gadad Table 1	Gadad Figures 2 & 4	Gadad Figure 3	Gadad Figures 5&6	Gadad S2
Controls	12	16	12	16	16	16	8	12	16
1990s Primate	12	12	16	12	12	12	8	12	12
2008	12	12	12	12	12	8	8	8	8
1990s Pediatric	12	N/A	12	12	12	N/A	N/A	N/A	N/A
MMR	12	12	15	15	15	N/A	N/A	N/A	5
TCV	12	12	12	12	12	N/A	N/A	N/A	5

CONTRADICTORY PHASE I & II FINDINGS

As seen above, the results from Phase I are opposite those of Phase II. In any finished document from Phase I (published papers, submitted papers, conference abstracts), adverse effects are consistently found across brain, behavior, and GI. In the two published papers and later conference abstracts from Phase II,

conclusions consistently "provide reassurance that TCVs do not contribute to the negative effects associated with ASD." (17) In the outcome measures which overlap the two phases and are reported in a finished document– amygdala size, neonatal reflexes, and discrimination reversal learning – the earlier studies show significant differences and the later studies do not, or at least not enough for the authors to declare an adverse effect.

A change of this magnitude should be fully explained. It is generally agreed that the study protocol for Phase II followed that of Phase I. Per Dr. Hewitson, "none of the study's procedures changed once her team moved from the pilot program to a larger sample" (16), that is, from Phase I to Phase II. Rather, the authors attribute the opposing findings to a sample size increase in Phase II. The Curtis paper states, "This discrepancy (with 2010) is most likely due to the larger number of animals in the present study providing more accurate estimates." We argue that sample size increase cannot explain the discrepancy and other factors relating to bias and confounding are at play.

Sample Size Explanation. Phase I had a sample size of 9-12 for the 1990s Primate group and 3-9 controls depending on the analysis being conducted. Given the significant effects reported in Phase I from the sample used, the Phase II sample size was set at similar levels: originally at 10 per group, later increased to 12, and as noted above, later still increased to 15 or 16 for some arms, and lowered to 5-8 for some arms depending on the analysis. The reason for the much larger total N of 79 in Phase II was to add more exposure groups, not to increase significantly the N per group. The fact that there were only 3-9 controls in the first phase, versus 12 or 16 in the second, cannot explain away the apparent treatment effects seen in the Phase I study; in fact, the logic goes the other way.

An increase in sample size leads to a finding that was not significant on a smaller sample to become significant, or to continue to not reach significance. It is almost never the case that a significant finding on a smaller sample disappears with a larger N as is being claimed. The only way that a significant result with a smaller N would go away with a larger N, with only random sample statistics to explain it, i.e. not biases nor confounders, would be if there was a statistically significant result that was false in the smaller N that happened by chance. If the study consistently uses alpha of 0.05 for example (i.e. p needs to be <= 0.05 to be considered significant) then for each outcome considered there is only a 1/0.05 =1/20 chance of this happening due to randomness. If that is the case, then repeating the experiment with a different sample would be unlikely to find the same outcome in the second experiment. Note however that if the first experiment's results are much stronger than required for p<0.05, for example p<0.001, then for each outcome there is only a 1/1000 chance that the results were obtained purely by chance.

The odds of observing <u>multiple</u> significant findings (which Phase I did) in a smaller sample, which disappears with a larger sample, are virtually nil. If there really is no effect, yet the first study with smaller N gets multiple significant findings which disappear in a second study that is otherwise identical except for larger N, then those findings in the first study would have to be all entirely due to chance, and the probability of that happening would be approximately the product of the p values of all the significant findings. If there were, for example, a study that examined four outcomes and all four had significant findings with p values of 0.05, 0.04, 0.01 and 0.01, the probability of all of this happening would be 0.0000002, i.e. 1/5 million. In other words, if the initial findings were due only to chance then they would be extremely unlikely to occur. It is vastly more likely that the first study's results represented true effects, unless they were caused by a bias or a confounder (see next section below).

In Phase I, for some of the analyses, one of the Ns was small, i.e., 3 animals in the controls. The control group may have a small SD and a small N while the treatment group may have a larger SD and a larger N. In that case, the small N for the control group might not make much difference. Even if the control group has the same SD as the treatment group and a much smaller N, the result would be less likelihood of

finding a statistically significant result - not more likely, yet Phase I found significance and Phase II did not.

Other Explanations: Confounding and Bias. The bottom line is that increased sample size is not a valid explanation for findings disappearing. If a second study following the same protocol tested the same outcomes and did not find them, then almost certainly something else changed between the studies - probably one or more confounders or biases like sample biases or misclassification of exposure or outcome. More information is needed to try to determine if any of these are at play. Possibilities include:

(a) Exposure misclassification in Phase II, for instance, animals added to the study later were analyzed with the wrong exposure group, which might underlie the inconsistencies in descriptions of Phase II sample sizes noted previously.

(b) The lineage of monkeys may have differed from Phase I to Phase II. While needing confirmation, it is possible that Phase I macaques were of Indian origin and those of Phase II were Chinese (private communication with investigators) or Indian-Chinese hybrids, since during the 2000s Indian macaques became scarce due to an Indian government export ban and primate centers recognized the need for increased genetic diversity through expanded breeding sources. (18) The lineage answer is important, because each population has a distinct genetic profile, which is known to impact research results. (19)

> "The inadvertent inclusion of rhesus macaques from different geographic origins, such as China and India, or their hybrids, in experimental groups can foster results that are ambiguous because the elevated genetic contribution to phenotypic variance can obscure or swamp the contribution of experimental treatment effects to the total phenotypic variance. These taxa, particularly those of full-Chinese and full-Indian origins, can exhibit profound phenotypic differences in physiologic and behavioral traits that are controlled by genetic mechanisms."

Identified genetic differences consist of divergence in mitochondrial DNA and in MHC (major histocompatibility complex) genes involved in immune system function and brain development, both of which have been implicated in autism. Each macaque population has been shown to respond differently to infection, making, for example, Indian macaques a superior model for AIDS research because they respond more like humans to HIV infection. If the Phase II study sample was genetically less susceptible to vaccine side effects, the group sample sizes and the reporting of mean values would be less likely to identify rare outcomes, whereas if Phase I animals were more susceptible, the chance of observing a side effect would be higher.

(c) Illness differentially impacted Phase I and II animals. In Phase I, a giardiasis broke out in the colony, which was treated with antibiotics. (12) Either infection or antibiotics may have made Phase I animals more susceptible to vaccination side effects, and thus a significant finding more likely to be detected.

There are countless more possibilities. The point is that the inconsistencies between the two phases should be discussed and investigated, including reporting on the origin of the animals across time, not dismissed through a spurious and inadequate statistical argument on sample size.

LOW POWER OF PHASE II TO DETECT AN EFFECT

The Gadad et al., 2015 paper concludes: "Our data strongly support the conclusion that childhood TCVs do not produce ASD-like neuropathology or behavioral changes in the nonhuman primate." This conclusion is not supported by the statistical power of the study to find an effect in a low incidence

population which is characteristic of autism, currently at 1-3% of children, depending on the study cited. As explained below, the sample size of Gadad and Curtis is too small to show a lack of an effect but not too small to show an effect like the ones seen in Phase 1. Declaring that not finding an effect in Phase II is evidence that there is not an effect is simply incorrect. Declaring further that one's data "strongly" prove no effect is overreach, particularly given the opposite findings from Phase I.

Power is the probability of observing a statistically significant effect if there really is a specified degree of true difference, that is, if the true difference in means of the population is as least as great as some pre-specified value. The randomness introduced by sampling means that there is a chance that a real difference, if there is one, would not show up in the results, which is a false negative. Avoiding false negative results is what power is all about. When the data do not show a significant difference between the groups, that does not mean that there is no difference, because of power. For example, if power in a study is 80%, then if there is a true significant difference there is only an 80% chance of finding it and 20% chance of missing it. If power is 5%, there is only a 5% chance of finding it and 95% chance of missing it. To be comparable to the conventional 5% threshold used for p values, we would need to have power of 95%, with only a 5% chance of a false negative. If a study is designed to be, or billed as, a safety study, i.e. one that shows a lack of an effect, it should have power of approximately 95% (at least 90%), and it may use a less stringent p value (positive finding) criterion. Even then, if it does not find a significant effect, that does not prove there is not one, only that it is unlikely to obtain this non-significant result if in fact there is a difference of the size assumed in the power calculation.

It is possible to estimate power after the fact, called post-hoc power analysis, using estimated or assumed values and values from the study. Here is one example for power with a difference in proportions using values that might approximate those of this study. Assume the study is able to detect a difference between groups where the exposed group has the outcome of interest in 2% of the population, the control group has the outcome in 1%, each group has N=12 samples, and statistical significance is declared at p<0.05. Then the Z value of power is:
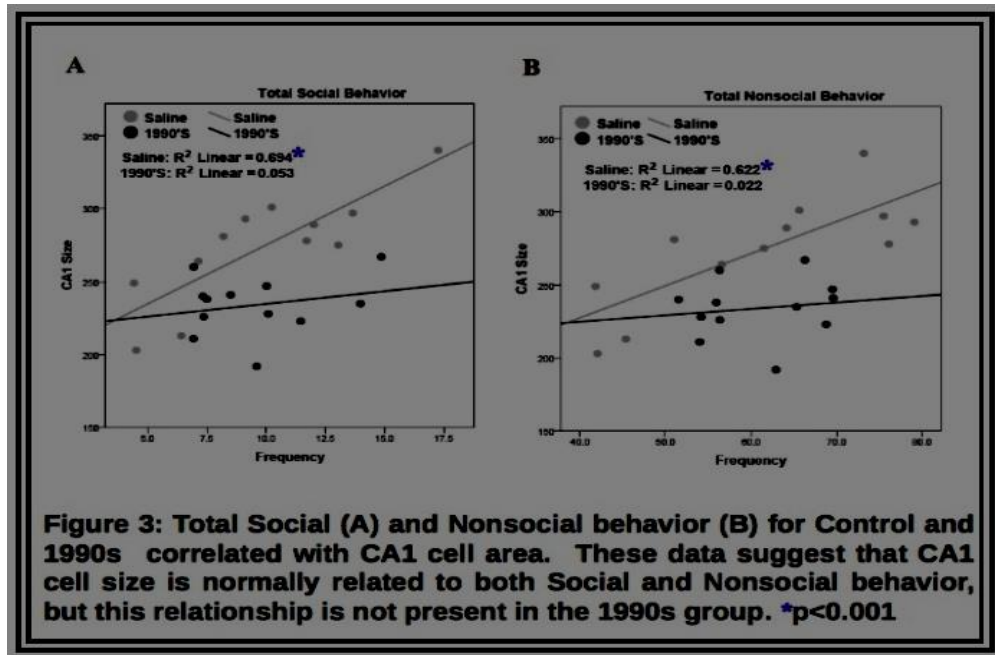
$$Z_{power} = \sqrt{12*(0.01)^2/(2*0.015*0.985)} - 1.96 = -1.76$$

and the power is about 4%. This means that if the true proportions of the outcome in the groups are 1% and 2% respectively there is only a 4% chance of observing a significant difference and a 96% chance of missing it. If either of the Ns is smaller, then the power gets smaller. To obtain a high probability of observing a significant effect, the difference between the groups would have to be many times larger than the difference assumed here in Phase II.

SELECTIVE REPORTING

Readers of the either Curtis or Gadad papers cannot know what tests the researchers ran and chose not to publish. For example, a finding presented at the 2013 SFN conference correlating CA1 size and behavior, which found significant positive correlation among controls but not exposed (Figure N), did not make it into the final Gadad publication. If the authors chose not to publish a finding that might conflict with their conclusion, that constitutes an important form of publication bias, which is a common concern with medical literature. If the study had published a detailed protocol in advance, as well as more supplementary information (which is not subject to space limitations), it would be easier to tell.

Figure N. 2013 SFN Poster section showing significant correlation between CA1 size and behavior for controls but not 1990s Primate.

Figure 3: Total Social (A) and Nonsocial behavior (B) for Control and 1990s correlated with CA1 cell area. These data suggest that CA1 cell size is normally related to both Social and Nonsocial behavior, but this relationship is not present in the 1990s group. *p<0.001

Researchers in other medical fields have raised questions about the publication process for clinical trials on efficacy or safety of medical interventions (20,21). One concern is that the process favors publication of trials whose results fit a pre-established view, while those with results that run counter do not get published, for a variety of reasons along the publication pathway and independent of methodological quality. These non-published or "invisible" trials are subsequently not considered in review papers or for practice and policy decision-making. As we noted above, after the first two publications, the many papers expected to be published from Phase 1 showing an adverse effect from vaccination were not published, while the Phase II papers concluding no adverse effects have been published in prestigious journals. Yet Phase I and Phase II followed the same protocols and model. There is no indication that Phase I is in any way deficient relative to Phase II. Publication bias may be a factor here.

The journals in which the two Phase II papers were published are Environmental Health Perspectives, a publication of NIEHS of the NIH, and PNAS which is the publication of the National Academy of Sciences, a federally chartered non-profit to advise the nation on issues related to science and inform public policy decisions (www.pnas.org). The review process at PNAS may reflect undisclosed bias, irrespective of intent.

The International Committee of Medical Journal Editors (ICMJE), in their "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication" states that "conflict of interest exists when an author (or the author's institution), reviewer, or editor has financial or personal relationships that inappropriately influence (bias) his or her actions (such relationships are also known as dual commitments, competing interests, or competing loyalties)." Moreover, the ICMJE further elaborates that "the potential for conflict of interest can exist regardless of whether an individual believes that the relationship affects his or her scientific judgment." (www.icmje.org/)

The Gadad paper in PNAS is termed a "Direct Submission". The PNAS website describes the process for a direct submission:

"Every published paper is peer reviewed and has been approved for publication by an NAS

member."

and

> "When a Direct Submission is received, the PNAS Editorial Board identifies an Academy member expert in the field of research described in the paper to serve as editor. The Member Editor may choose reviewers, guide modifications and revisions to the text, and decide whether the paper should be recommended for publication."

The Gadad paper identifies its PNAS Member Editor as Matthew State ("Edited by Matthew State, University of California, San Francisco, CA, and accepted by the Editorial Board August 9, 2015"). Dr. State's focus is genetics and child psychiatry (http://profiles.ucsf.edu/matthew.state) and he has published extensively on the prominent role of genetic mutations in autism etiology. He serves as chair of the scientific advisory board of the Autism Science Foundation, which has an official position that, considering all autism-vaccine research to-date, "the data show no relationship between vaccines and autism." (http://autismsciencefoundation.org/autismandvaccines.html). Paul Offit, who wrote a commentary in PNAS accompanying the Gadad paper, is on the board of the Autism Science Foundation. He is a developer and patent holder of a vaccine currently recommended in the U.S. childhood vaccination schedule.

CONCLUSIONS

Given the Phase II issues around sample size, inconsistent findings over time, and data reporting choices described in this review, the Gadad paper overreaches in its conclusion regarding vaccine safety: *"Our data strongly support the conclusion that childhood TCVs do not produce ASD-like neuropathology or behavioral changes in the nonhuman primate."*

1. Even if there were no biases and confounders in the study, it is not powered to show a lack of an effect for a condition like autism, which occurs in 1-3% of children.
2. Phase 1 showed large, significant results. No suggestions have been made by the authors that Phase I was deficient, and Phase II followed the same methodology and protocols.
3. There are significant differences between the Phase I and Phase II results, and between earlier and later versions of Phase II findings, and those discrepancies need to be explained. They cannot be explained by sample sizes. Misclassification, sample bias or other confounders should be explored.
4. There are discrepancies in the samples reported for the various versions of the papers, and these need to be explained.
5. The choice of data reporting approaches may be obscuring outlier effects. Such effects are likely in an animal model of autism and exposures, as the non-genetic risk for autism is hypothesized by the authors to be an uncommon or idiosyncratic response to environmental agents.
6. Across the known study documents – papers, abstracts, posters and reports for both phases - significant differences between exposed and unexposed have been observed, including brain pathology, learning and behavior problems, and gastrointestinal illness. These findings should be recognized in drawing conclusions on vaccine side effects.
7. Reporting bias including publication bias may be a factor in publication of Phase II and the strong conclusions of the Gadad paper, as well as the lack of publication of additional Phase I papers.

The autism-vaccine primate research spanned over 10 years, comprised two phases, many arms, and many outcome measures. The complexity of the effort cannot be adequately contained within a few journal

articles. Our review, and interpretation by others wishing to evaluate this research, is limited by lack of access to the raw data (both published and unpublished), all study protocols, and all study reports. To resolve the concerns raised here and allow verification, these materials should be posted publicly and immediately, even if the authors intend to conduct future additional analyses. Further, the unpublished papers from Phase I should be published, to correct reporting biases, and the authors should refrain from diminishing the importance of Phase I by claiming superiority of Phase II based on weak and statistically inaccurate arguments. Since Phase II of this research is being messaged as absolving vaccine regimens from having a role in autism risk, these recommendations have important implications for public health, vaccine safety, and autism prevention.

REFERENCES

1. http://ucdmc.ucdavis.edu/publish/news/newsroom/10214

2. Robin P. Goin-Kochel • Sarah S. Mire • Allison G. Dempsey. Emergence of Autism Spectrum Disorder in Children from Simplex Families: Relations to Parental Perceptions of Etiology. J Autism Dev Disord. 2014. DOI 10.1007/s10803-014-2310-8

3. Curtis B, Liberato N, Rulien M, Morrisroe K, Kenney C, Yutuc V, Ferrier C, Marti CN, Mandell D, Burbacher TM, Sackett GP, Hewitson L. Examination of the Safety of Pediatric Vaccine Schedules in a Non-Human Primate Model: Assessments of Neurodevelopment, Learning, and Social Behavior. Environmental Health Perspectives. Feb 18, 2015. http://ehp.niehs.nih.gov/1408257. Supplemental Material http://ehp.niehs.nih.gov/wp-content/uploads/123/6/ehp.1408257.s001.508.pdf

4. Bharathi S. Gadad, Wenhao Li, Umar Yazdani, Stephen Grady, Trevor Johnson, Jacob Hammond, Howard Gunn, Britni Curtis, Chris English, Vernon Yutuc, Clayton Ferrier, Gene P. Sackett, C. Nathan Marti, Keith Young, Laura Hewitson, and Dwight C. German. Administration of thimerosal-containing vaccines to infant rhesus macaques does not result in autism-like behavior or neuropathology. PNAS 2015; published ahead of print September 28, 2015, doi:10.1073/pnas.1500968112. http://www.pnas.org/content/early/2015/09/24/1500968112

5. American Academy of Pediatrics. Thimerosal in Vaccines—An Interim Report to Clinicians. Committee on Infectious Diseases and Committee on Environmental Health. Pediatrics Vol. 104 No. 3 September 1, 1999. pp. 570 -574. http://pediatrics.aappublications.org/content/104/3/570.full

6. Thimerosal (54-64-8) Nomination to the National Toxicology Program. Review of the literature, April, 2001. http://ntp.niehs.nih.gov/ntp/htdocs/chem_background/exsumpdf/thimerosal_508.pdf

7.   Department of Health and Human Services. 21 *Code of Federal Regulations*: 601.25. 1985.

8. Institute of Medicine. The Childhood Immunization Schedule and Safety: Stakeholder Concerns, Scientific Evidence, and Future Studies (2013). http://iom.nationalacademies.org/Reports/2013/The-Childhood-Immunization-Schedule-and-Safety.aspx

9. Hewitson, L., Lopresti, B. J., Stott, C., Mason, N. S., Tomko, J. Influence of Pediatric Vaccines on Amygdala Growth and Opioid Ligand Binding in Rhesus Macaque Infants: A Pilot Study. Acta Neurobiol Exp, 2010, 70:147-164.

10. Hewitson L, Houser LA, Stott C, Sackett G, Tomko JL, Atwood D, Blue L, White ER. Delayed acquisition of neonatal reflexes in newborn primates receiving a thimerosal-containing hepatitis B vaccine: influence of gestational age and birth weight. J Toxicol Environ Health A. 2010;73(19):1298-313. doi: 10.1080/15287394.2010.484709

11. .Hewitson L, Houser L, Sackett G, Stott C, Tomko J (2010b). Longitudinal cognitive development in rhesus infant macaques in response to vaccination: A pilot study, submitted. [cited in reference (9) above]

12. Laura Hewitson, Nithya Karthik, Carol Stott, Edwin Klein, Carlos Castro, Mario Rodriguez, Jaime Tomko, Sudhir Gupta, Carrie Redinger, and Andrew Wakefield. Evolution and Characteristics of Inflammatory Bowel Disease in Rhesus Macaques Exposed to Childhood Vaccine Schedule. Submitted per private correspondence.

13. Hewitson L, Houser L, Dettmer A, Sackett G. Discrimination Reversal Learning in Infant Primates Exposed to Low-Dose Thimerosal: A Pilot Study. Neurotoxicology and Teratology. Volume 33, Issue 4, July-August 2011, Page 501.

14. Hewitson-German poster abstract at the November 2013 annual Society for Neuroscience (SNF) conference of 2013. http://www.abstractsonline.com/Plan/ViewAbstract.aspx?sKey=d4097393-5c5e-4ed4-9ae6-398e149f7f10&cKey=25368e04-b3e5-47b5-ae2a-67d2891042ad&mKey=8d2a5bec-4825-4cd6-9439-b42bb151d1cf

15. Hewitson L, Gadad B, Li W, Grady S, Curtis B, Yutuc V, Ferrier C, Sackett G, German D. Does Administration of Thimerosal- Containing Vaccines to Infant Rhesus Macaques Result in an Autism-Like Neuropathology?. Neurtoxicology and Teratology. Volume 49. May-June 2015. Pages 135-136.

16. Interview with Jessica Firger, Newsweek. October 2, 2015.

17. Curtis B, Morrisroe K, Yutuc V, Ferrier C, Sackett G, Hewitson L. Influence of pediatric vaccines on social behavior in the rhesus monkey. Neurtoxicology and Teratology. Volume 43. May-June 2014. Page 90.

18. S Kanthaswamy, L Gill, J Satkoski, V Goyal, V Malladi, A Kou, K Basuta, L Sarkisyan, D George, DG Smith. The development of a Chinese-Indian hybrid (Chindian) rhesus macaque colony at the California National Primate Research Center (CNPRC) by introgression. Journal of Medical Primatology. 38(2):86-96. August 2008.

19. Sree Kanthaswamy, Alexander Kou, David Glenn Smith. Population Genetic Statistics from Rhesus Macaques (Macaca mulatta) in Three Different Housing Configurations at the California National Primate Research Center. J Am Assoc Lab Anim Sci. 2010 Sep; 49(5): 598–609. Published online 2010 Sep.

20. Peter Doshi, Kay Dickersin, David Healy, S Swaroop Vedula, Tom Jefferson. Restoring invisible and abandoned trials: a call for people to publish the findings. BMJ 2013; 346 doi: http://dx.doi.org/10.1136/bmj.f2865 (Published 13 June 2013)

21. Peter Doshi. No correction, no retraction, no apology, no comment: paroxetine trial reanalysis raises questions about institutional responsibility. BMJ 2015; 351 doi: http://dx.doi.org/10.1136/bmj.h4629 (Published 16 September 2015)