Emanuela Taioli MD PhD
Director, Institute for Translational Epidemiology
Associate Director for Population Science, Tisch Cancer Institute
Director, Center for the Study of Thoracic Diseases Outcome
Professor, Population Health Science and Policy, and Thoracic Surgery

New York, 11/05/2020

Review of "Health Effects Assessment: Potential Neurobehavioral Effects of Synthetic Food Dyes in Children".

Based on my expertise and experience, I am reviewing the findings, assumptions, or conclusions I agreed I could review with confidence - point n 1 of Attachment 2, and to the extent possible - point n 3 of Attachment 2.

Conclusion n. 1 of Attachment 2: After reviewing the epidemiological literature on the neurobehavioral effects of synthetic food dyes, OEHHA concludes that the data suggest an effect of artificial food dyes on children's neurobehavior.

Review:

The contributors did an outstanding job in summarizing and interpreting the available epidemiologic data on neurobehavioral effects of synthetic food dyes. I have conducted a literature review myself, and have not found additional epidemiologic studies that needed to be included. I reviewed the comments on statistical analysis and the discussion of statistical concerns and found them satisfactory.

This is an exhaustive review of randomized clinical trials (RCTs) on dietary dye use and neurobehavioral effects in children. The choice of RTC was made *a priori*, and was based on a series of premises that I embrace. After reading and reviewing the available RCTs, I would consider adding a brief review of available observational studies, mostly because I realized that the RCTs were conducted on convenience sample of children, thus have limitations related to selection bias and external validity, and may not be representative of what happens in the general population. The other issue to consider is that the epidemiologic data are really scarce anyway, and perhaps a look at observational studies can improve the understanding and interpretation of the observed associations. I also would stress that data is really old, the most recent study included was conducted in 2007. These last 13 years may have introduced significant changes in dietary patterns; I'm not sure if the dyes currently used are exactly the same in quality and quantity as it was 10-15 years ago. This latter aspect should probably be checked out and verified.

The review includes studies both on general population of children as well as attention deficit disorder (ADD) children; both sets of studies show effects on neurocognitive function. Most studies test several dyes together, thus making it impossible to point at one of them as the responsible for the association with neurocognitive effects. This may

not be a negative thing all together, since it may represent an opportunity to look at the real life situation of a diet where many different components and dyes are ingested at the same time. This situation gives to opportunity to assess the overall effect that such diet has on cognitive behavior.

When the studies are described, Im not clear on why summary estimates are not presented, together with tests for heterogeneity. This would allow calculating an overall estimate, as well as identifying important sub-group and conducting sensitivity analyses. It is hard to judge an effect if there is no quantitative assessment of the data, and the reasoning behind the lack of a formal meta-analysis is not clearly articulated. If a scientific rationale is behind this choice, it should be described in a paragraph, otherwise an attempt of meta-analysis should be performed.

One thing that should be addressed clearly is the gap in knowledge, and this should be clearly stated in the conclusion. For example the lack of data on genetic susceptibility, the lack of data on biomarkers such as DNA methylation before/after exposure, the missing link between short term and long term effects of these repeated brief exposures on brain development and function are all important gaps that need to be highlighted and filled as soon as possible.

Another aspect that should be highlighted is that animal studies suggest some possible mechanisms, such as oxidative stress, or binding of the dyes to proteins that regulate neurotransmitters function. Given the complete lack of information on the metabolism, internal dose, and biological effects of these food components in humans, the information from animal studies become key to guide the development and the direction of future human studies. The hypotheses generated by animal experiments about possible mechanisms through which these dyes act on neurological functions should be listed as priorities to be investigated in order to better understand the risk for children and the possible remediation measures.

I understand from the Nigg et al meta-analysis, a well conducted study although a little old by now, that there are also concerns among the experts about the effects of exposure to food dyes alone versus dyes plus preservatives, with the latter being more harmful, but this point has not been addressed, unless I missed it.

There are two ongoing studies that have not generated results yet, but should be mentioned in future plans or somewhere appropriate. These data collections are ongoing and will eventually contribute relevant information that reflect more recent dietary habits and food composition:
A two arm randomized controlled trial comparing the short and long term effects of an elimination diet and a healthy diet in children with ADHD (TRACE study). Rationale, study design and methods.
BMC Psychiatry. 2020 May 27;20(1):262. doi: 10.1186/s12888-020-02576-2.

Rationale and design of an international randomized placebo-controlled trial of a 36-ingredient micronutrient supplement for children with ADHD and irritable mood: The

Micronutrients for ADHD in Youth (MADDY) study. Contemp Clin Trials Commun. 2019 Oct 26;16:100478. doi: 10.1016/j.conctc.2019.100478.

I have some more detailed comments that could help improve the document:

Page 30 – the authors indicate that no exclusion of studies was made based on the number of participants. Usually a minimum number of participants is set, for example 10, especially here where we are looking at randomized trials. The exclusion of very small studies could address possible publication bias, and could reduce heterogeneity. Study quality: the process for assessing study quality should be described more in detail. The table of items used to assess quality is not derived from one of the validated scores published by NIH, as far as I can tell. There are many validated systems for quality assessment, some for descriptive studies, some for RCT. I suggest that, unless there is a good reason for using a personalized quality score, the authors should use a validated published system. If the authors decide that they will use the current list of items for quality scoring purposes, then there should be a section that describes how the items composing the list were chosen, and how the list was validated.

Pag 32 – there is some confusion between the concepts of confounders and the quality scores. For example, there is a comment on elimination diets studies been possibly more sensitive in showing neurobehavioral effects than other studies; however this hypothesis was not confirmed by preliminary analyses, therefore the item was not included in the quality assessment. I think that these are part of sensitivity analyses, where subgroups are studied, and should not affect the quality evaluation of the studies.

Pag 38 – I cant figure out why elimination studies were excluded from the sub-analyses. Instead I suggest that they should be analyzed separately, as additional information may derive from these studies

Pag 39 – range of participants is quite large, and perhaps a decision to limit to studies with > 10 participants would have helped. Same for the dye dose, the range is quite large. Again, some strategy earlier when defining the inclusion criteria may have helped here, or the decision of conducting some sub-analyses of certain doses that are more meaningful and representative of the average dietary usage

Pag 40, 41 – several important concepts are included in this short section, and perhaps they should be separated into paragraphs with subtitles. We read here about dose response, latency and age groups. All these issues should be described more in detail, including implications. For example, is the dose response showing effects at doses that are commonly used, or only at doses that are unrealistically high? Is dose response present in certain age groups but not others? How about race? This section is a little bit of the core of the results and needs to be expanded and interpreted with more details. If there is no information on issues such as the ones I described above, then it should be stated as a gap in knowledge. I think that pointing at issues that haven't been studied is as important as showing results of studies that were properly conducted.

Discussion: the discussion is very thorough and detailed. I have no major comments.

Pag 44 – Design issues: I wonder how one can affirm with certainty that a RCT conducted on a convenience sample is superior to an observational study. This concept brings up the idea I discussed earlier that perhaps observational studies should have been given more weight, given the scarcity of available data.

Susceptibility: this is a very relevant paragraph, and again should be expanded. The first issue just touched upon is that younger children seem to be more sensitive than older children. Why is that? And would the increased susceptibility in younger children translate in any long-term effect on brain development deriving from this early sensitivity? If there is no scientific literature, then the gap should be highlighted. Genetic polymorphisms: looks like a metabolic chain is involved in degradation and elimination of these dyes, and clearly germline variations may play a role in individual sensitivity. The questions are: what is the degradation pathways of these dyes in humans? What are the genes involved? What is the population frequency of variants in these genes? I have read some of the relevant sections in this document (although were not among those assigned to me), just to have a better idea on what is known, and data seem scarce. Again, this is an important piece of the puzzle and if the data is missing, it should be mentioned.

Publication bias: Just as a suggestion, I wonder if these trials were registered in the public database? If so, there should be a way to find out how many of them were registered but not published.

The concluding statement about publication bias is not very convincing: the fact that several high quality studies show a positive association does not preclude the fact that other good studies with negative results were not published. Unless I'm not getting the point here, this statement is not completely appropriate and should be revised.

As I mentioned in the overall comment, there is no formal assessment of heterogeneity described here, and I wonder why. In general a test for heterogeneity helps defining the variability of the results, and points at subgroups and sensitivity analyses to try to address heterogeneity.

The conclusion should mention if there are long-term effects on these sensitive children, or on children in general who were exposed to these dyes. It seems relevant to talk about chronic, long term neurobehavioral effects. If there is no long term follow-up of these children, it should be mentioned as a scientific gap. I feel that transient effects could be of less relevance than persistent, chronic effects.

Tables: I would include a classic PRISMA graph to show included and excluded papers, together with number of papers excluded for each reason, and details about the reasons for exclusion. We usually group the excluded papers by broad categories, such as no RTC, etc

Table 2.2: I wonder if some of the data can be recovered from these papers, for example the first paper says that a fraction of subjects was challenged; can we recover these cases that underwent a challenge, and use the results for this review?

Table 2.3a: can sex be added as an extra column? It seems important to know if studies were conducted mostly in males, females, or both.

Table 2.4: looks like a sub analysis of US and UK is worth, since the majority of studies were conducted in these two countries. I wonder if a meta-analysis of dose-response can be conducted, given the number of studies reporting on it, from table 2.4

CONCLUSIONS

This is an impressive review of the topic of synthetic food dyes and neurobehavioral effects in children. Although there are no extra articles that I found for this review, I had suggestions on how to handle in a more formal way the various steps of the review, from defining the PRISMA for inclusion and exclusion of the studies, to quality evaluation, to summary estimates and sources of heterogeneity, to publication bias. All these aspects were discussed in detail above. I also suggest that gaps in current literature be highlighted so that future directions can be easily delineated to the reader.

Based on my expertise and experience, I am reviewing the findings, assumptions, or conclusions I agreed I could review to the extent possible - point n 3 of Attachment 2:

Conclusion 3 – Our estimates of exposure indicate widespread exposure to artificial food dyes in children, that children are exposed to a larger amount per body weight than women, and that the highest exposures were from over-the-counter medications in a single day.

Review:
I have some general comments on this section, for which I have some tangential expertise. This is a commendable effort to assess exposure levels in children, and characterize risks by poverty level, race and ethnicity, and education of the mother. Table 6.2 includes a review of available literature, and highlights how limited is the published information on human exposure. The table includes two unpublished thesis and five publications. This is a very small amount of data, with many inherent limitations for each of the studies reported. None of the studies seems to have reported age and sex-adjusted estimates, for example. I suggest that these limitations and gaps in literature are included in the comments.
Another big limitation of this section of the document is that all the assessment are estimates based on the dietary questionnaires, not actual measures, and derived from the known levels of various dyes contained in various food items as well as common drugs. There is no actual measure of blood levels, urinary metabolites, other markers of metabolism and excretion that could highlight levels of exposure and variations in such levels according to the important covariates mentioned above, such as age, sex, and race. This is a real problem, because all the variability observed and presented here is attributed to variations in dietary intake and in food dye content, not in individual ability to absorb, metabolize and excrete the products and their metabolites. The latter process, which is under control of several genetic pathways, could contribute greatly to the observed variability even if the dye intake is the same. I suggest adding this comment to the Summary section (6.9.2)
Another limitation that could be easily fixed is that the results of the new NHANES analysis, specifically commissioned by CALEPA, are reported in a very descriptive way, and is currently comprised of a series of univariate analyses. The available data very likely would allow for a more advanced statistical analysis, for example age, race and sex-adjusted estimates, at a minimum. The interaction between sex and age could also be looked at in detail. These further analyses can still be conducted and added to the document. I suggest doing so, as they would greatly improve a section that is very scarce in relevant information, mostly because there is very little literature available on exposure.

CONCLUSIONS
The section on human exposure to synthetic dyes in food should stress the limited number of studies available in the literature, and the many limitations of what is available. Among the most striking gaps in literature, we highlighted the lack of measures of dye levels in various human compartments (urine, blood), of metabolic gene polymorphisms that could contribute and explain individual variability in response

to exposure, the very descriptive nature of the statistical analysis of the NHANES data generated under CALEPA request. The section in my view should underline the existing gaps in knowledge, since human exposure assessment is one of the key steps in the evaluation process of any possible toxic substance.