**Response to Comments on Walach, H., Weikl, R., Prentice, J., Diemer, A., Traindl, H., Kappes, A., & Hockertz, S. (2021). Experimental assessment of carbon dioxide content in inhaled air with or without face masks in healthy children: A randomized clinical trial. *JAMA Pediatrics*. doi:10.1001/jamapediatrics.2021.2659**

By Harald Walach, Ronald Weikl, Helmut Traindl & Andreas Diemer

Let us begin with thanking all those who have provided us with comments, as these allow us to clarify details and clear up misunderstandings. Let us also clarify: this research letter is a very condensed version, allowing only 600 words, of a more detailed report, which is available at doi:……

Let us also remind everyone of one important scientific principle: Facts are not constituted by single studies, but by multiple replications and discourse. This is the first peer-reviewed study of carbon dioxide content under face masks in children in a short measurement set-up. The measurements, we contend, are valid and were conducted by individuals with high content expertise (Dr. Ing. Helmut Traindl, a measurement engineer & Andreas Diemer, a physicist and physician). If someone doubts our results, the way to go is not to claim they are wrong without proof, but to produce better and different results.

We start from the first comment, which we quote and answer below the relevant passage. The original comments are in italics.

**Comment 1 by Eve Bloomgarden, Elisabeth Marnick, Alison Bernstein & Rebecca J. Heick**

*First, this is written by individuals with known bias against masks and other non-pharmaceutical interventions against COVID, as well as vaccinations. The lead author is a psychologist with no training pertaining to this area of research.*

First, arguments “ad hominem” are the worst kind of arguments. It is true that we were skeptical regarding the usefulness of such masks. But skepticism is a good precondition for studying a phenomenon. We suggest the authors of this comment apply their own skepticism against our results and conduct a better study. The lead author (HW) has ample experience with all kinds of experimental and clinical trials and about 200 peer reviewed papers to his name. A quick perusal of Pubmed shows, that none of the authors of this comment are senior enough, nor do they have a wide range of publications to their name to be able to judge the competence of our team. As is the case in every complex study there is always a range of different expertise necessary. In a clinical trial the statistician does not have to be a medical doctor to analyze and publish a study on an endocrinological treatment, say. In the same sense the lead author here has enough expertise to judge experimental set-ups in general, analyze data and interpret them. He and the rest of the team relied on the expertise of the measurement specialist, Dr. Ing. Traindl, who has ample experience with this kind of measurement. *Second, there are extensive methodological issues. They used a G100 CO2 incubator analyzer, which is made for measuring CO2 levels in incubators. There is no data supporting the use of G100 as a valid and accurate instrument for the type of measurement used in this study. They also did not address the amount of dead space within the mask that could be further making their measurements inconsistent or unreliable. It is also unclear how they could reliably distinguish between inhaled and exhaled air using the described methods.*

It is wrong to say that this kind of analyzer was not a valid instrument. A carbon dioxide analyzer analyzes carbon dioxide, no matter whether it is in the open, in an incubator, in a stable, or in a closed room. It contains access plugs for probes and for tubes which can be fit into incubators. It is exactly this feature which makes it also fit for measurement of carbon dioxide under masks. In order to do that a small tube needs to be fit between mask and face. The diameter needs to be small enough such as to not disturb the air exchange too much. And exactly this is possible with that kind of analyzer. The dead space volume was addressed by us as a potential reason for the back-inspiration of CO2. We could have measured that as well, but it would have required some more drastic measure, such as submersing the head with and without mask under a defined amount of water, which we did not want to do. To distinguish between inhaled and expired air, the aspiration pump of the measuring device, was switched on and off for every inhalation resp. expiration.

 *Third, their results presented are incomplete and difficult to interpret as many previous commenters have noted. Their trial protocol included in supplement 2 outlines that they would also collect blood oxygenation, heart rate and breathing frequency. Yet none of this data was included in this letter. Their results also show almost the same CO2 readings for both surgical masks and respirator FFP2 masks. If their findings were accurate you would expect a difference, given the differing filter capabilities. It is also unclear why their measurement of CO2 in baseline inhaled air is different than in ambient air.*

The results are not incomplete and very easy to interpret. The full results are in the full paper version. We reported those that are important. None of these secondary parameters changed much, and additional analyses with, for instance heart rate or breathing frequency as covariates did not change the results. Blood oxygenation was measured in the first 10 or so children, but as it had no variance whatsoever and did never fall below 99 to 98%% we did not further measure and analyze it. We discuss that in the full version. The reason is likely that we measured short term and the body has enough capacity to buffer lower oxygen for a while before blood oxygenation goes down. This would have probably to be measured invasively, and after a longer period of wearing a mask.

The data are available, including heart rate and breathing frequency, and everyone who has doubts can analyze the data for themselves. It is also not true that the results are the same for both types of masks. As a cursory glance on the table shows, there is a difference, and we said so. The difference is about a third of a standard deviation, which is a small, but not negligeable difference for such a short period of time. In order to see a statistical difference for this point, one would have to measure roughly 80 children and perhaps also longer than only 3 minutes. It was not our goal to find a difference here, but to describe the difference between no mask and mask, if there was any.

 *Fourth, there was also no discussion regarding whether these results are clinically meaningful. They’re using flawed measurements obtained with a device that was not designed for this purpose and stating that this demonstrates harm. There were no actual health metrics reported, like pO2 or pCO2 and no discussion of the actual data needed to confirm clinical significance. These would be more reliable indicators of potential harmful effects on children.*

We did discuss this, both in the letter and in the full paper. There is a large survey of more than 25.000 children of whom 68% are reporting side effects, which we quoted. The side effects reported are consistent with clinical hypercapnia, and we said this. The full link between short term rise of CO2 and clinical symptoms needs further exploration. But we contend: The burden of proof is on those who say that face masks do not confer harm. We remind everyone: it is not those who doubt the safety of any intervention to show that there are harms. The burden of proof is on those who advocate an intervention to show that it does no harm. We challenge the authors of this comment to prove their point. *Overall, this paper misinterprets inappropriately collected data leading to incorrect, distorted, and dangerous conclusions.*

Overall, the authors of this comment have neither shown that nor produced better data.

 *Lastly, we have real life evidence from millions of children who have been successfully wearing masks every day for months. We have extensive evidence that masking is an essential protective strategy to slow and prevent the transmission of SARS-CoV-2. If this study were accurate and reflective of the real world we would see adverse health events reported worldwide. We do not. We also have data from other peer reviewed papers that has shown no harmful effects. This is a study searching for a mechanism for a non-existent problem. The subject of this letter, regardless of conclusion, should have prompted intense scientific scrutiny prior to publication. The bell cannot be unrung, as this research letter is already being used as “scientific proof” that masks are harming our children, but a retraction should be strongly considered.*

Well, in that case US-children seem to be different from German children. We have many thousand cases where children report harm. We would ask the authors of this comment: Where is that evidence? Can you point to the studies? What evidence are you referring to? We do see adverse events reported, manifold. And we would like to know about those papers that have shown no harmful effects. We have quoted a systematic review of 109 studies, 44 of which entered into a quantitative synthesis 1. This review reports many side effects found and comes to the conclusion that the harms do not outweigh the benefits, except in specific situations but not in general terms. Apart from that our data are close to another measurement study conducted independent of us in Italy which arrives at around the same values, with different setups and measurement devices, also referred to by one of the other commentators 2.

We are glad the bell has rung. The next step would now be to prove us wrong by a study which is at least as good as this one and adds perhaps some features, such as long term measurement and blood oxygenation after a longer period of mask wearing to prove what is currently an unproven statement namely that prolonged mask wearing is harmless. It is likely not, and we have given a reason by valid experimental measurement.

**Comments 2 by Michael Braun and 3 by Sander Orent**

We are grateful for the comments here, which elucidate some facts that we were not able to discuss or did not discuss in our paper.

**Comment 4 by John Murphy**

*The lack of methodological detail makes it unclear as to why the researchers concluded that they were measuring the concentration of inhaled CO2 as opposed to the concentration of exhaled CO2 or a mix of both. But the real question the author's don't clearly address is "Why does it matter?". The paper begins by stating "The question whether nose and mouth covering increases carbon dioxide in inhaled air is crucial.", and infer that it the reason is that the jurisdiction's regulatory limit for the concentration of CO2 in ambient indoor air is 2000 ppm, and that mask use results in overexposure in relation to that limit, which presumably is considered hazardous. Such a line of reasoning is mistaken. Indoor ambient CO2 limits are not set because CO2 itself is hazardous above those values, but because elevated CO2 correlates with elevated constellations of a range of indoor air pollutants that cause increasing occupant discomfort as levels rise. The CO2 indoor air standard is ultimately a ventilation effectiveness guideline, not a health-based exposure limit. It should also be recognized that any re-breathed CO2 is endogenous, and such air is qualitatively different from the ambient indoor air for which CO2 is solely a proxy indicator of composition, and to which the CO2 indoor air guidelines only apply.*

As stated by Comment 3 the maximum value of CO2 indoors is not only set because of pollutants, but because too much CO2 is harmful in itself, if one is exposed to it for too long. It is well known that hypercapnia leads to various symptoms1, from headache to fatigue, all reported by the majority of children who answered the survey we quoted 3. It is obvious that such surveys are not representative. But we contend: Our data do matter, because they show that CO2 content rises quickly and sharply to undesirably high values and very likely remains there. Hence hypercapnia is likely to occur after a while and the symptoms reported by many are in agreement with that. Anybody doubting this should conduct a representative survey of a representative sample of children who wear masks, say for a whole school day to find out how many of them have symptoms and compare those symptoms with those known to be a consequence of hypercapnia, or else produce better measurements.

**Comment 5 by Preston Witherspoon**

*A review of the supplemental content appears to suggest that one of the measurements taken by researchers was the oxygen level of the blood. The concern with mask wearing among children is, realistically, two-fold: 1) that CO2 will be increased in the blood, and 2) that oxygen will be reduced in the blood.

While CO2 concentration is best assessed via bloodwork, which would be difficult to obtain from children in a timely manner after a short testing period, O2 concentration seems simpler data to collect. Current publication information does not list the O2 levels as measured by the researchers. Adding that information to the study may provide further context for the effects of masks on blood chemistry beyond an observation that inhaled air increases in CO2 content underneath a child's mask.*

As stated in our response to comment 1: We did measure this, but did not see any variance and stopped measurements after the first 10 children. Blood oxygenation did not change. This is likely due to the fact that the body has enough buffering capacity for a while. Blood oxygenation would probably have to be measured after a longer period of mask wearing. We have heard from parents who have done that themselves after a school day and found low values, but we are not aware of a study which has done that long term.

**Comment 6 by Harikrishnan Pandurangan**

This comment makes recommendations like wearing different types of masks or suggesting breaks. We think the comment is worthwhile considering, and these measures should be accompanied by appropriate evaluations.

**Comment 7 by Alejandro Keller**

*The authors use an CO2 incubator analyzer (0-20%) that has an accuracy of 1% of the range (i.e. 2000ppm) and a time response T^90<=20 seconds for CO2. This device is unsuitable for measuring the transient concentrations during the respiration processes. A normal respiration cycle has a duration of 3 to 4 seconds. Thus, it is impossible to separate the concentration of CO2 in inhaled and exhaled air using this device. This also explains why the authors measure average concentrations of around 2700ppm-CO2, way above the ambient value of 740ppm, even when no mask is present.*

This is a misreading: the measuring device has an accuracy of 1% of the **calibrated** range, which in the end is 0,05% of the measured values and hence it is still very accurate. The time response of the system is, as stated in our answer to comment 1, 20 seconds. As can be seen from the long version of our paper, we adjusted the measurements in between the phases of measuring joint air, inhalation and exhalation and allowed 30 seconds for the new type of measurement to be started. As we did not measure continuously, but only the phases of interest (i.e. either inhalation, or exhalation, or both) the argument made here does not apply. The measurement of CO2 without mask of 2700 ppm is due to the fact that the measurement tube was fixed to the space between nose and mouth. Hence the child sits within his or her own breathing bubble and the inhaled CO2 content is higher than that in ambient air. One could consider correcting for this fact by subtracting that difference of about 2000 ppm. But since this is irrelevant for our measurements under the mask we do not think it necessary. *The authors refer to the dead volume behind the mask as the main problem. The relevant data would be the comparison between this dead volume and the lung capacity and/or the volume of one respiration cycle. The lungs never collapse completely during respiration. Together with the rest of the respiratory airways, the respiratory system has also a dead volume that is much larger than the dead volume between the mask and the face. The comparison of these volume is of extreme importance for the discussion and may change the author's conclusions.*

This is an interesting aspect. But we think that the physiological dead volume of the respiratory system is natural, the additional one under the face mask is not. Besides, a major point that explains our results is that the mask prevents the proper exchange of gases with the ambient air between expiration and inhalation. Furthermore, as we said in our discussion, the whole volume of the airway passages in a child is smaller than that of an adult in comparison to the dead space volume of the mask.

1. Kisielinski K, Giboni P, Prescher A, et al. Is a Mask That Covers the Mouth and Nose Free from Undesirable Side Effects in Everyday Use and Free of Potential Hazards? *International Journal of Environmental Research and Public Health.* 2021;18(8):4344.

2. Oberrauch B, Adami M, Gutweniger U, et al. *Ist der Gebrauch von Mund-Nasen-Bedeckungen in der Gesamtbevölkerung eher schädlich als nützlich unter Berücksichtigung der CO2 Konzentration? Luftqualität während des Tragens von Mund-Nasen-Bedeckungen mit Mini-Review [Does the use of a mask covering mouth and nose confer benefit or harm on the population: Air quality while wearing a nose-mouth coverage and mini-review].* Bolzano2020.

3. Schwarz S, Jenetzky E, Krafft H, Maurer T, Martin D. Corona children studies "Co-Ki": First results of a Germany-wide registry on mouth and nose covering (mask) in children. *Research Square Preprint: Non-peer-reviewed preliminary publication.* 2021.