**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

We continue our responses to comments with the next comments, commenting point by point.

The original texts of the comments are in italics, our response in normal type.

**Lack of Outcome Comparisons**

**Matthew Loxton, MA Knowledge Management**

*The method states that "carbon dioxide content in inhaled air" was measured, but this is not an accurate statement since the amount of gas in the mask region is a small fraction of tidal volume, and the method of measurement is sensitive to the respiration rate.*

The carbon dioxide content in inhaled air was measured by a measurement tube fixed to the region between nose and upper lip. We distinguished three types of air: inhaled air, exhaled air and mixed air (joint inhaled and exhaled air), which we report. The distinction was made by a physician observing the breathing pattern of the child and activating the measurement only during the appropriate phase. That way inhalation and the air inhaled could be separated from exhalation and the mixed air. The air inhaled contains of course fresh air that enters the mask through the fabric, and this amount is dependent on the volume of the breath and the breathing frequency. And it contains air that is rebreathed from the dead space volume of the mask. The mixing of these different types of air is what is of interest. We did indeed measure inhaled air and its CO2 content. The fact that it is so high seems to indicate that quite a quantity of the previously exhaled air remains in the mask and is rebreathed by mixing with fresh air. We documented and analyzed breathing frequency as well, but did not see much difference here, likely because the experiment was short term and the children were seated.

*Secondly, to know if this measurement is of any clinical significance, both blood oxygenation and blood CO2 levels would need to be measured and compared to masked respiration and unmasked respiration.*

We only partially agree. We did measure blood oxygenation, which did not change over the short period. Blood oxygenation would likely take much longer to show a difference, as the body has various buffering mechanisms. We were not interested in the distal end of breathing physiology but in the proximal part. There are clear regulations for a workplace, for indoor ambient air, for the wearing of FFP2 masks under working conditions. They are all put in place because of ample previous knowledge that a CO2 level beyond a certain point is unhealthy. This has been set at 2.000 ppm by the task force of the German Environmental Office that screened all the literature1. We think it is enough to use this as a reference point and this we did.

*Without these tests, any measurement of CO2 levels in the mask region is merely a curiosity, and not indicative of a risk factor.*

We disagree. If internationally accepted and well validated norms of CO2 levels of inhaled air are violated, this poses a health hazard in itself. Why else would we have such norms? Apart from that, the frequency of symptoms reported and their phenomenology support a causal link between high CO2 content in inhaled air and symptomatology. We agree that a clear proof of a causal link would require a more diligent study of breathing physiology, likely also long term. But this was not our goal.

**Morgan Robinson, BSc, MSc, PhD Candidate | University of Waterloo**

*One wonders how CO2 would be retained behind the mask if for instance other gases, such as O2, are not impeded by the mask as would be suggested by no change in O2 saturation.*

The fact that O2-saturation does not change does not mean that O2 is freely flowing through the fabric. It means that the body has ample capacity to buffer a lack of oxygen and a surplus of CO2 in breathed air. The argument goes the other way round: if CO2 content in inhaled air rises so quickly, this is a sign that the exchange of gas is not working quickly enough. To our knowledge there is one study (a MD thesis from the Technical University of Munich) which measured blood-CO2 concentration non-invasively in 15 surgeons and operating theater personnel wearing surgical masks during operations. After 30 minutes a significant rise of CO2 in blood was seen.2

*It could be that the CO2 is dissolved in the exhaled water vapour. The solubility of CO2 in water is about 33x greater than that of oxygen. This water vapour does tend to build up in the mask.*

This would be an interesting study which we did not do. We only measured CO2 content in the inhaled air and we think it unlikely that CO2 that is dissolved in water is a source of CO2 in inhaled air.

**Amy Houtrow, MD, PhD, MPH | University of Pittsburgh**

*As a member of the editorial boards for three journals I understand the challenges in evaluating the scientific rigor and validity of submitted manuscripts. In my experience it has become increasingly difficult to secure high quality peer reviews. Nonetheless, the fact that this manuscript made it through editorial screening, peer-review, and approval is concerning. The need for the study was not well justified, the premise biased, the methodology flawed and thus the resultant data unhelpful and the conclusions inappropriate. I fear that those with political incentive to demonize mask-wearing will use this study as ‘scientific evidence’ to undermine public health strategies.*

The need for the study was justified: We have no knowledge about the CO2 content of inhaled air under face masks. We have data on complaints, which we cited. We do not know what biased premises are: Premises are always biased to a degree, because you have to start from some assumption, and this is a bias. You can have the assumption that wearing masks is safe and has no consequence on the health of children. This is a clear bias without evidence. You can have the assumption that masks are a problem. This is also a clear bias with no data. We wanted to find out and did a study to measure CO2 in inhaled air. Where is the bias? The methodology was not flawed: We took great care to measure exactly what we wanted to measure: Carbon dioxide in inhaled air under face masks. Whether data are helpful or not depends on what one expects from data. If they do not meet our expectations, we call them unhelpful. But data are data, helpful or not. And we have to make decisions. Our conclusions are based on our data, and if other researchers find other data the conclusions will change accordingly. If anything, everybody challenging our data and conclusion should produce better and possibly different data.

**Robert Rosen, D.O. | Private Practice, Dermatology**

*The subtitle "Randomized Clinical Trial" implies comparison between clinical interventions. It is unclear what participants were randomized to, and no significant difference was found between 2 unidentified types of masks. The subtitle is bound to mislead casual (e.g. Twitter, Facebook) readers. Further the discussion references hypercapnia, defined as elevated CO2 concentration in blood, but they include no blood measurements of C02 nor any evidence their results could be extrapolated to that.*

This is correct. The subtitle came from the journal’s proofreading editor and was compulsory, because the journal does not have a category of “experimental measurement study in healthy children”, as it was originally termed. We randomized participants to the sequence of masks, as explained. We introduced hypercapnia as a potential explanation which would have to be studied further by direct measurements, we agree. This was done in another study 2 , which did find elevated CO2 concentration in blood after 30 minutes of mask wearing in surgeons.

***Flawed Methodology***

***Julius Cheng, MD | University of Rochester***

*Questions have been raised about the use of an environmental CO2 measuring device (PCE-CMM 10), and not a medically calibrated end-tidal CO2 device.*

This is a misreading: we used this device for monitoring ambient air, and not the CO2 content of the inhaled air. And for ambient air measurement this device (PCE-CMM 10) is exactly the right one.

*This is especially relevant since the headline-grabbing mean CO2 of "between 13 120 (384) and 13 910 (374) ppm" is higher than the stated measuring range of 400-5000 ppm by the device's manufacturer.*

We did not use this device for measuring the inhaled CO2 content. The one we used for measuring the CO2 content (G100 CO2 analyzer) has a measurement range from 0 to 20 vol% for CO2 (i.e. 0 to 200.000 ppm) with an accuracy of 0,1% and was thus well able to establish the data we found.

*The authors did not measure the actual dead space behind any of the masks.*

No, but that was neither necessary nor intended. We intended to measure the CO2 content of inhaled air and this is what we did.

*And, the measured respiratory rate of the participants did not change. If the participants were getting clinically hypercapnic, you would think there would be some change in respiratory rate.*

This is true, but irrelevant for our argument. We would not assume that clinical hypercapnia was induced by a short-term measurement (and if that would have been expected, we doubt the ethics committee would have allowed the study). We used hypercapnia as a potential causal link between elevated CO2 levels in inhaled air and clinical symptoms at the distal end of the purported causal chain in the discussion.

**Flawed Design, Unsafe Conclusions**

**Trisha Greenhalgh, FMedSci | University of Oxford**

*Walach et. al’s Research Letter [1] contains major flaws which cast doubt on the paper’s scientific integrity.

Most importantly, the main outcome cannot test the authors’ hypothesis. The authors used a surrogate outcome - carbon dioxide (CO2) levels inside the mask - as a proxy for CO2 levels inhaled by a child. This is actually a measure of ‘dead space’ CO2. Most of the air entering the child’s lungs will be inhaled from the ambient air outside the mask (in the same way a swimmer using a snorkel breathes in a small amount of ‘dead space’ air from within the snorkel, but most air comes from outside it).*

This is a misunderstanding. We did not use CO2 levels inside the mask as a proxy for CO2 level inhaled. We did not measure the CO2 level inside the mask, but the CO2 level of inhaled air, which is a mix of incoming fresh air and air trapped in the dead space volume of the mask. The inhaled air will have a large part of fresh air, we agree, but it will contain some mix of rebreathed CO2. Our aim was to find out how high the CO2 content of the inhaled air is. Now we know: it is too high.

 *Second, CO2 levels were not correctly measured. The Geotech G100 is designed to be used with incubators. It has a response time of 20 seconds (T90 = 20 seconds, i.e. the instrument takes 20 seconds to reach 90% of a stable reading) [2]. In comparison, a capnograph [3], which would have been an appropriate instrument to use in this study, stabilizes in milliseconds. The study claims to have measured CO2 in inhaled air, but the authors provide no basis for distinguishing between inhaled and exhaled air, using the G100, at 1.5 cm from nose and face.*

What is an appropriate instrument is dependent of what one wants to know. If one wants to know the pressure of tidal end volume of CO2 in the lungs then capnography is certainly the standard measure. But we were not interested in the physiology of breathing but simply in the amount of CO2 contained in inhaled air. And for this, the equipment and the measurement set-up were right. It is correct: the response time is about 20 seconds, but only for the measurement of a new and different set of gases, depending on the length of the tube. The actual response time of the equipment is 1 second, and with our tube it was 15 seconds. We adjusted for that by a break of 30 seconds between different measurements. We apologize if we did not clearly describe the set-up. This was due to the word limitation. We also provide a longer version, in which it is clearly described: We made sure we measured inhaled (or exhaled) air, by triggering the measurements only when a child inhaled (or exhaled). This was done by a medical doctor who carefully observed the movements of the thorax and abdomen and started the measurement, when inhalation started.

*The authors report no meaningful physiological variables, such as blood oxygen saturation or arterial partial pressure of CO2 (pCO2). Measurement of transient increases in CO2 around the breathing zone (even if accurate) are not indicative of harm. Although Supplementary material reports that oximeters were used during the study and pulse rates appear to have been measured by oximeters, oxygen saturation level data are not reported*.

We measured oxygen saturation level, but this did not change and was always between 98% and 99%. This is reported in the long version of the paper. We were not interested in documenting the full causal chain from high CO2 in the inhaled air to the distant physiological effects. This would have required a considerably longer test set-up, as the body has ample buffering capacities for a while and as this would have required different instruments to measure pCO2. As it was unclear, whether CO2 is elevated in inhaled air, we focused our attention an that. We disagree: if the level of inhaled CO2 is beyond accepted safety levels, this is an indication for potential harm, because the levels are derived from a broad knowledge of association between harms and CO2 levels. This is comparable to alcohol level in blood and safety of driving: There is abundant evidence that beyond certain levels safety of driving might be compromised, even though in an individual case the causal link might be difficult to establish. Nevertheless, we generally accept that certain blood alcohol levels are indicative of potential harm.

*The authors’ causal claims regarding physiological impact are wholly unsupported by the data.*

We did not make claims about physiological impact. We used common sense and common knowledge that certain symptoms, as supported by the reviews we cited, are frequently seen as consequence of mask wearing. We used the known complaint of children from a large survey. And we reasoned that elevated CO2 in inhaled air could explain this. This is not a causal claim, it is a hypothesis, which needs further testing.

*The science of mask-wearing should be based on good data. Some of the authors have publicly associated themselves with, and taken funding from, non-science-based campaign groups [4]. We respectfully suggest that not only is this study scientifically flawed, but its publication puts the reputation of JAMA Pediatrics and its peer-review process at risk.*

We contend that these data are good. They are certainly not final, as no single study is able to speak a final word. This is always achieved by multiple studies, careful scrutiny of all the evidence and discourse. The purpose of this study was to provide the scientific community and the public with an important and as yet unknown piece of the puzzle: the content of carbon dioxide in inhaled air under masks. Anybody doubting our results can replicate or refute it. Our study is a replication of already existing data 3 and replicates the findings of this earlier study quite closely. It is wrong to claim that Stefan Hockertz has taken funding from non-science-based campaign groups. We respectfully disagree: Neither was this study flawed, nor was the peer-review process compromised.

***Misrepresentation of Endogenously Produced CO2***

***Prof W Meier-Augenstein, Dr. rer. nat. | Robert Gordon University, Aberdeen***

*The values of 3.846 and 3.847% CO2 measured behind masks might actually represent CO2 endogenously produced as a result of the human body’s metabolism. In adults, body weight-adjusted endogenous CO2 production at rest is about 1.4 mmol/L which equates to a CO2 content of 5.03% in exhaled breath. The aforementioned values taken from the Table of this article are entirely consistent with what one would expect in breath exhaled by children into an enclosed yet not leak-tight space. Chances are that measurements of exhaled breath in a space formed by a hand cupped over the mouth and nose will have yielded values similar to those between mask and mouth and nose. Furthermore, to compare values of endogenous CO2 in exhaled breath as a result of the body’s energy expenditure at rest with threshold values for CO2 in enclosed spaces appears ill thought-out. It’s like comparing threshold concentrations for (e.g.) NOx in ambient air with NOx values measured directly at the exhaust of a car.*

This is an interesting idea. However, we do not think it is correct. For one, we see clearly different CO2 values in different conditions (inhalation, exhalation, mix). Furthermore, we did not compare the levels of CO2 in exhaled breath to threshold values, but levels of CO2 in inhaled air. These we compared to exactly those thresholds that were developed for levels of CO2 in inhaled air. Third, it might be true that the CO2 level we measured might reflect the body’s metabolism. However, the difference is: if a child exhales into a cupped container and then the CO2 content is measured, the next breath the child will take will not be from that container, but from fresh air. Under the masks, and this is what our measurement establishes, a considerable amount of this exhaled CO2 will be re-inhaled, as the fresh air mixes with this. So, the images Prof. Meier-Augenstein uses are interesting, but do not apply in our case.

1. Umweltbundesamt. Gesundheitliche Bewertung von Kohlendioxid in der Innenraumluft [Health assessment of carbon dioxide in air within closed rooms]. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz.* 2008;51(11):1358-1369.

2. Butz U. *Rückatmung von Kohlendioxid bei Verwendung von Operationsmasken als hygienischer Mundschutz an medizinischem Fachpersonal [Respiration of carbon dioxide when using surgical mask as hygienic mouth covering in medical personnel]*. München: Klinikum Rechts der Isar, Technische Universität München; 2005.

3. 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.