Carbon dioxide rises beyond acceptable safety levels in children under nose and mouth covering: Results of an experimental measurement study in healthy children

Harald Walach [a,*], Helmut Traindl [b], Juliane Prentice [c], Ronald Weikl [d], Andreas Diemer [e], Anna Kappes [f], Stefan Hockertz [g]

[a] *Change Health Science Institute, Berlin, Germany*
[b] *Traindl-consult, Vienna, Austria*
[c] *Psychotherapeutic Practice, Müllheim, Germany*
[d] *Obstetric, Gynecological and General Practice, Passau, Germany*
[e] *General Practice, Gernsbach, Germany*
[f] *Anna Kappes, Psychotherapeutic Practice for Children and Youths, Müllheim, Germany*
[g] *Tpi Consult GmbH, Bollschweil, Germany*

ARTICLE INFO

ABSTRACT

Nose and mouth covering (NMC) has been made compulsory for children in many countries during the Covid-19 pandemic. We wanted to determine the average CO_2 levels in inhaled air with NMC in children between age 6 and 17. We used short term measurements under surgical masks and FFP2 masks according to European norm EN 149, compared to baseline in an experimental, intra-individually controlled study over 25 min. CO_2 content was measured every 15 s using an automated dual-wavelength infrared CO_2 measurement device (G100, Geotech, Leamington Spa, UK) over 25 min in a short-term experimental setting, with children seated. After baseline measurement children were provided with two types of commonly worn NMC: surgical masks and FFP2–masks in randomized sequence for 3 min each. We kept ambient CO_2-levels below 1000 parts per million (ppm) through frequent ventilation. We measured breathing frequency and pulse as potential physiological moderator variables. Forty-five children, 25 boys, 20 girls, with a mean age of 10.7 years (standard deviation 2.6) were measured. We measured 13,100 ppm (SD 380) under surgical mask and 13,900 ppm (SD 370) under FFP2 mask in inhaled air. A linear model with age as a covariate showed a highly significant effect of the condition ($p < 1*10^{-9}$). We measured 2,700 ppm (SD 100) CO_2 at pre-baseline and 2,800 ppm (SD 100) at post-baseline, a non-significant small difference. Appropriate contrasts revealed that the change was due to the masks only and the difference between the two types of masks was small and not significant. Wearing of NMC (surgical masks or FFP2- -masks) raises CO_2 content in inhaled air quickly to a very high level in healthy children in a seated resting position that might be hazardous to children's health.

1. Introduction

Since the WHO has alarmed the world to the SARS-CoV2 pandemic in March 2020 most governments tried to stop the spread of the novel corona virus. Many governments have made the wearing of nose and mouth covering (NMC), or face masks, compulsory for children in school. The evidence-base for such a procedure to prevent infection is mixed. Two recent systematic reviews studying different types of NMC, such as surgical and FFP2/N95 respirators reach the conclusion that wearing face masks does not prevent infections by influenza virus, which is very similar to SARS-CoV2 (Jefferson et al., 2020; Xiao et al., 2020). Some data support the wearing of NMC in general contexts, but practically none for children (Kappstein, 2020). A review of non-randomized studies concludes that a significant benefit cannot be excluded (Chu et al., 2020). However, the first pragmatic randomized study comparing the suggestion to wear NMC in public with no recommendation found that the effect is

small and not significant (Bundgaard et al., 2020): Of 6.000 participants 42 or 1,8% were infected in the experimental group, and 53 or 2,1% in the control group. When comparing those that actually did wear the masks the effect was even smaller. A recent comprehensive review concluded that the effects of facemasks for preventing the spread of SARS-CoV2 is robustly documented by for a work context (Herby et al., 2022). Positive effects of NMC for preventing infections in community settings are likely small and probably only useful in high incidence environments and only if included in a comprehensive strategy (Kiselinski et al., 2021; Matuschek et al., 2020; World Health Organization (WHO), 2020). Robust results that the wearing of NMC, especially by children, would help prevent the spread of SARS-CoV2 did not exist when we started to embark on the planning of this study during early 2021.

Against this background the question whether NMC increases carbon dioxide in breathed air is becoming important. The first large scale German survey in parents and children, the Co-Ki-study of the University Witten/Herdecke using data of 25,930 children has shown that children report side effects frequently (Schwarz et al., 2021): 68% of parents report that their children have problems. Most frequently they report irritation, tension and stress (60% of parents report this), headaches (53%), difficulties concentrating (50%), fatigue and sleepiness (30%). It is possible that a high content of carbon dioxide in inhaled air might be causal for those symptoms and complaints. Wearing of NMC is associated with headache in health care workers (Ong et al., 2020), which is also one of the side effects of mask wearing according to WHO guidance (World Health Organization (WHO), 2020). Short term exposure to carbon dioxide contents of 1.000 ppm is associated with decline in concentration and cognitive problems (Azuma et al., 2018).

The normal content of carbon dioxide in breathed air in the open is about 0.04 vol % (i.e. 400 parts per million/ppm). 0.2 vol% or 2,000 ppm are acceptable for closed rooms according to the German federal environmental office (Umweltbundesamt, 2008). This is at the same time the cut off for children and pregnant women, which is considered safe (Umweltbundesamt, 2008).

Maximum concentration at the work place for healthy adults during 8 h of work and 40 h per week as a time-weighted average is considered 0.5 vol% or 5,000 ppm. This limit is accepted in many countries, for instance in Germany (Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung, 2021) or in the United States (Centers for Disease Control and Prevention (CDC), 2019).

To the best of our knowledge there are no solid peer-reviewed data on carbon dioxide concentration in inhaled air under NMC, especially for children. There are two studies that measured end tidal CO_2 pressure (PetCO$_2$) in children wearing face masks using capnographs. One study measured 47 healthy children for 60 min, with and without exertion (Lubrano et al., 2021). While there were no significant changes, PetCO$_2$ fell by 0.5 mm Hg in the younger children and by 1 mm Hg in the older children after mild exertion. The other study measured 106 children over a period of 45 min using two different masks and a scheme of mild exertion. They found a rise in PetCO$_2$ by 3.2 mm Hg, which is a clinically relevant standardized mean difference (d) of one standard deviation in a resting condition, and a rise by 3.8 mm Hg under slight exertion, equivalent to an effect size of d = 1.3 (Goh et al., 2019). While the outcome parameters were clinical safety limits which were not violated and physiological distress signals which were not seen, this study shows that physiological parameters change. But none of the studies measured the actual carbon dioxide content in inhaled air under a face mask. Ing. Dr. Traindl, coauthor of this study, has made some pilot measurements in 3 persons and found 3–5% CO_2 in the accumulated air in the dead space volume under NMC (30,000–50,000 ppm). One of these volunteers was a 13-year old child, and here CO_2-concentrations were steadily measured at 3.4–5.0 vol% (34,000–50,000 ppm) (Traindl, 2020). Measuring the dead space volume of the face mask allowed to estimate the CO_2-concentration in inhaled air. This yielded an estimate of 0.8–1.3 vol % (i.e. 8,000 to 13,000 ppm) of CO_2 in inhaled air. A team from South-Tyrol/Italy conducted measurements in November 2020 in

24 volunteers using different types of NMC and clarified discrepancies to a study that had been conducted by the official government of the autonomic region in Bolzano (Oberrauch et al., 2020). The results reported by Oberrauch are considerably higher than those reported by the government. This is obviously due to the fact, that the governmental working group of the region of Bolzano had subtracted the environmentally measured carbon dioxide values from measures of CO_2 under the masks, which led to an artificially lowered result. The data of the South Tyrolian study (Oberrauch et al., 2020) regarding the influence of different types of NMC on CO_2-content of inhaled air range from 3,143 ppm for baseline without mask to 7,292 ppm (0.7 vol%) with surgical masks and 15,000 ppm (1.5 vol%) with FFP2 masks. These were results with adults and a few children.

This is the reason why we wanted to measure in a well-controlled, experimental study in volunteer children carbon dioxide content in inhaled air with and without different types of NMC to find out whether raised values are found under different conditions and how CO_2 content changes in inhaled air under NMC.

2. Method

2.1. Participants

Participants were children at school age, whose parents have shown interest in the study and were willing to give consent for their children to participate. Children also gave their own consent. The children were healthy, free from infections or neurological diseases, had no psychological disorders that would produce problems while wearing a face mask and had no medically indicated exception from the compulsory NMC mandate for school children that was effective in Germany at the time of measurement.

Participation was strictly on a volunteer basis and no remuneration was given. An informed consent and information leaflet for children was presented and informed consent of children and their parents was sought. The study was approved by the Ethics Committee of the University Witten/Herdecke (Registration Number 22/2021).

2.2. Design and measurements

2.2.1. Design and outcome measure

The design of the study was an intra-individually controlled experiment, i.e. a study where each person acts as his or her own control and is measured under each condition in randomized sequence. It started with a baseline measurement before experimental measures without NMC. After the baseline, children wore a surgical mask and a FFP2 mask in randomized, balanced order. Finally, it was concluded with a post-baseline measurement after the experimental measurements with NMC.

The main outcome was the carbon dioxide content of the inhaled air, both under normal conditions without mask (baseline, post-baseline), and under NMC condition. We also measured the CO_2-concentration in mixed inhaled/exhaled air and exhaled air.

2.2.2. Method of measurement

We measured in a short-term experimental protocol the CO_2-concentration in inhaled air in the facial area without NMC and under NMC. Our goals were to.

- measure the CO_2-concentration under different NMC
- see whether CO_2-concentration in inhaled air would be increased by the accumulation of CO_2 in the dead space volume of the face mask, and thus to
- measure the CO_2-concentration both in inhaled and exhaled air
- find out whether the CO_2-concentration under NMC would be different from baseline and if so, if the measured CO_2-concentration would violate accepted safety norms.

The method of measurement followed the prescriptions of the European Norm EN 149 for the measurement of respiratory protective devices (Deutsches Institut für Normierung, 2009). We used a tube that conducted the air from a probe to the analyzer with a delay of approximately 20 s. This time delay was taken into account, when defining different phases for the analysis of the measurements. The measurement of the specific air of interest, for instance inhaled air only, was started manually by a physician who observed the breathing pattern of the child and triggered the pump only, when a breathing cycle started, for instance inhalation, and stopped the pump, when it finished. The measurement tube was fixed to the upper lip of the child between nostril and mouth using a flexible band that was adapted to the head size of the child about 1,5 cm distant from the nostril, and remained in place throughout the measurements. The measurements lasted approximately 25 min for each child. Apart from time for preparation, 3 min measurement were taken for baseline carbon dioxide in inhaled air without face mask. Nine minutes measurement for each type of mask were allowed, 3 min for measuring carbon dioxide content under the face mask in joint inhaled and exhaled air, 3 min for measuring carbon dioxide during inhalation and 3 min during exhalation.

For the acquisition of the baseline, carbon dioxide during inhalation without mask was measured. The measurement of the respective breathing phases was initiated by a medical doctor (RW) who observed the breathing patterns of the child carefully, and triggered the aspiration mechanism, a pump that is integrated in the measurement device, once the target phase (inhalation, exhalation) began and ended the measurement, when the target phase was over. This assured that only the particular type of air that was intended for measurement, for instance inhaled air only for 3 min, was collected in the measurement tube and forwarded to the measurement sensor.

During the first 3 min under the face mask the mixture of inhaled and exhaled air that collects under the mask (called "joint air") was measured. Then, after a 30 s waiting period to allow for the adaptation of the system to the new measurement, CO_2 was measured exclusively during inhalation for another 3 min. And, after a second measurement break of 30 s, CO_2 was measured exclusively during exhalation. At the final minute of each NMC-measurement block, pulse and breathing frequency were measured, as well as blood oxygenation. The face mask was changed, which took around another 30 s and the same sequence as before was carried out (see e-Fig. 1 in the Supplement for a sample measurement protocol).

While the sequence of masks was counterbalanced and randomized, the sequence of the measurements for one condition was always CO_2 content in joint air first, then inhaled and finally exhaled air.

Each child was provided with a fresh set of masks. Masks by different producers were used randomly to cover an adequate and practical range of masks used in the community and to avoid any potential producer bias (see e-Table 1 in the Supplement).

2.2.3. Instrumentation

The inhaled air was measured with a G100 CO_2 incubator analyzer, (Geotech, Leamington Spa, UK). This device measures CO_2 content of the air via dual wavelength infrared measurement every second. The specifications of the instrument are given in Table 1 and can be found in the data sheet and the operating manuals (http://www.ybux.eu/wp-content/uploads/2018/09/Geotech-G100-Datasheet.pdf; https://www.apc.co.nz/site/associatedprocess/G100_G110_G150_Manual.pdf; http://www.tridinamika.com/wp-content/uploads/2016/12/ADM-operating-manual.pdf; all accessed on 22nd Sept 2021).

The CO_2-content of ambient air was measured with a second, independent device.

(PCE-CMM 10 by PCE). CO_2 content was always kept well under 1,000 ppm or 0.1 vol %. The specifications of this device can be found in the data sheet (https://www.pce-instruments.com/deutsch/messtechnik/messgeraete-fuer-alle-parameter/arbeitsschutzmessgeraet-pce-instruments-arbeitsschutzmessgeraet-pce-cmm-10-det_5890067.htm;

Table 1
Requirements for the measurement of carbon dioxide content in breathed air and comparison with the specifications of the G100 analyzer used.

	Requirements	Instrument Specification G100
Measurement range (full scale)	0 - 5 vol%	0 - 20 vol% (0–200,000 ppm)
Accuracy	±0.1 vol%	±1% of measurement range after calibration; at calibration 5.0 vol% the device has a accuracy of approximately 0.1 vol%. In the calibration certificate with a 2.5 vol%-certified gas an accuracy of 0.064 vol% is confirmed. Display-accuracy: 0.1 vol%
Response time	1–2 s	The response time of the CO_2 sensor is approximately 1 s. The response time of the whole system – from the tube-opening to the sensor – is dependent on the length of the tube and was less than 20 s in our case

Conversion factor: 1.0 vol% = 10,000 ppm.

accessed 22nd Sept 2021). They are presented in Table 2.

It should be noted that our measurement set-up is similar to that used by the technical gauging of norm values by the German Office of Standards for technical norms of FFP masks (DIN EN 149) (Deutsches Institut für Normierung, 2009). Results of such measurements have led to the current work-place regulations that allow the wearing of FFP2 masks only for 75 min, after which a break of 30 min is required, exactly because the CO_2 content collects in the mask and the exchange of air is not good enough due to the resistance of the material.

While the measurement apparatus for measuring ambient air is specifically designed for this purpose with a measurement range between 0 and 5,000 ppm, the apparatus used for kmeasuring CO_2 under the mask is designed for a higher measurement range (0–200,000 ppm).

The measurement equipment we used is medically certified to measure gases in medically relevant contexts, such as incubators. It has a sensitivity range and a precision that is sufficient for our purpose of CO_2 measurement. The measurement range of this apparatus is between 0 and 20 vol%. As the system has a response delay of 1 s, which rises to 20 s if a measurement hose is attached, we took care of this. We only measured one type of gases at a time, for instance inhaled air. By manually controlling the type of air that was pumped to the measurement sensor during the respective phases, we could make sure that only the type of air that was intended for measurement was directed to the measurement sensor. By disregarding data of a 30 s duration between those phases we allowed for the system to adapt and to make sure that only the type of air intended for a particular measurement phase was considered. We averaged data across phases and types of air to control for individual and time variance.

2.3. Protocol and deviations

The measurement protocol was published in advance and is available at the Open Science Foundation platform at https://osf.io/yh97a/?view_only=df003592db5c4bd1ab183dad8a71834f.

There were the following deviations from the original protocol which were due to simplifications and time restraints: The experimental

Table 2
Requirements for the measurement of ambient carbon dioxide concentration and comparison with the specification of the PCE-CMM 10 measurement instrument.

	Requirements	Specification of PCE-CMM 10
Measurement range (full scale)	400 ppm–2,000 ppm (0.04–0.2 vol%)	400 ppm–5,000 ppm (0.04–0.5 vol%)
Accuracy	Ca.50 ppm	+ (5% + 50 ppm) between 400 und 2,000 ppm Display-accuracy: 1 ppm

measurements were approximately 18 min in length instead of 15. The blood oxygenation measurements were not carried on after the measurements of the first children had revealed that blood oxygenation never dropped below 98% and was nearly always at 99%, making this variable superfluous. We did not carry out measurements of temperatures and of the breathing volume, as we did not expect reliable results with a face mask. Also, anticipated measurements of the breathability of the material were not done, as these were initially intended for "community masks" which were not used.

2.4. Controls, randomization and quality assurance

Blinding was considered unnecessary, as the measurements are objective. Measurements were conducted exclusively with calibrated and producer-certified apparatuses. The measuring engineer has ample experience in using the apparatuses and has conducted a pilot study of carbon dioxide under masks and piloted all procedures extensively. He is a court-certified, oathbound authorized expert for the measurement of the burden of indoor air with carbon dioxide and methane. Data were documented in real time by written documentation and data capture via the instruments used (data tracing, screen snap shots). Although the device took measures every second, we used only measurements every 15 s, because this assured that the whole period of one experiment of 25 min duration could be documented on one screen (see e-Fig. 1). The data of one sequence of 3 min measurement (i.e. joint air, inhaled air, exhaled air, 12 to 15 measurements per sequence) were then averaged for statistical analysis.

The breathing through the fabric of the mask changed the breathing rhythm in some participants. They either inhaled or exhaled more deeply or both, or had a flatter and quicker breathing pattern. Our measurement design allowed for the smoothing of these changes through averaging across the periods (see Supplement, Figure e−1 for a typical measurement pattern).

The sequence of masks was randomized and randomization was stratified by age of children (below and above age 10). Randomization was conducted using randomizer.org. Two sets of random numbers were prepared, for children up to 10 years of age and older. A coin toss decided whether even or odd numbers meant first surgical or first FFP2 masks. Accordingly, cards with the sequence written on it were put in sealed opaque envelopes with sequential numbering of the child and the age category written on it.

Hygiene rules were followed according to regulations. Personnel was tested to be free of SARS-CoV2.

2.5. Statistics – power analysis

2.5.1. Power analysis

We based our analysis on existing data (Oberrauch et al., 2020). We assumed that we will measure 3,000 ppm (or 0.3 vol%) CO_2 at baseline (inhaled air without mask), i.e. a value which is slightly above current accepted norms because 1000 ppm was expected to be ambient air and a higher value was expected because exhaled CO_2 remains in traces in the vicinity of the face for a while. Thus, this is a conservative estimate. We assumed further that masks will produce values between 5,000 ppm and 12,000 ppm CO_2 in inhaled air. The table of raw-data from (Oberrauch et al., 2020) allowed us to calculate the mean for CO_2-content of breathed air without masks as 3,143 ppm, with surgical mask of 7,292 ppm, as well as a standard deviation of 2,500 ppm for surgical masks, and 1,000 ppm for no masks. This results in standardized mean differences (calculated with the larger SD for a conservative estimate) of d = 1.6. In order to secure such a strong effect with 90% power 7–9 children would have been sufficient per comparison, i.e. 18 children altogether. We used a safety factor of 2 and aimed at 40 to 50 children to be included.

2.5.2. Handling of missing data and data treatment

There were few missing data. In some cases, children stopped their experiment early and were not willing to allow for a post-measurement baseline. Such missing values were not interpolated. Sometimes a phase of measurement, for instance inhalation under surgical mask, was shorter than other phases, or those of other children. However, in each and every case, there were enough data to calculate a phase-specific average. Data were averaged over each of the phases (baseline, mask 1 joint mixed air, mask 1 inhalation, mask 1 exhalation, mask 2 joint mixed air, mask 2 inhalation, mask 2 exhalation, baseline post).

2.5.3. Statistical analysis

The statistical analysis used a linear model with a within-subjects factor, called time-factor, or "time" for short. As the mask-type was counterbalanced, a check was run whether there was a sequential effect using a simple *t*-test and visual inspection. There were no differences between the sequences, and hence the sequence was not entered into the model as categorical predictor. Preconditions of linear modeling were checked and met. Since some of the children were not able or willing to stay until the post-baseline measurement, this was discarded from further analysis, because the missing data would have reduced power. There was no numerical and statistical difference between the baseline and the post-baseline (see e-Fig. 2 in the Supplement). Correlations of predictors such as age, breathing frequency, pulse frequency, ambient CO_2 levels were inspected via scatterplots. The only potential predictor was age which was negatively correlated with CO_2 content of inhaled air, i.e. the CO_2 in inhalted air content was larger for younger children, and this was used as a covariate in the linear model (See Fig. 1). All analyses were calculated using Statistica Version 13.3.

3. Results

Forty-five children or their parents called in to participate in the study. Due to organizational restraints – the experiment was tightly timed – and because after three days 45 participants, the figure stipulated in the protocol, were measured, we stopped recruitment. No child was excluded because of a medical condition or exclusion criteria. Children were included in sequential order as they called in for the study. The mean age was 10.73 years (standard deviation 2.63; range 6–17). Twenty children were girls, 25 were boys.

Results are presented in Table 3. Fig. 1 presents the correlation scatterplot of carbon dioxide under FFP2 masks vs. age.

The linear model over time, i.e. the intra-individual sequences, with age as covariate is presented in Fig. 2.

Linear modeling with age as a significant covariate (covariate age: F = 5.6; p = .022; partial eta^2 = 0.11; interaction age*time: F = 4.09; p < .02; partial eta^2 = 0.08) revealed a strong effect of condition (F = 32.9; p < 1 *10^{-9}; partial eta^2 = 0.43). Contrasts showed that the effect is due to the difference between baseline and both masks jointly. Contrasts between the two types of masks were not significant (F = 2.38; p = .13). Residuals were normally distributed and the linearity assumption was met. Linear models of the other carbon dioxide measures - in exhaled air, in joint inhaled and exhaled air and an average of all three – reveal the same pattern of very steep rise from baseline and no difference between the two types of masks, with FFP2 masks showing slightly higher values and were all highly significant. E-Fig. 3 represents this pattern from the data in Table 1 as a Box-and-Whisker-Plot.

There were no significant effects in breathing frequency and in pulse, although a slight increase both in breathing frequency and pulse was visible (e-Table 2). Oxygen saturation of the blood remained always at 98–99%.

4. Discussion

The primary goal of this study was to find out, whether children breathing under a face mask – a surgical mask and an FFP2 mask – would
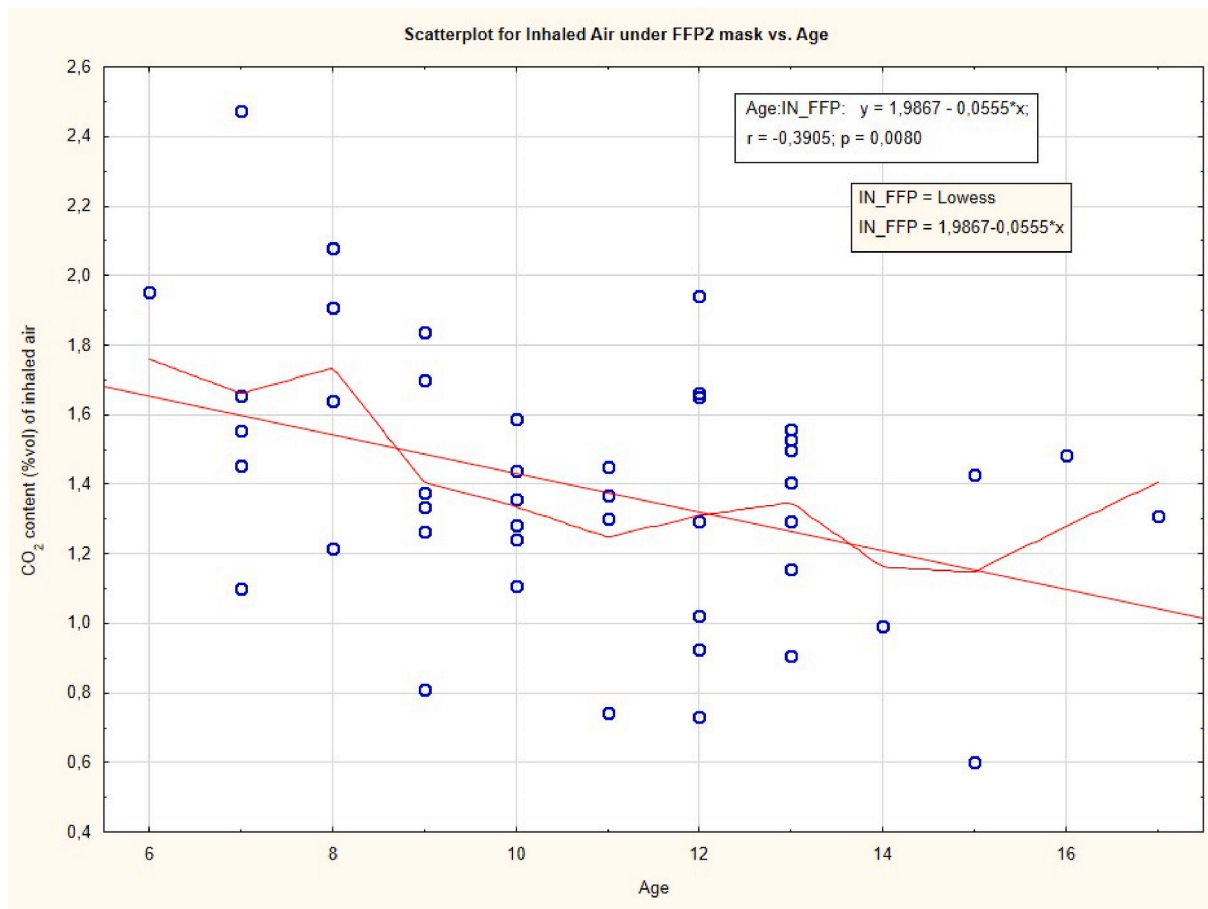
Fig. 1. Scatterplot of CO_2 content in inhaled air under FFP2 mask vs. Age with locally weighted scatterplot smoother to demonstrate approximate linearity.

Table 3

CO_2 values (vol %) under different conditions: means, (standard deviation), [95% confidence intervals], median, minima and maxima, n); * - main outcome.

	Mean (SD)	Median	Minimum	Maximum
Baseline Pre (n = 45)	0.270 (0.110) [0.230; 0.300]	0.230	0.1	0.630
Baseline Post (n = 39)	0.280 (0.100) [0.250; 0.320]	0.260	0.1	0.520
*Inhaled Surgical Mask (n = 45)	1.300 (0.380) [1.200; 1.430]	1.300	0.580	2.550
*Inhaled FFP2 (n = 45)	1.400 (0.370) [1.300; 1.500]	1.370	0.6	2.500
Joint Exhaled and Inhaled Surgical Mask (n = 45)	2.650 (0.490) [2.500; 2.800]	2.750	1.30	3.40
Exhaled Surgical Mask (n = 44)	3.850 (0.680) [3.640; 4.00]	4.100	1.800	4.750
Joint Inhaled and Exhaled FFP2 mask (n = 45)	2.700 (0.400) [2.600; 2.800]	2.750	1.70	3.400
Exhaled FFP2 (n = 45)	3.850 (0.550) [3.700; 4.00]	4.000	2.600	5.20
Ambient Air CO_2 Content	0.075 (0.003) [0.070; 0.075]	0.075	0.070	0.080

Conversion factor: 1.0 vol% = 10,000 ppm.

be exposed to carbon dioxide levels in inhaled air beyond those assumed safe under current regulations in Germany. We deliberately used a still setting in which children were not exposed to any physical or mental workload that would increase their demand on oxygen supply. Even under conditions of sitting still for approximately 18 min with NMC, we measured strong increases in the carbon dioxide of the inhaled air under the face mask. The increases were numerically large and statistically highly significant. The results were very robust. Thus, carbon dioxide is accumulated in the mask and is inhaled back. This increases carbon dioxide in inhaled air under NMC to levels that violate accepted safety norms for carbon dioxide. We were clearly able to distinguish between carbon dioxide content in inhaled air, in exhaled air, and in the joint inhaled and exhaled air which speaks to the validity of our results.

Our findings have been corroborated since by a different research group that used measurement tubes inserted into the noses of adults (Rhee et al., 2021). These findings support the potential causal link between symptoms, such as headaches and fatigue (Ong et al., 2020; World Health Organization (WHO), 2020), and raised carbon dioxide content in inhaled air under NMC. To our knowledge there are no other measurement studies that would invalidate or contradict our data.

4.1. Limitations

Limitations of our data can be considered the fact that we only measured sedentary children. Because of time restraints we could not conduct a more extended measurement with various conditions, such as physical exercise, or relaxed reading. Instead, all children were just measured seated. While some of them brought a book and read during the measurement, others simply observed the experiment. Further work might consider a more extended period of measurement time, real life monitoring or measurement after exertion. Also, long term measurements after prolonged mask-wearing after a full school day should be instantiated to see whether oxygen saturation of the blood is affected long term, which was not affected during the short period of our experiment.

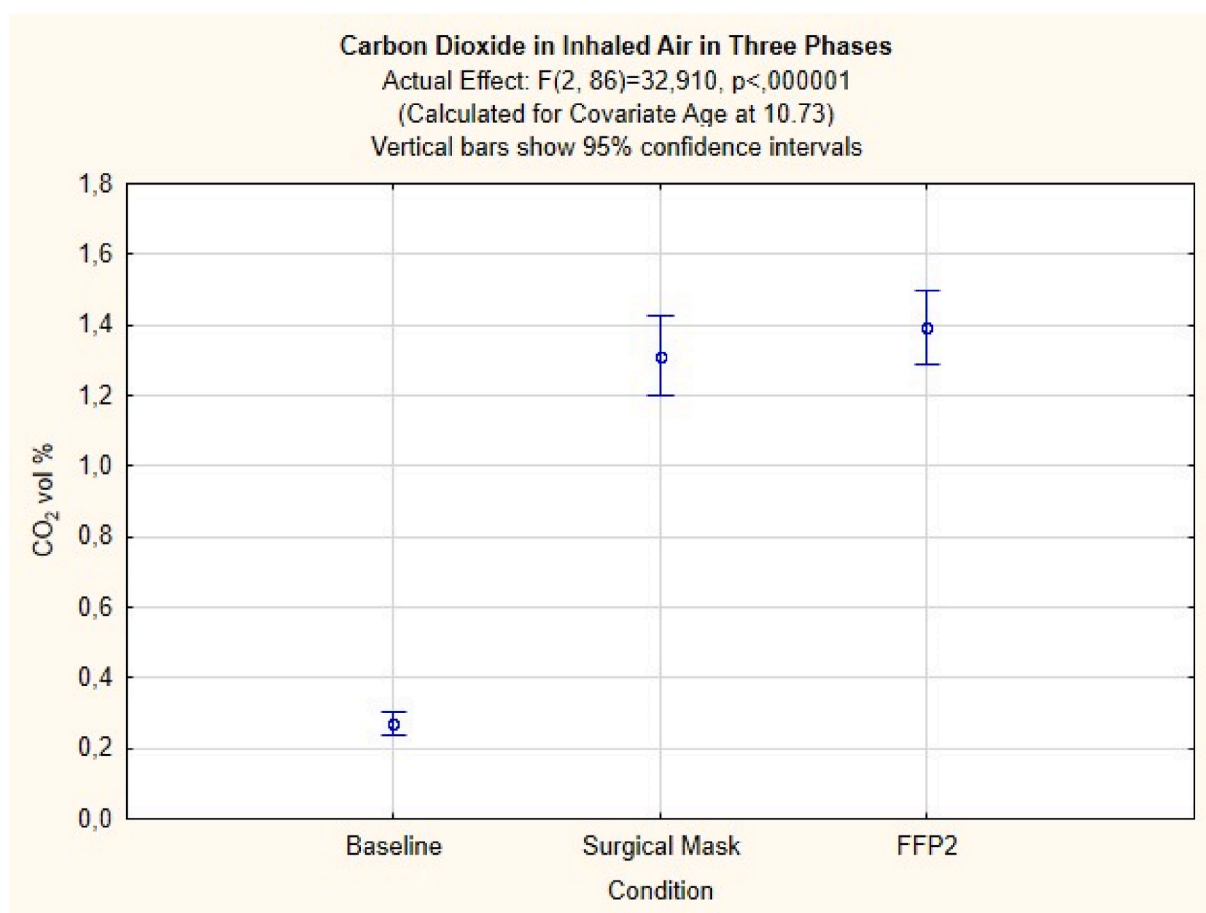Our baseline measurements were comparatively high, although the

Fig. 2. Average carbon dioxide content (vol%) in inhaled air during baseline (no mask) and after 6 min breathing under a surgical and a FFP2 mask; results of a linear model with age as covariate.

ambient CO_2 content was kept well under 0.1 vol % and was on average 750 ppm. This is due to the fact that we measured on the face, between upper lip and nostril, and exhaled CO_2 lingers on in traces until the next inhalation, producing higher measurements. As we were interested in CO_2 content of inhaled air under NMC this elevated baseline value exerts a conservative effect, decreasing potential differences. Hence, it cannot invalidate our findings.

Another point for potential critique is that we could only exploit every 15th measurement that the apparatus provided. This was due to the fact that we used screen capture as a safe and visual method of capturing the data, as it allowed for immediate feedback about the behavior of the measurement device, and the screen size had to be resized to allow all measurements in one display. But as can be seen from the sample screen (e-Fig. 1) the measurements reached stability very soon and a higher sampling frequency of measurements would have produced more stability if anything. But as already those measurements which we took gave very stable results, this limitation does in no way invalidate our results. It would have clearly been better to use both methods in parallel but this was impossible for logistical and technical reasons.

Other potential confounders were clarified in pilot measurements, such as the potential suction or pressure during breathing that might produce errors. These pressures were measured beforehand with a high-resolution manometer that showed maximally 5 Pa pressure changes. Furthermore, the sensor of the G100 makes sure that the air transported to the sensor is always of sufficient volume, as the flow rate is 100 ml/min. If the pressure and the volume had been too low an error message would have resulted, which never occurred. Another potential measurement error could have been the reaction time of the medical doctor

operating the pump. However, this potential error is quite irrelevant. The mean duration of inspiration was 1.36 s. Assuming a delay of 0.2 s in operating the pump when inhalation starts or ends, this would result in a transfer of a volume of about 0.33 ml of air. The delay at the initial phase of the inhalation will result in 0.33 ml of air less that is pumped out of the dead space volume to the sensor, i.e. a reduction in CO_2. But the same will occur at the end of the inhalation phase, where about the same amount of air that is potentially part of the exhalation will be pumped to the sensor. Thus, these two errors average out. Altogether, this amounts to a potential error of 0.15 vol %, which is nearly one order of magnitude smaller than the changes we found. Hence, errors are not a sufficient explanation for our findings.

Along the same lines one might argue that the actual inflow of CO_2 is not continuous but follows a pattern, whereby initially a higher carbon dioxide content is re-inhaled, while at the end of an inhalation cycle more oxygen and thus less carbon dioxide will come in. Hence one would have to model the sinusoidal variation of the CO_2 content in the air. While this is technically speaking correct, it was not possible to model such a more complicated pattern with the data we have. Also, we think that this will introduce only a slight error, as mentioned above, and that averaging across the whole phase is a robust and valid procedure. Even if one were to use this argument to reduce the measured CO_2 values, as indicated in the previous paragraph, we would still see CO_2 values that are much too high.

One might argue that after each exhalation the CO_2 content of the air under the face mask is much higher and hence the actually measured CO_2 content at the beginning of the next inhalation phase is an underestimation of the true inhaled CO_2 content because of the limitation of the air flow in the instrument. While this might be theoretically an issue,

we do not think that this is a large effect. First, we always discarded the first 30 s of measurement during the new phase to get rid of such artifacts. Second, if such an artifact had befallen our measurements, it would mean that our results are systematically underestimating the true CO_2 content and would in truth be higher than they already are. For all practical purposes, this would not make much difference.

4.2. Explanations

One explanation of our results is the accumulation of air loaded with CO_2 in the dead space volume of the mask. Exhaled carbon dioxide gets trapped there and mixes with freshly inhaled air. Because of the geometry of the mask one can assume that the dynamics of the turbulence leads to a mixing of exhaled and inhaled air. This has been confirmed in several studies (Butz, 2005; Rhee et al., 2021). Thus, when inhaling under NMC the CO_2 trapped in the mask mixes with incoming air. However, the fresh air is coming in mainly at the margins of the mask, depending on its form and fit. Our results tally comparatively well with the findings by Oberrauch and colleagues (Oberrauch et al., 2020): While we found 1.4 vol % CO_2 in median with FFP2 masks, Oberrauch found 1.5 vol %. Our results with surgical masks are around 1.3 vol %, while Oberrauch found 1.15 vol % with community masks made of textiles, and a group of volunteers that ranged from 7 to 80 years of age. Also, the recent study by Martellucci et al. (2022), who used capnography and measured below the lips, i.e. measured the mixed air, found very similar results for FFP2 masks: in children, the mean CO_2 content inhaled under FFP2-masks was 12.847 ppm (range 10.774–14.920). As these authors measured the turbulent mixture or air, which can be assumed to be more turbulent in surgical masks, their measurement of CO_2 under surgical masks is lower than ours. We measured the inhaled air directly, while Martellucci and colleagues calculated the inhaled CO_2 content. It should be noted that the measurements of the DIN EN 149 norm define 1.0 vol % as upper limit for adults (Deutsches Institut für Normierung, 2009).

In addition, it is very reasonable to assume that in younger children the dead space volume of the mask is larger than in older children, due to the comparatively smaller size of their faces compared to the size of the face mask. Consequently, we can assume that the mask collects more CO_2 in younger children than in older children. This hypothesis seems a reasonable additional potential explanation, which would, however, have to be ascertained by measurements. Also, younger children usually have a higher respiratory rate, which increases the amount of CO_2 inhaled, according to the data of Martellucci et al. (2022). Breathing volume increases with age. Thus, the ratio of dead space volume to breathing volume changes with age, such that it is larger in younger children and smaller in older children. These potential factors can make it plausible, why we see a larger amount of inhaled CO_2 in younger children than in older children (see Fig. 1), which produces a clear negative linear correlation between age and CO_2 content in inhaled air. This is also corroborated by our "Back to the Envelope Calculation" (see below). Thus, the fact that in younger children the inhaled carbon dioxide content is higher is both an indirect validation of our measurement and a worrying signal. For younger children the continuous exposure to high carbon dioxide contents that exceed safety limits by a factor of 8–12 is very worrying.

One might wonder, why the difference between FFP2 masks and surgical masks is very small and not statistically significant. We assume that this has to do with a combination of the dead-space volume, which is larger in FFP2 masks, the geometry of the masks, the different head-size of the children and the fact that the fresh air enters via the fringes in surgical masks. The combination of these factors likely leads to the fact that the actual dead space volume of air available for breathing is comparatively similar. This might explain, why there is little difference between the masks. But only precise turbulence analyses could actually inform us better. Our study only aimed at measuring the CO_2 content of inhaled air under masks, not at elucidating causal processes.

4.3. Back-to-the-envelope calculation

Using some approximations, measurements of Xu and colleagues (Xu et al., 2015), knowledge of breathing volume, and our results one can calculate the expected CO_2 concentration under FFP2 masks in a "back-to-the-envelope" calculation and compare these with the measured results. The CO_2 concentration in the dead space volume of the face masks (joint inhaled and exhaled air) is between 1.62 and 3.42 vol %. The dead space volume is around 100 ml (Xu et al., 2015). The breathing volume of children is around 7 ml/kg (Marcus et al., 2002). The CO_2 concentration of ambient air near the face is about 0.3 vol % as measured in our study. Using these parameters one can calculate expected CO_2 concentration and compare it with measured concentrations. We have done this with 10 randomly chosen participants of various ages and with differing values of CO_2 concentration and present these calculations in e-Table 3 in the Supplement. The calculations agree with the measurements quite well, and by the same token our assumption that the measured CO_2 concentrations in inhaled air under face masks reflect the rebreathing of accumulated CO_2 in the dead space volume is strengthened.

4.4. Comparison of our results with legal norms and other studies

A value of 5000 parts per million (ppm) or 0.5 vol percent of CO_2 is considered the maximum exposure level of the German healthy-at-work regulations for adult workers during the day, as a time-weighted average over 8 h per day and 5 days per week (Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung, 2021). Similar norms exist in the USA, UK and other European Countries (Centers for Disease Control and Prevention (CDC), 2019). For children and other persons not actively working, lower norms are recommended (Tappler et al., 2017; Umweltbundesamt, 2008). Such regulations state that CO_2 concentrations over 2000 ppm (i.e. 0.2 vol %) are "not acceptable" (Umweltbundesamt, 2008).

We measured between 13,000 and 13,750 ppm of CO_2 in median in inhaled air under surgical and FFP2 masks, which is by a factor 6 higher than the 2000 ppm that is already deemed "inacceptable" for indoor air which is identical to inhaled air by the German Federal Environmental office. This limit of 2000 ppm of CO_2 is by a factor 5 higher than the CO_2 content in normal air (400 ppm). What we measured is an average value of inhaled air during 3 min of measurement and after 6 min of wearing each mask. It is safe to assume that later measurements would have not produced lower values, although it would be interesting to learn what longer time monitoring would result in.

Children under normal conditions in schools wear such masks often for hours. This high content of CO_2 in inhaled air may explain why in a survey in more than 25.000 children 68% of the parents report impairments and problems (such as irritability, headache, difficulty concentrating, less happiness, reluctance to go to school/kindergarten, malaise, impaired learning and drowsiness or fatigue) (Schwarz et al., 2021). Most of them can be understood as consequences of elevated CO_2 levels in inhaled air, which might lead to functional and physiological impairments (Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung, 2021). Similar findings were reported by health workers that had to wear N95 face masks due to Covid-19 regulations (Ong et al., 2020).

Two studies in children wearing NMC did not find clinically relevant altered end tidal CO_2 pressure (Goh et al., 2019; Lubrano et al., 2021). One of them, however, found a relevant change of 3.2 mm Hg, or 3.8 mm Hg respectively, after 45 min wearing NMC. No study evaluated long-term breathing physiology.

A transcutaneous measurement study in medical personnel wearing surgical masks for 30 min confirmed that surgical masks lead to a re-inhalation of carbon dioxide and to an elevated partial pressure of CO_2 which is not compensated by altered breathing patterns (Butz, 2005). Although we did see occasionally changed breathing, overall

breathing frequency did not change during our experiment. A recent study using transcranial Doppler monitoring found increased end-tidal CO2 pressure under N95 respirator masks in health care workers that could be relieved by a special powered air purifying system (Bharatendu et al., 2020). This study shows that such masks do indeed change physiological parameters. While the goal of this study was to demonstrate the effect of a specially powered air purifying system, such systems are complex, expensive and in fact developed for special contexts such as health care workers in high-risk settings.

A recent review summarizing 109 experimental studies, 44 of them quantitatively, concluded that there was ample evidence for adverse effects of wearing face masks (Kisielinski et al., 2021). They are prone to induce what is now called mask induced exhaustion syndrome (MIES) with headache, fatigue, dizziness as its main symptoms, similar to what Schwarz and colleagues found (Schwarz et al., 2021). Our findings are in support of this and can explain why: Even after a short period of time carbon dioxide rises under the mask to unacceptable levels. This is because the fabric prevents free exchange of air and because of the dead-space volume of the masks which collects the exhaled CO_2 and provides it for re-inhalation mixing it with fresh air entering the mask through the fabric and from the fringes of the mask. The process is illustrated in e-Fig. 4.

A recent meta-review summarized evidence of 16 randomized controlled trials and 16 meta-analyses studying effects of face masks in the community (Liu et al., 2021). 14 of 16 studies show no effect, and 8 meta-analyses are critical or unsupportive, while 8 present cautious conclusions. The review itself concludes that data are not convincing. To our knowledge no further measurements studies exist measuring carbon dioxide content in inhaled air in children. It should also be noted that some of the more frequently cited studies in support of face masks are modeling studies that start from unplausible assumptions of close proximity to a viral load of a patient who is deadly ill (Bagheri et al., 2021; Ueki et al., 2020), a situation that is hardly ever realistic in community settings or for children, for that matter.

We did not see any change in blood-oxygenation, which was measured non-invasively using optical methods. This is likely due to the short time frame of our measurements. This was long enough to demonstrate the rise of CO_2 in the inhaled air, but not to see a change in blood oxygenation. It would be interesting to produce such measurements after prolonged wearing of face masks and when actual symptoms are reported, which we did not see.

6. Conclusion

In conclusion we have produced experimental data that show that carbon dioxide content in inhaled air rises on average to 13,000 to 13,750 ppm no matter whether children wear a surgical or an FFP2 mask. This is far beyond the level of 2,000 ppm considered the limit of acceptability and beyond the 1,000 ppm that are normal for air in closed rooms. This estimate is rather on the low side, as we only measured this after a short time without physical exertion. Decision makers and law courts should take this into consideration when establishing rules and guidance to fight infections.

Sponsoring

Data sharing

Data will be provided publicly after acceptance at https://osf.io/yh97a/?view_only=df003592db5c4bd1ab183dad8a71834f and until then for reviewers or on reasonable request.

Author contribution statement

HW is the guarantor of the study. He wrote the protocol, analyzed the data and wrote the first draft of the paper. RW is one of two medical consultants of the study. He generated the idea, helped organize the study and contributed to data collection, interpretation of the results and writing. JP was one of the originators of the study, helped organize it, helped with data collection and contributed to writing and interpreting the data. AD is the second medical consultant. He helped collecting the measurement data and contributed to protocol writing and writing of the paper, as well as to interpreting the data. HT is the measurement engineer. He provided the technical expertise and collected the measurement data. He produced the original data and helped with data preparation. He also contributed to interpreting and writing. AK organized the locale for the study, supervised the children and their parents, managed the recruitment process and contributed to data collection, as well as interpretation of the data and writing. SH is the senior author. He was one of the initiators of the study, provided immunological and general scientific expertise, contributed to organization, data collection, interpretation of the data and to writing.

Ethics statement

The study was conducted according to the regulations of Helsinki, was submitted to the ethics committee of the University Witten/Herdecke, which approved it under Reg. Nr 22/2021. All children participated at their own free will and could withdraw from the experiment at any time and gave written informed consent, as well as their parents, if children were under age 16.

Funding

Role of the funding source

Declaration of competing interest

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2022.113564.

References

Azuma, K., et al., 2018. Effects of low-level inhalation exposure to carbon dioxide in indoor environments: a short review on human health and psychomotor performance. Environ. Int. 121, 51–56.

Bagheri, G., et al., 2021. An upper bound on one-to-one exposure to infectious human respiratory particles. Proc. Natl. Acad. Sci. Unit. States Am. 118, e2110117118.

Bharatendu, C., et al., 2020. Powered Air Purifying Respirator (PAPR) restores the N95 face mask induced cerebral hemodynamic alterations among Healthcare Workers during COVID-19 Outbreak. J. Neurol. Sci. 417, 117078.

Bundgaard, H., et al., 2020. Effectiveness of adding a mask recommendation to other public health measures to prevent SARS-CoV-2 infection in Danish mask wearers. Ann. Intern. Med. 174, 335–343.

Butz, U., 2005. 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]. Klinikum Rechts der Isar. Technische Universität München, München. MD.

Centers for Disease Control and Prevention (CDC), 2019. The National Institute for Occupational Safety and Health (NIOSH). Carbon Dioxide. Standards and Regulations. CDC, Atlanta.

Chu, D.K., et al., 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 395, 1973–1987.

Deutsches Institut für Normierung, 2009. Respiratory Protective Devices - Filtering Half Masks to Protect against Particles - Requirements, Testing, Marking; German Version EN 149:2001+A1:2009. Beuth, Berlin.

Goh, D.Y.T., et al., 2019. A randomised clinical trial to evaluate the safety, fit, comfort of a novel N95 mask in children. Sci. Rep. 9, 18952.

Herby, J., et al., 2022. A Literature Review and Meta-Analysis of the Efffects of Lockdowns on COVID-19 Mortality. Studies in Applied Economics 200. Johns Hopkins Institute für Applied Economics, Global Health, and the Study of Business Enterprise, Baltimore.

Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung, 2021. Datenblatt Kohlendioxid [Datasheet Carbon Dioxide]. IFA Gestis Stoffdatenbank. DGUV - Geutsche Gesetziche Unfallversicherung, Berlin, 2021.

Jefferson, T., et al., 2020. Physical interventions to interrupt or reduce the spread of respiratory viruses. Cochrane Database Syst. Rev. https://doi.org/10.1002/14651858.CD006207.pub5. CD006207.pub5.

Kappstein, I., 2020. Mund-Nasen-Schutz in der Öffentlichkeit: keine Hinweise für eine Wirksamkeit. Krankenhaushygiene Up2date 15, 279–297.

Kisielinski, K., et al., 2021. Is a mask that covers the mouth and nose free from undesirable side effects in everyday use and free of potential hazards? Int. J. Environ. Res. Publ. Health 18, 4344.

Liu, I.T., et al., 2021. Evidence for Community Cloth Face Masking to Limit the Spread of SARS-CoV-2: A Critical Review. Cato Institute Working Paper, vol. 64. Cato Institute, Washington DC.

Lubrano, R., et al., 2021. Assessment of respiratory function in infants and young children wearing face masks during the COVID-19 pandemic. JAMA Netw. Open 4, e210414.

Marcus, R.J., et al., 2002. Pulmonary volume recruitment restores pulmonary compliance and resistance in anaesthetized young children. Pediatric Anesthesia 12, 579–584.

Martellucci, C.A., et al., 2022. Inhaled CO2 concentration while wearing face masks: a pilot study using capnography. medRxiv, 2022.05.10.22274813.

Matuschek, C., et al., 2020. Face masks: benefits and risks during the COVID-19 crisis. Eur. J. Med. Res. 25, 32.

Oberrauch, B., et al., 2020. 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]. Bolzano.

Ong, J.J.Y., et al., 2020. Headaches associated with personal protective equipment – a cross-sectional study among frontline healthcare workers during COVID-19. Headache. The Journal of Head and Face Pain 60, 864–877.

Rhee, M.S.M., et al., 2021. Carbon dioxide increases with face masks but remains below short-term NIOSH limits. BMC Infect. Dis. 21, 354.

Schwarz, S., et al., 2021. Coronakinderstudien "Co-Ki": erste Ergebnisse eines deutschlandweiten Registers zur Mund-Nasen-Bedeckung (Maske) bei Kindern [Corona children studies "Co-Ki": first results of a Germany-wide registry on mouth and nose covering (mask) in children]. Monatsschr. Kinderheilkd. 169, 353–365.

Tappler, P., et al., 2017. Richtlinie zur Bewertung der Innenraumluft. Kohlenstoffdioxid als Lüftungsparameter [Guidance for Evaluation of Indoor Air. Carbon dioxide as a Parameter for Ventilation]. Bundesministerium für Land und Forstwirtschaft. Umwelt und Wasserwirtschaft, Wien.

Traindl, H., 2020. Kohlendioxid-Messungen der Luft unter MNS-Masken [Carbon dioxide measurement of air under standard mouth-nose protection masks]. Traindl-consult, Vienna.

Ueki, H., et al., 2020. Effectiveness of face masks in preventing airborne transmission of SARS-CoV-2. mSphere 5 e00637-20.

Umweltbundesamt, 2008. Gesundheitliche Bewertung von Kohlendioxid in der Innenraumluft [Health assessment of carbon dioxide in air within closed rooms]. Bundesgesundheitsblatt - Gesundheitsforsch. - Gesundheitsschutz 51, 1358–1369.

World Health Organization (WHO), 2020. Mask Use in the Context of Covid-19. WHO, Geneva.

Xiao, J., et al., 2020. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings - personal protective and environmental measures. Ermerging Infectious Diseases 26, 967–975.

Xu, M., et al., 2015. Estimating the dead space volume between a headform and N95 filtering facepiece respirator using microsoft kinect. J. Occup. Environ. Hyg. 12, 538–546.