International Journal of Hygiene and Environmental Health 232 (2021) 113684

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

A Phased Approach for preparation and organization of human biomonitoring studies

Ulrike Fiddicke [a,*], L. Kim Pack [a], Hanna Tolonen [b], Ovnair Sepai [c], Marta Esteban López [d], Argelia Castaño [d], Greet Schoeters [e], Marike Kolossa-Gehring [a]

[a] German Environment Agency (UBA), Berlin, Germany
[b] Department of Public Health Solutions, Finnish Institute for Health and Welfare (THL), Helsinki, Finland
[c] Public Health England, UK
[d] National Centre for Environmental Health, Instituto de Salud Carlos III, Madrid, Spain
[e] VITO Health, Flemish Institute for Technological Research (VITO), Mol, Belgium

ARTICLE INFO

Keywords:
HBM4EU
Study development
Study protocol
Planning in phases
Project management
PM²

ABSTRACT

Background: Human biomonitoring (HBM) studies like other epidemiological studies are costly and time-consuming. They require the administration of questionnaires and collection of biological samples, putting substantial burden on the participants which may result in low participation rates. The growing importance of HBM studies in epidemiology, exposure assessment and risk assessment underline the importance of optimizing study planning, designing and implementation thus minimizing the above-mentioned difficulties.
Methods: Based on frameworks from survey design and fieldwork preparation of the European Joint Program HBM4EU, the German Environment Surveys and the COPHES/DEMOCOPHES twin projects combined with elements of project management strategies, a Phased Approach has been developed, introducing a step-by-step guideline for the development of epidemiological studies.
Results: The Phased Approach splits the process of developing a study into six phases: Phase 0 (Scoping and Planning): All aspects that are necessary to conduct a study are compiled and put on the agenda for decision-making. Phase 1 (Preparation and Testing): Instruments (e.g. questionnaires), materials (e.g. guidelines, information), and ethics and data management issues, needing thorough preparation and testing before a study can start. Phase 2 (Initiation): Organization and acquisition of necessary equipment and engaging and training personnel. Phase 3 (Implementation): All procedures that require temporal proximity to the start date of fieldwork, such as obtaining contact information of invitees. Phase 4 (Fieldwork and Analysis): Involvement of participants and chemical analysis of the collected samples. Phase 5 (Results and Evaluation): Final procedures leading to closure of the project, such as providing and communicating results.
Conclusions: The separation of the planning and conduct of human biomonitoring studies into different phases creates the basis for a structured procedure and facilitates a step-by-step approach reducing costs, warranting high participation rates and increasing quality of conduct. Emphasis is put on a comprehensive scoping phase ensuring high quality of the study design, which is indispensable for reliable results.

1. Introduction

Early epidemiological studies often focused on infectious diseases and death; modern epidemiology has a broader application. Environmental epidemiology studies can be classified into two categories: descriptive and analytical. Descriptive studies include case reports, ecological studies, and cluster studies while analytical studies are based on more detailed data from individuals that can be used to control for confounding, they are usually costlier and more labor-intensive.

Human biomonitoring (HBM) studies are environmental epidemiological studies and can be either descriptive or analytical in design. They include not only response to questionnaires, but also collection of biological samples such as urine or blood from the participants (Angerer et al., 2007). They aim at assessing the exposure levels of a population and establishing the exposure sources by applying statistical procedures on the laboratory results and the provided questionnaire responses

* Corresponding author. Umweltbundesamt, Corrensplatz 1, D-14195, Berlin, Germany.
E-mail address: ulrike.fiddicke@uba.de (U. Fiddicke).

https://doi.org/10.1016/j.ijheh.2020.113684
Received 18 November 2020; Accepted 13 December 2020
Available online 26 December 2020

(Kolossa-Gehring et al., 2012; National Research Council, 2006). Some studies additionally include the collection of environmental samples such as tap water or indoor air or dust samples from the homes of the participants (Kolossa-Gehring et al., 2012). Collecting materials and data has to follow a structured and quality assured procedure because every single step of the study has an impact on the overall quality of the conclusions (Calafat and Needham, 2009; Heffernan et al., 2020). Each one of these single steps needs to be decided upon because it can enhance or deteriorate the survey results and it needs (restricted) resources. Therefore, optimization of processes and resources is necessary. A so called "total survey design perspective" that considers all important aspects of study design can serve as orientation (Fowler, 2014).

Epidemiological studies that include participants in fieldwork are costly. What makes HBM studies even more costly and complex is the need for collecting biological matrices, transport and conservation, the chemical analyses, and related issues e.g. control of contamination and ethical considerations. As HBM gains strength, for example in the field of risk assessment (Angerer et al., 2007; Bates et al., 2005; Choi et al., 2015; Sobus et al., 2015) an increase in the numbers of HBM-studies is expected. At the same time, it has become more and more difficult to recruit volunteer participants for these time-consuming studies (Tolonen, 2015; Blair and Czaja, 2014).

HBM or other epidemiological studies require a wide variety of expertise such as project management, questionnaire development, data management, ethics, data protection, general and participant communication, logistics of materials and biological samples, chemical analysis, toxicology, statistics, personnel for the study conduct (interviewer, field staff, in some cases medical staff) etc. (Choi et al., 2014; Tolonen, 2016).

Successful conduct of a study relies on the development and execution of a well-organized research plan which includes the research question (incl. literature research), development of methods or instruments, study implementation as well as management and analysis of data and last but not least interaction and communication with participants (Weber and Cobaugh, 2008). Support for organizing and managing the resources for a defined project within a defined time period is provided by project management (Berkun, 2005). Project management includes establishing a multidisciplinary project team, defining, in balance with available financial resources, the project scope, developing a timeline, directing project activities, managing problems, and tracking the progress and ensuring quality in the whole process. For the development of a timeline the study needs to be broken down into steps or clusters of tasks, to which resources have to be allocated (Weber and Cobaugh, 2008). Applying these principles of project management to the planning of a study can lead to a better organized and effective study (Weber and Cobaugh, 2008).

Some support for the planning of epidemiological studies is provided by epidemiological associations when they explain which elements have to be generally considered or have to be described in a study protocol. According to the International Epidemiologic Association (IEA), "… the [study] *protocol is the cornerstone of any epidemiological research project*" (IEA 2007) which makes it the main instrument for implementing an epidemiological, or human biomonitoring, study. The study protocol - also called study plan, handbook, or book of operations – starts with a chapter about the purpose and the underlying hypotheses of the study. Following chapters refer to the study design, the study (target) population and the sampling frame to be applied and details about planned activities and analyses. Furthermore, administrative issues, ethical and legal considerations and possible problems and limitations are described (IEA 2007). There are standards of good scientific practice for study conduct available and should be followed by all epidemiologic research initiatives (IEA, 2007). Such standards have been published e.g. by the American Chemical Manufacturers Association (Cook, 1991) and by the German Society for Epidemiology (DGEpi) (DGEpi, 2008; Hoffmann et al., 2019). Whenever a country or institution is striving to perform an epidemiological study, e.g. a health and/or nutrition survey or a human biomonitoring survey including collecting biospecimen, it is well-advised to respect these guidelines.

However, while these standards provide useful guidelines, they do not inform about the roadmap for the planning of such studies which would be helpful considering the complexity of the studies and the different issues to be regarded. We propose here the Phased Approach which complements the existing literature on details of study design, fieldwork, ethics or data management (Ahrens and Pigeot, 2014) in providing a precise, practically and usefully structured, step-by-step approach for planning an epidemiological HBM-study in detail.

2. Methods

The Phased Approach was developed in the framework of the European Joint Program HBM4EU (HBM4EU, 2020). It is based on experiences of the German Environment Agency (UBA) which has conducted environmental surveys regularly since 1985 (Schulz et al., 2007). The experiences of the author and co-authors were already incorporated in the ESBIO project (Development of a coherent approach to human biomonitoring in Europe 2005–2007) (EU, 2013a) and the twin projects COPHES/DEMOCOPHES (EU 2013b; COPHES 2013), a European coordination action on HBM, and now in the Horizon 2020 project HBM4EU (HBM4EU, 2020).

In HBM4EU to some extent data from newly developed studies in different countries will be used. To facilitate data interpretation, this data should be collected using a standardized study protocol. In reality a 'one fits all' study protocol does not exist. Even though the general design can be standardized to a significant extent each country still has to plan and coordinate an HBM-study on its own (Joas et al., 2012; Becker et al., 2014; Casteleyn et al., 2015). These facts were the reason to develop the Phased Approach. The Phased Approach was based on the aforementioned experiences and established through a profound literature research, including other fields of surveys, e.g. clinical trials, complemented with information from project management. It compiles all necessary steps for study conduct and puts them into the chronological order.

For project management several commercially available project management tools exist but also the European Commission has developed an online available project management tool, called PM^2, available for free. It is based on standards and methodologies of globally accepted project management best practices and also on experiences from several European Commission projects (European Commission, 2018). As the idea of the Phased Approach was developed in the frame of the European HBM4EU-Project the PM^2 project management tool was selected to serve as a model for the Phased Approach.

In the field of project management, it is common to operate projects in several – often five or six - phases to facilitate a proper follow-up of the different tasks involved. They start with an initiation or project conception phase, followed by a definition/planning phase and moving on to the development/project launching and execution phase also called implementation/performance phase. Finally, there is the follow-up or project closing phase (Baars, 2006).

PM^2 includes 4 project phases: 1) Initiating Phase; 2) Planning Phase; 3) Executing Phase and 4) Closing Phase (European Commission, 2018: The PM^2 Methodology Guide v3.0). The activities of monitoring and controlling are part of each of the four phases. A description of the content of the five phases is shown in Table 1.

Besides the content of the single phases also their intertwining is of interest.

Fig. 1 demonstrates that in each phase one activity is predominant (e.g. in the Initiating Phase activities of initiating are predominant) but that these phase-related activities will also be executed in neighboring phases, i.e. there is an overlap of activities or some activities are continued in other phases, e.g. initiating activities are repeated in the Planning Phase. A project moves on to the next phase when the goals of the current phase have been reached (ideally, results are reported

Table 1
Descriptions of the project phases in the PM^2 project management tool.

Phase	Description of content
1. Initiating Phase	Define the desired outcomes. Create a Business Case. Define the project scope. Get the project off to a good start.
2. Planning Phase	Assign the Project Core Team (PCT). Elaborate the project scope. Plan the work.
3. Executing Phase	Coordinate the execution of project plans. Produce deliverables.
4. Closing Phase	Capture Lessons Learned and post-project recommendations. Close the project administratively.
5. Monitor & Control	Oversee all project work and management activities over the duration of the project: monitor project performance, measure progress, manage changes, address risks and issues, identify corrective actions etc.

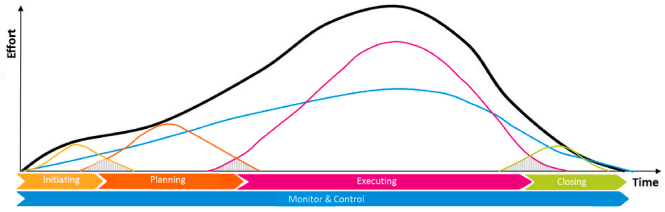


Fig. 1. The PM^2 project lifecycle: indicative phase overlapping and cumulative effort (source: The PM^2 Methodology Guide v3.0).

formally in a review). Although the Phased Approach intends to provide a step-by-step approach, an overlapping of some parts of a phase to the neighboring phases cannot completely be avoided and is partially necessary.

For the Phased Approach six phases have been developed. This was necessary to strengthen the step-by-step approach which is really helpful not to get lost in the manifold topics an HBM study comprises. As pointed out by the IEA (IEA, 2007) the cornerstone of epidemiological projects is the 'study protocol', an instrument widely used for epidemiological studies. It is a compilation and precise description of each issue necessary for the implementation, application, and evaluation of a study. Therefore, the Phased Approach is linked to the study protocol. Within this Phased Approach the key words of epidemiological studies (study design, selection of participants, recruitment, biological samples, fieldwork, questionnaires, analysis, interpretation and communication, and quality control) (Choi et al., 2014) are addressed as well.

3. Results

3.1. Overview of the Phased Approach

The Phased Approach breaks down the process of mapping a study into the 6 phases of (0) scoping and planning (1), preparation and testing, (2) initiation, (3) implementation, (4) conducting the fieldwork and starting with analysis and (5) reporting and evaluation procedures to close the project. This way, a step-by-step procedure becomes apparent and facilitates the development of an individualized study concept and the implementation of the study.

The numbering of the phases starts with "0" because there all general decisions necessary for the study conduct will be taken - this decision process is regularly not part of a study protocol where the Phased Approach is anchored.

Fig. 2 presents a rough overview of the 6 phases which are explained in more detail in the following text.

As an example, Fig. 3 provides an overview of a possible distribution of the workload (effort) for the six phases during a four-year period. In other studies, the duration of the phases will vary according to the survey design.

3.2. Phase 0 – Scoping and Planning

The Phase 0 sets the basis for all aspects of the study. Therefore, it is necessary to take the time for compiling all that is needed later on for the study and to prepare necessary decisions.

Decisions that have to be taken refer to three main questions:

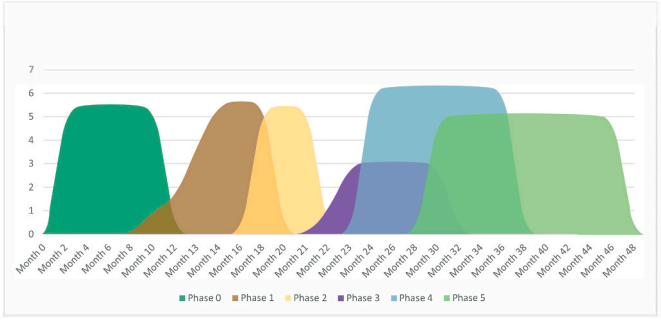


Fig. 3. Example of a possible distribution of the workload (effort) for the six phases during a four-year period.

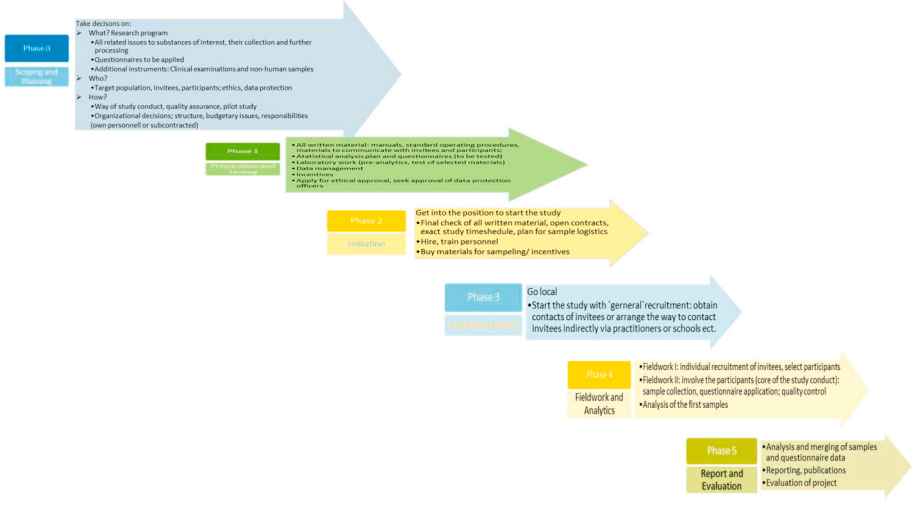


Fig. 2. Overview of the 6 phases of the Phased Approach.

a) **What?** Scope and hypothesis of the study and planned research program, including the substances to be analyzed, and instruments, i.e. human and additional environmental samples, clinical examination and questionnaires (Table 2)
b) **Who?** Target population and selected invitees (Table 3)
c) **How?** Organization and organizational structure of the study, including the distribution of responsibilities and resources (Table 5).

These questions need to be answered in detail, to disclose their interconnections, and to facilitate a proper study conduct.

This description of Phase 0 starts with a closer look first on the What? then on the Who? and finally on the How? questions. But the parts of the How? question dealing with responsibility and budget may be necessary to be decided in advance: who will take all these decisions, who will have to be included for taking specific decisions, which research program is possible with the available budget? All these decisions belong to Phase 0, but without knowing what exactly the study consists of budget is difficult to estimate, therefore, in this text the What? questions precede the How? questions.

Table 2
What? Things to consider and decide pertaining substances, questionnaires, and other additional instruments. Main question: Who?.

Substances of interest	Questionnaires (to elucidate exposure pathways and/or medical history)	Clinical examinations (e.g. lung function tests, weight and height measurements) and additional environmental samples of the homes of the participants (e.g. indoor or outdoor air samples, tap water, dust)
• Specified biomarkers (matrix/analyte): limits of quantification in the target population, time frame for sample collection • Analytical method: LOQ, certified reference material and standards, costs, etc. • Sample volume: for single analysis, repetitions and biobanking • Sample collection and storing material: material, volume, preservatives, control for background contamination, stability for biobanking (including labels), quality assurance aspects. • Processing of the samples during fieldwork (after receiving them from the participants) • Sample conservation and shipment • Sample reception and aliquoting process: number/volume of aliquots, labelling, criteria for acceptance/rejection of the samples • Qualified laboratories for the analyses (different labs may need different volumes)	• Content (e.g. living environment, diet/smoking behaviour, occupation, leisure time activities, health history, medication, socio-economic background; for substances with short half-lives questions that cover the last 24–72 h past urine collection) • Supporting questionnaires: recruitment questionnaire (meeting inclusion/exclusion (= eligibility) criteria, availability, history of contact); sampling questionnaires (i.e. period of sampling, recent exposure related to the target biomarker, etc.); interview guide (explaining the questions and possible answers for the interviewers or participants); non-responder questionnaire; satisfaction questionnaire. • Way of application (face-to-face interview, telephone interview, self-administered, paper and pencil, web based) (also view "How?")	• necessary research equipment (devices, sample collecting containers, communication material, personnel) • necessary validation procedures for new instruments (e.g. new questionnaires or devices) • to whom shall which instrument be applied (to all participants or subgroups?) • time needed for the run of these procedures during fieldwork

Table 3
Who? Target population.

Population aspects	Decisions related to involved population
Target population	general population, risk-exposure groups (occupational, hot-spot) children, adults, etc.
Definition of participants	Age, sex, number; eligibility criteria
Degree and direction of representativeness	National, Regional, local, and according to sex and socio-economic status; oversampling of a subgroup
Data to be collected (and stored?)	Organizational aspects of data management; data protection
Ethics committee for authorization	Requirements of the selected ethics committee, procedure, and duration of the decision procedures

3.2.1. *Main question: What?*

An HBM study is usually conducted either to obtain the distribution of certain substances of interest in the population or to estimate the exposure level of a specific target population. For both questions the substance(s) of interest is a starting point for all further decisions, as target substances have implications on the study design, the questionnaires, and the fieldwork. For example, the substance impacts the matrix of the biological sample to collect and the half-life of the target chemical in the selected matrix impacts the time of sample collection which again has consequences for the fieldwork (collection procedures differ for different matrices). Also, statistical power calculations used to detect the necessary number of participants for a representative sample are based on the selected substances. More detailed aspects connected to the substances of interest are shown in Table 2. Another What? - also shown in Table 2 - concerns the questionnaires that regularly accompany HBM samples, e.g. to elucidate the exposure pathways of the target substances. When additional instruments, e.g. clinical examinations, or environmental samples are included in the study, additional issues in the "What?"-category have to be thought of.

3.2.2. *Main question: Who?*

The decision on what to study is closely connected to the question on target population (whom). Some countries may start their decision cascade with the decision on the target population and continue with the decisions on the substances, both decisions are closely related.

Table 3 provides an overview of accompanying questions to the matter of the involved population.

To be able to apply the results of the study to the target population, not only the sample of invitees needs to be representative for the target population but also those who actually participate (the participants) have to represent the target population which can only be achieved if the sample was representative for the target population. Decisions pertaining the intended representativeness of the study have on the one hand a great influence on the way the fieldwork can be performed and on the other hand a high impact on data interpretation. Table 4 gives examples of how to obtain a representative sample in different population groups.

A good sampling method for selection of individuals if no population registry for a random sample is available or applicable is e.g. the stratified clustered multi-stage design. Via this design, geographical areas (stratification) are selected within a country. This can be applied if e.g. pregnant women shall be included and be approached when they arrive at the clinic before delivery. Within each of the geographical areas, primary sampling units (PSU), in this example the clinics (other examples: general practitioners, schools, work registries) are selected randomly, (or proportional to the number of individuals in these PSU). Furthermore, individuals are selected randomly within the PSU.

Once the target population is defined, in order to assess if an invitee can be included in the study as a participant or not, the definition of eligibility criteria is necessary. As this may have influence on the necessary efforts to obtain the desired number of participants, the decision on eligibility criteria is important at an early stage of planning of

Table 4
Methods for obtaining a representative sample in different population groups and their sampling frames.

Target population	Sampling frame (to select from the list of …)	Methods for obtaining a representative sample
General population of - adults - with or without children - or only children/ adolescents (separated by sex and/or age)	Population register (country, regional) Or stratified clustered multi-stage design	a) Obtain a random sample, keep track of non-responders and dropouts b) Extract from ongoing study
Vulnerable population (pregnant, newborns, seniors, etc.)	Patient files of clinics/ doctors/midwives	Obtain a random sample, keep track of non-responders and dropouts
Occupational population	Employment records, branch organizations, large cohorts	a) Prepare a list of eligible sampling units (workplaces) for random sample b) Extract from large database/cohort
Children/adolescents (different age groups)	Kindergartens/day care centers, or their groups Schools, vocational schools, or classes	Prepare a list of eligible sampling units (schools, day care centers) for random sample

Table 5
How? General decisions on way of study conduct, and on organizational structure, responsibilities, and budget.

Study conduct, general decisions	General decisions on organizational structure, personnel, and subcontracting
• Study design (cross-sectional, cohort, etc.)	• Organizational structure (Study Owner; Principal Investigator; Project Manager, managing team, etc.)
•Sampling frame for the participants (how shall participants be contacted?)	• Budgetary issues and allocation of financial, material, and human resources
• Timing and duration of the fieldwork of the study (season (how many seasons?)	• Personnel/responsibility for the preparation of all the written documents (like study protocol, Standard Operating Procedures (SOPs) for each instrument that is applied or developed, the data management and communication material for the participants, collaborating organizations and the general public
• Repeated (?) involvement of individual participants in the study	• Personnel/responsibility for organizational and pre-sampling procedures (e.g. mailing issues, preparation of sampling equipment/vessels)
• Manner to apply selected instruments (e.g. questionnaires); collection and drop-off of biological samples; additional physical measurements	• Personnel for execution of the fieldwork (nurses and/or professional interviewers), temporarily or permanent staff?
• Place of study conduct/participant involvement (participant homes; place of productive hours like schools, workplace, or the official head quarter for the study (examination center like in schools, clinics, town halls) or mobile centers	• Necessity for subcontracting (e.g. execution of fieldwork, laboratory work
• Incentives for participants (if decided to provide, what kind? reimbursement for travel costs and/or for spending time and samples): Information on individual and general study results; small gifts and certificates for participation	
• Quality assurance of the whole program and single steps (internal, external?)	
• Pilot Study – which instruments and processes shall be tested?	

the study. The eligibility criteria should consider general aspects of the study design, e.g. the age of the participants, and others that can affect the results such as illness, language difficulties, hot spot areas or working in specific enterprises.

A last but most important aspect pertaining the Who? is the compliance with ethical standards and data protection issues. The national and e.g. European regulations on ethical, legal and data protection issues have to be ascertained in an early state of the study planning to warrant proper compliance.

3.2.3. Main question: How?

Decisions on the How? already go deep into the details of a study. They involve the two aspects of 1) way of study conduct and general decisions necessary for this (based on the decision on the research program), and 2) organizational decisions (Table 5).

The type of **study** has big implications on each aspect of the study conduct but also on the scientific significance especially if elucidating causality is aimed at. Cross-sectional studies, for example, can answer policy questions related to the actual exposure levels of the target population but cannot be used to answer questions on causality (Kleinbaum et al., 1982). Decisions on the **sampling frame** detail how invitees can be contacted. The **timing** (seasonal aspects) and duration of the fieldwork have implications on the representativeness and on organizational aspects of the study. The target groups and their respective occupation have to be reflected (e.g. a study planned to recruit school children in schools should not start during holidays). Preferably, a study should cover all four seasons to avoid seasonal variation of the outcome measures. If this is not possible, potential biases have to be taken into account. Time needed for training activities of the field staff before and during the fieldwork has to be considered appropriately.

The decisions on the research program (What?) include decisions on the **instruments** to be applied, but it is also necessary to decide about the way **how** they shall be applied and how much time is needed, e.g. to answer the questionnaires. This duration has implications on the duration of the whole fieldwork. The more instruments shall be applied the more burden is put on the participants and this may also influence the participation rate (Mindell, 2015). Burden put on participants is an ethical aspect. Therefore, a balance between the necessities of the research program and the practical feasibilities of the study is necessary.

Decisions on the **place** of direct contact to the participants, private homes, or centralized examination rooms, also have implications on the organization of the fieldwork. It is advisable to offer alternative possibilities to comfort the participants, though it has to be checked if this is appropriate for the study instruments (home visits are necessary if e. g. indoor air or drinking water are to be collected by the field staff).

In order to obtain high participation rates, it is advisable to identify potential obstacles in the enrolment process early on and think about ways how these could be addressed. Progress of the recruitment should also be monitored throughout the fieldwork and adapted if necessary. If ethics committee allows, additional **incentives** can be offered such as financial and in-kind compensations. Incentives can be provided unconditional or conditional to participation. Type and format of incentives should be addressed in the first information for the participants (invitation).

Besides the quality control aspects pertaining the analytical and pre-analytical aspects of the samples and sample collection, **quality** control and quality assurance have to be applied for all instruments and the study as a whole. Internal quality control, e.g. for the fieldwork, can be performed by personnel of the unit of the institute or division not involved in the study. Additional external experts for quality control are especially advisable for larger research programs. Time for hiring and budget implications for these external experts have to be considered, too.

To warrant high quality results, it is advisable to conduct a **pilot study** to prove the interaction and functionality of the main instruments with a small group of participants. Conducting a pilot study will include

additional time and budget as all decision steps for the main study have to be repeated (in parallel) for the pilot study. But the pilot study provides information on the feasibility of the actions and acceptability of measurements by participants and thus has implications to the costs and the quality of the results of the main study to a considerable extent (Becker et al., 2014; Biemer and Lyberg 2003).

Additionally, and finally, decisions are also necessary on **organizational issues**, pertaining the structure and the personnel responsible for the main parts of the study (Table 5).

To have a clear **structure** and a clarification of the sharing of **responsibility** is extremely important for the success of a complex study (Tolonen, 2016). According to project management, the establishment of a project team is a necessary activity to effectively plan and execute a project (Weber and Cobaugh, 2008). This supports the inclusion of different perspectives while setting up the research plan and planning the study. It plays an important role in ensuring the successful conduct of the project and for the concrete tasks, e.g. the preparation of all necessary (written) material, but also for decisions in case of unforeseen events. Allocation of human resources is, like material and financial **resources,** included in budgetary decisions. Mostly the **budget** has a steering (often limiting) function on the structure, the number of participants involved, the instruments applied, the duration and organization of the fieldwork and the **subcontracts** or additional permanent or temporary **staff** required.

The How? Questions also tackles the question on **responsible personnel** for developing the written material. Indispensable for the quality of an epidemiological study is the proper elaboration of written instructions, summarized in the **study protocol** and detailed in the Standard Operating Procedures (**SOPs)** and **questionnaires**.

It will be necessary to develop diverse **communication** material, as communication is key to ensure a successful contact to the participants and to reach acceptable participation rates (Exley et al., 2015; Tolonen, 2015). These might not only be written material but can also take the shape of information meetings for potential participants, informational videos, social media presence or websites. For this or, e.g. for a potential new corporate design for the study (a logo), subcontracts/tender procedures may have to be set up.

Project management guidelines recommend preparing a timeline for the whole project.

This could also be part of the Phase 0 or a first action of Phase 1.

After these final decisions have been taken, Phase 0, the Scoping/Planning Phase, ends. All theoretical considerations are now done. At this point the next phase starts with the elaboration of more details.

3.3. Phase 1: Preparation and Testing

The main focus of Phase 1 is on arranging those instruments and materials that need preparation and/or testing before they can be used for the study. If not already done in Phase 0, a first task of Phase 1 is the fixing of a **timeline** for the project. Also, it has to be safeguarded that the progress of each task is trackable (with indicators, milestones) to be able to steer the process.

The first emphasis should then be put on the **written materials** because much of this is needed for the **approval of the study by an ethics committee** and the acceptance of the data protection officer. Both have to be approached in this phase and their approval carefully reviewed as it is a prerequisite for each study. Additionally, translations into different languages might be necessary which will also take some time.

The preparation of a **Fieldwork Manual**, a collection of all written materials that are necessary for the fieldwork should start.

The development of **questionnaires** is in itself a complex and time-consuming endeavor. Therefore, it is helpful to rely on already used and tested or validated questionnaires or questions whenever possible. A **statistical analysis plan** can help to streamline the questions to the research plan. Each question needs a justification or rationale for being asked and this background information should be put together in an **Interviewer Manual** if the questions are administered by interview. Otherwise this background information can be used for the elaboration of the information material for the participants (FAQs). Before questions or questionnaires can be applied to participants it is necessary to test them with at least 15 volunteers (Perneger recommends 30 (Pernegger, 2015)) not only for the validation of the questions, but also to get an impression on the time needed for answering. This participant burden is of interest for ethic committees. If questionnaires are administered by computer assistance (CAPI/CATI/CASI/CAWI), respective codebooks for the programming of the questionnaires for computer programs have to be developed.

Another part of the **communication material** that should be prepared in this phase are all other information materials about the study itself (information leaflet, flyer etc.) and instructions for study participant, e.g. handling and storage of the samples the participants have to take themselves, such as the first morning urine samples.

First ideas for the communication of the study to the target population should also be developed in this phase.

Suitable **communication** is key when aiming at ensuring a successful contact to the participants and to reach acceptable participation rates. Therefore, already at this planning stage implications of the communication aspects are important to reflect (Exley et al., 2015; Fiddicke et al., 2015).

Besides these written materials, **other preparatory work** should be started in this phase as shown in Table 6.

3.4. Phase 2: Initiating

Phase 2 comprises the concretization of the work ahead. All what is needed to be able to start with the fieldwork has to be turned into practice, initiated, e.g. material has to be bought or laboratories contracted, and final decisions regarding fieldwork, laboratory work and data management have to be made and respective measures taken. At this point, all prerequisites for the study, like ethics authorization and data protection issues are already solved.

Table 7 provides a detailed overview of the tasks of Phase 2.

The **Fieldwork Manual** (and other written material) has to be checked for necessary updates (e.g. printing errors etc., but attention not to get into conflict with the already received ethical authorization) so it can be used for the **training** of the field staff. Training workshops provide an overview on the study itself, its background, the background of

Table 6
Preparatory work to be completed in Phase 1. After all this preparatory work is finalized, or at least well on track, the next, more detailed work can start in the Phase 2.

Data management and incentives	Laboratory work
• Develop a data management plan • Creating a database for the contact details and the recruitment procedure • Creating a separate database for the questionnaire data and analytical results •If non-monetary incentives are planned their design should be prepared • For monetary incentives, a reception sheet should be prepared	• According to QA/QC-criteria qualified laboratories for the selected biomarkers have to be contacted or a tender process be prepared. • Test the planned collection materials on their usability, e.g. tubes and vessels are free of contaminants, labels are suited for the sampling and storing conditions (temperature, humidity, etc.) • Prepare a sample reception protocol to be filled in by involved laboratories, necessary to control the integrity of the packaging and the conditions of the sample tubes and vessels (respect the General Data Protection Regulation) • **Database** of aliquots: Create a database including the sample ID code, aliquot ID code, sampling date, freezing date, type of sample, aliquots remaining after analysis, location in the bio bank, etc.

Table 7
Overview of the tasks of Phase 2.

Remaining issues to get into the position to start the study
• Updates of the Fieldwork Manual with all SOPs, communication material and questionnaires.
• Engage qualified interviewers/fieldwork staff.
•Organize and perform the training of the interviewers/fieldwork staff.
• Decision on exact starting date and duration of the fieldwork, provide a route plan for the visit of included cities (sampling or study locations).
• Organize/purchase the incentives which have been selected for the participants (books, bags, etc. with study logo) and a reception sheet for monetary incentives.
•Purchase material for the sampling of the matrix to be collected (sample vessels, aliquot tubes) and material for the field staff (laboratory equipment, office, and dispatch material). If necessary, prepare the material for sampling (clean with acid solution, label it, etc.). Also, material for sample transport to the laboratory or biobank have to be considered.
• Fix relation to targeted laboratories, sign contracts. Define the date and delivery format for the analytical results: type of file, units, report about the internal quality controls applied, etc.
• Provide packing lists and prepared material for the field staff.

the questions and specific topics and explain the details of the workflow (how to plan and conduct the interview, how to take samples, sample aliquoting, transport, sample reception, filling out all documents involved in the sampling procedure etc.) The training should also provide hands on training, e.g. pertaining the samples and sampling and conducting interviews.

Phase 2 also includes the creation of a detailed **fieldwork schedule** for the field staff with start date, end date and route plans concerning the involved study locations (towns, cities, or regions where the fieldwork will take place).

The decisions pertaining the **incentives or compensation,** be it monetary or non-monetary, will need concretization in terms of buying (e.g. small gifts) or designing (certificate).

The fieldwork logistics have to be planned and prepared including obtaining required devices and materials, their temporal storage, and transport to and from the fieldwork location.

Laboratory equipment plays an important part in this phase. If the collected samples shall be (partly) processed directly in the field, it is necessary to ensure that minimum laboratory equipment will be available, e.g. refrigerator, centrifuge to warrant optimal conservation, and appropriate facilities to avoid the contamination of the samples.

If transport of samples is foreseen, the packaging must fulfil the regulations (local and general) concerning the shipping of biological material. If the transport will be done by couriers, the details what exact their service covers needs to be checked in advance to prevent loss of samples.

If external laboratories are involved (selected in Phase 1), contracts need to be negotiated, set up and signed. These contracts should also include the date of finalizing the analyses and the date and format for reporting the results.

To further support the field staff or interviewers, a checklist with all materials and devices for a study visit can be compiled.

All databases necessary and a data management plan have already been developed in Phase 1. Final test runs for these plans and databases can still be done in this Initiation Phase.

3.5. Phase 3: Implementation

Phase 3 deals with the 'general' recruitment. It pertains all that is needed to reach the target group that shall be involved in the planned study. It only begins some two-to-three weeks before the start of the fieldwork, which is defined as the direct involvement of the participants. What exactly is necessary in this phase varies according to the way volunteers will be included in the study.

In general, for population based representative studies, an invitation to all persons of the selected sample of the sampling frame fitting to the target population is necessary. Often contact details are obtained from sampling frames, like population registries of selected towns or lists of pupils of selected schools. Another way is to contact selected schools to prepare teachers to be the promotor of the study in selected classes or to inform the parents via the school internet or special events that their children will be invited. Physicians, midwives, or selected maternities are sometimes approached to reach special groups of probable participants. The professionals need to agree and be prepared to forward invitations to the persons that fit to the target group. Finally, as a result of this 'general' recruitment, one of the two options has to be accomplished either a list of potential participants to whom invitations can be sent shall be available or counterparts for direct participant contact are established – so the sample of the invitees is prepared.

Timing of this 'general' recruitment is a sensible issue. It should not be done too long in advance before the fieldwork shall start because invitees might change addresses or approached counterparts are no longer available. This risk increases with the timely distance between searching for addresses and sending individual invitations.

If addresses of individuals have been collected, they then can be transferred to the databases set up in Phase 2. Additionally, each individual is assigned a study-specific ID number which serves for pseudonymization of results. During this process and from here on out, data protection always has to be ensured.

During this Implementation Phase also the visit at the first study location (i.e. the region or city where the study will take place) is prepared. For example, the place or rooms that serve as examination centers have to be rented as well as rooms for field staff. Examination centers serve the needs for the study, i.e. consist of separate rooms (as waiting room or reception, room for interviews, room for exercises, sanitary facilities, etc.) and can be in schools, town halls, or clinics etc. They should be easily accessible and are necessary in each study location. If overnight stays of the fieldwork staff are necessary, this has also to be arranged.

This is also a good period of time to raise awareness for the study and start communication with the target or additionally the general public about the study at the sampling location with the aim to increase the participation rate.

Finally, the laboratories that will analyze the biological samples have to be informed about the upcoming start of the fieldwork so that they are ready to start with the first analysis.

3.6. Phase 4: Fieldwork and Analysis

Fieldwork starts with the direct contact to the individual participants. The first part of fieldwork (Fieldwork I) comprises the individual recruitment and starts directly after the general recruitment has provided the contact details of those who shall be invited or has informed about the way how to approach them.

Invitations can now be sent, or direct contact be established. Invitees shall be provided different ways to articulate their acceptance or rejection of the invitation (e.g. personally, via phone, a reply card, email, or online tool). For organizational issues and to have the ability to calculate response rates it is necessary to keep track on the answers. Volunteers need to be contacted and checked for eligibility criteria. Those who will finally be included in the study, the participants, will be informed on their participation and probably sent detailed information or additional materials necessary for the study. Appointments for personal interviews or clinical examinations will be fixed.

Individual recruitment can be very time consuming, as it is necessary to grant several days for response and provide several reminders – how many should be detailed in the Fieldwork Manual.

In those cases where invitations have been sent but invitees did not react the field staff can still try to recruit them personally. For statistical reasons (representativeness) it is important to try very hard to reach each randomly selected participant (see Fieldwork II below).

In case a potential participant refuses to participate, a non-responder

questionnaire should be administered to be able to compare non-participants with participants to check for potential response bias.

During the whole fieldwork procedure – until the participants have received their individual results - a phone number and an email address of the general study office as contact address has to be provided.

The second part of fieldwork (**Fieldwork II**) is the core element of the study. Here the participants are directly involved in the study: samples collected, questionnaires are applied, clinical examinations are carried out and afterwards incentives and results provided - all depending on the instruments decided upon in Phase 0 and precisely laid down in the Fieldwork Manual.

An example for a schedule for the fieldwork procedures at study locations could be: Upon arrival at the study location the examination center is furnished with study equipment and devices. Before the visit of every single participant, the material (such as interviewer identity card, papers, laptop, additional sampling vessels, incentives etc.) is checked, prepared, and stocked up. On arrival of a single participant, the field staff checks and accepts the declaration of informed consent from the participant. Only then the other instruments of the study are applied, e. g. the questionnaire can be filled out, measurements and samples taken. In case not all parts of the fieldwork could be completed at once, an additional visit in the next few days should be offered to the participant.

If the participant involvement takes place in the participants' homes special care has to be taken. Respect for the residents and close observation of household etiquette is strongly recommended to avoid negative effects on participation rates.

Each visit needs to be well documented with details concerning duration, completion, handovers and consent, and all procedures accompanied by quality control measures to warrant high quality of the received results.

Procedures for **handling of the samples** collected from the participants have been laid down in the Standard Operation Procedures (Phase 1 and 2). Now, in this phase those SOPs have to be followed closely, e.g. aliquoting of the matrices, preparing for shipment or storage.

If the study includes more than one study location all procedures will have to be repeated.

In parallel, the laboratories can **start analyzing samples** to be prepared to report the results back to the participants (as laid down in the ethics documents).

The managing site of the study (Principal Investigator, Survey Office) has to supervise the work of the field staff and to provide help and advice if necessary. Furthermore, internal quality control for fieldwork has to be established to check for signs of differential participation and to compare with the target population. It is further required to organize and conduct additional trainings for the field staff to maintain the quality of the fieldwork.

Another important issue pertains the safeguarding of data protection. E. G. in Europe, the GDPR sets strict rules on the handling of sensitive personal data. Safeguarding compliance with the GDPR is a prerequisite for every study to be conducted in Europe.

3.7. Phase 5: Reporting and Evaluation – Leading to Closure

After the fieldwork is completed in one study location and the results of the sample analyses have been checked, the participants have a right (e.g. based on GDPR) to have their individual results reported back to them, preferably accompanied by some explanation if the values are prone to raise concern. After the fieldwork and the sample analyses are completed in all study locations results of the different instruments have to be merged and checked, participants who can count as a case have to be defined, i.e. who provided the information defined as indispensable (e.g. on smoking, alcohol/drugs use, diet, number of years living in the sampling location etc.). In the following, the cases are statistically analyzed according to the research questions. Only then advice for the (general) public and policies can be provided.

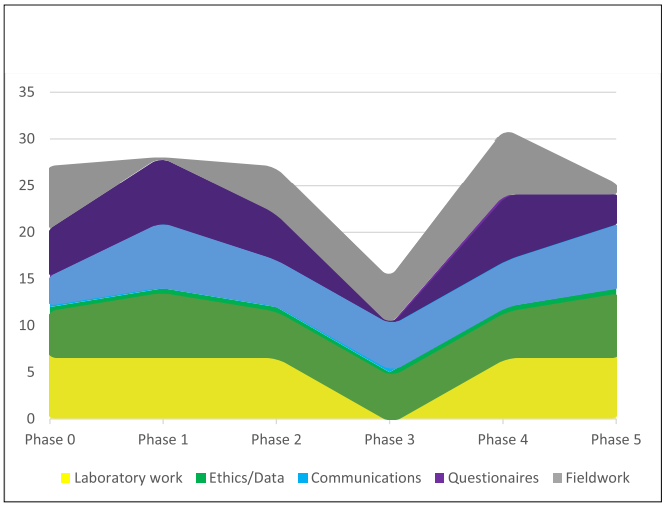


Fig. 4. Example for a schematic overview on the distribution of efforts of the different parts involved in study conduct.

Each part of the study should finally be evaluated – to support the conduct of following studies.

This phase is to close the project – all remaining activities, like publishing results, reporting back to funding organizations and finalizing contractual obligations are part of this last phase.

The workload of the different parts involved in a study, like laboratory work, ethics and data management activities, preparation and use of communication materials, questionnaire development and application and fieldwork preparation and conduct are unequally distributed across the different phases. Fig. 4 provides an example for a schematic overview on the distribution of efforts of the different parts.

4. Discussion

Epidemiological studies consist of many elements and have high budgetary implications. These rise even more if it concerns a human biomonitoring study which includes not only the interview of volunteers but also biospecimen to be taken and analyzed. Given the wide range of biomarkers, biological matrices, analytical methods and influencing factors, the study design is a critical aspect in the efficient use of HBM in risk assessment (Choi et al., 2014). The planning of such a study should therefore be comprehensive and include all aspects such as research subjects, materials to be included, target population, selection of participants, communication with all interested groups (stakeholders), biological analyses, data management, scientific publications, policy advice and first of all ethical and legal considerations. It is possible to find scientific publications on all of these single aspects and there are some guidelines available for good epidemiological practice. Also, several countries that run regular HBM programs have informed about their national strategy on human biomonitoring Belgium (Reynders et al., 2017; Schoeters et al., 2012), Canada (Haines et al., 2017), Czech Republic (Cerna et al., 2007, 2012), Denmark (Thomsen et al., 2008), France (Fréry et al., 2012), Germany (Kolossa-Gehring et al., 2012; Schulz et al., 2007). But none of these publications provide details on the complex procedure of planning a study or inform about an advisable

order for the planning of all the involved aspects.

The Clinical Trials Transformation Initiative (CTTI, 2015) has developed quality-by-design principles for clinical trials to remove errors with meaningful impact on the credibility of the results. They aim at promoting critical thinking about trial design to prospectively build quality into the scientific and operational design of clinical trials rather than relying on retrospective scientific reviews (Meeker-O'Connell et al., 2016). These principles also include the study protocol design which contributes to high quality results if it is designed appropriately.

A good example for the planning of studies is available for the European Health Examination Survey (EHES). The EHES Manual (Tolonen, 2016) provides guidelines for the implementation of standardized national health examination surveys (HES) in European countries but can also be a reliable source for studies in other fields. EHES proposes a circular planning procedure, consisting of 6 parts: (1) defining the scope of the study; (2) planning and preparation; (3) pre-testing and piloting; (4) final survey design, planning and preparing; (5) fieldwork and data collection; and (6) data file construction, analysis and reporting. Quality assurance aspects link these individual parts. After all is done, the planning for the next study can start in the same order. Tolonen also recommends links to project management tools (Tolonen, 2016).

A similar attempt has been undertaken by Statistics Canada 'Survey Methods and Practices' providing insights into survey practices (Statistics Canada 2003 and 2010). Like the EHES Manual (Tolonen, 2016) and the publication by Becker et al. (2014), all provide valuable information on alternative approaches for epidemiological studies, providing background information for an information-based decision making on nearly all aspects of study conduct.

Further in-depth information on study-design is available, e.g. in Blair and Czaja 2014 and Flower 2014 (Blair and Czaja, 2014; Flower, 2014). Both books elaborate on general and specific issues on survey design for research and provide a comprehensive overview of the sources of error and the range of methodological issues in survey data collection. The US National Research Council (National Research Council, 2006) describes in detail considerations in the design of biomonitoring studies and provides with this highly valuable information on several aspects of biomonitoring studies which could help the decision-making processes.

Conducting a study involves many activities and several of them have to be started in parallel. None of the above-mentioned publications promote a structured concept or process for the planning of a study. To avoid not considering important aspects up front - and sometimes it might be too late to change things during the run of a study - the idea to develop a stringent concept for the planning of epidemiological/HBM studies arose and was turned into reality with the Phased Approach.

As stated above, HBM studies need experts from different fields, i.e. multi- or interdisciplinary teams will work together. This in itself may pose some problems to the success of the project. To overcome such problems, Tobi and Kampen developed a Methodology for Interdisciplinary Research (MIR) framework (Tobi and Kampen, 2018) based on a process approach which puts the common goal of the researchers at the center of the project thus creating different phases of the process: first the research team discusses about the different parts of the design of their study and only then the study is executed. During the first discussion of the conceptual design of the study the 'why' (research objective) and 'what' (research question) of the research is decided. Then, the team discusses the technical design of the study ('how': study design; instrument selection or design; sampling plan; analysis plan). The execution of the work (including fieldwork) only starts after decisions of the complete research design have been taken (Tobi and Kampen, 2018). This concept is closely related to the handbook of project management from Baars (2006), the definition phase is also addressed as "What?" and the Design phase with "How?". The MIR framework and the Phased Approach have the technical part – answering the question of "How?" in common, it deals mainly with fieldwork. Whereas the "What?" describes the research question (MIR framework) and the research program (Phased Approach), respectively. This discrepancy can be explained with the fact that the Phased Approach does not include an extra phase for elucidating the research question as it was developed for studies conducted in the frame of HMB4EU where the research question is to a great extent already decided upon on the level of the HBM4EU project. MIR framework and Phased Approach do have in common that first decisions have to be taken (Phase 0) and only afterwards the execution of the study is prepared and started. The Phased Approach was originally developed for (preferably population representative) HBM studies conducted in the frame of the European Joint Program HBM4EU though it can be adapted for other epidemiological studies. Conducting studies in the frame of HBM4EU has the consequence that the scope of and the research question for the intended study has to be within the scope of HBM4EU, i.e. is restricted to certain substances and age groups. Additionally, the time frame for the study is set. In consequence the project conception phase, defining the scope and developing the research questions is no single phase in the Phased Approach but is included in the scoping/planning phase (Phase 0). This can be adapted if the Phased Approach is used in other contexts.

5. Conclusion

Complex (HBM) studies need considerable financial, material, and human resources. This obliges those responsible for the planning of a study to search for methods that support the ideal allocation of the available resources. A study can be dealt with like a project and therefore tools of project management can be applied. Project management, e.g. PM^2 of the European Commission, separates the work within the project into several phases, like initiating, planning, executing, and closing phase which serves a structured procedure and a suitable allocation of resources. Several years of experiences in planning different HBM studies lead to the introduction of project management tools into study design. This is done with the aim to support new researchers in the field of HBM studies in conceptualizing resource efficient studies. Resources in terms of personnel and applied instruments can best be made use of if the concept of a study is thoroughly thought of. Therefore, the proposed Phased Approach lists all aspects which may occur in the later stages of a study already in the first phase as decision points. Only after all decisions are taken, the concrete work for the planned study shall start in a step-by-step approach. Following the Phased Approach can lead to a well-planned study resulting in highly valuable quality assured results with optimized allocation of resources.

Declaration of competing interest

Declarations of interest: none.

Acknowledgment

This document has been created for the HBM4EU project. HBM4EU has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 733032.

This Phased Approach was developed in the framework of the Horizon 2020 project HBM4EU (WP7 "Survey design and fieldwork preparation") (www.hbm4eu.eu) and based on experiences from the development of the COPHES/DEMOCOPHES projects (Becker et al., 2014; Fiddicke et al., 2015; Den Hond et al., 2015; European Commission, 2013b; COPHES, 2013) and the German GerES studies (Schulz et al., 2007).

References

Ahrens, W., Pigeot, I. (Eds.), 2014. Handbook of Epidemiology. Springer, New York. https://doi.org/10.1007/978-0-387-09834-0. New York, NY.

Angerer, J., Ewers, U., Wilhelm, M., 2007. Human biomonitoring: state of the art. Int. J. Hyg Environ. Health 210, 201–228. https://doi.org/10.1016/j.ijheh.2007.01.024.

Baars, W., 2006. Handbook of Project Management 1.2. https://www.projectmanagement-training.net/articles-tools/book/. (Accessed 6 November 2020).
Bates, M.N., Hamilton, J.W., LaKind, J.S., Langenberg, P., O'Malley, M., Snodgrass, W., 2005. Workgroup report: biomonitoring study design, interpretation, and communication-lessons learned and path forward. Environ. Health Perspect. 113 (11), 1615–1621. https://doi.org/10.1289/ehp.8197.
Becker, K., Seiwert, M., Casteleyn, L., Joas, R., Joas, A., Biot, P., Aerts, D., Castaño, A., Esteban, M., Angerer, J., Koch, H.M., Schoeters, G., Den Hond, E., Sepai, O., Exley, K., Knudsen, L.E., Horvat, M., Bloemen, L., Kolossa-Gehring, M., 2014. A systematic approach for designing a HBM Pilot Study for Europe. Int. J. Hyg Environ. Health 217, 312–322. https://doi.org/10.1016/j.ijheh.2013.07.004.
Berkun, S., 2005. The Art of Project Management. O'Reilly, Sebastopol, CA.
Biemer, P.P., Lyberg, L., 2003. Introduction to Survey Quality. Wiley, Hoboken, N.J.
Blair, J., Czaja, R.F., Blair, E.A., 2014. Designing Surveys. A Guide to Decisions and Procedures, 3nd edition. Sage Publications, Inc., California, USA.
Calafat, A.M., Needham, L.L., 2009. What additional factors beyond state-of-the-art analytical methods are needed for optimal generation and interpretation of biomonitoring data? Environ. Health Perspect. 117 (10), 1481–1485. https://doi.org/10.1289/ehp.0901108.
Casteleyn, L., Dumez, B., Becker, K., Kolossa-Gehring, M., Den Hond, E., Schoeters, G., Castaño, A., Koch, H.M., Angerer, J., Esteban, M., Exley, K., Sepai, O., Bloemen, L., Horvat, M., Knudsen, L.E., Joas, A., Joas, R., Biot, P., Koppen, G., Dewolf, M.C., Katsonouri, A., Hadjipanayis, A., Černá, M., Krsková, A., Schwedler, G., Fiddicke, U., Nielsen, J.K., Jensen, J.F., Rudnai, P., Közepésy, S., Mulcahy, M., Mannion, R., Gutleb, A.C., Fischer, M.E., Ligocka, D., Jakubowski, M., Reis, M.F., Namorado, S., Lupsa, I.R., Gurzau, A.E., Halzlova, K., Jajcaj, M., Mazej, D., Tratnik Snoj, J., Posada, M., López, E., Berglund, M., Larsson, K., Lehmann, A., Crettaz, P., Aerts, D., 2015. A pilot study on the feasibility of European harmonized human biomonitoring: strategies towards a common approach, challenges and opportunities. Environ. Res. 141, 3–14. https://doi.org/10.1016/j.envres.2014.10.028.
Černá, M., Krsková, A., Čejchanová, M., Spěváčková, V., 2012. Human biomonitoring in the Czech Republic: an overview. Int. J. Hyg Environ. Health 215, 109–119. https://doi.org/10.1016/j.ijheh.2011.09.007.
Černá, M., Spěváčková, V., Batáriová, A., Šmíd, J., Čejchanová, M., Očadlíková, D., Bavorová, H., Beneš, B., Kubínová, R., 2007. Human biomonitoring system in the Czech Republic. Int. J. Hyg Environ. Health 210, 495–499. https://doi.org/10.1016/j.ijheh.2007.01.005.
Choi, J., Mørck, T.A., Polcher, A., Knudsen, L.E., Joas, A., 2014. Review of the State of the Art of Human Biomonitoring for Chemical Substances and its Application to Human Exposure Assessment for Food Safety. EFSA supporting publications. https://doi.org/10.2903/sp.efsa.2015.EN-724, 2015:EN- 724.
Cophes, 2013. Consortium to Perform Human Biomonitoring on a European Scale. http://eu-hbm.info/. (Accessed 6 November 2020).
Ctti, 2015. Clinical trials transformation initiative quality by design project. In: CTTI Recommendations: Quality by Design. https://www.ctti-clinicaltrials.org/files/ctti_quality_by_design_recommendations_final_1jun15_1.pdf. (Accessed 6 November 2020).
Cook, R.R., 1991. Overview of good epidemiologic practices. J. Occup. Med. 33 (12), 1216–1220.
Den Hond, E., Govarts, E., Willems, H., Smolders, R., Casteleyn, L., Kolossa-Gehring, M., Schwedler, G., Seiwert, M., Fiddicke, U., Castaño, A., Esteban, M., Angerer, J., Koch, H.M., Schindler, B.K., Sepai, O., Exley, K., Bloemen, L., Horvat, M., Knudsen, L.E., Joas, A., Joas, R., Biot, P., Aerts, D., Koppen, G., Katsonouri, A., Hadjipanayis, A., Krskova, A., Maly, M., Mørck, T.A., Rudnai, P., Kozepesy, S., Mulcahy, M., Mannion, R., Gutleb, A.C., Fischer, M.E., Ligocka, D., Jakubowski, M., Reis, M.F., Namorado, S., Gurzau, A.E., Lupsa, I.R., Halzlova, K., Jajcaj, M., Mazej, D., Tratnik, J.S., López, A., Lopez, E., Berglund, M., Larsson, K., Lehmann, A., Crettaz, P., Schoeters, G., 2015. First steps toward harmonized human biomonitoring in Europe: demonstration project to perform human biomonitoring on a European scale. Environ. Health Perspect. 123 (3), 255–263. https://doi.org/10.1289/ehp.1408616.
European Commission, Digit, Centre of Excellence in Project Management (CoEPM2), 2018. The PM^2 project management methodology Guide 3.0. https://op.europa.eu/en/publication-detail/-/publication/ac3e118a-cb6e-11e8-9424-01aa75ed71a1/language-en/format-PDF/source-108080890. (Accessed 6 November 2020).
EU, European Commission, Cordis, 2013a. EU research results: 'esbio'. https://cordis.europa.eu/search?q=%27esbio%27&p=1&num=10&srt=Relevance:decreasing. (Accessed 6 November 2020).
EU, European Commission, Cordis, 2013b. EU research results: 'cophes'. https://cordis.europa.eu/search?q=%27cophes%27&p=1&num=10&srt=Relevance:decreasing/. (Accessed 6 November 2020).
Exley, K., Aerts, D., Biot, P., Casteleyn, L., Kolossa-Gehring, M., Schwedler, G., Castaño, A., Angerer, J., Koch, H.M., Esteban, M., Schindler, B.K., Schoeters, G., Den Hond, E., Horvat, M., Bloemen, L., Knudsen, L.E., Joas, R., Joas, A., Sepai, O., 2015. Pilot study testing a European human biomonitoring framework for biomarkers of chemical exposure in children and their mothers: experiences in the UK. Environ. Sci. Pollut. Res. Int. 22 (20), 15821–15834. https://doi.org/10.1007/s11356-015-4772-4.
Fiddicke, U., Becker, K., Schwedler, G., Seiwert, M., Joas, R., Joas, A., Biot, P., Aerts, D., Casteleyn, L., Dumez, B., Castaño, A., Esteban, M., Angerer, J., Koch, H.M., Schoeters, G., Den Hond, E., Sepai, O., Exley, K., Knudsen, L.E., Horvat, M., Bloemen, L., Katsonouri, A., Hadjipanayis, A., Cerna, M., Krsková, A., Jensen, J.F., Nielsen, J.K., Rudnai, P., Közepésy, S., Gutleb, A.C., Fischer, M.E., Ligocka, D., Kamińska, J., Reis, M.F., Namorado, S., Lupsa, I.R., Gurzau, A.E., Halzlová, K., Mazej, D., Tratnik, J.S., Rivas, T.C., Gómez, S., Berglund, M., Larsson, K., Lehmann, A., Crettaz, P., Dewolf, M.C., Burns, D., Kellegher, A., Kolossa-Gehring, M., 2015. Lessons learnt on recruitment and fieldwork from a pilot European human biomonitoring survey. Environ. Res. 141, 15–23. https://doi.org/10.1016/j.envres.2014.08.039.
Fowler, F.J., 2014. Survey Research Methods/Floyd J. Fowler, Jr, – fifth ed. Center for Survey Research, University of Massachusetts, Boston.
Fréry, N., Vandentorren, S., Etchevers, A., Fillol, C., 2012. Highlights of recent studies and future plans for the French human biomonitoring (HBM) programme. Int. J. Hyg Environ. Health 215, 127–132. https://doi.org/10.1016/j.ijheh.2011.08.008.
German Society for Epidemiology (DGEpi), 2008. Guidelines and Recommendations to Assure Good Epidemiologic Practice (GEP). Long Version. With Revisions after Evaluation (April 2004). With supplement by implementation rule for Good Practice Secondary Data Analysis (GPS). https://www.dgepi.de/assets/Leitlinien-und-Empfehlungen/Recommendations-for-good-Epidemiologic-Practice.pdf. (Accessed 6 November 2020).
Haines, D.A., Saravanabhavan, G., Werry, K., Khoury, C., 2017. An overview of human biomonitoring of environmental chemicals in the Canadian Health Measures Survey: 2007-2019. Int. J. Hyg Environ. Health 220 (2 Pt A), 13–28. https://doi.org/10.1016/j.ijheh.2016.08.002.
Hbm4EU, 2020. Coordinating and Advancing Human Biomonitoring in Europe to Provide Evidence for Chemical Policy Making. https://www.hbm4eu.eu/. (Accessed 6 November 2020).
Heffernan, A.L., Gomez-Ramos, M.J., Symeonides, C., Hare, D.J., Vijayasarathy, S., Thompson, K., Mueller, J.F., Ponsonby, A.L., Sly, P.D., Group, B.I.S.I., 2020. Harmonizing analytical chemistry and clinical epidemiology for human biomonitoring studies. A case-study of plastic product chemicals in urine. Chemosphere 238, 124631. https://doi.org/10.1016/j.chemosphere.2019.124631.
Hoffmann, W., Latza, U., Baumeister, S.E., Brünger, M., Buttmann-Schweiger, N., Hardt, J., Hoffmann, V., Karch, A., Richter, A., Schmidt, C.O., Schmidtmann, I., Swart, E., van den Berg, N., 2019. Guidelines and recommendations for ensuring good epidemiological practice (GEP): a guideline developed by the German society for epidemiology. Eur. J. Epidemiol. 34, 301–317. https://doi.org/10.1007/s10654-019-00500-x.
International Epidemiologic Association (Iea), 2007. Good Epidemiological Practice (GEP), IEA Guidelines for Proper Conduct in Epidemiologic Research. Retrieved from. http://ieaweb.org/IEAWeb/Content/IEA_Publications.aspx Adobe pdf. (Accessed 6 November 2020). "Good Epidemiology Practices (GEP)" (last accessed.
Joas, R., Casteleyn, L., Biot, P., Kolossa-Gehring, M., Castano, A., Angerer, J., Schoeters, G., Sepai, O., Knudsen, L.E., Joas, A., Horvat, M., Bloemen, L., 2012. Harmonised human biomonitoring in Europe: activities towards an EU HBM framework. Int. J. Hyg Environ. Health 215, 172–175. https://doi.org/10.1016/j.ijheh.2011.08.010.
Kleinbaum, D.G., Kupper, L.L., Morgenstern, H., 1982. Epidemiologic Research: Principles and Methods. Wiley, New York.
Kolossa-Gehring, M., Becker, K., Conrad, A., Schroter-Kermani, C., Schulz, C., Seiwert, M., 2012. Environmental surveys, specimen bank and health related environmental monitoring in Germany. Int. J. Hyg Environ. Health 215 (2), 120–126. https://doi.org/10.1016/j.ijheh.2011.10.013.
Meeker-O'Connell, A, Glessner, C, Behm, M, Mulinde, J, Roach, N, Sweeney, F, Tenaerts, P, Landray J, M, et al., 2016. Enhancing clinical evidence by proactively building quality into clinical trials. Clinical Trials 13 (4), 439–444. https://doi.org/10.1177/1740774516643491.
Mindell, J.S., Giampaoli, S., Goesswald, A., Kamtsiuris, P., Mann, C., Mannisto, S., Morgan, K., Shelton, N.J., Verschuren, W.M., Tolonen, H., Group, H.E.S.R.R., 2015. Sample selection, recruitment and participation rates in health examination surveys in Europe–experience from seven national surveys. BMC Med. Res. Methodol. 15, 78. https://doi.org/10.1186/s12874-015-0072-4.
National Research Council, 2006. Human Biomonitoring for Environmental Chemicals. The National Academies Press, Washington, DC. https://doi.org/10.17226/11700.
Perneger, T.V., Courvoisier, D.S., Hudelson, P.M., Gayet-Ageron, A., 2015. Sample size for pre-tests of questionnaires. Qual. Life Res. 24 (1), 147–151. https://doi.org/10.1007/s11136-014-0752-2.
Reynders, H., Colles, A., Morrens, B., Mampaey, M., Coertjens, D., Koppen, G., Schoeters, G., Loots, I., Chovanova, H., Winderickx, W., Van Campenhout, K., 2017. The added value of a surveillance human biomonitoring program: the case of FLEHS in Flanders (Belgium). Int. J. Hyg Environ. Health 220, 46–54. https://doi.org/10.1016/j.ijheh.2016.09.013.
Schoeters, G., Den Hond, E., Colles, A., Loots, I., Morrens, B., Keune, H., Bruckers, L., Nawrot, T., Sioen, I., De Coster, S., Van Larebeke, N., Nelen, V., Van de Mieroop, E., Vrijens, J., Croes, K., Goeyens, K., Baeyens, W., 2012. Concept of the Flemish human biomonitoring programme. Int. J. Hyg Environ. Health 215, 102–108. https://doi.org/10.1016/j.ijheh.2011.11.006.
Schulz, C., Conrad, A., Becker, K., Kolossa-Gehring, M., Seiwert, M., Seifert, B., 2007. Twenty years of the German Environmental Survey (GerES): human biomonitoring–temporal and spatial (West Germany/East Germany) differences in population exposure. Int. J. Hyg Environ. Health 210, 271–297. https://doi.org/10.1016/j.ijheh.2007.01.034.
Sobus, J.R., DeWoskin, R.S., Tan, Y.M., Pleil, J.D., Phillips, M.B., George, B.J., Christensen, K., Schreinemachers, D.M., Williams, M.A., Cohen Huba, E.A., Edwards, S.W., 2015. Uses of NHANES biomarker data for chemical risk assessment: trends, challenges, and opportunities. Environ. Health Perspect. 123 (10), 919–927. https://doi.org/10.1289/ehp.1409177.
Thomsen, M., Knudsen, L.E., Vorkamp, K., Frederiksen, M., Bach, H., Bonefeld-Jorgensen, E.C., Rastogi, S., Fauser, P., Krongaard, T., Sorensen, P.B., 2008. Conceptual framework for a Danish human biomonitoring program. Environ. Health 7 (Suppl. 1). https://doi.org/10.1186/1476-069X-7-S1-S3. S3.

Tobi, H., Kampen, J.K., 2018. Research design: the methodology for interdisciplinary research framework. Qual. Quantity 52 (3), 1209–1225. https://doi.org/10.1007/s11135-017-0513-8.

Tolonen, H., 2016. EHES Manual. Part A. Planning and Preparation of the Survey, second ed. National Institute for Health and Welfare, Directions 2016_13. URN, ISBN 978-952-302-700-8. URL: http://urn.fi/.

Tolonen, H., Ahonen, S., Jentoft, S., Kuulasmaa, K., Heldal, J., 2015. European health examination pilot project. Differences in participation rates and lessons learned about recruitment of participants - the European health examination survey pilot project. Scand. J. Publ. Health 43 (2), 212–219. https://doi.org/10.1177/1403494814565692.

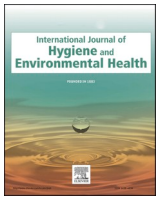
Weber, R.J., Cobaugh, D.J., 2008. Developing and executing an effective research plan. Am. J. Health Syst. Pharm. 65 (21), 2058–2065. https://doi.org/10.2146/ajhp070197.

International Journal of Hygiene and Environmental Health 232 (2021) 113682

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh



Cost effectiveness of community led total sanitation in Ethiopia and Ghana

Jonny Crocker [a,*], David Fuente [b], Jamie Bartram [c,d]

[a] Department of Global Health, University of Washington, 325 9th Avenue (Box 359931), Seattle, WA, 98104, USA
[b] School of Earth, Ocean and Environment, University of South Carolina, Columbia, SC, USA
[c] Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, USA
[d] University of Leeds, England

ARTICLE INFO

Keywords:
CLTS
Sanitation
Open defecation
Cost efficiency
Cost effectiveness
Economic

ABSTRACT

We conducted cost effectiveness analyses of four different CLTS interventions implemented in Ethiopia and Ghana. In each country, a pilot approach in which additional local actors were trained in CLTS facilitation was compared to the conventional approach. Data were collected using bottom-up costing, household surveys, and observations. We assessed variability of cost effectiveness from a societal perspective for latrine ownership and latrine use outcomes in different contexts. Cost effectiveness ranged from $34–$1897 per household ($5.85–$563 per person) gaining access to a private latrine or stopping open defecation, depending on the intervention, context, and outcome considered. For three out of four interventions, CLTS appeared more cost effective at reducing open defecation than at increasing latrine ownership, although sensitivity analysis revealed considerable variation. The pilot approaches were more cost effective at reducing open defecation than conventional approaches in Ethiopia, but not in Ghana. CLTS has been promoted as a low-cost means of improving the ownership and use of sanitation facilities. In our study, the cost of CLTS per household gaining latrine access was slightly higher than in other studies, and the cost of CLTS per household stopping OD was slightly lower than in other studies. Our results show that aggregate measures mask considerable variability in costs and outcomes, and thus the importance of considering and reporting context and uncertainty in economic analysis of sanitation interventions.

1. Introduction

In 2010, the United Nations recognized sanitation as a human right (United Nations General Assembly, 2010). The Sustainable Development Goals (SDGs) include behavioral and infrastructure targets for sanitation (UN General Assembly, 2015), that require 892 million people switching from open defecation (OD) to latrine use, and 2.3 billion people gaining access to a private latrine (WHO/UNICEF, 2017). Open defecation and lack of safe sanitation are associated with increased diarrhea and mortality, and negative social impacts among women and children (Bisung and Elliott, 2017; Prüss-Ustün et al., 2019; Wolf et al., 2018).

Interventions for changing sanitation behavior and increasing access to private latrines span from no-subsidy demand generation to latrine provision to microfinance. Economic analysis can be used to identify economically attractive development programs (Hutton et al., 2014; Trémolet et al., 2010; Whittington et al., 2012). Community-led total sanitation (CLTS) is an approach for generating demand for sanitation that has quickly reached wide-scale implementation in over 50 countries since it was introduced in 2000. The rapid expansion of CLTS preceded evidence generation, and was driven largely by perceptions that it was cheap and fast (Attanasio et al., 2004). Evidence on the effectiveness of CLTS is growing, with many recent published evaluations (USAID, 2018; Venkataramanan et al., 2018). These show that CLTS can reduce OD and increase access to private latrines (Pattanayak et al., 2009; Pickering et al., 2015); however, it can also be ineffective (Guiteras et al., 2015). Evidence on the cost and cost effectiveness of CLTS has not kept pace with evidence on effectiveness. There are only a few prior studies that report economic outcomes for CLTS or a related intervention (Briceño and Chase, 2015; Trémolet et al., 2010; Woode et al., 2018). A recent study modeling benefit-cost ratios for CLTS relied mostly on assumptions for estimating CLTS costs, due to the lack of empirical data available (Radin et al., 2019). They found that CLTS can be cost-beneficial in many but not all situations.

Evaluations of sanitation interventions tend to report aggregate effect estimates, even for multi-site studies. The impact of sanitation

* Corresponding author.
E-mail address: crockerj@uw.edu (J. Crocker).

https://doi.org/10.1016/j.ijheh.2020.113682
Received 7 April 2020; Received in revised form 29 October 2020; Accepted 4 December 2020
Available online 24 December 2020

interventions on latrine coverage, use, and health outcomes varies dramatically between studies, and the setting in which interventions occur can be a determinant of the outcomes (Garn et al., 2017; Schmidt, 2015). There have been several global economic analyses in the water, sanitation, and hygiene (WaSH) sector (Hutton et al., 2007). However, local context conditions the cost and outcomes of WaSH programs (Whittington, 2015). Thus, these global averages do not provide useful information about the relative performance of sanitation investments in specific contexts. Ideally, sanitation economic evaluations should study variation between sites, settings, and intervention delivery mechanisms (Kremer and Zwane, 2007). This would yield a more nuanced sense of where programs are likely to succeed or meet particular investment criteria.

This paper presents the results of a retrospective (ex-post) cost effectiveness analysis to highlight the importance of incorporating local context and uncertainty in economic analysis. We report the cost effectiveness of four different CLTS interventions in five regions in Ethiopia and Ghana. This study had two research questions: 1) what is the cost effectiveness of CLTS in reducing OD and increasing latrine ownership, and 2) to what extent does cost effectiveness vary between contexts and with uncertainty in key parameters. We examine costs from a societal perspective, report cost effectiveness for two different outcome measures (reductions in OD and increases in latrine ownership), and report how cost effectiveness varies between different regions. This study follows the CHEERS reporting guideline (Husereau et al., 2013).

2. Methods

2.1. Program description

Four different CLTS interventions were implemented: in Ethiopia, (1) health extension worker and kebele leader-facilitated CLTS ("HEW CLTS"), and (2) teacher-facilitated CLTS ("Teacher CLTS"); and in Ghana, (3) NGO-facilitated CLTS ("NGO CLTS"), and (4) NGO-facilitated CLTS with additional training for natural leaders ("CLTS + NL Training") (Table 1). A kebele is the lowest administrative unit in Ethiopia, comprising 20–30 villages and approximately 5000 people in rural areas. Natural leaders (NLs) are motivated community members who encourage others to construct latrines and change sanitation-related behaviors. Facilitation comprised visits to study villages by facilitators to conduct the three typical stages of CLTS: pre-triggering (community entry), triggering, and follow-up, which involves monitoring a community's progress and guiding them toward eliminating OD.

The two interventions in Ethiopia lasted 12 months, and the two in Ghana lasted 18 months. The interventions in Ethiopia took place in the Oromia and Southern Nations, Nationalities, and Peoples (SNNP) regions, and in Ghana in the Central, Upper West, and Volta regions. Interventions 1 and 2 in Ethiopia began with Plan International (Plan) training local actors who then led CLTS facilitation. Interventions 3 and 4 in Ghana were facilitated by Plan. Plan's interest in this study was to evaluate if pilot interventions (interventions 2 and 4) would improve the cost effectiveness of the conventional CLTS interventions in Ethiopia and Ghana, which were facing challenges as CLTS implementation was scaled up. There were several pre-existing enabling factors, including supportive national governments, local governments tasked with implementing CLTS, and a strong partnership between Plan and UNC (Saywell and Crocker, 2019). Further program details are available in implementation narratives and situational assessments on the project website (waterinstitute.unc.edu/clts).

Table 1
Differences between interventions analyzed.

Intervention	Facilitation
HEW CLTS (Ethiopia)	Health extension workers and kebele leaders
Teacher CLTS (Ethiopia)	Teachers
NGO CLTS (Ghana)	NGO staff
CLTS + NL Training (Ghana)	NGO staff and natural leaders

This table describes the main distinguishing factor between each intervention. Other differences are more thoroughly described in a previous publication (Crocker et al., 2017b).

2.2. Study design

Study design, data collection, rationale for village selection, and costing methods are described in previous publications (Crocker et al, 2016a, 2016b, 2017a, 2017b). Briefly, the study in Ethiopia used a quasi-experimental design, in which six kebeles (165 villages) were matched on latrine access and population, then manually assigned to receive CLTS facilitated by either HEWs and kebele leaders, or by teachers. The study in Ghana used a cluster-randomized design, in which all 58 project villages received CLTS, and 29 of the villages were randomly selected to receive NL training as an add-on activity. A survey of a representative sample of households (2182 in Ethiopia, 1594 in Ghana) was used to measure sanitation outcomes and household spending on latrines. Surveys in Ethiopia were administered immediately before implementation began, and again 2 years later (1-year after implementation ended). Surveys in Ghana were administered 2.5 years after implementation began (1-year after implementation ended). In Ghana, household recall on latrine age was used to estimate baseline coverage. Surveys covered demographics and WaSH indicators, and included latrine observations. All data collection was conducted in local languages by an experienced independent contractor. Printed surveys were used in Ethiopia, and SurveyCTO software on Nexus tablets was used in Ghana.

Costs were assessed using a bottom-up, activity based costing method. Costs were measured (rather than estimated). Plan used checklists to track their implementation activities in detail (including management, training, and facilitation). Financial expenditures and receipts were reviewed to generate unit costs (e.g. for training venue rental, transportation, and person-time). Household surveys were used to collect data on community members' time and money spent on sanitation during the study period. Costs were categorized as program cost (management, training, facilitation), and local costs (local actor time, community member time, and latrine spending). The management category includes time spent planning and preparing for the interventions. Training and facilitation include transportation costs. Local actors' unpaid time was monetized using value-of-time estimates generated through review of local wages and survey results. No discount rate was used as costs were evaluated over a 2-year period (2012–2014). Ethical clearance was obtained from the UNC Institutional Review Board, and from the appropriate institutions in Ethiopia and Ghana. Informed consent was obtained from all study participants.

2.3. Analysis

Cost effectiveness of the CLTS interventions was calculated using two outcome measures (reduced OD, and increased latrine ownership). Latrine use included use of communal, shared, and private latrines. Self-reported private latrine use was validated by observing latrines. Households who reported using latrines that were full, unstable, or could not be observed were categorized as OD. Similarly, households were categorized as owning a latrine only if it had stable flooring and was not full on the day of surveying. The difference between self-reported behaviors and observation of a usable latrine are reported in a prior publication (Crocker et al., 2017a). Others have also pointed out that self-reported latrine use is an imperfect proxy for latrine use, although it is used for measuring the SDG indicators (Thomas et al., 2018).

The assessment of cost effectiveness of the CLTS interventions takes a

societal perspective by including both program and local costs (Fuente et al., 2012). Program costs were those borne by Plan, which included management, training, and facilitation. Program costs were assessed through implementation tracking of Plan's CLTS interventions, and review of financial expenditures. Local costs were those borne by local actors (district government, teachers, health workers, community members), comprising the economic value of their time, and households' financial expenditures on latrines. Local costs were assessed through a combination of Plan's implementation tracking, and through household and local actor surveys.

To arrive at cost effectiveness, program and local costs were divided by the number of households stopping OD or gaining latrine ownership. Cost effectiveness values should be interpreted as the societal cost to convert a household to latrine use or latrine ownership.

$$CE_i = \text{Cost} / \text{Outcome}_i$$

Where i = OD or latrine ownership. Sensitivity of cost effectiveness to uncertainty about the costs and outcomes of the programs was assessed using Monte Carlo analysis, using assumptions summarized in the supplement. Costs were measured, not estimated, so a uniform distribution with cost range of ± 30% from base values presented in Table 1 was chosen as a conservative approach for sensitivity analysis. For outcomes, normal distributions using standard deviations taken from impact evaluation data were used. Standard deviations for changes in open defecation in Ghana were not available, so 50% of the point estimates were used for the Monte Carlo analysis as a conservative approach, which is a larger range than used in previous studies (Radin et al., 2019; Whittington et al., 2012). Monte Carlo analysis was implemented using Oracle Crystal Ball and 1000 draws of the parameter values.

3. Results

3.1. Study population

Large differences existed between the Ethiopia and Ghana study populations. The Ethiopia study population had more people per household, indications of lower wealth, smaller villages, lower access to an improved water supply, and higher latrine ownership than the Ghana study population. Within both countries, there were regional differences; the Oromia region in Ethiopia and the Upper West region in Ghana showed larger households, and lower wealth, access to improved water supply, and latrine ownership (Table 2). A breakdown of these variables by intervention is in the appendix.

3.2. Cost effectiveness

Table 3 presents cost and outcome data and cost effectiveness for the four CLTS interventions by region for two outcome measures. For the full study population across both Ethiopia and Ghana in aggregate, CLTS cost $358.87 per household stopping open defecation, and $530.63 per household gaining ownership of a usable latrine. This equates to $75.27 *per person* stopping open defecation, and $111.29 *per person* gaining ownership of a usable latrine (data not shown).

In Ethiopia, the pilot intervention was more cost effective than conventional CLTS for both outcomes (Table 3, comparing column B to C). However, in Ghana, the pilot intervention was *less* cost effective than conventional CLTS for both outcomes (comparing column H to I). Both of the CLTS interventions in Ethiopia were more cost effective than those in Ghana (comparing columns B and C to H and I). For example, the two interventions in Ethiopia cost $218.26 and $150.65 per household stopping OD, while the two interventions in Ghana cost $442.26 and $479.11 per household stopping OD. In most cases, it was cheaper to reduce OD than it was to increase latrine ownership, with conventional CLTS in the Central and Volta regions being the exceptions (comparing the last two rows). This was due to some of the reduction in OD coming from latrine sharing, as discussed in previous publications (Crocker et al., 2016a, 2016b). Incremental cost effectiveness ratios are presented in Table A3 in the supplement.

In Ethiopia, both CLTS interventions were more cost effective in the Oromia region. In the SNNP region, open defecation and latrine ownership did not improve, which would yield negative cost effectiveness ratios so they are not presented in Table 3. In Ghana, the CLTS interventions were more cost effective at reducing OD in the Upper West region, and more cost effective at increasing latrine ownership in the Volta region where latrine ownership increased more than latrine use (columns L, M, N, O).

Fig. 1 summarizes the Monte Carlo sensitivity analysis for cost effectiveness of the CLTS interventions in Ghana and Ethiopia for two outcome measures (OD and latrine ownership). Cost effectiveness was sensitive to uncertainty about the costs and outcomes of the relative interventions. For example, the 95% confidence intervals from the Monte Carlo simulations do not indicate that one intervention was more cost effective than the other for either of the outcome measures.

4. Discussion

Since its first introduction in Bangladesh in 1999, CLTS has been

Table 2
Household and respondent characteristics in villages receiving CLTS in Ethiopia and Ghana, by region.

Variable	Ethiopia			Ghana			
	All regions	Oromia	SNNP	All regions	Central	Upper West	Volta
Female respondent	76%	74%	77%	71%	68%	95%	62%
Five or more years of education	18%	13%	21%	42%	57%	15%	75%
Household size	5.8	6.1	5.6	4.9	3.3	6.4	3.6
Number of children per household	0.9	1.0	0.9	0.6	0.5	1.1	0.5
Metal roof	21%	11%	29%	91%	97%	72%	94%
Own radio	26%	30%	24%	49%	43%	49%	55%
Own television	1%	0%	1%	37%	46%	19%	38%
Years family lived in village	22	25	21	28	27	36	25
Years family lived in current house	14	6	18	14	14	16	15
Access to improved water supply	21%	4%	33%	59%	65%	41%	60%
Baseline latrine ownership	78%	51%	98%	21%	24%	10%	25%
Baseline open defecation	45%	70%	27%	49%[c]	34%[c]	96%[c]	36%[c]
Village size	34	29	38	116	164	68	123
Villages with prior WaSH projects	0%	0%	0%	72%	100%	45%	79%
Villages with prior subsidized latrines	0%	0%	0%	28%	33%	15%	37%

All Ghana values are taken from the 1.5-year follow up household census and survey, and describe the two treatment groups at that time, except for baseline private latrine ownership, which is based on recall of how old their latrines were. The education variable assumes that respondents who have completed primary education in Ghana have spent at least five years in education. Baseline surveys were not used in Ghana, so baseline open defecation was based on the conservative assumption that decreases in open defecation were equivalent to increases in latrine ownership. "Access" meaning the improved water supply is within 30 min walking distance.

Table 3
Cost effectiveness of four CLTS interventions in Ethiopia and Ghana.

	Variable	Both countries all regions	Combined Ethiopia regions		Oromia region		SNNP region		Combined Ghana regions		Central region		Upper West region	Volta region		
		All CLTS interventions	CLTS	Teacher-CLTS	CLTS	Teacher-CLTS	CLTS	Teacher-CLTS	CLTS	CLTS + natural leader training	CLTS	CLTS + natural leader training	CLTS	CLTS + natural leader training	CLTS	CLTS + natural leader training
	Column ID	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**
	Households	12,217	1624	3838	651	1586	973	2252	3443	3312	1463	1495	808	540	1172	1277
	People	58,248	9829	21,216	4013	9153	5816	12,063	14,263	12,940	4670	5041	5208	3474	4385	4425
Outcomes	Change in open defecation	−13%	−10%	−11%	−48%	−56%	15%	21%	−9%	−22%	−6%	−16%	−13%	−53%	−9%	−15%
	Change in usable latrine ownership	9%	4%	6%	43%	53%	−22%	−27%	9%	15%	7%	4%	8%	33%	12%	19%
	Households stopping open defecation	1605	168	421	310	895	−141	−474	298	718	89	237	108	287	100	195
	Households gaining ownership of a usable latrine	1086	70	229	282	841	−211	−612	305	481	102	58	634	179	139	243
Costs	Management cost	$79,660	$8690	$14,867	$4345	$7434	$4345	$7434	$26,958	$29,145	$8797	$9525	$9308	$10,037	$8853	$9582
	Training cost	$210,850	$17,617	$33,229	$9979	$17,279	$7638	$15,951	$4076	$155,928	$1199	$38,427	$928	$62,874	$1949	$54,627
	Facilitation cost	$169,595	$4891	$6225	$2820	$1118	$2071	$5107	$73,428	$85,052	$28,573	$32,281	$18,270	$23,209	$26,584	$29,562
	Local actor time cost	$12,583	$1926	$3863	$1084	$1889	$842	$1974	$1242	$5552	$314	$1525	$372	$1998	$555	$2030
	Community time cost	$21,529	$2546	$4151	$1435	$1886	$1110	$2265	$5847	$8985	$1502	$3661	$1452	$2028	$2893	$3297
	Hired labor cost	$21,942	$0	$0	$0	$0	$0	$0	$7132	$14,810	$1096	$1835	$2191	$3859	$3845	$9115
	Hardware cost	$59,901	$1070	$1013	$911	$633	$159	$380	$13,101	$44,718	$4319	$23,501	$1282	$2452	$7501	$18,765
	Total cost	$576,061	$36,739	$63,348	$20,573	$30,239	$16,166	$33,110	$131,783	$344,190	$45,800	$110,754	$33,803	$106,458	$52,180	$126,978
Cost effectiveness	**Cost per household stopping open defecation**	**$358.87**	$218.26	$150.65	$66.44	$33.79	*NA*		$442.26	$479.11	$512.01	$467.50	$311.86	$371.43	$521.12	$651.62
	Cost per household gaining usable latrine ownership	**$530.63**	$522.02	$276.53	$73.03	$35.95			$432.11	$715.29	$447.24	$1897.01	$531.31	$593.22	$375.54	$521.80

Sanitation outcomes worsened in the SNNP region, yielding misleading negative cost effectiveness ratios, so they are not presented in this table. Changes in open defecation and latrine ownership are for a 2.5-year period for Ghana, and a 2-year period for Ethiopia (baseline to 1-year after the interventions ended). Cost effectiveness estimates assume open defecation would not have changed in the absence of the CLTS interventions. Management, training, and facilitation costs were borne by Plan. Local actor and community time are economic costs. Hired labor and hardware costs were borne by households. "Usable" latrines were not full and had stable flooring on the day of surveying. Changes in open defecation and latrine ownership in Ghana are based on recall of latrine age. A previous publication explores variation in cost within and between countries(Crocker et al., 2017b)

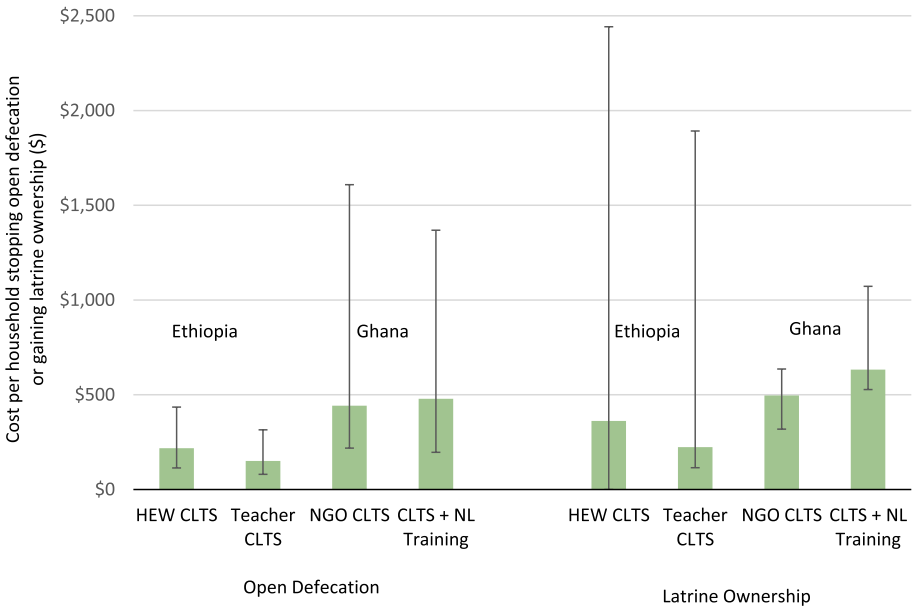


Fig. 1. Cost effectiveness of four different CLTS interventions per household for two outcome measures (stopping open defecation or gaining ownership of a private latrine) in Ethiopia and Ghana. Error bars represent 95% confidence intervals from Monte Carlo simulations. Lower-bound confidence interval for latrine ownership results in Ethiopia are truncated at zero.

promoted as a low cost means of improving the ownership and use of improved sanitation facilities (Kar and Chambers, 2008). In our study, the cost of CLTS per household gaining latrine access was $530.63 in aggregate, which is slightly higher than related interventions. An evaluation of a World-Bank funded program in Tanzania reported the cost of gaining access to a latrine as $194 per household for a sanitation-promotion program, and $491 for a sanitation and hygiene promotion program (Briceño and Chase, 2015). Another study reported a ratio of $344 per household gaining latrine access in Senegal (Trémolet et al., 2010). Our study found the cost of CLTS per household stopping OD was $358.87 in aggregate. We are not aware of previous studies reporting the cost per household stopping OD for a CLTS intervention. A study of a small hygiene promotion program in four villages in Ghana saw costs per household changing any hygiene behavior (latrine use, handwashing, or safe water use) of $532, which is higher than the cost per household stopping OD in our study (Woode et al., 2018). They used improvement in any of three WaSH behaviors as their outcome, making direct comparison less conclusive. Care should be taken when interpreting these direct comparisons between studies, due to methodological differences. For example, two of the three cited studies used top-down costing, which may result in underestimated costs and lower cost effectiveness ratios.

The pilot CLTS interventions yielded improved sanitation outcomes over conventional CLTS in both Ethiopia and Ghana—a greater portion of households within project communities stopped open defecation and gained ownership of a usable private latrine. Additionally, pilot CLTS was more cost effective in aggregate than the conventional interventions at reducing OD and increasing latrine ownership in Ethiopia. For the full sample (all regions and interventions; Table 3, column A), OD reduced by 13% while latrine use increased by 9%, indicating that CLTS was 44% more cost effective at reducing OD than increasing latrine use. This may indicate that for CLTS, moving households up the sanitation ladder (WHO/UNICEF, 2017) is cheaper for the lowest rung on the ladder (stopping OD) than for higher rungs that require owning a private latrine.

Nevertheless, there was considerable variability in the cost effectiveness of the interventions, depending on intervention, country, region, and outcome measure. For example, in Ethiopia neither intervention improved sanitation outcomes in the SNNP region. This yields an aggregate cost effectiveness in Ethiopia of $151–$218 per household stopping OD, compared to a much more attractive cost effectiveness of $34–$66 per household stopping OD when considering the Oromia region alone. NL training in Ghana was more cost effective than conventional CLTS in the Central region using OD as an outcome measure. NL training was less cost effective than conventional CLTS in the Upper West and Volta regions. In an ex-ante cost effectiveness analysis, these results could be interpreted as a justification for implementing the NL intervention in the Central region and the conventional CLTS intervention in the Upper West and Volta regions. However, both latrine ownership and use outcomes were substantially (4X) higher with the inclusion of NL training in the Upper West region. CLTS outcomes have varied with implementation differences elsewhere as well (Harter et al., 2019). The higher spending in more intensive CLTS programs that include training more people within villages can improve latrine ownership and use within those villages and likely justify the costs, especially given emerging evidence that high neighborhood-level sanitation coverage is important for health impacts (Fuller and Eisenberg, 2016; Harris et al., 2017).

Cost effectiveness differed between Ethiopia and Ghana. This was largely driven by differences in costs, not in outcomes. It is also important to note that the quality of latrines resulting from CLTS interventions tended to be higher in Ghana (Crocker et al., 2017a). We do not think our results would justify prioritizing investing in CLTS in one country over another. We instead compare cost effectiveness between Ethiopia and Ghana to discourage others from directly extrapolating cost or cost effectiveness results between countries. We are not aware of any prior studies that report empirical economic data from the same sanitation interventions in multiple countries.

Cost effectiveness versus just effectiveness provide contrasting results in Ghana. The conventional CLTS intervention was more cost effectiveness, in aggregate and in each region, than the NL intervention in Ghana. However, the NL training had a greater impact on latrine ownership and OD in aggregate and in the Upper West and Volta regions than conventional CLTS, but that increase in outcomes came at a cost. The NL intervention had a greater impact particularly on reductions in OD through stimulating greater latrine sharing. Both of these outcomes were assessed 1-year after the interventions ended, which does not necessarily indicate longer-term sustainability of the outcomes. This

highlights the fact that for interventions that yield multiple outcomes, the results of cost effectiveness analyses are sensitive to the choice of the outcome measure (Boardman et al., 2010). Our study thus reinforces the limitations of using cost effectiveness analyses to guide decision making. Given the importance of high sanitation for health impact, and that sanitation has been recognized as a human right, the most cost effective intervention may not be preferred by policy makers if a less cost effective intervention yields higher coverage. Cost effectiveness analysis can reduce the uncertainty of the consequences of a given decision, but cannot reflect the values of those that will be affected by the decision and the constraints of those making the decision (Wilkinson et al., 2016). Thus, cost effectiveness can serve as a tool but not the sole determinant for decision making.

In addition to variability observed within and across countries, the relative cost effectiveness of the interventions was sensitive to uncertainty about the costs and outcomes of the respective interventions. The Monte Carlo simulations show considerable overlap in the 95% confidence intervals for cost effectiveness of the respective CLTS interventions for both outcome measures. Thus, in addition to considering contextual factors that may influence outcomes, decision makers must be mindful of uncertainty about both costs and the effectiveness of interventions. In situations where costs, outcomes, or both are subject to considerable uncertainty, decision makers should be cautious when using point estimates to guide decision making and consider the distribution of potential outcomes.

4.1. Limitations

This study has a number of limitations. Cost effectiveness values are from Ghana and Ethiopia, and may not be generalizable to other contexts. Changes in open defecation and latrine ownership in Ghana are based on recall of latrine age. Outcomes were measured at the household rather than individual level, which does not account for intra-household variation. Households reporting using a latrine observed to be unusable were categorized as OD, which may overestimate OD. Local costs incorporate assumptions about value-of-time. For further details and limitations of the study designs and data collection, refer to prior publications that go into more detail on measuring outcomes and cost (Crocker et al, 2016a, 2016b, 2017b).

4.2. Conclusion

There is limited cost effectiveness evidence for sanitation interventions, and the existing evidence generally uses top-down costing methods, and reports aggregate outcomes. We contribute to filling this evidence gap by performing cost effectiveness analysis of CLTS interventions in Ethiopia and Ghana using bottom-up costing, and report results disaggregated by intervention, outcome, and geographical region.

Both the pilot interventions (Teacher CLTS and CLTS + NL Training) involved training more local actors than was conventional in Ethiopia and Ghana. In Ethiopia, the pilot intervention was 30% more cost effective at reducing open defecation. In Ghana, the pilot intervention was 8% less cost effective at reducing open defecation, but had two-to three-times more impact. The NL intervention may be more appealing to policy-makers as reaching high community-level coverage and use of sanitation is both important and challenging (Fuller and Eisenberg, 2016; Garn et al., 2017; Harris et al., 2017). We conclude that the extra cost of building local capacity in the pilot interventions was justified by improved outcomes, and improved effectiveness for the primary outcome targeted by the approach (defecation practices).

Our results demonstrate the extent to which aggregate measures can mask considerable variability in costs and outcomes across countries and regions. We demonstrate the need for economic analyses that attend to local contextual factors and suggest that exploring variability of economic results is more valuable than considering solely aggregate results (Jeuland and Pattanayak, 2012). Instead of reporting just aggregate results, studies should report outcomes for different geographic regions within the study, and should report demographic characteristics of study participants by region as well. Point estimates of economic measures need to be interpreted with care. Both the costs and outcomes of interventions are subject to variability and uncertainty which should be reflected in economic analysis.

Cost effectiveness analysis requires the analyst to choose a single outcome measure. The focus on individual outcomes inhibits comparing multiple potential outcomes that might be considered (Berman, 1982; Briscoe, 1984; Fuente et al., 2012). For CLTS, in addition to the two outcomes we present, this includes handwashing practices, solid waste management, time savings from stopping OD, and improvements in well-being associated with enhanced social status and safety for women. In such instances, cost effectiveness does not provide sufficient guidance to policy makers on how to proceed. While cost effectiveness analysis can and should be used to inform public health policy and programming, it should not be treated as a standalone tool for determining resource allocation.

Acknowledgements

This research is a result of a subagreement to the University of North Carolina at Chapel Hill from Plan International USA, which received a grant from the Bill & Melinda Gates Foundation (OPP1028953). The authors would like to thank: the study participants; Plan International USA for grant management; and Plan International Ghana and Ethiopia for implementation. The authors declare no competing financial interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2020.113682.

References

Attanasio, O.P., Meghir, C., Szekely, M., 2004. Using randomized experiments and structural models for scaling up: evidence from the PROGRESA evaluation. In: Bourguignon, F., Pleskovic, B. (Eds.), 16th Annual World Bank Conference on Development Economics (ABCDE): Accelerating Development. The World Bank, Washington, DC, pp. 315–339. https://doi.org/10.1596/0-8213-5388-8.

Berman, P.A., 1982. Selective primary health care: is efficient sufficient? Soc. Sci. Med. 16, 1054–1059.

Bisung, E., Elliott, S.J., 2017. Psychosocial impacts of the lack of access to water and sanitation in low- and middle-income countries: a scoping review. J. Water Health 15, 17–30. https://doi.org/10.2166/wh.2016.158.

Boardman, A.E., Greenberg, D.H., Vining, A.R., Weimer, D.L., 2010. Cost-benefit Analysis: Concepts and Practice, fourth ed. Prentice Hall, Upper Saddle River, NJ.

Briceño, B., Chase, C., 2015. Cost-efficiency of rural sanitation promotion: activity-based costing and experimental evidence from Tanzania. J. Dev. Effect. 1–12 https://doi.org/10.1080/19439342.2015.1105848.

Briscoe, J., 1984. Water supply and health in developing countries: selective primary health care revisited. Am. J. Publ. Health 74, 1009.

Crocker, J., Abodoo, E., Asamani, D., Domapielle, W., Gyapong, B., Bartram, J., 2016a. Impact evaluation of training natural leaders during a community-led total sanitation intervention: a cluster-randomized field trial in Ghana. Environ. Sci. Technol. 50, 8867–8875. https://doi.org/10.1021/acs.est.6b01557.

Crocker, J., Geremew, A., Atalie, F., Yetie, M., Bartram, J., 2016b. Teachers and sanitation promotion: an assessment of community-led total sanitation in Ethiopia. Environ. Sci. Technol. 50, 6517–6525. https://doi.org/10.1021/acs.est.6b01021.

Crocker, J., Saywell, D., Bartram, J., 2017a. Sustainability of community-led total sanitation outcomes: evidence from Ethiopia and Ghana. Int. J. Hyg Environ. Health 220, 551–557. https://doi.org/10.1016/j.ijheh.2017.02.011.

Crocker, J., Saywell, D., Shields, K.F., Kolsky, P., Bartram, J., 2017b. The true costs of participatory sanitation: evidence from community-led total sanitation studies in Ghana and Ethiopia. Sci. Total Environ. 601–602, 1075–1083. https://doi.org/10.1016/j.scitotenv.2017.05.279.

Fuente, D., Ozdemir, S., Allaire, M., Whittington, D., 2012. Application of Cost Effectiveness Analysis in the Water and Sanitation Sector.

Fuller, J.A., Eisenberg, J.N.S., 2016. Herd protection from drinking water, sanitation, and hygiene interventions. Am. J. Trop. Med. Hyg. 95, 1201–1210. https://doi.org/10.4269/ajtmh.15-0677.

Garn, J.V., Sclar, G.D., Freeman, M.C., Penakalapati, G., Alexander, K.T., Brooks, P., Rehfuess, E.A., Boisson, S., Medlicott, K.O., Clasen, T.F., 2017. The impact of

sanitation interventions on latrine coverage and latrine use: a systematic review and meta analysis. Int. J. Hyg Environ. Health 220, 329–340. https://doi.org/10.1016/j.ijheh.2016.10.001.

Guiteras, R., Levinsohn, J., Mobarak, A.M., 2015. Encouraging sanitation investment in the developing world: a cluster-randomized trial. Science 348 (80-), 903–906. https://doi.org/10.1126/science.aaa0491.

Harris, M., Alzua, M.L., Osbert, N., Pickering, A.J., 2017. Community-level sanitation coverage more strongly associated with child growth and household drinking water quality than access to a private toilet in rural Mali. Environ. Sci. Technol. 51, 7219–7227. https://doi.org/10.1021/acs.est.7b00178.

Harter, M., Lilje, J., Mosler, H.-J., 2019. Role of implementation factors for the success of community-led total sanitation on latrine coverage. A case study from rural Ghana. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.9b01055.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., Augustovski, F., Briggs, A.H., Mauskopf, J., Loder, E., 2013. Consolidated health economic evaluation reporting standards (CHEERS) statement. Cost Eff. Resour. Allocation 11.

Hutton, G., Haller, L., Bartram, J., 2007. Global cost-benefit analysis of water supply and sanitation interventions. J. Water Health 5, 481–502. https://doi.org/10.2166/wh.2007.009.

Hutton, G., Rodriguez, U.-P., Winara, A., Anh, N.V., Phyrum, K., Chuan, L., Blackett, I., Weitz, A., 2014. Economic efficiency of sanitation interventions in southeast Asia. J. Water, Sanit. Hyg. Dev. 4, 23. https://doi.org/10.2166/washdev.2013.158.

Jeuland, M.A., Pattanayak, S.K., 2012. Benefits and costs of improved cookstoves: assessing the implications of variability in health, forest and climate impacts. PloS One 7. https://doi.org/10.1371/journal.pone.0030338 e30338.

Kar, K., Chambers, R., 2008. Handbook on Community-Led Total Sanitation. Plan UK and Institute of Development Studies.

Kremer, M., Zwane, A.P., 2007. Cost-effective Prevention of Diarrheal Diseases: a Critical Review. No. 117.

Pattanayak, S.K., Yang, J.-C., Dickinson, K.L., Poulos, C., Patil, S.R., Mallick, R.K., Blitstein, J.L., Praharaj, P., 2009. Shame or subsidy revisited: social mobilization for sanitation in Orissa, India. Bull. World Health Organ. 87, 580–587. https://doi.org/10.2471/BLT.08.057422.

Pickering, A.J., Djebbari, H., Lopez, C., Coulibaly, M., Alzua, M.L., 2015. Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. Lancet Glob. Heal. 3, e701–e711. https://doi.org/10.1016/S2214-109X(15)00144-8.

Prüss-Ustün, A., Wolf, J., Bartram, J., Clasen, T.F., Cumming, O., Freeman, M.C., Gordon, B., Hunter, P.R., Medlicott, K., Johnston, R., 2019. Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: an updated analysis with a focus on low- and middle-income countries. Int. J. Hyg Environ. Health. https://doi.org/10.1016/j.ijheh.2019.05.004.

Radin, M., Jeuland, M., Wang, H., Whittington, D., 2019. Benefit-cost Analysis of Community-Led Total Sanitation: Incorporating Results from Recent Evaluations (No. 13). Guidelines for Benefit-Cost Analysis Project.

Saywell, D., Crocker, J., 2019. Process learning on partnerships: building functioning research and practice organizational relationships. Waterlines 38, 3–19. https://doi.org/10.3362/1756-3488.18-00014.

Schmidt, W.-P., 2015. Comment: seven trials, seven question marks. Lancet Glob. Heal. 3, e659–e660. https://doi.org/10.1016/S2214-109X(15)00182-5.

Thomas, E., Andrés, L.A., Borja-Vega, C., Sturzenegger, G. (Eds.), 2018. Innovations in WASH Impact Measures. World Bank, Washington, DC. https://doi.org/10.1596/978-1-4648-1197-5.

Trémolet, S., Kolsky, P., Perez, E., 2010. Financing On-Site Sanitation for the Poor: a Six Country Comparative Review and Analysis. Washington, DC.

UN General Assembly, 2015. Draft outcome document of the United Nations summit for the adoption of the post-2015 development agenda. https://doi.org/10.1093/oxfordhb/9780199560103.003.0005.

United Nations General Assembly, 2010. Resolution 64/292. The Human Right to Water and Sanitation.

USAID, 2018. An Examination of CLTS's Contributions toward Universal Sanitation.

Venkataramanan, V., Crocker, J., Karon, A., Bartram, J., 2018. Community-led total sanitation: a Mixed-methods systematic review of evidence and its quality. Environ. Health Perspect. 126 https://doi.org/10.1289/EHP1965.

Whittington, D., 2015. Water and Sanitation Perspective Paper: Benefits and Costs of the Water and Sanitation Targets for the Post-2015 Development Agenda. Copenhagen. Post-2015 Consensus.

Whittington, D., Jeuland, M.A., Barker, K., Yuen, Y., 2012. Setting priorities, targeting subsides among water, sanitation, and preventative health interventions in developing countries. World Dev. 40, 1546–1568. https://doi.org/10.1016/j.worlddev.2012.03.004.

WHO/UNICEF, 2017. Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baselines. Geneva.

Wilkinson, T., Sculpher, M.J., Claxton, K., Revill, P., Briggs, A., Cairns, J.A., Teerawattananon, Y., Asfaw, E., Lopert, R., Culyer, A.J., Walker, D.G., 2016. The international decision support initiative reference case for economic evaluation: an aid to thought. Value Health 19, 921–928. https://doi.org/10.1016/j.jval.2016.04.015.

Wolf, J., Hunter, P.R., Freeman, M.C., Cumming, O., Clasen, T.F., Bartram, J., Higgins, J. P.T., Johnston, R., Medlicott, K., Boisson, S., Prüss-Ustün, A., 2018. Impact of drinking water, sanitation and hand washing with soap on childhood diarrhoeal disease: updated meta-analysis and regression. Trop. Med. Int. Health 23, 508–525. https://doi.org/10.1111/tmi.13051.

Woode, P.K., Dwumfour-Asare, B., Nyarko, K.B., Appiah-Effah, E., 2018. Cost and effectiveness of water, sanitation and hygiene promotion intervention in Ghana: the case of four communities in the Brong Ahafo region. Heliyon 4. https://doi.org/10.1016/j.heliyon.2018.e00841 e00841.

International Journal of Hygiene and Environmental Health

doi: 10.1016/S1438-4639(21)00012-2

International Journal of Hygiene and Environmental Health 232 (2021) 113668

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

Risk assessment for irritating chemicals – Derivation of extrapolation factors

Inge Mangelsdorf [a,b,*], Katrin Schröder [b], Sylvia E. Escher [b], Marike Kolossa-Gehring [c], Malgorzata Debiak [c]

[a] Toxicology Consulting, Hamburg, Germany
[b] Fraunhofer Institute for Toxicology and Experimental Medicine, Hannover, Germany
[c] German Environment Agency, Berlin, Germany

ARTICLE INFO

Keywords:
Indoor air
Sensory irritation
Extrapolation factor
Risk assessment
Alarie test
Controlled human exposure studies

ABSTRACT

Irritation of the eyes and the upper respiratory tract are important endpoints for setting guide values for chemicals. To optimize the use of the often-limited data, we analysed controlled human exposure studies (CHS) with 1–4 h inhalation of the test substance, repeated dose inhalation studies in rodents, and Alarie-Tests and derived extrapolation factors (EF) for exposure duration, inter- and intraspecies differences.

For the endpoint irritating effects in the respiratory tract in rodents, geometric mean (GM) values of 1.9 were obtained for the EF for subacute→subchronic (n = 16), 2.1 for subchronic→chronic (n = 40), and 2.9 for subacute→chronic (n = 10) extrapolation. Based on these data we suggest an EF of 2 for subchronic→chronic and of 4 for subacute→chronic extrapolation. In CHS, exposure concentration determines the effects rather than exposure duration. Slight reversible effects during 4 h exposure indicate that an EF of 1 can be considered for assessing chronic exposures.

To assess species extrapolation, 10 chemicals were identified with both, reliable rat inhalation studies and CHS. The GM of the ratio between the No Observed Adverse Effect Concentration (NOAEC) in rats and humans was 2.3 and increased to 3.6 when expanding the dataset to all available EF (n = 25). Based on these analyses, an EF of 3 is suggested to extrapolate from a NOAEC in a chronic rat study to a NOAEC in a CHS.

The analysis of EFs for the extrapolation from a 50% decrease in respiratory frequency in the Alarie test in mice (RD_{50}) to a NOAEC in a CHS resulted in a GM of 40, for both, the reliable (n = 11) and the overall dataset (n = 19). We propose to use the RD_{50} from the Alarie test for setting guide values and to use 40 as EF.

Efs for intraspecies differences in the human population must account for susceptible persons, most importantly for persons with chemical intolerance (CI), who show subjective signs of irritation at low concentrations. The limited data available do not justify to deviate from an EF of 10 – 20 as currently used in different regulatory settings.

1. Introduction

Sensory irritation of the eyes or the upper respiratory tract is often the cause of complaints about quality of indoor air, and the critical effect, when e.g. the German Committee for Indoor Air Guide Values derives guide values for indoor air (Fromme et al., 2019), or when occupational exposure levels are established (Brüning et al., 2014; Kuwabara et al., 2007).

1.1. Cytotoxic and sensory irritation

Generally, two types of irritation are distinguished: sensory irritation and cytotoxic irritation.

Sensory irritation is mediated mainly by the TRPA1 receptor on free nerve endings of the trigeminal nerve, which extends to the epithelia of the upper respiratory tract and the eye (Bessac and Jordt, 2010; Doty et al., 2004). A wide variety of chemicals or their metabolites are able to bind to TRPA1, e.g. acrolein, formaldehyde, hydrogen peroxide, naphthalene, and styrene (Achanta and Jordt, 2017; Bautista et al., 2006;

* Corresponding author. Toxicology Consulting, Hamburg, Germany.
E-mail address: toxconsult@posteo.de (I. Mangelsdorf).

https://doi.org/10.1016/j.ijheh.2020.113668
Received 28 September 2020; Received in revised form 19 November 2020; Accepted 24 November 2020
Available online 14 December 2020

Bessac and Jordt, 2008; Lanosa et al., 2010). Receptor activation results in an influx of calcium into the nerve cells and release of mediators such as substance P or calcitonin gene-related peptide (Bascom et al., 1997; Bessac and Jordt, 2008, 2010), which then initiate e.g. vasodilation, leukocyte infiltration or mast cell degranulation i.e. neurogenic inflammation (Bautista et al., 2013). TRPA1 knockout mice do not exhibit sensory irritation, when exposed to styrene or naphthalene (Lanosa et al., 2010). In humans, activation of the nerve cells causes pungent, tingling, prickling, stinging, and burning sensations (Bessac and Jordt, 2008; Hummel et al., 2003). In addition, physiological responses are observed such as lacrimation, eye blinking, reduced respiratory rate, and reduced nasal flow. All these events have been summarized as an adverse outcome pathway for sensory irritation with binding to TRPA1 as molecular initiating event (Martinez and Eling, 2019).

Sensory irritation can be assessed in CHS, where volunteers are exposed for up to 4 h to low levels of potentially irritating chemicals and subjective symptoms of nose and eye irritation as well as objective findings like eye blinking rate, tear film stability or secretion of mediators in nasal lavage fluid are recorded. Criteria for adequate performance of these studies are described by Nielsen and Wolkoff (2017). The review by (Doty et al., 2004) describes these endpoints and provides their strengths and limitations.

An animal test that detects sensory irritation, is the Alarie test with mice. Exposure to sensory irritants and subsequent binding to the TRPA1 receptor causes a vagal-mediated reduction in respiration rate. The concentration of a chemical that reduces the respiratory rate to 50 % (RD_{50}) is determined as measure of sensory irritation (Alarie, 1981a). Results of Alarie tests are available for a wide number of chemicals (Alarie, 2016).

Recently also an *in vitro* assay has been developed with fibroblasts stably transfected with the TRPA1 receptor. Ca influx as measure of TRPA1 binding of a chemical correlated well with results from the Alarie test, thus offering a promising approach for assessing sensory irritation with *in vitro* methods in the future (Martinez and Eling, 2019).

In contrast to sensory irritation, cytotoxic irritation occurs at higher concentrations. Eye irritation is characterized by redness of the conjunctivae, oedema, and necrosis. In the nose, cytotoxic irritation involves e.g. cell proliferation, metaplasia, necrosis, cell death and tumours (Harkema et al., 2006). These effects can be observed in histopathological examinations in subacute, subchronic or chronic toxicity studies with rodents.

1.2. Risk assessment for irritating chemicals

There are several publications that propose approaches for risk assessment of irritating chemicals and propose EFs, which involve exposure duration (ECETOC, 2010; Kalberlah et al., 2002), adjustment to continuous exposure (Shusterman et al., 2006), species extrapolation from rodent studies to controlled human exposure studies (Brüning et al., 2014), and species extrapolation from the Alarie-Test with mice (Kuwabara et al., 2007). Furthermore there are publications on risk assessment for irritating chemicals in indoor air (Nielsen and Wolkoff, 2017; Nielsen et al., 2007), and guidance documents of different institutions (ECETOC, 2010; ECHA, 2012; Ad hoc Arbeitsgruppe Innenraumrichtwerte, 2012). A commonly recognised science-based approach for the quantitative risk assessment of irritating chemicals is unfortunately missing.

1.3. Objectives

To evaluate these proposals and to develop a sound basis for deriving guide values for the German Committee for Indoor Air Guide Values, we derived EFs for irritating substances (Mangelsdorf, 2018; Schröder et al., 2015). This work involved a review of the literature on CHS, analysis of inhalation studies with rats and mice collected in the Fraunhofer RepDose® Database, and of Alarie tests with mice as well as on studies addressing mechanistic questions.

2. Materials and methods

2.1. Derivation of EFs for exposure duration with the RepDose® database

The RepDose® database (www.fraunhofer-repdose.de, (Bitsch et al., 2006)) was used for the analysis of EFs for exposure duration for irritating effects in rodent inhalation studies. RepDose® is a database on repeated-dose toxicity studies with rodents with oral and inhalation exposure. The database contains information on study design, affected organs, No Observed Effect Concentrations (NOECs) of the studies, and Lowest Observed Effect Concentrations (LOECs) for all observed effects. For deriving extrapolation factors for study duration a similar approach as in Batke et al. (2011) was applied. Briefly, high quality inhalation studies - preferably according to OECD Guidelines 407, 413, 452, with subacute, subchronic and chronic duration were identified in the database. For these studies it was determined, which type of effect was observed at the LOEC, either local or systemic. Effects in the eyes or in the respiratory tract (site of first contact) were considered as local effects while all other effects were considered as systemic effects. The extrapolation factor was calculated by dividing the NOEC or LOEC of the shorter study duration by the NOEC or LOEC of the longer study duration, either for chemicals with local effects at the LOEC, systemic effects at the LOEC or for all chemicals. For some chemicals, there was more than one inhalation study available for the same species and study duration. Accordingly, more than one EF could be derived. To avoid a chemical-triggered bias, the median of all EFs per chemical was calculated and used for further analyses. Statistical analyses were performed with the software Statistica. To characterize the distribution function, geometric means (GM), medians, 10th and 90th percentiles are provided. The experimental EFs followed a lognormal distribution, therefore two GM values are provided: GM_e – experimental data from the logarithmic data set, and GM_f – the fitted value based on the assumption of a lognormal distribution of the original EF values.

2.2. Literature review

CHS (search term “controlled human exposure studies”, “chamber”) were searched in PubMed and LIVIVO using search terms referring to irritation and specific effects/endpoints, e.g. “(sensory)irritation”, “nasal lavage”, “rhinometry”, “eye blink frequency”. Additional controlled human exposure studies were identified from the references in the publications obtained. Studies were selected according to the following criteria: investigation of sensory irritation (i.e. subjective symptoms or objective parameters), at least one test concentration plus a control group, at least 5 persons exposed, and at least 1-h exposure. Investigations with mixtures and with particles were excluded for the derivation of EFs for species extrapolation. The literature search covers publications until August 2020.

For the selected human studies, experimental details NOAECs, Lowest Observed Adverse Effect Concentrations (LOAECs) and effects at the LOAEC were recorded. The NOAEC was defined as the concentration, at which only very slight subjective symptoms of irritation of the eyes or the upper respiratory tract (assessed in questionnaires) occurred and no objective signs of irritation such as changes in eye blinking rate, rhinomanometry, cytokines in nasal lavage fluid. The symptom ratings were considered as slight, when they were below 1 (somewhat, scale 0–5 in the Swedish Performance Evaluation System), below 70 (weak, scale 1–1000 in the Labelled Magnitude Scale), below 1 (hardly at all, scale 1–7 in the Visual Analog Scale). This collection of human studies was used for correlation of NOAECs with rat inhalation studies, with RD_{50} from the Alarie test, for deriving EFs from rat inhalation studies and results of the Alarie test and for evaluating intraspecies differences.

Rat inhalation studies with durations of at least 90 days were

Table 1
Extrapolation factors for exposure duration for local and systemic effects in inhalation studies with rats and mice.

Shorter study duration[a]	Longer study duration[a]	Effect at LOEC	N	GM_e	Median	10th	90th	GM_f
Short	Subacute	Local	9	9.0	8.8	1.9	40.0	9.6
		Systemic	15	2.2	2.1	0.2	25.5	2.2
		Any effect	18	2.8	4.9	0.2	25.5	3.0
Short	Subchronic	Local	22	3.7	4.0	0.5	20.0	3.8
		Systemic	35	2.8	2.5	0.7	20.0	3.1
		Any effect	41	3.0	2.5	0.6	16.3	3.2
Short	Chronic	Local	16	4.6	2.4	0.3	97.1	4.7
		Systemic	28	3.6	3.0	0.6	51.8	3.9
		Any effect	29	3.3	2.9	0.3	51.8	3.6
Subacute	Subchronic	Local	16	2.1	2.0	0.4	10.0	2.1
		Systemic	34	1.6	1.7	0.1	9.7	1.6
		Any effect	44	1.7	1.5	0.4	10.0	1.8
Subchronic	Chronic	Local	40	1.7	1.4	0.5	11.9	1.9
		Systemic	59	1.4	1.0	0.3	9.8	1.6
		Any effect	71	1.5	1.5	0.4	9.8	1.7
Subacute	Chronic	Local	10	2.8	3.7	0.5	9.6	2.9
		Systemic	18	1.8	2.8	0.1	18.7	1.9
		Any effect	23	1.8	3.2	0.2	13.4	1.9

[a] Short 10–19 days, subacute 20–35 days, subchronic 78–122 days, chronic >350 days, N number of chemicals, GM geometric mean, GM_e GM from experimental (logarithmised) dataset, GM_f GM from fitted original distribution.

searched for those chemicals, for which controlled human exposure studies were available. For this search, reviews on the respective substances from national and international institutions were used as well as the database of registered substances of the European Chemicals Agency (ECHA). Inhalation studies with rats were described similarly as the human studies. The lowest NOAEC or LOAEC from the study with the longest duration was used for derivation of a species extrapolation factor rat to human. If only a subchronic study was available, the NOAEC was divided by a factor of 2 to adjust for the lower NOAECs in studies with longer study durations according to our analyses with the RepDose® database described above.

Results of Alarie tests were obtained from a data collection by Alarie (2016).

Correlations between NOAECs in controlled human exposure studies and rat inhalation studies and results from Alarie tests were analysed. If a LOAEC but no NOAEC was available in the corresponding human or rat study, the respective LOAECs were extrapolated to NOAECs by using a default factor of 3 (ECETOC, 2010; ECHA, 2012), which reflects the dose spacing most often used in toxicological studies. It is thus assumed, that the next lower dose level would be the NOAEC. For correlating results of the Alarie tests with human short term exposure studies, $RD_{50} \times 0.03$ was used, as proposed by Alarie (1981b). In addition, EFs from NOAECs of rat studies to NOAECs in human studies were derived, and from RD_{50} from Alarie tests to NOAECs in human studies. If NOAEC or LOAEC was missing in one of the paired studies or in both, the resulting EF was considered as "less reliable". The resulting EF distributions were characterised by their geometric means, medians, 10^{th} and 90^{th} percentiles.

To analyse differences of the human subpopulation, CHS investigating the influence of gender, age, smoking, allergic rhinitis, asthma, and chemical intolerance were evaluated. If possible, an EF accounting for differences between a sensitive subgroup and a control group was derived.

3. Results and discussion

3.1. EFs for exposure duration

3.1.1. Rodent subacute, subchronic and chronic inhalation studies

We investigated, whether EFs for exposure duration for the endpoint irritation in inhalation studies in rodents differ from those, usually used in risk assessment (ECHA, 2012). Table 1 shows the EFs derived for different study durations for irritating effects in rodents with the Fraunhofer RepDose® -database. The geometric mean (GM_f) of the fitted distribution for subacute→subchronic extrapolation was 2.1, for subchronic→chronic extrapolation 1.9 and for subacute→subchronic extrapolation 2.9. If 2.1 would be multiplied by 2.9 the combined EF for subacute to subchronic extrapolation would be 4 instead of 2.9. As both, the extrapolation subacute→subchronic and chronic bears some uncertainty due to the low number of studies, we propose to use the higher value of 4 for subacute→chronic extrapolation and 2 for subchronic→chronic extrapolation.

Table 2
Effects triggering the LOAEC in rat inhalation studies.

Organ/epithelium affected	N	Chemical
No histopathological findings in respiratory tract	10	1,1,1-trichloroethane, 1,1,1-trifluoroethane, 1,1,1,3,3-pentafluoropropane, 1-methoxy-2-propanol-2, 2-ethyl hexanol, ethyl benzene, 2-propanol, methanol, 2-butanone, 1-methyl-2-pyrrolidone, 2-methoxy-2-methylpropane
Nose: olfactory epithelium	10	Acetaldehyde, chlorine, dioxane, ethyl acetate, ethyl acrylate, formaldehyde, methyl methacrylate, 1-butyl acetate, styrene, toluene
Nose: respiratory epithelium	7	2-Propanol, acrolein, chlorine, 1,4-dioxane, formaldehyde, methyl methacrylate, naphthalene, toluene
Nose: squamous epithelium	2	Chlorine, hydrogen peroxide

Higher EFs of 3.8 and 4.7 were obtained, when instead of a subacute study (20–35 days) a "short" study (10–19 days) was used as basis for extrapolation to subchronic or chronic exposure. Studies with 14 days duration are often dose finding studies for subsequent studies with longer duration, e.g. within the US National Toxicology Programme (NTP). In these studies, the scope of examinations is lower than in a 28-day subacute study following OECD guidelines (and therefore they have a lower sensitivity (higher NOAECs) than studies according to OECD guidelines. Our results confirm earlier findings from Kalberlah et al. (2002), who obtained a factor of 3.2 for subacute→subchronic, and 6.6 for subacute→chronic extrapolation for irritating chemicals, when using 14-day studies from the NTP. In conclusion, when considering extrapolations for exposure duration, 14-day dose finding studies should be differentiated from 28-day studies according to OECD guidelines and should probably not be used for risk assessment due to their limited scope of examination.

There seems to be a specific sensitivity of the respiratory tract to prolonged exposure to locally acting chemicals at cytotoxic concentrations. The EFs obtained for systemic effects were consistently lower than for local irritating effects in our analysis of inhalation studies (Table 2),

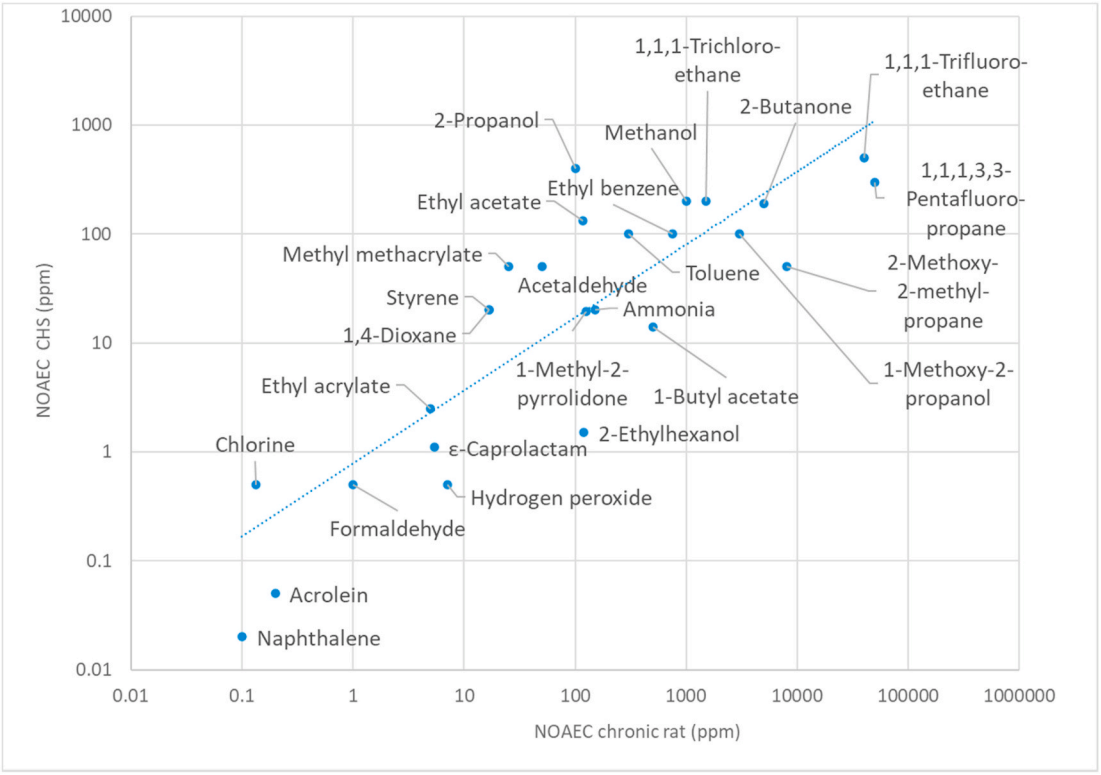


Fig. 1. Correlation between NOAECs in rat inhalation studies (≥90 d) and CHS.

but were comparable with analyses of EFs for exposure duration with much larger datasets for the oral route with the RepDose® database (Batke et al., 2011; Escher et al., 2020). From all these analyses a systemic EF of 2 for subchronic→chronic and 3 for subacute→chronic extrapolation can be derived, which is less conservative than the currently applied values of 2 for subchronic→chronic and 6 for subacute→chronic extrapolation (ECHA, 2012; Fromme et al., 2019).

3.1.2. Human studies

For risk assessment of chemicals in indoor air, lifelong exposure for 24 h per day is assumed. Consequently, CHS with long durations should form the basis for risk assessment, which are, however not available.

Only one CHS with repeated exposures was identified. During 5 days exposure to ethyl acrylate (4 h per day), symptom ratings for eye irritation slightly decreased from day to day indicating adaptation, while eye blinking frequency was the same at each exposure day (Kleinbeck et al., 2020). There was a complete recovery from day to day.

In some studies, it was investigated, whether previous exposure at workplaces leads to higher sensitivity in chamber studies compared to previously unexposed controls. For naphthalene and toluene higher symptom ratings or objective findings were found in persons with previous exposure at workplaces (Blaszkewicz et al., 2016; Dalton et al., 1997; Orbaek et al., 1998). This may indicate, that the workers have developed cytotoxic damage from exposures to rather high concentrations at the workplaces. In contrast, there were lower symptom ratings after preexposure to acetone and ammonia (Dalton et al., 1997; Hoffmann et al., 2004; Ihrig et al., 2006) indicating adaptation. From these few studies, it can be concluded that both, aggravation, and adaptation is possible after repeated exposures, possibly depending on the exposure concentrations at the workplace and the toxicity of the respective chemials.

Some additional information can be obtained from analysing the course of the concentration-time curve in CHS. Generally, the intensity of the subjective symptoms for irritation of eye and nose as well as objective parameters like eye blinking frequency increased during the first 1–2 h of exposure, irrespective of chemical investigated (Cain et al., 2010; Ihrig et al., 2006; Kleinbeck et al., 2017; Pacharra et al., 2017; van Thriel et al., 2005). The increase was steeper for higher exposure concentrations, compared to lower exposure concentrations (Claeson and Lind, 2016; Dwivedi et al., 2015; Kleinbeck et al., 2017). In several studies, after 1–2 h a plateau was reached (Cain et al., 2010; Claeson and Lind, 2016; Dwivedi et al., 2015; Hempel-Jorgensen et al., 1999; Sundblad et al., 2004; van Thriel et al., 2002, 2005, 2007). Sometimes the intensity of ratings even decreased after 1–2 h (Ernstgard et al., 2006b; Iregren et al., 1993). But also an increase during the whole 4 h of exposure was found (Kleinbeck et al., 2017, 2020). These results indicate that also during an exposure session both, aggravation, and adaptation is possible.

In addition, CHS with acrolein, ethylacrylate, ethylhexanol, formaldehyde, and 1-octanol show that the concentration of a chemical determines the effect primarily and not the duration of exposure. In these studies, exposure conditions with intermittent peaks led to higher irritation (objective findings as well as subjective symptoms) than the same time weighted average (TWA) concentration under continuous exposure (Baelum et al., 1990; Claeson and Lind, 2016; Kiesswetter et al., 2005; Kleinbeck et al., 2017; Lang et al., 2008; Sucker et al., 2019; van Thriel et al., 2003a, 2005, 2007).

Based on the studies described above, exposure duration is not an important parameter at low exposure concentrations, because at low concentrations slight effects are reversible. Furthermore exposure concentration is more important for determining the NOAEC. Therefore it is concluded that no EF is needed for exposure duration based on a NOAEC in a CHS.

3.2. EFs for species extrapolation

3.2.1. Extrapolation rat inhalation → humans

We evaluated whether NOAECs in inhalation studies with rats may predict NOAECs in CHS on sensory irritation. 25 chemicals were identified, for which both, CHS and rat studies of at least 90 days duration

Table 3
EF for extrapolation from a NOAEC in a chronic rat study to a NOAEC in a CHS.

Dataset	n	GM	Median	10th	90th
Reliable	10	2.3	3.1	0.7	7.0
less reliable	15	4.9	5	0.5	64.0
All	25	3.6	3.5	0.5	36.0

were available.

A correlation was found for the NOAECs in these rat studies with the NOAECs in the controlled human exposure studies with a coefficient of determination $r^2 = 0.74$ (Fig. 1).

When assessing the correlations of NOAECs in CHS with NOAECs in rat inhalation studies, the study design related limitations must be considered. The NOAECs of the human studies often were the only concentration tested; thus the "real" NOAEC may be higher. For ethical reasons, the tested concentrations in CHS had to be chosen low enough to be considered as safe, i. e usually at or below occupational exposure limits. On the other side, there were also several human studies, where no NOAEC could be derived. For our correlation we used a factor of 3 to extrapolate from a LOAEC to a NOAEC irrespective of effect or effect size detected at the LOAEC.

Another source of uncertainty is that in most CHS, the authors did not report NOAECs and LOAECs. We had to decide and to be consistent between studies from different research groups, and to consider differences in terminology and grading of symptoms. So, our assignments bear considerable uncertainty.

Despite these limitations, the correlation was reasonably good, indicating that the potency of a sensory irritant in humans can be predicted by rat inhalation studies.

Table 2 shows the effects that triggered the LOAEC in the animal studies for those chemicals that were investigated in humans with respect to sensory irritation. The nose was the most sensitive organ in the rat for all those chemicals, where details from histopathological examination were available. Several chemicals affected more than one epithelium in the nose, effects in the lung were less sensitive. Thus, effects in histopathological examinations of the nose (cytotoxic irritation) seem to be a good predictor for sensory irritation in humans.

For obtaining an interspecies EF from rats to humans, the data pairs of human and rat studies were subdivided into two groups in Table 3: those with "reliable" NOAEC (9 chemicals) with both, NOAEC and LOAEC in the human and the rat studies, and the "less reliable" data pairs due to the presence of only NOAEC or LOAEC in either the human studies, or the rat studies or both (16 chemicals). The chronic NOAEC from the rat inhalation studies was divided by the NOAEC of the human controlled exposure studies. For the chemicals with "reliable" NOAEC a geometric mean of 2.3 was obtained as EF. The 90th percentile was 7. For the "less reliable" chemicals, the corresponding figures were higher.

Based on the "reliable dataset", and taking into consideration the uncertainty due to the small dataset of 10 chemicals a species EF of 3 for higher susceptibility in CHS compared to findings in histopathology in the nose in chronic rat studies may be applied. This factor is in agreement with the factor of 3 obtained in an earlier analysis by Brüning et al. (2014) of 8 chemicals. It is also in the same order of magnitude as the factor of 2.5 that is often assigned for species differences due to other interspecies differences, e.g. from toxicodynamic differences (ECHA, 2012). The investigations above were with studies with rats. For risk assessment purposes it is assumed that the same EF of 3 is also applicable for inhalation studies with mice.

3.2.2. *Extrapolation Alarie test → humans*

A similar analysis as above was performed also for the Alarie test. 19

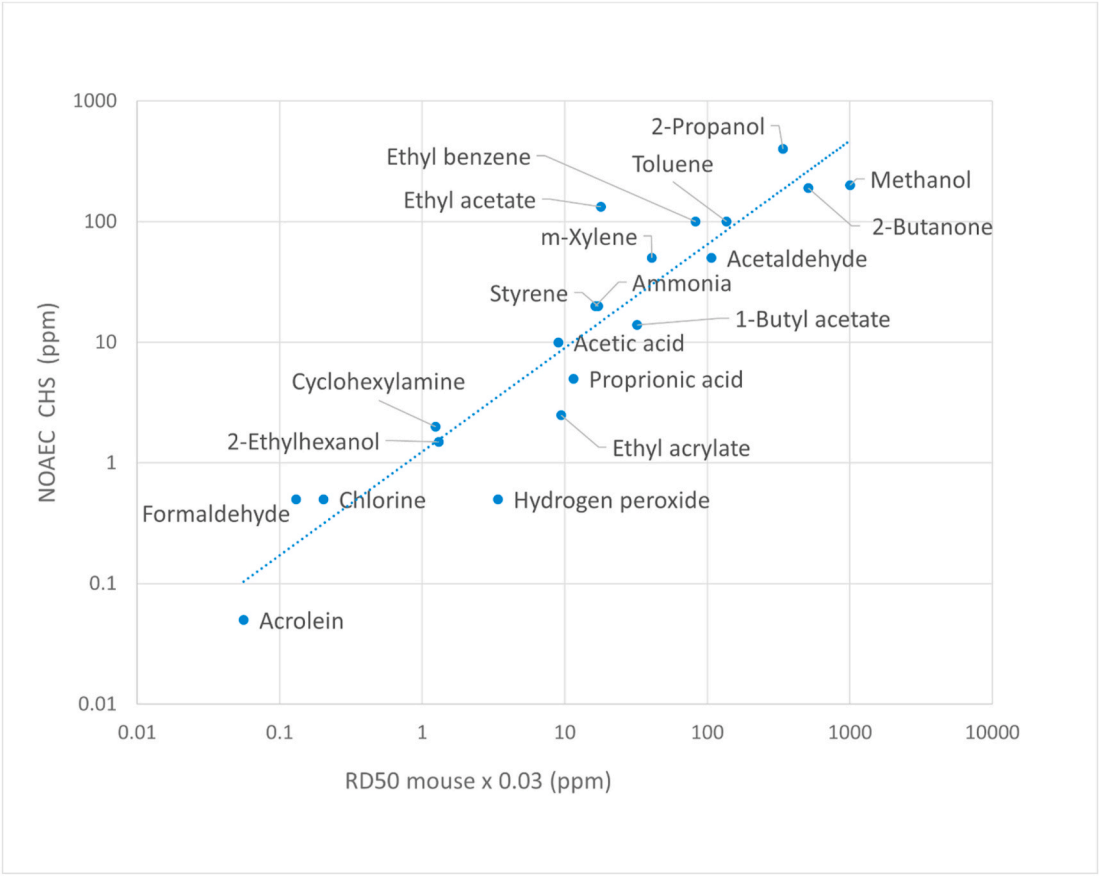


Fig. 2. Correlation between $RD_{50} \times 0.03$ from Alarie tests with mice and NOAECs from CHS.

Table 4
EF for extrapolation from RD_{50} to a NOAEC in a CHS.

Dataset	n	GM	Median	10th	90th
Reliable	12	39.4	30.0	13.3	126
Less reliable	8	38.1	36.8	20.0	113
All	20	38.8	30.0	12.4	134

chemicals were identified with both, human studies and Alarie tests in mice. Fig. 2 shows the correlation between $RD_{50} \times 0.03$ from the Alarie test and the NOAEC in human studies. $RD_{50} \times 0.03$ was chosen, as originally proposed by Alarie (1981b). With a coefficient of determination of $r^2 = 0.87$, the correlation between the RD_{50} of the Alarie test and the NOAECs of the human studies on sensory irritation is quite good. There are some outliers. For ethyl acetate, the reason can be explained: the dose response curve in the Alarie test was much steeper than for other chemicals (Alarie, 1981b) and therefore, the factor of 0.03 overestimates the toxicity in the Alarie test.

Our analysis confirms earlier findings, where $RD_{50} \times 0.03$ correlated well with irritation thresholds (Cometto-Muniz and Cain, 1994), occupational exposure limits (Alarie, 1981a, b; Kupczewska-Dobecka et al., 2006; Schaper, 1993) and human studies on irritation (Kuwabara et al., 2007). The quality of the correlation in our analysis was slightly better than in the study by Kuwabara et al. (2007) with 25 chemicals and $r^2 = 0.80$. Our analysis included more recent publications and the focus was narrower than by Kuwabara et al. who also included studies, where lung function or airway hyperresponsiveness were investigated as the only endpoint and studies with exposure durations as low as 1 min.

Our results confirm the proposal by Nielsen and Wolkoff (2017) to use the Alarie Test to derive guide values for irritating effects of chemicals in indoor air.

$RD_{50} \times 0.03$ as originally proposed by Alarie (1981b) for occupational exposure limits, could also be expressed as an EF of 33 in analogy to the rat studies. Based on CHS we derived a slightly higher EF of 40 (Table 4), which should be applied for extrapolation from a RD_{50} value to a NOAEC in a CHS.

For obtaining an interspecies EF from the RD_{50} in mice to humans, chemicals with reliable NOAECs in controlled human exposure studies (11 chemicals) were distinguished from less reliable studies (8 chemicals), where either NOAEC or LOAEC in the human studies was missing. For the chemicals with reliable NOAEC a geometric mean of 39.4 was obtained as EF, the 90th percentile was 126. In contrast to the rat inhalation studies, there was only a small difference between reliable and less reliable datasets (Table 4).

In conclusion, it is proposed to apply an EF of 40 to extrapolate from a RD_{50} in the Alarie test with mice to a NOAEC in a CHS.

The Alarie test can, however, not be used as the only starting point for setting a guide value, because systemic effects that may occur at lower dose levels than the irritating effects that are detected in the Alarie test, are not covered.

There are two possibilities: A subacute, subchronic or chronic rodent inhalation study is available in addition to an Alarie Test. In this case, the guide values derived after applying appropriate EFs from Alarie test and rodent inhalation study should be compared. Unless there are specific reasons - the lower value should be used.

If no rodent inhalation study is available, we recommend performing a route-to-route extrapolation from oral rodent studies with subacute, subchronic or chronic duration as described in ECHA (2012). In a previous research project (Schröder et al., 2016) we could show that the GM_f for the ratio of NOAECs of oral and inhalation studies for the same chemical and species was 2 for systemic effects and 3.4 for chemicals with irritating effects. As it is not known a priori, whether a chemical is irritating or not, we recommend taking a conservative factor of 4 for the extrapolation. The resulting value then should be compared with the result of the Alarie test, and – as above - the lower value shall be used to derive the guide value. Route-to-route extrapolation without supporting Alarie test is not considered as sufficient.

3.3. EF for intraspecies differences

In the following, results from CHS are summarized that provide information on possible special sensitivity of some population subgroups. No suitable CHS were identified that investigated possible special sensitivity of children, elderly people, or smokers.

3.3.1. Atopy, allergic rhinits, asthma

Only a limited number of CHS investigated the influence of atopy, allergic rhinitis or asthma on sensory irritation. Compared to controls, no differences in ratings of sensory irritation were found in atopic persons (with elevated IgE for frequent allergens in Sweden) exposed to 2-ethylhexanol (Ernstgard et al., 2010), persons with allergic rhinitis exposed to ammonia (Blaszkewicz et al., 2016; Pacharra et al., 2017), or asthmatics exposed to formaldehyde (Green et al., 1987). In a study with exposure to a mixture of ozone and limonene, the scores of sensory irritation were even lower in mild asthmatics compared to controls (Fadeyi et al., 2015).

Small differences between controls and persons with allergy, however, may not have been detected due to the fact, that also controls may be allergic. A recent study revealed, that from 105 young volunteers who considered themselves as non-allergic, almost half were atopic and 10% had airway disease or bronchial hyperreactivity (van Kampen et al., 2020).

The possible special sensitivity of asthmatics to irritating chemicals with respect to changes in lung function such as reductions of the forced expiratory volume in 1 s (FEV_1) and specific airway resitance (SR_{aw}) was evaluated in a review by Johansson et al. (2016). For 8 of 19 chemicals a somewhat higher sensitivity of asthmatics was reported. The highest factor of 3 was obtained for acetaldehyde, >.2 for sulfur dioxide, all other "positive" chemicals were >1, but could not be better quantified, due to lack of suitable data.

In our analyses, in all studies investigated lung function was less sensitive than nose or eye irritation: with acrolein (Dwivedi et al., 2015), ammonia (Hoffmann et al., 2004; Sundblad et al., 2004), butyl acetate (Iregren et al., 1993), acetic acid (Ernstgard et al., 2006a), formaldehyde (Lang et al., 2008), and 2-propanol (Ernstgard et al., 2002). A higher susceptibility of persons with allergy therefore may appear at higher concentrations of the chemicals tested for more severe effects. The relevance of the findings of Johansson et al. (2016) is therefore questionable as not the most sensitive endpoint was investigated.

3.3.2. Sex

Several studies show that women are somewhat (no quantification possible) more susceptible than men, especially at low concentrations of irritating chemicals and with respect to subjective symptoms (Ernstgard et al., 2002, 2012; Pacharra et al., 2016; Sucker et al., 2019). In some studies, differences were seen even without exposure (Ernstgard et al., 2002, 2012; Pacharra et al., 2016). There were also differences between sexes in processing of irritant stimuli as assessed by event related potentials during exposure to carbon dioxide stimuli (Lundstrom et al., 2005). In one study, where subjective symptom ratings of eye irritation were compared to objective eye blinking frequency, only subjective symptom ratings were higher in women than in men but not eye blinking frequency (Sucker et al., 2019). In the sham exposure no differences in symptom rating between sexes were seen.

A different susceptibility of the sexes is made plausible by findings that show that sex hormones can regulate the expression of TRP receptors and modulate intracellular signalling pathways that activate TRP receptors (Artero-Morales et al., 2018). Furthermore, women generally suffer more often from eye symptoms then men, probably due to hormonal influences that make the tear film less stable, especially in women above 40 years (Wolkoff, 2017).

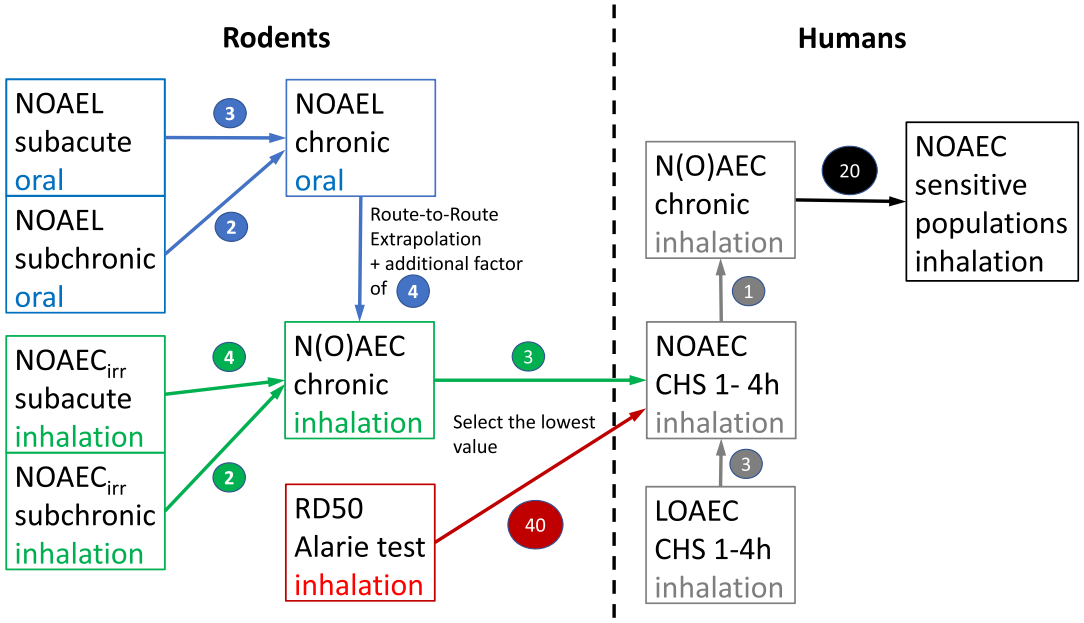


Fig. 3. Proposals for EFs for risk assessment or irritating chemicals in indoor air.

3.3.3. Chemical intolerance

In several controlled exposure studies, persons with chemical intolerance (CI) or multiple chemical sensitivity (MCS), which is a more severe form of CI (Dantoft et al., 2015), reacted more sensitive to exposure with irritant chemicals than controls without CI. Subjective symptoms were increased (Claeson and Andersson, 2017; Österberg et al., 2003; van Thriel et al., 2002, 2003b, 2005, 2007), sometimes already before start of exposure (Andersson et al., 2016; Österberg et al., 2004). Objective findings such as changes in nasal flow, heart rate variability or pulse rate were detected only in few studies in persons with CI (Andersson et al., 2016; Wiesmüller et al., 2002). From some of these studies, factors accounting for special sensitivity of persons CI/MCS could be derived. Differences in sensitivity compared to healthy controls ranged from a factor of 1–3 for most of the studies, up to about 20.

It has been discussed, whether subjective symptoms of persons with CI reflects sensory irritation, or whether odour causes subjective symptoms. One study showed that when the odour of an irritating chemical (acrolein) was masked by an odorous chemical (heptane), persons with CI reacted only to the mixture of the irritating and odourous substance but not to the odorous substance alone (Claeson and Andersson, 2017). It can, however, not be excluded that the mixture was more unpleasant and therefore caused the effects. In another study, exposure of the eyes with perfume caused eye irritation symptoms in persons with MCS, although they could not perceive an odour (Millqvist et al., 1999), and there was no difference found between persons with CI/MCS compared to controls with respect to their ability to detect odours (Azuma et al., 2016; Blaszkewicz et al., 2016; Claeson and Andersson, 2017; Österberg et al., 2003). As mentioned above, in some studies, also objective findings were reported, which would indicate physiological reactions and not just a consequence of perceptions.

3.3.4. Negative expectations

The possibility that negative expectations about the exposure could cause symptoms was investigated in several studies by applying a specific questionnaire on negative affectivity (Ihrig et al., 2006; Mueller et al., 2013; Nordin et al., 2017; Triebig et al., 2016; Ziegler et al., 2008). There was no difference in symptom ratings between persons with and without negative affectivity.

3.3.5. Conclusion

In conclusion, the limited data available on intraspecies differences in CHS do not justify to derive an EF which would deviate from the precautionary EF of 10–20 as currently used in different regulatory settings for the general population (ECHA, 2012; Fromme et al., 2019).

4. Conclusion – concept for risk assessment

Based on the evaluations described above, the EFs shown in Fig. 3 are proposed. CHS are preferred for the derivation of guide values for indoor air. If CHS are not available, rodent repeated dose studies and Alarie Tests can be used. In case of the Alarie test it must be confirmed that irritating effects are the most sensitive endpoint. A NOAEC from systemic effects can be determined by route-to-route extrapolation from studies with oral application.

While the factors for exposure duration and route-to-route extrapolation from rodent studies are reliable as they were derived from datasets with reasonable size, the EFs for species extrapolations, extrapolation LOAEC→NOAEC in CHS and to chronic exposure in humans are still preliminary. The factor for interindividual differences is a precautionary factor.

As it has been shown that a wide variety of chemicals can bind to the receptors on the nerve endings that mediate sensory irritation, additive effects of different chemicals should be considered when assessing exposure situations in indoor air.

Boxes: Endpoints, bullets with figures: size of EFsRoute-to-route extrapolation should only be used in a weight of evidence approach together with other info. The factor of 20 for human sensitive populations is a precautionary factor, due to insufficient data.

Declaration of competing interest

The authors declare that they have no conflicts of interest.

Acknowledgement

The work was financially supported by grants from the German Environment Agency (UBA, FKZ 40191 and UFOPLAN No 3717 61 204 0).

References

Achanta, S., Jordt, S.E., 2017. TRPA1: acrolein meets its target. Toxicol. Appl. Pharmacol. 324, 45–50.

Alarie, Y., 1981a. Bioassay for evaluating the potency of airborne sensory irritants and predicting acceptable levels of exposure in man. Food Chem. Toxicol. 19, 623–626.

Alarie, Y., 1981b. Dose-response analysis in animal studies: prediction of human responses. Environ. Health Perspect. 42, 9–13.

Alarie, Y., 2016. Update of the Schaper Database of RD50 and TLV Values.

Andersson, L., Claeson, A.S., Dantoft, T.M., Skovbjerg, S., Lind, N., Nordin, S., 2016. Chemosensory perception, symptoms and autonomic responses during chemical exposure in multiple chemical sensitivity. Int. Arch. Occup. Environ. Health 89, 79–88.

Artero-Morales, M., González-Rodríguez, S., Ferrer-Montiel, A., 2018. TRP channels as potential targets for sex-related differences in migraine pain. Front Mol Biosci 6, 1–13.

Azuma, K., Uchiyama, I., Tanigawa, M., Bamba, I., Azuma, M., Takano, H., Yoshikawa, T., Sakabe, K., 2016. Association of odor thresholds and responses in cerebral blood flow of the prefrontal area during olfactory stimulation in patients with multiple chemical sensitivity. PloS One 11, e0168006.

Baelum, J., Lundqvist, G.R., Molhave, L., Andersen, N.T., 1990. Human response to varying concentrations of toluene. Int. Arch. Occup. Environ. Health 62, 65–71.

Bascom, R., Meggs, W.J., Frampton, M., Hudnell, K., Killburn, K., Kobal, G., Medinsky, M., Rea, W., 1997. Neurogenic inflammation: with additional discussion of central and perceptual integration of nonneurogenic inflammation. Environ. Health Perspect. 105 (Suppl. 2), 531–537.

Batke, M., Escher, S., Hoffmann-Doerr, S., Melber, C., Messinger, H., Mangelsdorf, I., 2011. Evaluation of time extrapolation factors based on the database RepDose. Toxicol. Lett. 205, 122–129.

Bautista, D.M., Jordt, S.E., Nikai, T., Tsuruda, P.R., Read, A.J., Poblete, J., Yamoah, E.N., Basbaum, A.I., Julius, D., 2006. TRPA1 mediates the inflammatory actions of environmental irritants and proalgesic agents. Cell 124, 1269–1282.

Bautista, D.M., Pellegrino, M., Tsunozaki, M., 2013. TRPA1: a gatekeeper for inflammation. Annu. Rev. Physiol. 75, 181–200.

Bessac, B.F., Jordt, S.E., 2008. Breathtaking TRP channels: TRPA1 and TRPV1 in airway chemosensation and reflex control. Physiology 23, 360–370.

Bessac, B.F., Jordt, S.E., 2010. Sensory detection and responses to toxic gases: mechanisms, health effects, and countermeasures. Proc. Am. Thorac. Soc. 7, 269–277.

Bitsch, A., Jacobi, S., Melber, C., Wahnschaffe, U., Simetska, N., Mangelsdorf, I., 2006. REPDOSE: a database on repeated dose toxicity studies of commercial chemicals–A multifunctional tool. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 46, 202–210.

Blaszkewicz, M., Brüning, T., Bünger, J., Kleinbeck, S., Pacharra, M., Schäper, M., Sucker, K., van Thriel, C., 2016. Abschlussbericht zum Vorhaben „Irritative Effekte: zeitextrapolation, intra- und interindividuelle Unterschiede“ (FP-0326). Leibnitz-Institut für Arbeitsforschung an der TU Dortmund.

Brüning, T., Bartsch, R., Bolt, H.M., Desel, H., Drexler, H., Gundert-Remy, U., Hartwig, A., Jackh, R., Leibold, E., Pallapies, D., Rettenmeier, A.W., Schluter, G., Stropp, G., Sucker, K., Triebig, G., Westphal, G., van Thriel, C., 2014. Sensory irritation as a basis for setting occupational exposure limits. Arch. Toxicol. 88, 1855–1879.

Cain, W.S., Dourson, M.L., Kohrman-Vincent, M.J., Allen, B.C., 2010. Human chemosensory perception of methyl isothiocyanate: chemesthesis and odor. Regul. Toxicol. Pharmacol. 58, 173–180.

Claeson, A.S., Andersson, L., 2017. Symptoms from masked acrolein exposure suggest altered trigeminal reactivity in chemical intolerance. Neurotoxicology.

Claeson, A.S., Lind, N., 2016. Human exposure to acrolein: time-dependence and individual variation in eye irritation. Environ. Toxicol. Pharmacol. 45, 20–27.

Cometto-Muniz, J.E., Cain, W.S., 1994. Sensory reactions of nasal pungency and odor to volatile organic compounds: the alkylbenzenes. Am. Ind. Hyg. Assoc. J. 55, 811–817.

Dalton, P., Wysocki, C.J., Brody, M.J., Lawley, H.J., 1997. Perceived odor, irritation, and health symptoms following short-term exposure to acetone. Am. J. Ind. Med. 31, 558–569.

Dantoft, T.M., Andersson, L., Nordin, S., Skovbjerg, S., 2015. Chemical intolerance. Curr. Rheumatol. Rev. 11, 167–184.

Doty, R.L., Cometto-Muniz, J.E., Jalowayski, A.A., Dalton, P., Kendal-Reed, M., 2004. Assessmento of upper respiratory tract and ocular irritative effects of volatile chemicals in humans. Crit. Rev. Toxicol. 34, 85–142.

Dwivedi, A.M., Johanson, G., Lorentzen, J.C., Palmberg, L., Sjogren, B., Ernstgard, L., 2015. Acute effects of acrolein in human volunteers during controlled exposure. Inhal. Toxicol. 27, 810–821.

ECETOC, 2010. Guidance on Assessment Factors to Derive a DNEL.

ECHA, 2012. Charcterisation of Dose [concentration]-Response for Human Health, Guidance on Information Requirements and Chemical Safety Assessment, 1 ed, Version 2.

Ernstgard, L., Gullstrand, E., Lof, A., Johanson, G., 2002. Are women more sensitive than men to 2-propanol and m-xylene vapours? Occup. Environ. Med. 59, 759–767.

Ernstgard, L., Iregren, A., Sjogren, B., Johanson, G., 2006a. Acute effects of exposure to vapours of dioxane in humans. Hum. Exp. Toxicol. 25, 723–729.

Ernstgard, L., Iregren, A., Sjogren, B., Svedberg, U., Johanson, G., 2006b. Acute effects of exposure to hexanal vapors in humans. J. Occup. Environ. Med. 48, 573–580.

Ernstgard, L., Norback, D., Nordquist, T., Wieslander, G., Walinder, R., Johanson, G., 2010. Acute effects of exposure to 1 mg/m(3) of vaporized 2-ethyl-1-hexanol in humans. Indoor Air 20, 168–175.

Ernstgard, L., Sjogren, B., Johanson, G., 2012. Acute effects of exposure to vapors of hydrogen peroxide in humans. Toxicol. Lett. 212, 222–227.

Escher, S.E., Mangelsdorf, I., Hoffmann-Doerr, S., Partosch, F., Karwath, A., Schroeder, K., Zapf, A., Batke, M., 2020. Time extrapolation in regulatory risk assessment: the impact of study differences on the extrapolation factors. Regul. Toxicol. Pharmacol. 112, 104584.

Fadeyi, M.O., Tham, K.W., Wu, W.Y., 2015. Impact of asthma, exposure period, and filters on human responses during exposures to ozone and its initiated chemistry products. Indoor Air 25, 512–522.

Fromme, H., Debiak, M., Sagunski, H., Rohl, C., Kraft, M., Kolossa-Gehring, M., 2019. The German approach to regulate indoor air contaminants. Int. J. Hyg Environ. Health 222, 347–354.

Green, D.J., Sauder, L.R., Kulle, T.J., Bascom, R., 1987. Acute response to 3.0 ppm formaldehyde in exercising healthy nonsmokers and asthmatics. Am. Rev. Respir. Dis. 135, 1261–1266.

Harkema, J.R., Carey, S.A., Wagner, J.G., 2006. The nose revisited: a brief review of the comparative structure, function, and toxicologic pathology of the nasal epithelium. Toxiclogic. Pathol. 34, 252–269.

Hempel-Jorgensen, A., Kjaergaard, S.K., Molhave, L., Hudnell, H.K., 1999. Time course of sensory eye irritation in humans exposed to N-butanol and 1-octene. Arch. Environ. Health 54, 86–94.

Hoffmann, J., Ihrig, A., Triebig, G., 2004. Expositionsstudie zur arbeitsmedizinischen Bedeutung Ammoniak-assoziierter gesundheitlicher Effekte. Arbeitsmed. Sozialmed. Umweltmed. 39, 390–401.

Hummel, T., Mohammadian, P., Marchl, R., Kobal, G., Lotsch, J., 2003. Pain in the trigeminal system: irritation of the nasal mucosa using short- and long-lasting stimuli. Int. J. Psychophysiol. 47, 147–158.

Ihrig, A., Hoffmann, J., Triebig, G., 2006. Examination of the influence of personal traits and habituation on the reporting of complaints at experimental exposure to ammonia. Int. Arch. Occup. Environ. Health 79, 332–338.

Iregren, A., Lof, A., Toomingas, A., Wang, Z., 1993. Irritation effects from experimental exposure to n-butyl acetate. Am. J. Ind. Med. 24, 727–742.

Johansson, M.K., Johanson, G., Oberg, M., 2016. Evaluation of the experimental basis for assessment factors to protect individuals with asthma from health effects during short-term exposure to airborne chemicals. Crit. Rev. Toxicol. 46, 241–260.

Kalberlah, F., Föst, U., Schneider, K., 2002. Time extrapolation and interspecies extrapolation for locally acting substances in case of limited toxicological data. Ann. Occup. Hyg. 46, 175–185.

Kiesswetter, E., Thriel, C., Schaper, M., Blaszkewicz, M., Seeber, A., 2005. Eye blinks as indicator for sensory irritation during constant and peak exposures to 2-ethylhexanol. Environ. Toxicol. Pharmacol. 19, 531–541.

Kleinbeck, S., Schäper, M., Pacharra, M., Lehmann, M.L., Golka, K., Blaszkewicz, M., Brüning, T., van Thriel, C., 2020. A short-term inhalation study to assess the reversibility of sensory irritation in human volunteers. Arch. Toxicol. 94, 1687–1701.

Kleinbeck, S., Schaper, M., Zimmermann, A., Blaszkewicz, M., Bruning, T., van Thriel, C., 2017. Prediction of human sensory irritation due to ethyl acrylate: the appropriateness of time-weighted average concentration x time models for varying concentrations. Arch. Toxicol.

Kupczewska-Dobecka, M., Socko, R., Czerczak, S., 2006. RD50 value as the criterion for setting maximum admissible levels of occupational exposure to irritants in Poland. Int. J. Occup. Saf. Ergon. 12, 95–99.

Kuwabara, Y., Alexeeff, G.V., Broadwin, R., Salmon, A.G., 2007. Evaluation and application of the RD50 for determining acceptable exposure levels of airborne sensory irritants for the general public. Environ. Health Perspect. 115, 1609–1616.

Lang, I., Bruckner, T., Triebig, G., 2008. Formaldehyde and chemosensory irritation in humans: a controlled human exposure study. Regul. Toxicol. Pharmacol. 50, 23–36.

Lanosa, M.J., Willis, D.N., Jordt, S., Morris, J.B., 2010. Role of metabolic activation and the TRPA1 receptor in the sensory irritation response to styrene and naphthalene. Toxicol. Sci. 115, 589–595.

Lundstrom, J.N., Frasnelli, J., Larsson, M., Hummel, T., 2005. Sex differentiated responses to intranasal trigeminal stimuli. Int. J. Psychophysiol. 57, 181–186.

Mangelsdorf, I., 2018. Erstellung eines Risikobewertungskonzepts für lokal reizende Stoffe in der Innenraumluft auf der Grundlage von Nagetierstudien. Umweltbundesamt, Berlin.

Martinez, J.M., Eling, T.E., 2019. Activation of TRPA1 by volatile organic chemicals leading to sensory irritation. ALTEX 36, 572–582.

Millqvist, E., Bengtsson, U., Lowhagen, O., 1999. Provocations with perfume in the eyes induce airway symptoms in patients with sensory hyperreactivity. Allergy 54, 495–499.

Nielsen, G.D., Wolkoff, P., 2017. Evaluation of airborne sensory irritants for setting exposure limits or guidelines: a systematic approach. Regul. Toxicol. Pharmacol. 90, 308–317.

Nielsen, G.D., Wolkoff, P., Alarie, Y., 2007. Sensory irritation: risk assessment approaches. Regul. Toxicol. Pharmacol. 48, 6–18.

Orbaek, P., Osterberg, K., Akesson, B., Bergendorf, U., Karlson, B., Seger, L., 1998. Suprathreshold intensity and annoyance reactions in experimental challenge to toluene and n-butyl acetate among subjects with long-term solvent exposure. Scand. J. Work. Environ. Health 24, 432–438.

Österberg, K., Orbaek, P., Karlson, B., Akesson, B., Bergendorf, U., 2003. Annoyance and performance during the experimental chemical challenge of subjects with multiple chemical sensitivity. Scand. J. Work. Environ. Health 29, 40–50.

Pacharra, M., Kleinbeck, S., Schaper, M., Blaszkewicz, M., Golka, K., van Thriel, C., 2017. Does seasonal allergic rhinitis increase sensitivity to ammonia exposure? Int. J. Hyg Environ. Health 220, 840–848.

Pacharra, M., Kleinbeck, S., Schaper, M., Juran, S.A., Hey, K., Blaszkewicz, M., Lehmann, M.L., Golka, K., van Thriel, C., 2016. Interindividual differences in chemosensory perception: toward a better understanding of perceptual ratings during chemical exposures. J. Toxicol. Environ. Health 79, 1026–1040.

Schaper, M., 1993. Development of a database for sensory irritants and its use in establishing occupational exposure limits. Am. Ind. Hyg. Assoc. J. 54, 488–544.

Schröder, K., Escher, E., Licht, O., Simetska, N., Mangelsdorf, I., 2015. Time extrapolation factors for local effects in inhalation studies. Umweltbundesamt Berlin.

Schröder, K., Escher, S.E., Hoffmann-Dorr, S., Kuhne, R., Simetska, N., Mangelsdorf, I., 2016. Evaluation of route-to-route extrapolation factors based on assessment of repeated dose toxicity studies compiled in the database RepDose®. Toxicol. Lett. 261, 32–40.

Sucker, K., Hoffmeyer, F., Monsé, C., Jettkant, B., Berresheim, H., Rosenkranz, N., Raulf, M., Bünger, J., Brüning, T., 2019. Ethyl acrylate: influence of sex or atopy on perceptual ratings and eye blink frequency. Arch. Toxicol. 93, 2913–2926.

Sundblad, B.M., Larsson, B.M., Acevedo, F., Ernstgard, L., Johanson, G., Larsson, K., Palmberg, L., 2004. Acute respiratory effects of exposure to ammonia on healthy persons. Scand. J. Work. Environ. Health 30, 313–321.

van Kampen, V., Hoffmeyer, F., Monsé, C., Raulf, M., Brüning, T., Bünger, J., Sucker, K., 2020. Discrimination between atopic, allergic, and asthmatic volunteers for human exposure studies on sensory irritation. Adv. Exp. Med. Biol. 1279, 27–35.

van Thriel, C., Haumann, K., Kiesswetter, E., Blaszkewicz, M., Seeber, A., 2002. Time courses of sensory irritations due to 2-butanone and ethyl benzene exposure: influences of self-reported multiple chemical sensitivity (sMCS). Int. J. Hyg Environ. Health 204, 367–369.

van Thriel, C., Kiesswetter, E., Schaper, M., Blaszkewicz, M., Golka, K., Juran, S., Kleinbeck, S., Seeber, A., 2007. From neurotoxic to chemosensory effects: new insights on acute solvent neurotoxicity exemplified by acute effects of 2-ethylhexanol. Neurotoxicology 28, 347–355.

van Thriel, C., Kiesswetter, E., Schaper, M., Blaszkewicz, M., Golka, K., Seeber, A., 2005. An integrative approach considering acute symptoms and intensity ratings of chemosensory sensations during experimental exposures. Environ. Toxicol. Pharmacol. 19, 589–598.

van Thriel, C., Seeber, A., Kiesswetter, E., Blaszkewicz, M., Golka, K., Wiesmuller, G.A., 2003a. Physiological and psychological approaches to chemosensory effects of solvents. Toxicol. Lett. 140–141, 261–271.

van Thriel, C., Wiesmuller, G.A., Blaszkewicz, M., Golka, K., Kiesswetter, E., Seeber, A., Bachert, C., 2003b. Intranasal effects in chemically sensitive volunteers: an experimental exposure study. Environ. Toxicol. Pharmacol. 14, 129–137.

Wiesmüller, G.A., Van Thriel, C., Steup, A., Bachert, C., Clinic, E.N., Blaszkewicz, M., Golka, K., Kiesswetter, E., Seeber, A., 2002. Nasal function in self-reported chemically intolerant individuals. Arch. Environ. Health 57, 247–254.

Wolkoff, P., 2017. External eye symptoms in indoor environments. Indoor Air 27, 246–260.

International Journal of Hygiene and Environmental Health 232 (2021) 113671

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

Seroprevalence of SARS-CoV-2 antibodies among hospital workers in a German tertiary care center: A sequential follow-up study

Thomas Theo Brehm [a,b,*,1], Dorothee Schwinge [a,1], Sibylle Lampalzer [a], Veronika Schlicker [a,b], Julia Küchen [a], Michelle Thompson [a], Felix Ullrich [a], Samuel Huber [a], Stefan Schmiedel [a,b], Marylyn M. Addo [a,b], Marc Lütgehetmann [b,c], Johannes K. Knobloch [b,c], Julian Schulze zur Wiesch [a,b,1], Ansgar W. Lohse [a,b,1]

[a] I. Department of Internal Medicine, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246, Hamburg, Germany
[b] German Center for Infection Research (DZIF), Partner Site Hamburg-Lübeck-Borstel-Riems, Germany
[c] Institute of Medical Microbiology, Virology and Hygiene, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246, Hamburg, Germany

ARTICLE INFO

Keywords:
SARS-CoV-2
COVID-19
Seroprevalence
Hospital workers
Healthcare workers
Germany

ABSTRACT

We sequentially assessed the presence of SARS-CoV-2 IgG antibodies in 1253 hospital workers including 1026 HCWs at the University Medical Center Hamburg-Eppendorf at three time points during the early phase of the epidemic. By the end of the study in July 2020, the overall seroprevalence was 1.8% (n = 22), indicating the overall effectiveness of infection control interventions in mitigating coronavirus disease 2019 (COVID-19) in hospital workers.

1. Introduction

Health care workers (HCWs) are at the front line of the coronavirus disease 2019 (COVID-19) pandemic response and disproportionally at risk of contracting severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) due to occupational exposure to droplets, aerosols and contaminated surfaces (Razzini et al., 2020). Reports about high infection rates of HCWs in many countries have illustrated the challenges of HCW protection, especially with regard to the risk of infections from pre-symptomatic COVID-19 patients (Arons et al., 2020; Zhan et al., 2020). Besides their personal health risk, infected HCWs may contribute to critical staff shortages and pose a risk for vulnerable patients and fellow HCWs. Serological surveillance of hospital workers allows for an estimation of the overall infection rate and for early identification of professional groups and hospital environments at increased risk for contracting COVID-19. This approach is paramount to evaluate and adjust infection control measures to mitigate nosocomial transmission of SARS-CoV-2. By now, several studies have demonstrated high seroprevalence in HCWs in countries most affected by COVID-19 (Garcia-Basteiro et al., 2020; Sotgiu et al., 2020; Rudberg et al., 2020; Houlihan et al., 2020; Martin et al., 2020; Moscola et al., 2020; Stubblefield et al., 2020) but data from German HCWs are scarce (Korth et al., 2020). By 17th July 2020, a total of 5232 cases had been confirmed amongst the 1.9 million inhabitants of the city of Hamburg, which constituted one of the highest infection rates in Germany (Robert Koch Institute, 2020). At the University Medical Center Hamburg-Eppendorf, more than 170 COVID-19 patients were treated during the same time period. The aim of our study was to longitudinally assess the SARS-CoV-2 seroprevalence and seroconversion rates in hospital workers at our tertiary care center at three time points during the early phase of the epidemic.

2. Material and methods

2.1. Recruitment of the study population

The study protocol was reviewed and approved by the Ethics Committee of the Medical Council of Hamburg (PV 7298). Participants were recruited by informing employees of the University Medical Center Hamburg-Eppendorf both in person and via an internal email newsletter and written informed consent was obtained by all study participants

* Corresponding author. I. Department of Internal Medicine, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany.
E-mail address: t.brehm@uke.de (T.T. Brehm).
[1] contributed equally.

https://doi.org/10.1016/j.ijheh.2020.113671
Received 1 September 2020; Received in revised form 20 November 2020; Accepted 25 November 2020
Available online 30 November 2020

prior to recruitment. Both at baseline (Screening period 1 (SP 1): 20th March – 9th April) and during the first (SP 2: 20th April 20 – 8th May) and second (SP 3: 22nd June – 17th July) follow-up visit serum samples were drawn. Study participants were contacted by phone and email to remind them of the follow-up visits to reduce selection bias.

2.2. Data collection

On the day of recruitment, we collected demographic and general work-related data using a standardized questionnaire. Hospital workers were classified as HCWs if they reported regular occupational contact with patients. Study participants were asked to report only their main clinical role. Since some employees work in different departments in parallel, study participants could to assign themselves to different locations of work. In addition, we used another questionnaire to assess known and possible past contact to COVID-19 patients with or without sufficient personal protective equipment (PPE) as well as the presence of symptoms during the prior 4 weeks at SP 1 and SP 2 and during the prior 8 weeks at SP 3 respectively. Fever, cough, and dyspnea were classified as typical symptoms while rhinorrhea, sore throat, headache, stomach pain, joint and muscle pain, nausea, and diarrhea and were considered as uncharacteristic symptoms. Questionnaires were available both paper-based and on online REDcap electronic data capture tools hosted at our center. Study participants with positive SARS-CoV-2 IgG were contacted again by phone to assess the probable source of infection and to ask if nasopharyngeal swabs for RT-PCR had been performed.

2.3. ELISA

At all three screening periods, serum samples were drawn from all study participants. A semi-quantitative SARS-CoV-2 immunoglobin (Ig) G enzyme-linked immunosorbent assay (ELISA) targeting the S1-Domain of the S-protein spike protein subunit (Euroimmun Medizinische Labordiagnostika, Lübeck, Germany) was performed according to the manufacturer's protocol. Results are evaluated by calculation of a ratio of the extinction of the serum sample over the extinction of the calibrator. According to the manufacturer, a ratio <0.8 is considered negative, ≥ 0.8 and < 1.1 borderline, and ≥ 1.1 positive. The manufacturer reports a specificity 99.6% and sensitivity of 94.4% 10 days after the onset of symptoms. However, we used a more stringent cut-off value of ≥ 1.5 for positive results, which has been shown to display a specificity of 100% (Pflüger et al., 2020) to account for the low prevalence environment. Given the longitudinal character of our study, we do not expect a relevant loss of test sensitivity. Study participants were classified as seropositive in all future sampling periods if they had an IgG antibody titer ≥ 1.5 at least once, even if antibody titers waned over time.

2.4. Statistical analyses

Seroprevalence of antibodies against SARS-CoV-2 at different time points was calculated as proportions with 95% confidence interval (CI). Continuous variables were expressed as mean and interquartile range (IQR) and compared with student's t-test. Categorical variables were expressed as number (%) and compared by Chi-square or Fisher's exact test. P values less than 0.05 were considered statistically significant. Figures were designed using GraphPad Prism version 8 for macOS (GraphPad Software, La Jolla, California, USA). All other analyses were performed using SPSS, version 21.0 (IBM Corp., Armonk, New York, USA).

3. Results

3.1. Characterization of the study population

A total of 1253 individuals were included during SP 1, which represent around 11% of all employees at the University Medical Center Hamburg-Eppendorf. The majority of study participants were HCWs (n = 1026) who were recruited at different departments of our hospital including regular wards (n = 332), outpatient clinics (n = 234), the intensive care unit (n = 130), and the emergency department (n = 63) (Table 1). The remainder were researchers, administrative staff, or belonged to other occupational employee groups not directly involved in patient care. While we did not recruit a strictly representative sample of hospital workers at our institution, the relative proportion of different professional groups in our study cohort (35.4% nurses, 21.9% medical doctors, 42.7% others) fairly well matched the overall distribution at our hospital (30.1% nurses, 25.8% medical doctors, 44.1% others). Information on whether study participants had been knowingly in close contact with COVID-19 patients was available for 1163 (92.8%) study participants (Table 2). Of those, 25.4% (n = 295) reported to having been in direct contact with COVID-19 cases. A total of 226 individuals were only involved in the care of patients with diagnosed SARS-CoV-2 infections and were thus equipped with appropriate PPE. Another 69 study participants had contact with patients who were later diagnosed with COVID-19, which is why these HCWs did not wear adequate PPE at the time of exposure. Occupational contact to infected colleagues was reported by 12.0% (n = 140), community contact by 3.0% (n = 35) of participants.

3.2. Serological results

A total of 23 (1.8%) participants missed the first follow-up visit, but all study participants attended the second follow up visit. At the initial screening period, 0.8% (n = 10, 95% confidence interval [CI] 0.3 to 1.3) of the study participants were found to be SARS-CoV-2 seropositive (Fig. 1). Another ten individuals showed seroconversion at SP 2, giving a seroprevalence of 1.6% (95% CI 0.9 to 2.3). At SP 3, two more individuals had developed anti-SARS-CoV-2 IgG antibodies, thus the overall seroprevalence was 1.8% (n = 22, 95% CI 1.0 to 2.5) by the end of the study.

No significant differences in age or the relative distribution of sex were observed between seronegative and seropositive study participants. Seropositivity was increased in physicians (3.3% vs 1.3%; p = 0,04), while none of the other professional categories and none of the different locations of work showed any association with the presence of

Table 1
Characterization of the study population.

	Total	Seropositive	p
Total, n	1253	22 (1.8)	
Age			
Median	36	33	0.33
IQR	29; 48	29; 40	
Sex			
Male	308	8 (2.5)	0.21
Female	934	14 (1.5)	0.23
Diverse	11	0 (0)	1.0
Clinical role, n (%)			
Nurse	444	10 (2.3)	0.37
Physician	275	9 (3.3)	0.04
Medical technician	105	0 (0)	0.25
Medical student	73	2 (2.7)	0.37
Physiotherapist	15	0 (0)	1.0
Other HCW	114	1 (0.3)	0.25
Non-HCW	227	1 (0.4)	0.16
Location of work, n (%)			
Regular ward	332	9 (2.7)	0.14
Outpatient clinic	234	5 (2.1)	0.58
Intensive care unit	130	2 (1.5)	1.0
Operating room	111	1 (0.9)	0.71
Emergency department	63	1 (1.6)	1.0
Other	354	4 (1.1)	0.35

Baseline characterization of the study population and association of variables between seropositive and seronegative individuals; IQR = interquartile range.

Table 2
Contact to COVID-19 cases reported by the study participants.

	Total	Seropositive, n (%)	p
Any known contact	417	18 (4.3)	<0.001
- Patients, only with PPE	226	5 (2.2)	0.79
- Patients, without PPE	69	6 (8.7)	0.001
- Colleagues	140	3 (2.1)	0.74
- Community	35	4 (11.4)	0.003

Contact to COVID-19 cases reported by seropositive and -negative study participants and association of variables between seropositive and seronegative individuals. Information on contact to COVID-19 cases was available for 1163 study participants.

antibodies to SARS-CoV-2.

The most probable source of infection could be identified for the majority of seropositive study participants (Table 2). Six HCWs presumably got infected when caring for pre-symptomatic patients who were later diagnosed with COVID-19 and had thus not used appropriate PPE at the time of exposure. Five HCWs most likely contracted SARS-CoV-2 while caring for COVID-19 patients with adequate PPE. Other probable infection routes in our cohort were close contact with infected colleagues (n = 3), community transmission in high-risk regions (n = 3), and infected household contacts (n = 1). The remainder of four seropositive employees did not report any known contact to COVID-19 patients, and the probable source of infection could not be established. Seroprevalence was higher in study participants who reported contact with COVID-19 patients (4.3%; p < 0.001). The subgroup of study participants who reported caring for pre-symptomatic COVID-19 and did not use appropriate PPE at the time of exposure showed the highest seroprevalence (5.4%; p = 0.001). Of note, no increased rate of anti-SARS-CoV-2 IgG-antibodies was observed in the subgroup of hospital workers who reported providing direct clinical care to diagnosed COVID-19 patients with appropriate PPE (2.2%; p = 0.79) or contact to SARS-CoV-2 infected colleagues (2.1%; n = 0.74). However, study participants who reported delivering care for pre-symptomatic COVID-19 and did therefore not use appropriate PPE at the time of exposure had a significantly higher seroprevalence (5.4% vs 27.3%; p = 0.001). Moreover, community contact with COVID-19 cases was a significant risk factor for seropositivity in our study cohort (2.7% vs 18.2%, p = 0.003).

3.3. Information on nasopharyngeal swabs of seropositive individuals

Out of all seropositive study participants, a total of 40.9% (n = 9) had been diagnosed with SARS-CoV-2 infection by RT-PCR and thus had been placed under quarantine during the study period. Another six seropositive HCWs reported that they had at least one nasopharyngeal swab performed due to symptoms or close contact with a COVID-19 case but were tested negative at that time. The remainder of seven seropositive individuals had not been tested by RT-PCR at all.

3.4. Symptoms

Complete information on symptoms throughout the study period was provided by 1094 (87.3%) of the study participants (Table 2). The overall rate of hospital workers who reported typical SARS-CoV-2 symptoms did not differ between seronegative and seropositive individuals (54.6% vs 63.6%; p = 0.52) (Table 3). Only fever (9.9% vs. 36.4%; p = 0.001) and muscle aches (22.6% vs 59.1%; p < 0.001) were more commonly reported in hospital workers with anti-SARS-CoV-2 antibodies.

4. Discussion

The first results of our ongoing study demonstrate a low overall SARS-CoV-2 seroprevalence of 1.8% (n = 22, 95% CI 1.0 to 2.5) in a

Table 3
Symptoms reported by seropositive and -negative study participants.

	Seronegative	Seropositive	p
Typical symptoms, n (%)	597 (54.6)	14 (63.6)	0.52
- Fever, n (%)	*108 (9.9)*	*8 (36.4)*	*0.001*
- Cough, n (%)	*511 (46.7)*	*11 (50.0)*	*0.83*
- Dyspnea, n (%)	*254 (23.2)*	*8 (36.4)*	*0.12*
Rhinorrhea, n (%)	736 (67.3)	16 (72.7)	0.65
Sore throat, n (%)	576 (52.7)	11 (50.0)	0.83
Headache, n (%)	717 (65.5)	15 (68.2)	1.0
Abdominal pain, n (%)	297 (26.9)	8 (36.4)	0.34
Muscle aches, n (%)	247 (22.6)	13 (59.1)	<0.001
Nausea, n (%)	211 (19.3)	8 (36.4)	0.06
Diarrhea, n (%)	299 (27.3)	8 (36.4)	0.34

Respective symptoms as stated by the study participants at SP 1 or SP 2 for the preceding month and at SP 3 for the preceding 2 months. Complete information was provided by a total of 1094 seronegative and 22 seropositive participants.

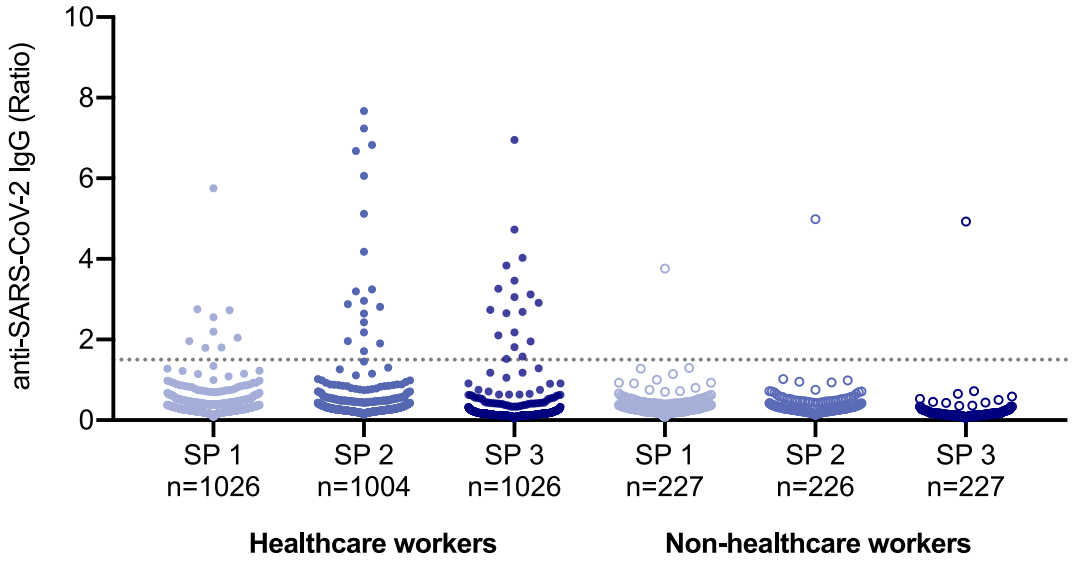


Fig. 1. Serologic results of the study population.

cohort of 1253 hospital workers at the University Medical Center Hamburg-Eppendorf at the early phase of the COVID-19 epidemic. Remarkably, seroprevalence was not significantly increased in HCWs directly involved in patient care compared to other hospital workers. These low overall infection rates are despite more than 170 patients with SARS-CoV-2 infections that had been treated at our center during the study period. Notwithstanding considerable exposure to COVID-19, infection rates in our study were substantially lower than in HCWs from Belgium (12.6%) (Martin et al., 2020), Spain (11.2%) (Garcia--Basteiro et al., 2020), Italy (14.4%) (Sotgiu et al., 2020), Sweden (19.1%) (Rudberg et al., 2020) the United Kingdom (10.8%–43.5%) (Martin et al., 2020; Houlihan et al., 2020) and the United Stated of America (7.6%–13.7%) (Moscola et al., 2020; Stubblefield et al., 2020). Those marked differences in infection rates may be partially explained by local differences in community transmission: population seroprevalence has been estimated to be 8.0% in Belgium, 5.5% in Spain, 4.6% in Italy, 3.7% in Sweden, 5.1% in the United Kingdom, and only 0.85% in Germany by May 2020 (Flaxman et al., 2020). However, our results confirm the findings of a previous cross-sectional study that detected anti-SARS-CoV-2 in 1.6% (n = 5) of 316 HCWs at another German tertiary care hospital (Korth et al., 2020) and are likely also an indicator that local infection control interventions were effective in protecting HCWs from COVID-19. A timeline of the most important infection control measures that were undertaken at our institution and in Hamburg, Germany are listed in Table 4.

It has been shown that stringent use of appropriate PPE by HCWs when providing direct care to patients with suspected or confirmed SARS-CoV-2 infections effectively prevents transmission (Wang et al., 2020). Indeed, we did not observe a significant association between the rate of anti-SARS-CoV-2 IgG-antibodies and contact with COVID-19 patients when using appropriate PPE. However, five HCWs in our study presumably contracted SARS-CoV-2 despite the use of appropriate PPE when caring for COVID-19 patients. We cannot determine with absolute certainty whether those HCWs contracted COVID-19 while caring for infected patients, and whether infection control standards were maintained by these individuals at all times. Our observations nevertheless demonstrate that awareness for SARS-CoV-2 infection is crucial even when appropriate PPE is used.

Moreover, another six HCWs presumably contracted SARS-CoV-2 infections when delivering care for pre-symptomatic patients who were not suspected COVID-19 cases at the time of exposure. Indeed, such contacts with yet undiagnosed patients without appropriate PPE were associated with seropositivity. Our observations are in line with the results of previous studies, which have demonstrated that pre-symptomatic patients play a pivotal role in both community and healthcare facility associated transmission of SARS-CoV-2 (Arons et al., 2020). To detect those pre-symptomatic patients pre-symptomat, universal admission screening for all hospitalized patients has been suggested (Sutton et al., 2020) and was also established at our institution by 20 April.

While exposure- and symptom-based SARS-CoV-2 RT-PCR is the mainstay of HCW surveillance, previous studies have demonstrated potential challenges and limitations of symptom-based screening efforts due to a lack of both sensitivity and specificity of symptoms suggestive of COVID-19 (Black et al., 2020). Our findings highlight those limitations. Indeed, only 68.2% of seropositive study participants had prior RT-PCR testing for SARS-CoV-2, and only 40.9% had a positive test result. Fever, cough, or dyspnea - symptoms generally considered suggestive of COVID 19 – were reported by only 63.3% of individuals with anti-SARS-CoV-2 antibodies. On the other hand, 54.6% of seronegative individuals reported at least one of those symptoms. Fever, but not cough or dyspnea were significantly associated with the detection of anti-SARS-CoV-2 antibodies. While testing of only symptomatic HCWs might miss a substantial number of COVID-19 cases, expanding RT-PCR surveillance to all asymptomatic HCWs might be a powerful, but resource- and cost-intensive tool to reducing the risk of nosocomial transmission.

Table 4
Infection control interventions at the University Medical Center Hamburg-Eppendorf and in Hamburg, Germany.

Interventions at the University Medical Center Hamburg-Eppendorf	
27 February	Opening of an on-campus COVID-19 testing clinic
11 March	Cancellation of all meetings not directly related to patient care
18 March	Mandatory wearing of face masks in the emergency department
19 March	Implementation of visitor restrictions
22 March	Mandatory wearing of face masks in all clinical settings
10 April	Mandatory wearing of FFP2 masks at all oncology and hematology wards and clinics
20 April	Universal RT-PCR admission screening of patients
11 May	Universal RT-PCR screening of employees caring for COVID-19 patients or vulnerable patients
Interventions in Hamburg, Germany	
16 March	Closure of educational facilities
22 March	Stay at home order
27 April	Mandatory wearing of face masks

Infection control interventions at the University Medical Center Hamburg-Eppendorf and in Hamburg, Germany.

Our study is subject to a number of limitations. Firstly, since various infection control measures were sequentially adopted and the study did not include a control group, we are not able to answer the question which of the described interventions are the most effective in preventing healthcare transmission. However, our results suggest that the overall approach in our institution was effective in mitigating SARS-CoV-2 transmission to HCWs. This is also supported by the fact that after the adoption of all infection control interventions, only two more seroconversions were observed in the two months between SP 2 and SP 3. Secondly, when analyzing serological data, we did not use the cut-off $\geq$1.1 as provided by the manufacturer, but an optimized IgG ratio of $\geq$1.5 to ensure maximum specificity of test results. Given the longitudinal character of our study and the expected increase of IgG ratios over time in infected individuals, we did not expect any relevant loss of test sensitivity. When analyzing serological data with the cut-off value of $\geq$1.1, the overall seroprevalence was 2.8% (n = 36). Thirdly, we did not recruit a strictly representative sample of hospital workers at our institution and participation was voluntary, which limits the overall generalizability of our results.

5. Conclusions

In conclusion, the SARS-CoV-2 seroprevalence of hospital workers at our tertiary care hospital in Germany was 1.8% during the early phase of the COVID-19 pandemic. Local infection control interventions appear to be generally effective in mitigating nosocomial transmission of SARS-CoV-2. However, our study also highlights the challenges and limitations of those measures. Awareness for SARS-CoV-2 infections remains crucial for HCWs on COVID-19 wards as well as for other hospital workers without known contact to SARS-CoV-2 infected patients.

Declaration of competing interest

AWL had a material transfer agreement with Euroimmun GmbH. Ten ELISA plates as well as technical help were provided by the company. All other authors declare no support from any organisation for the submitted work. All authors declare no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgments

We thank Sabrina Kreß, Jennifer Wigger, Martina Schulz, Corinna Eggers, Angelika Schmidt, Silke Kummer, Robin Woost, Nils Dittberner and Marcus Wurlitzer for excellent technical assistance. We thank all study participants and departments of the University Medical Center Hamburg-Eppendorf for active participation in the study. TTB, MML, ML, JKK, JSzW and AWL are funded by the German Center for Infection Research (DZIF). MMA is funded by the DZIF TTU 1.702 and the DZIF TTU 1.921 (First Steps to Fight SARS-CoV-2 (FITS)).

References

Arons, M.M., Hatfield, K.M., Reddy, S.C., Kimball, A., James, A., Jacobs, J.R., Taylor, J., Spicer, K., Bardossy, A.C., Oakley, L.P., Tanwar, S., Dyal, J.W., Harney, J., Chisty, Z., Bell, J.M., Methner, M., Paul, P., Carlson, C.M., McLaughlin, H.P., Thornburg, N., Tong, S., Tamin, A., Tao, Y., Uehara, A., Harcourt, J., Clark, S., Brostrom-Smith, C., Page, L.C., Kay, M., Lewis, J., Montgomery, P., Stone, N.D., Clark, T.A., Honein, M. A., Duchin, J.S., Jernigan, J.A., 2020. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. N. Engl. J. Med. 382, 2081–2090.

Black, J.R.M., Bailey, C., Przewrocka, J., Dijkstra, K.K., Swanton, C., 2020. COVID-19: the case for health-care worker screening to prevent hospital transmission. Lancet 395, 1418–1420.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H.J.T., Mellan, T.A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J.W., Monod, M., Ghani, A.C., Donnelly, C. A., Riley, S., Vollmer, M.A.C., Ferguson, N.M., Okell, L.C., Bhatt, S., 2020. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 584, 257–261.

Garcia-Basteiro, A.L., Moncunill, G., Tortajada, M., Vidal, M., Guinovart, C., Jiménez, A., Santano, R., Sanz, S., Méndez, S., Llupià, A., Aguilar, R., Alonso, S., Barrios, D., Carolis, C., Cisteró, P., Chóliz, E., Cruz, A., Fochs, S., Jairoce, C., Hecht, J., Lamoglia, M., Martínez, M.J., Mitchell, R.A., Ortega, N., Pey, N., Puyol, L., Ribes, M., Rosell, N., Sotomayor, P., Torres, S., Williams, S., Barroso, S., Vilella, A., Muñoz, J., Trilla, A., Varela, P., Mayor, A., Dobaño, C., 2020. Seroprevalence of antibodies against SARS-CoV-2 among health care workers in a large Spanish reference hospital. Nat. Commun. 11, 3500.

Houlihan, C.F., Vora, N., Byrne, T., Lewer, D., Kelly, G., Heaney, J., Gandhi, S., Spyer, M. J., Beale, R., Cherepanov, P., Moore, D., Gilson, R., Gamblin, S., Kassiotis, G., McCoy, L.E., Swanton, C., Crick, C.-C., Hayward, A., Nastouli, E., Investigators, S., 2020. Pandemic peak SARS-CoV-2 infection and seroconversion rates in London frontline health-care workers. Lancet 396, e6–e7.

Robert Koch Institute, 2020. Coronavirus disease 2019 (COVID-19) daily situation report of the robert Koch Institute, 17.07.2020, 01.09.20. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-07-17-en.pdf?_blob=publicationFile.

Korth, J., Wilde, B., Dolff, S., Anastasiou, O.E., Krawczyk, A., Jahn, M., Cordes, S., Ross, B., Esser, S., Lindemann, M., Kribben, A., Dittmer, U., Witzke, O., Herrmann, A., 2020. SARS-CoV-2-specific antibody detection in healthcare workers in Germany with direct contact to COVID-19 patients. J. Clin. Virol. 128, 104437.

Martin, C., Montesinos, I., Dauby, N., Gilles, C., Dahma, H., Van Den Wijngaert, S., De Wit, S., Delforge, M., Clumeck, N., Vandenberg, O., 2020. Dynamics of SARS-CoV-2 RT-PCR positivity and seroprevalence among high-risk health care workers and hospital staff. J. Hosp. Infect. 106 (1), 102–106.

Moscola, J., Sembajwe, G., Jarrett, M., Farber, B., Chang, T., McGinn, T., Davidson, K.W., 2020. Prevalence of SARS-CoV-2 antibodies in health care personnel in the New York city area. J. Am. Med. Assoc. 324, 893–895.

Pflüger, L.S., Bannasch, J.H., Brehm, T.T., Pfefferle, S., Hoffmann, A., Nörz, D., van der Meirschen, M., Kluge, S., Haddad, M., Pischke, S., Hiller, J., Addo, M.M., Lohse, A. W., Wiesch, J.S.z., Peine, S., Aepfelbacher, M., Lütgehetmann, M., 2020. Clinical evaluation of five different automated SARS-CoV-2 serology assays in a cohort of hospitalized COVID-19 patients. J. Clin. Virol. 130, 104549 [published online ahead of print, 2020 Jul 17].

Razzini, K., Castrica, M., Menchetti, L., Maggi, L., Negroni, L., Orfeo, N.V., Pizzoccheri, A., Stocco, M., Muttini, S., Balzaretti, C.M., 2020. SARS-CoV-2 RNA detection in the air and on surfaces in the COVID-19 ward of a hospital in Milan, Italy. Sci. Total Environ. 742, 140540.

Rudberg, A.S., Havervall, S., Månberg, A., Jernbom Falk, A., Aguilera, K., Ng, H., Gabrielsson, L., Salomonsson, A.C., Hanke, L., Murrell, B., McInerney, G., Olofsson, J., Andersson, E., Hellström, C., Bayati, S., Bergström, S., Pin, E., Sjöberg, R., Tegel, H., Hedhammar, M., Phillipson, M., Nilsson, P., Hober, S., Thålin, C., 2020. SARS-CoV-2 exposure, symptoms and seroprevalence in healthcare workers in Sweden. Nat. Commun. 11, 5064.

Sotgiu, G., Barassi, A., Miozzo, M., Saderi, L., Piana, A., Orfeo, N., Colosio, C., Felisati, G., Davì, M., Gerli, A.G., Centanni, S., 2020. SARS-CoV-2 specific serological pattern in healthcare workers of an Italian COVID-19 forefront hospital. BMC Pulm. Med. 20, 203.

Stubblefield, W.B., Talbot, H.K., Feldstein, L., Tenforde, M.W., Rasheed, M.A.U., Mills, L., Lester, S.N., Freeman, B., Thornburg, N.J., Jones, I.D., Ward, M.J., Lindsell, C.J., Baughman, A., Halasa, N., Grijalva, C.G., Rice, T.W., Patel, M.M., Self, W.H., 2020. Seroprevalence of SARS-CoV-2 among frontline healthcare personnel during the first month of caring for COVID-19 patients - nashville, Tennessee. Clin. Infect. Dis., ciaa936 [published online ahead of print, 2020 Nov 19].

Sutton, D., Fuchs, K., D'Alton, M., Goffman, D., 2020. Universal screening for SARS-CoV-2 in women admitted for delivery. N. Engl. J. Med. 382, 2163–2164.

Wang, X., Ferro, E.G., Zhou, G., Hashimoto, D., Bhatt, D.L., 2020. Association between universal masking in a health care system and SARS-CoV-2 positivity among health care workers. J. Am. Med. Assoc. 324 (7), 703–704 [published online ahead of print, 2020 Jul 14].

Zhan, M., Qin, Y., Xue, X., Zhu, S., 2020. Death from covid-19 of 23 health care workers in China. N. Engl. J. Med. 382 (23), 2267–2268.

International Journal of Hygiene and Environmental Health 232 (2021) 113689

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

Urinary phthalates and body mass index in preschool children: The MIREC Child Development Plus study

Jillian Ashley-Martin [a], Linda Dodds [b], Tye E. Arbuckle [a], Bruce Lanphear [c], Gina Muckle [d], Warren G. Foster [e], Pierre Ayotte [f], Angelika Zidek [g], Elizabeth Asztalos [h], Maryse F. Bouchard [i], Stefan Kuhle [b,*]

[a] Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, K1A 0K9, Canada
[b] Perinatal Epidemiology Research Unit, Dalhousie University. 5850-5980 University Ave, Halifax, NS, B3K 6R8, Canada
[c] Faculty of Health Sciences. Simon Fraser University. 8888 University Drive Burnaby, B.C, V5A 1S6, Canada
[d] Laval University and CHU Research Centre, Quebec City, Quebec, Canada
[e] McMaster University Health Sciences Center, Hamilton, Ontario, Canada
[f] Institut National de Santé Publique du Québec (INSPQ), Quebec, Canada
[g] Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario, Canada
[h] University of Toronto, Toronto, ON, Canada
[i] Department of Environmental and Occupational Health, University of Montreal, Canada

ARTICLE INFO

Keywords:
Phthalates
Body mass index
Children
Biomonitoring
Skinfolds
Cohort study
Chemical mixtures

ABSTRACT

Childhood exposure to phthalates, a class of chemicals with known reproductive and developmental effects, has been hypothesized to increase the risk of obesity, but this association is not well understood in preschool children. We examined the association between urinary concentrations of phthalate metabolites and concurrently measured body mass index (BMI) and skinfolds among children between the ages of two and five years.

We collected anthropometric measures and biomonitoring data on approximately 200 children enrolled in the Maternal-Infant Research on Environmental Chemicals Child Development Plus study. We measured 22 phthalate metabolites in children's urine and used the 19 metabolites detected in at least 40% of samples. Our primary outcome was BMI z-scores calculated using the World Health Organization growth standards. Skinfold z-scores were secondary outcomes. We used multivariable linear regression to evaluate the association between tertiles of phthalate concentrations and each anthropometric measure. We also used weighted quantile sum regression to identify priority exposures of concern.

Our analytic sample included 189 singleton-born children with complete anthropometric data. Children with concentrations of the parent compound di-n-butyl phthalate (∑DnBP) in the third tertile had 0.475 (95% CI: 0.068, 0.883) higher BMI z-scores than those in the lower tertile. ∑DnBP was identified as a priority exposure in the weighted quantile sum regression BMI model.

In this population of Canadian preschool aged children, we identified DnBP as a potential chemical of concern in regard to childhood obesity. Future research with serial phthalate measurements and anthropometric measurements in young children will help confirm these findings.

1. Introduction[1]

The global prevalence of childhood obesity has increased dramatically in recent decades (GBD Obesity Collaboration, 2014). Obesity is a multifactorial condition rather than a simple caloric imbalance (Keith et al., 2006; Romano et al., 2014). Early life exposure to endocrine-disrupting chemicals, such as phthalates, is hypothesized to be one of the pathways leading to the development of childhood obesity (Grun and Blumberg, 2009; Heindel et al., 2015). Experimental data has shown that phthalate exposure may induce weight gain by stimulating

* Corresponding author. IWK Health Centre, PO Box 9700 5850-5980 University Ave, Halifax, NS B3K 6R8, Canada.
E-mail addresses: jillian.ashley-martin@canada.ca (J. Ashley-Martin), l.dodds@dal.ca (L. Dodds), Tye.arbuckle@canada.ca (T.E. Arbuckle), blanphear@sfu.ca (B. Lanphear), gina.muckle@psy.ulaval.ca (G. Muckle), fosterw@mcmaster.ca (W.G. Foster), pierre.ayotte@inspq.qc.ca (P. Ayotte), Angelika.zidek@canada.ca (A. Zidek), Elizabeth.asztalos@sunnybrook.ca (E. Asztalos), Maryse.bouchard@umontreal.ca (M.F. Bouchard), stefan.kuhle@dal.ca (S. Kuhle).

https://doi.org/10.1016/j.ijheh.2021.113689
Received 13 August 2020; Received in revised form 30 December 2020; Accepted 30 December 2020
Available online 11 January 2021

peroxisome proliferator-activated receptors (PPAR), decreasing androgenic activity, and increasing adipocyte number and size (Biemann et al., 2012; Casals-Casas and Desvergne, 2011; Desvergne et al., 2009; Hao et al., 2012; Hurst and Waxman, 2003). Findings from epidemiological studies in adults and children have been equivocal (Goodman et al., 2014), and hindered by the methodological challenges of multiple correlated exposures and inability to establish temporality. Furthermore, insufficient research has been dedicated to measuring the association between phthalate exposure and anthropometry during windows of developmental susceptibility such as early childhood.

Phthalates are synthetic organic chemicals that are commonly used as plasticizers, solvents and additives in numerous consumer products (Wang et al., 2019). Low molecular weight (LMW) phthalates, which contain between one to four carbon chains, are commonly found in personal care products and cosmetics, adhesives, enteric coated medications, and toys. High molecular weight (HMW) phthalates, which contain greater than five carbon chains, are commonly found in plastics, food packaging, and polyvinyl products (Wang et al., 2019; Wittassek et al., 2011). These chemicals readily leach out of these products and are absorbed by humans via ingestion or dermal absorption. Ingestion of both food and dust is one of the primary routes of exposure to many phthalates in young children (Weiss et al., 2018; Wormuth et al., 2006); the phthalates are then converted into monoesters or oxidative metabolites in the gastrointestinal tract, conjugated with glucuronide or sulfate and excreted via urine (Wang et al., 2019; Wittassek et al., 2011).

Young children tend to have higher phthalate exposure than adults due to their specific behaviors (e.g., mouthing and crawling), higher food requirements per unit body mass and ventilation rate. Furthermore, their rapid growth and under-developed detoxification mechanisms put young children at increased risk to the potential adverse effects of phthalate exposures (Braun, 2017; Gow et al., 2001; Wormuth et al., 2006). Still, the evidence from human studies on the association between phthalate exposure and measures of body fat in early childhood is scant and conflicting (Boas et al., 2010; Deierlein et al., 2016; Goodman et al., 2014; Maresca et al., 2016; Shoaff et al., 2017; Vafeiadi et al., 2018; Zhang et al., 2014).

In a 2014 systematic review of seven studies, including six cross-sectional studies, Goodman et al. concluded that the evidence was insufficient to determine whether childhood phthalate exposure is associated with childhood obesity (Goodman et al., 2014). More recently, longitudinal studies remain inconsistent with positive (Deierlein et al., 2016; Shoaff et al., 2017; Vafeiadi et al., 2018), null (Maresca et al., 2016; Shoaff et al., 2017), and inverse (Vafeiadi et al., 2018; Zhang et al., 2014) associations being reported. Of note, only three identified studies have examined the potential obesity-related effects of phthalate exposure in children younger than five years of age (Boas et al., 2010; Shoaff et al., 2017; Vafeiadi et al., 2018). Statistically significant positive associations in these studies were restricted to a specific subgroup defined by the child's age (Shoaff et al., 2017) and sex (Vafeiadi et al., 2018).

The objective of the present study was to determine the association between childhood urinary phthalate concentrations and concurrently measured body mass index (BMI) and skinfold measurements in a population of Canadian preschool children with extensive biomonitoring data. Given previous epidemiological (Vafeiadi et al., 2018; Yang et al., 2017; Zhang et al., 2014) and experimental (Grun and Blumberg, 2009; Hao et al., 2012) evidence of sex-specific obesity related effects of phthalate exposures, our secondary objective was to explore whether the association between phthalate exposure and anthropometric outcomes differed between boys and girls. To account for the challenges of analyzing multiple correlated exposures, we supplemented traditional regression models with weighted quantile sum regression (Carrico and Gennings, 2015; Tanner et al., 2019).

2. Methods

2.1. Study population

The Maternal-Infant Research on Environmental Chemicals (MIREC) study is a national-level pregnancy cohort of 2001 women from 10 Canadian cities (Vancouver, Edmonton, Winnipeg, Sudbury, Ottawa, Kingston, Toronto, Hamilton, Montreal, and Halifax) (Arbuckle et al., 2013). Briefly, women were recruited between 2008 and 2011 in their first trimester and followed through delivery. Compared to the national average, MIREC participants tended be of higher socioeconomic status, and have a lower rate of smoking, and lower BMI (Arbuckle et al., 2013). A follow-up study of child growth, neurodevelopment, and exposure biomonitoring was conducted in six study sites between 2013 and 2015 (Vancouver, Toronto, Hamilton, Kingston, Montreal, and Halifax) (MIREC-Child Development Plus (MIREC-CD Plus)). This follow-up group was similar to the overall MIREC population in terms of key characteristics such as pre-pregnancy BMI, smoking, and education. Biomonitoring of phthalate metabolites was conducted on 200 children between ages of 2 and 5 years of age. A spot random urine sample was collected from each participating child, during a home visit, unless the participant requested a clinic visit. Parents were reminded that urine would be collected at this visit and they might want to give their child a beverage before the appointment time. The study was reviewed by the Health Canada Research Ethics Board, CHU de Quebec Ethics Board, as well as by ethics committees at all recruitment sites, and parents signed an informed consent form for their child's participation.

2.2. Exposures: child urinary phthalate metabolite concentrations

We measured 22 phthalate metabolites in children's urine samples that were collected at the time of the biomonitoring visit (Suppl Table 1). To create summary indices of parent compounds, we calculated molar sums for metabolites of parent compounds by dividing metabolites by their molecular weight, summing all metabolites of a specific parent compound and multiplying the sum by 1000. We calculated molar sums for DEHP, DiNP, DnBP, DIDP, DiBP; the corresponding metabolites are shown in Suppl Table 1. Four metabolites (MMP, MEP, MBzP, MCPP) were distinct and treated as individual analytes (Table 1). We excluded three metabolites (MCHP, MOP, MNP) from the analyses that were detectable in less than 40% of samples (Suppl Table 1).

Analysis of the urine samples was performed at the Centre de toxicologie du Québec of the Institut National de Santé Publique du Québec. In brief, following an enzyme deconjugation step, the analytes were extracted using a robotized liquid-liquid extraction unit at pH 3 with a mixture of hexane:ethyl acetate. The extracts were brought to dryness, taken up in a solution of acetonitrile:water (25:75), and analyzed by an ultra-performance liquid chromatograph (Waters Acquity, Waters Limited Mississauga, ON, Canada) and tandem mass spectrometer (Waters Xevo TQ-S with MassLynx software version 4.1) in the multiple reaction monitoring mode with an "electrospray" source operated in the negative mode. Machine readings were used for all values below the limit of detection. To account for heterogeneity in urinary dilution, phthalate concentrations were standardized for urine specific gravity (SG), measured using a refractometer, according to the following formula: $P_c = P_i\ [(SG_m - 1)/(SG_i - 1)]$, where P_c = SG standardized metabolite concentration (μg/ml), P_i = observed metabolite concentration, SG_i = specific gravity of the urine sample, and SG_m = median SG for the cohort (Just et al., 2010). Descriptive statistics are presented for SG-standardized phthalate concentrations. To account for urinary dilution in multivariable models, SG was entered as a covariate with the raw phthalate concentrations. This approach facilitates comparison of crude and unadjusted models.

In the MIREC study, first trimester urinary concentrations of 11 phthalate metabolites were measured and descriptive statistics have been previously reported for the seven metabolites with a sufficient

Table 1
Urinary concentrations of specific gravity-standardized molar sum phthalate metabolite compounds (nmol/L) and individual metabolites (μg/L) among children enrolled in MIREC-CD Plus.

	N	Minimum	25th percentile	Median	75th percentile	Maximum
Parent Compounds						
∑DEHP	187	22.3	106	155	229	2023
∑DiNP	180	0.0030	10.1	19.1	33.6	707
∑DnBP	181	17.3	66.9	100	156	2595
∑DIDP	174	0.750	3.96	6.58	9.61	179
∑DiBP	187	11.1	62.1	89.8	141	1039
Metabolites						
MBzP	187	0.540	3.64	8.34	17.4	201
MEP	187	2.58	7.60	11.5	21.8	1020
MMP	186	0.729	2.68	3.74	5.89	72.5
MCPP	187	0.160	1.34	1.87	3.03	39.2
MCPP	187	0.160	1.34	1.87	3.03	39.2
2-OH-MiBP	187	1.13	5.72	8.58	12.7	76.8
MCMHP	187	0.009	2.77	4.01	6.39	44.3
MCiNP	179	0.251	0.999	1.43	2.16	885
MCiOP	184	0.001	1.02	1.69	3.26	92.3
MEHHP	187	1.79	8.02	12.5	17.8	200
MEHP	187	0.337	1.06	1.86	2.51	69.3
MEOHP	187	1.30	6.36	9.11	14.4	149
MHBP	181	0.493	2.28	3.53	5.27	99.6
MHiDP	182	0.001	0.102	0.290	0.491	39.6
MHiNP	183	0.002	1.13	2.15	3.72	83.1
MOiDP	186	0.001	0.129	0.262	0.576	59.6
MOiNP	187	0.002	0.714	1.39	2.31	34.7
MiBP	187	1.42	8.00	12.2	19.0	194
MiNP	187	0.001	0.137	0.362	0.865	13.4
MnOP	186	0.001	0.001	0.001	0.002	0.284

Abbreviations for all phthalate parent compounds and metabolites are provided in Supplemental table 1.

percentage of samples with detectable concentrations (Arbuckle et al., 2014). In the present study, we assess correlations between maternal and child phthalate concentrations, compare descriptive statistics with the child concentrations, and determine whether adjustment for these prenatal exposures influences results of the associations between child phthalate concentrations and anthropometric measures (Engle and Wolff, 2013). The maternal phthalate metabolites were standardized for SG in the same manner as described for the childhood phthalate metabolites.

2.3. Outcomes: child adiposity

Our primary outcome measure was the body mass index (BMI); secondary outcomes were triceps skinfold (TSF) and subscapular skinfold (SSF) z-scores. Trained research personnel followed a detailed protocol to measure child's weight with a calibrated scale (Seca model 874 [Seca Corporation, Hanover, MD, USA]), standing height using a calibrated stadiometer (Seca model 217 [Seca Corporation, Hanover, MD, USA]), and SSF and TSF using a Lange skinfold caliper (Cambridge Scientific Instruments, Cambridge, MA, USA). All measurements were completed during the home visit and the average of two measured values was used. If the two measures differed by a pre-determined value, research personnel performed a third measurement. A third measurement was performed in 24%, 10%, 3%, and less than 1% of individuals for weight, height, TSF, and SSF respectively. Research personnel were blinded to maternal and child phthalate exposure concentrations. Age- and sex-specific z-scores for BMI, TSF, and SSF were calculated for each child based on WHO Child Growth Standards (WHO, 2015).

2.4. Statistical analysis

We calculated counts and percentages for the study population. Pearson correlation coefficients were calculated for all log2-transformed summary indices and metabolites measured in children and in mothers. Due to the skewed distributions of all exposures, we log2-transformed all exposures prior to use in regression models or calculation of correlation coefficients.

We used three approaches to evaluate the association between phthalates and each anthropometric outcome. First, we developed individual linear regression models for tertiles of each phthalate parent compound (DEHP, DiNP, DnBP, DIDP, DiBP) and metabolite (MCPP, MBzP, MMP, MEP) measured at the time of the MIREC-CD Plus study visit. We categorized each exposure variable in this way to account for potential non-linear associations. Separate models were developed for each outcome. Confounders, identified via causal diagrams, included maternal postnatal weight status (Voerman et al., 2019), household income (Garí et al., 2019; Lewin et al., 2017), and maternal age (Lewin et al., 2017) (Suppl Figure 1). We entered these variables into all models, along with urine SG. We assessed effect modification by stratifying models by sex and calculating p values of products terms for sex and each exposure. To account for missingness in the covariates and the exposures, multiple imputation with chained equations was used (m = 30, 30 iterations) (van Buuren and Groothuis-Oudshoorn, 2011).

Second, we used weighted quantile sum (WQS) regression to evaluate the association of the phthalate mixture on each anthropometric outcome using the gWQS package in R (Renzetti et al., 2019). This approach estimates the body burden of the chemical mixture, identifies priority chemicals among a set of correlated exposures and has been shown to have superior performance to regularization techniques such as elastic net, particularly when dealing with multiple correlated exposures (Carrico and Gennings, 2015). A WQS index was created using our primary phthalate exposures (5 parent compounds, 4 metabolites) and each exposure was assigned a weight that reflects the strength of that phthalate-outcome association. Briefly, using 40% of the data designated as the training set, individual exposures were categorized into tertiles and assigned a weight between 0 and 1 using a non-linear modeling optimization algorithm. To increase the stability of the estimates, average weights were calculated based on analysis performed in 75 bootstrapped samples. Due to the unidirectional nature of WQS regression, we constrained the β coefficient of the index to be positive in the bootstrap samples based on our hypothesis that phthalates would be positively associated with childhood anthropometric measures;

however, the coefficients were not constrained to be positive when weights were estimated in the test data. We then modelled the association between the WQS index and each outcome using the data designated as the test set (60%). These models included specific gravity to account for urinary dilution. We report weights for chemicals determined to be of concern based on a weight greater than 11% (100/9 exposures) as this value is higher than expected by chance (Tanner et al, 2019, 2020). To further account for the relatively small sample size and potential instability of results due to division of data into training and test sets, we used repeated holdout validation to randomly partition the data and repeated the WQS estimation 100 times. Based on the distribution of results, we calculated mean weights for each exposure (Tanner et al, 2019, 2020). As the WQS index does not have an intuitive quantitative interpretation, the WQS model was primarily used to qualitatively identify phthalate metabolites that are associated with the child's BMI and skinfold thickness.

The third step of the analysis was to evaluate dose-response relationships using generalized additive models for parent phthalate compounds or phthalate metabolites that were identified as a key exposure in both the linear regression and WQS models. In these models, we trimmed the data to exclude the top and bottom 3% of observations to prevent the curves from being overly influenced by extreme observations.

To assess the potential confounding effect of prenatal phthalate exposure, we conducted a sensitivity analyses adjusting for maternal first trimester urine phthalate concentrations. This analyses was restricted to the phthalates where both maternal and child data was available (MCPP, MBzP, MEP, MBP, ∑DEHP) and was accomplished by adding the log2-transformed maternal phthalate exposure in each model for the corresponding child phthalate. For example, log2-transformed maternal MCPP was added to the model examining the association between childhood MCPP and BMI.

Analyses were conducted in SAS version 9.4 (SAS Institute, Cary, NC, USA) and R version 3.5.2 (R Core Team, 2019).

3. Results

Of the 200 children with phthalate biomonitoring data, 192 were singleton births, and 189 had complete height and weight data and were included in the analyses. Consistent with other reports from the MIREC Study, the study population is of moderate to high socioeconomic status and mothers tended to be, on average, of normal pre-pregnancy BMI, non-smokers, and born in Canada. Children were, on average, 32 months (standard deviation (SD): 2.7) at the time of the MIREC-CD Plus study visit (Table 2). Mean (SD) BMI, TSF, and SSF z-scores were 0.55 (0.87), 1.0 (1.4), and −0.15 (1.6), respectively.

Eighteen out of the twenty-two metabolites measured in MIREC-CD Plus participants were detectable in at least 80% of children (Suppl Figure 1). Descriptive statistics for the five summed parent compounds and four individual metabolites measured in children are reported in Table 1. Descriptive statistics for the metabolites measured in mothers are presented in Suppl Table 2. Median urinary MCPP and MBzP concentrations were nearly twice as high in children compared to their mothers (MCPP mother = 0.94 μg/L, child = 1.9 μg/L; MBzP mother = 4.5 μg/L, child = 8.3 μg/L) whereas the median urinary MBP concentrations in mothers was approximately two-thirds that of the childs (mother = 13 μg/L, child = 19 μg/L). In contrast, the median urinary MEP concentration in the mother was more than double that of the child (mother = 26 μg/L, child = 12 μg/L). Concentrations of certain DEHP metabolites were also higher in children than mothers (MEHHP mother = 9.4 μg/L, child = 12 μg/L; MEOHP mother = 6.5 μg/L, child = 9.4 μg/L) (Table 1 and Suppl 2).

We observed moderate to strong correlations among metabolites and parent compounds in children, with correlation coefficients ranging from 0.387 (MBzP and MEP) to 0.843 (MCPP and ∑DiNP). In contrast, we observed weak correlations between maternal and child phthalate concentrations; correlation coefficients ranged from −0.001 (child ∑DIDP and maternal MBzP) to 0.165 (child MEP and maternal MEP) (Suppl Table 6).

In the adjusted linear regression models (Fig. 1, Suppl Table 3), children with ∑DnBP concentrations in the highest tertile had 0.475 (95% CI: 0.068, 0.883) higher BMI z-scores than those in the lowest tertile; no other phthalate metabolite or parent compound was associated with BMI z-scores in the adjusted models. In the sex-specific models (Suppl Table 6), no statistically significant associations were found, and there was no evidence of effect modification based on a product term p-

Table 2
Characteristics of MIREC-CD Plus study participants (n = 189).

Parental	N	%
Maternal age at time of delivery (years)		
≤29	54	29
30–34	73	39
≥35	62	33
Household income (CAD)		
<50 000	32	17
50 000–100 000	80	43
>100 000	73	40
Maternal Education		
High school or less	4	2.1
College courses of diploma	48	26
University degree	135	72
Maternal Country of Birth		
Canada	152	80
Other	37	20
Maternal pre-pregnancy BMI (kg/m^2)		
<25	115	67
25–29	28	16
≥30	29	17
Maternal postnatal BMI (kg/m^2)		
<25	108	58
25–29	41	22
≥30	38	20
Maternal prenatal smoking		
Current	6	3.2
Former or quit when pregnant	51	27
Never	127	70
Paternal BMI (kg/m^2)		
<25	52	28
25–29	88	47
≥30	46	25
Child		
Age at follow-up (mths) (mean (SD))	32	2.7
Gestational age (wks) (mean (SD))	39	1.4
Sex		
Male	83	44
Female	106	56

Missing income (n = 4), education (n = 2), pre-pregnancy BMI (n = 17), postnatal BMI (n = 2).
Paternal BMI (n = 3).
Abbreviations: BMI body mass index, CAD Canadian dollar, mths months, SD standard deviation, wks weeks.

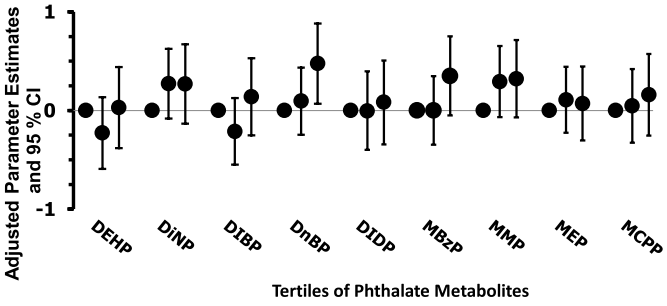


Fig. 1. Adjusted[1] parameter estimates and 95% confidence intervals (CI) for the association between tertiles of childhood urinary phthalate parent compounds (nmol/L), metabolites (μg/L), and BMI z-scores (n = 189).

value <0.05.

In the adjusted TSF z-score models, children with ∑DiNP, MMP, and MEP concentrations in the second tertile had higher TSF z-scores than children with levels in the referent category (∑DiNP 2nd tertile β = 0.623 (95% CI: 0.068, 1.178); MMP 2nd tertile β = 0.678 (95% CI: 0.113, 1.242); MEP 2nd tertile β = 0.551 (95% CI: 0.030, 1.071)) (Fig. 2, Suppl Table 4). We observed no statistically significant associations in the SSF z-score models (Fig. 3, Suppl Table 5). The magnitude and direction of effect in the sex-stratified models were largely consistent with the findings in the overall population (Suppl Tables 7–8). We observed no effect modification by sex in either the TSF or SSF models. We observed no material change in results for each of the outcomes when adjusted for prenatal phthalate exposures (data not shown).

According to the WQS regression models with repeated holdout validation (data not shown), ∑DnBP, ∑DIDP, and ∑DEHP accounted for 23%, 20%, and 13%, respectively, of the weight of the overall index of phthalate exposure in the BMI z-score models (range of weights 0.0–0.23). In the TSF models, MMP, ∑DnBP, and ∑DIDP accounted for 26%, 14%, and 14% of weights (range of weights: 0.03–0.26). The exposures of concern identified in the SSF z-score models were ∑DIDP (24%), ∑DnBP (20%), and MMP (12%) (range of weights: 0.04–0.24).

Based on the results of the categorical and the WQS models, we visualized the association between 1) log2-transformed ∑DnBP and BMI z-scores and 2) log2-transformed MMP and TSF z-scores using generalized additive models. Both models were characterized by a positive linear relationship with considerable imprecision (Fig. 4).

4. Discussion

In this population of Canadian preschool children, we evaluated concentrations of 22 phthalate metabolites in relation to concurrently measured BMI and skinfolds thickness. Compared to children with ∑DnBP concentrations in the referent tertile, those with concentrations in the highest tertile had moderately increased BMI z-scores. These results were consistent with the WQS analysis which identified ∑DnBP as a primary exposure of concern. In the TSF models, MMP was identified as a priority exposure in the WQS analysis and second tertile MMP concentrations were positively associated with TSF z-scores. We did not find effect modification by sex in any of the models. Our study provides a comprehensive assessment of phthalate metabolites in young children. Due to the number of measured metabolites, it is difficult to provide an overall summary comparison with other biomonitoring studies. Out of the LMW metabolites, median MEP and OH-MiBP concentrations in MIREC-CD Plus were lower than observed in children of similar ages enrolled in NHANES (CDC, 2009), GerES V (Schwedler et al., 2020), and a Polish study (Garí et al., 2019). With regard to HMW metabolites, median MEHHP, MEOHP, and MECPP concentrations were higher than observed in NHANES. A comparison table is provided in the supplemental material (Suppl Table 10). Moreover, the short half-life of these metabolites creates challenges to interpreting spot urine samples from different study populations and a more accurate comparison would include 24 h urine samples or repeated measurements over time.

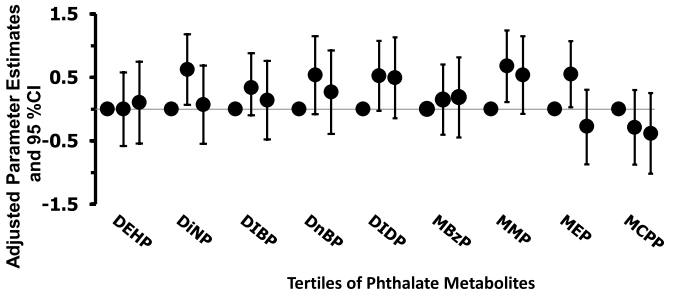


Fig. 2. Adjusted parameter estimates and 95% confidence intervals (CI) for the association between tertiles of childhood urinary phthalate parent compounds (nmol/L), metabolites (μg/L), and TSF z-scores (n = 186).

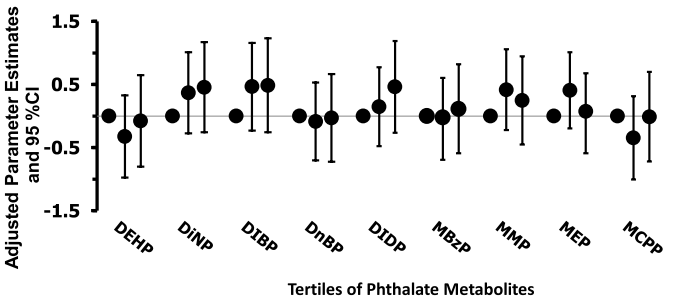


Fig. 3. Adjusted parameter estimates and 95% confidence intervals (CI) for the association between tertiles of childhood urinary phthalate parent compounds (nmol/L), metabolites (μg/L), and SSF z-scores (n = 184).

Children are most commonly exposed to DnBP via ingestion of food due to its ubiquitous presence in food packaging (Environment and Climate Change Canada, 2017; Fisher et al., 2019; Wormuth et al., 2006). In contrast, a primary route and source of exposure to DMP, the parent compound of MMP, is inhalation of indoor air resulting from the use of DMP as a solvent in building materials, such as paints and sealants (Wormuth et al., 2006). Children may also be exposed to DMP via dermal absorption of topical products (diaper creams, lotions), contact with toys (Environment and Climate Change Canada, 2015) (Fisher et al., 2019), and through diet (Correia-Sá et al., 2018). Previous epidemiological studies that examined the potential effects of these phthalates on childhood adiposity are equivocal due, in part, to heterogeneity in timing of measurement of both the exposures and the outcomes as well as different study designs and analytical methods. Despite this heterogeneity, our findings are consistent with one longitudinal and two cross-sectional studies of positive associations between concentrations of ∑DnBP, or its metabolite MBP, and measures of adiposity in children from Chinese and Mexican American populations between the ages 5 and 15 (Harley et al., 2017; Wang et al., 2013; Zhang et al., 2014). Exposure levels in MIREC-CD Plus participants (median MBP = 19 μg/L) tended to be lower than these studies (median MBP 30–43 μg/L in children ages 8–12 years of age) (Wang et al., 2013; Zhang et al., 2014). Our findings are also in line with reported positive associations between summary measures of LMW phthalates (which include both MBP and MMP) and childhood obesity (Buser et al., 2014; Deierlein et al., 2016; Teitelbaum et al., 2012). Some of these associations were only observed in subgroups of sex or weight status (Deierlein et al., 2016; Teitelbaum et al., 2012; Zhang et al., 2014). For example, authors of a study of children from New York City reported that the sum of LMW phthalates was associated with BMI and waist circumference only among children who were overweight (Teitelbaum et al., 2012). In a cross-sectional study of Chinese children, MBP and the sum of LMW phthalates were positively associated with BMI in boys, but not girls, older than 10 years of age (Zhang et al., 2014).

Our observed association between DnBP and increased BMI is further

[1] Abbreviations: BMI body mass index, DEHP di(2-ethylhexyl) phthalate, DiNP Di-iso-nonyl phthalate, DnBP Di-n-butyl phthalate, DIDP Diisodecyl phthalate, DiBP Diisobutyl phthalate, HMW high molecular weight, ICC intraclass correlation coefficient, LMW low molecular weight, MBzP monobenzyl phthalate, MEP monoethyl phthalate, MIREC Maternal-Infant Research on Environmental Chemicals, MMP mono-methyl phthalate, MCPP Mono-3-carboxypropyl phthalate, MOP Mono-n-octyl, MNP mono-isononyl phthalate, MCHP mono-cyclo-hexyl phthalate, PPAR peroxisome proliferator-activated receptors, SG specific gravity, SSF scapular skinfolds, TSF triceps skinfolds, WQS weighted quantile sum.

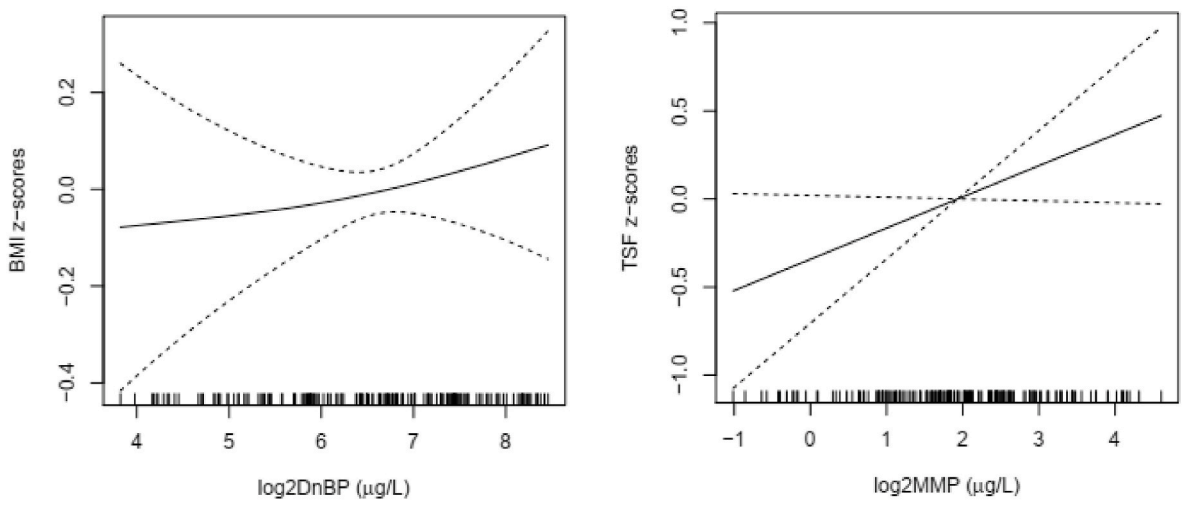


Fig. 4. Generalized additive models for associations and their 95% confidence intervals between log2DnBP and BMI z-scores and log2MMP and TSF z-scores.

supported by experimental data demonstrating that DnBP exhibits developmental and reproductive toxicity (Foster et al., 2000; Wittassek et al., 2011). Specifically, DnBP may promote adiposity via its anti-androgenic effects and subsequent influence on the estrogen/androgen ratio (Mylchreest et al., 1999). Other potential mechanisms underlying an association between DnBP and obesity are exposure-induced changes in testosterone, insulin, and thyroid levels (Mylchreest et al., 1999; Wittassek et al., 2011). Although other phthalates, such as DEHP, may disrupt lipid metabolism and promote adipogenesis via activation of PPAR-alpha and gamma receptors, DnBP may not operate through this pathway (Hurst and Waxman, 2003). MMP may modestly activate the PPAR-alpha pathway in a mouse model but not in humans (Hurst and Waxman, 2003). Further experimental research could help elucidate whether DnBP and DMP, the parent compound of MMP, have greater potential for inducing obesogenic effects than other phthalates.

Our study benefited from the extensive biomonitoring data and the use of two different analytical approaches that yielded consistent results. Although many studies have evaluated associations between phthalates and childhood growth, the present study is one of the first to examine these associations in children as young as 24 months. The extensive biomonitoring data available in the MIREC study also enabled us to evaluate more metabolites than any other identified biomonitoring study and to characterize how exposure patterns differed between pregnancy and early childhood. Consistent with previous data (CDC, 2009; Garí et al., 2019; Wormuth et al., 2006), we found that children had higher concentrations of certain metabolites than their mothers. These notable differences in exposure levels are compatible with the observed weak correlations between the maternal and child phthalate concentrations and confirm the variability in exposure patterns from the in utero to early life time periods. Additionally, this non-uniform pattern in exposure demonstrates the challenge in identifying a biomarker that is reflective of the critical window of exposure for childhood growth. Similar to our findings, authors of a Polish cohort study observed weak correlations between maternal third trimester and child phthalate concentrations (Garí et al., 2019). In contrast, a cross-sectional study of biospecimens from 17 European studies observed strong correlations between maternal and children's phthalate concentrations (Den Hond et al., 2015). The stronger correlation in this study could be due to the concurrent measurement of maternal and child phthalate concentrations and lack of influence from temporal trends in phthalate exposure concentrations (Zota et al., 2014).

The three primary limitations of our findings are the use of one spot urine sample, the small sample size, and the concurrent measurement of childhood phthalate metabolites and outcome measures. Phthalates are non-persistent chemicals and thus subject to intra-individual variability over time. However, MBP has been shown to be one of the most stable metabolites over time with an intraclass correlation coefficient for prenatal measurements ranging from 0.35 (Fisher et al., 2015) to 0.57 (Casas et al., 2018). A study of between-day and between-season phthalate variability in children reported intraclass correlation coefficients of 0.58 and 0.36 for MBP (Casas et al., 2018). Although the availability of only one childhood measurement precluded us from determining sensitive windows of exposure in the very early childhood years, we did consider the influence of in utero exposures by adjusting for certain prenatal phthalates. Future analyses of MIREC data will enable greater insight into this question by analyzing a wider range of prenatal phthalates and determining their association with children's anthropometric outcomes. At time of this writing, maternal urine samples are being re-analyzed to quantify levels of 22 phthalate metabolites. In addition to temporal variation, metabolite concentrations may be influenced by interindividual variations in urinary flow rate (Aylward et al., 2014; Hays et al., 2015). We addressed this potential source of misclassification bias by standardizing and adjusting metabolite concentrations for specific gravity.

It is also worth noting that our results are subject to type 1 error due to the multiple comparisons examined. It is possible, for example, that the statistically significantly results observed in middle tertiles of the TSF models are a result of chance rather than a true association. On the other hand, non-monotonic relationships have been observed between phthalate exposures and health outcomes (Hill et al., 2018). It is not possible within the present study to confirm or deny the presence of true non-linear relations; thus, we interpret our findings of positive associations between MMP and TSF z-scores with caution. The consistency in results between the traditional and weighted quantile sum regression models does, however, lessen concern that the observed positive associations between $\sum$DnBP and BMI and between MMP and TSF z-scorse are solely due to chance.

We cannot rule out the potential for reverse causality in the observed positive associations, given the cross-sectional data collection of child exposure and anthropometric data. Food packaging may serve as a source of exposure to DnBP as well as a surrogate of food consumption. A study of 112 children and adolescents aged 4–18 years found that adoption of a healthy diet low in packaged and processed food was associated with a reduction in phthalate metabolite levels (Correia-Sá et al., 2018). High consumption of processed food, which necessitates more packaging than fresh food, may contribute to risk of obesity (Poti et al., 2017), but we did not have the capacity to incorporate data on childhood dietary patterns or intake of sugar, fat, or processed foods in this study. It is, therefore, difficult to know whether associations

between ∑DnBP and BMI are real or an artifact of consumption and exposure patterns. It is also possible that we did not have the statistical power to detect associations and that the results are, therefore, subject to a type 2 error - particularly in the sex-specific analyses. Median ∑DnBP concentrations were higher in boys (110 nmol/L) than girls (89.9 nmol/L), whereas MMP concentrations were fairly similar (boys 4.48 nmol/L, girls 3.49 nmol/L). These modest differences may also reflect variation in the small sample and warrant further inquiry into potential sex-specific patterns of exposure and differences in potential exposure-related growth patterns.

MIREC study participants tend to have a healthier pregnancy profile (i.e. lower smoking rates and pre-pregnancy BMI) and are of higher socioeconomic status than the overall Canadian population (Arbuckle et al., 2013). Median concentrations of MnBP and MMP in MIREC-CD Plus participants were, however, comparable to concentrations in 3–6 year old children enrolled in the Canadian Health Measures Survey (Health Canada, 2019). Given the ubiquity of phthalate exposures across all socioeconomic groups, the sociodemographic composition of the MIREC cohort is not a notable barrier to the generalizability of our results. Our findings do, however, have limited generalizability to non-Caucasian racial and ethnic groups who may have heightened susceptibility to the growth-related effects of phthalate exposures (Trasande et al., 2013).

5. Conclusions

In this population of Canadian preschool-aged children, we identified DnBP as a potential chemical of concern with regard to childhood obesity. Future research with serial phthalate measurements and anthropometric measurements will help determine whether the observed findings are indicative of an obesogenic effect of these ubiquitous exposures. Follow-up studies being conducted in the MIREC study platform will facilitate this type of research.

Acknowledgements

We would like to acknowledge the MIREC Study Group as well as the MIREC study participants and staff for their dedication. We would also like to acknowledge Health Canada's Chemicals Management Plan for funding the MIREC-CD Plus study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2021.113689.

Funding sources

The MIREC-CD Plus study was funded by Health Canada's Chemicals Management Plan.

Ethics review

The study was reviewed by the Research Ethics Board at Health Canada (Ottawa, ON), the Research Ethics committee at St Justine's Hospital (Montreal, QC, Canada), and ethics committees of each participating study site. Parents signed a written informed consent form prior to their and their child's participation in this study.

References

Arbuckle, T.E., Davis, K., Marro, L., Fisher, M., Legrand, M., Leblanc, A., Gaudreau, E., Foster, W.G., Choeurng, V., Fraser, W.D., 2014. Phthalate and bisphenol A exposure among pregnant women in Canada - results from the MIREC study. Environ. Int. 68, 55–65. https://doi.org/10.1016/j.envint.2014.02.010.

Arbuckle, T.E., Fraser, W.D., Fisher, M., Davis, K., Liang, C.L., Lupien, N., Bastien, S., Velez, M.P., von Dadelszen, P., Hemmings, D.G., Wang, J., Helewa, M., Taback, S., Sermer, M., Foster, W., Ross, G., Fredette, P., Smith, G., Walker, M., Shear, R., Dodds, L., Ettinger, A.S., Weber, J.-P., D'Amour, M., Legrand, M., Kumarathasan, P., Vincent, R., Luo, Z.-C., Platt, R.W., Mitchell, G., Hidiroglou, N., Cockell, K., Villeneuve, M., Rawn, D.F.K., Dabeka, R., Cao, X.-L., Becalski, A., Ratnayake, N., Bondy, G., Jin, X., Wang, Z., Tittlemier, S., Julien, P., Avard, D., Weiler, H., Leblanc, A., Muckle, G., Boivin, M., Dionne, G., Ayotte, P., Lanphear, B., Séguin, J.R., Saint-Amour, D., Dewailly, E., Monnier, P., Koren, G., Ouellet, E., 2013. Cohort profile: the maternal-infant research on environmental chemicals research platform. Paediatr. Perinat. Epidemiol. 27, 415–425. https://doi.org/10.1111/ppe.12061.

Aylward, L.L., Hays, S.M., Smolders, R., Koch, H.M., Cocker, J., Jones, K., Warren, N., Levy, L., Bevan, R., 2014. Sources of variability in biomarker concentrations. J. Toxicol. Environ. Health B Crit. Rev. 17, 45–61. https://doi.org/10.1080/10937404.2013.864250.

Biemann, R., Navarrete Santos, Anne, Navarrete Santos, Alexander, Riemann, D., Knelangen, J., Blüher, M., Koch, H., Fischer, B., 2012. Endocrine disrupting chemicals affect the adipogenic differentiation of mesenchymal stem cells in distinct ontogenetic windows. Biochem. Biophys. Res. Commun. 417, 747–752. https://doi.org/10.1016/j.bbrc.2011.12.028.

Boas, M., Frederiksen, H., Feldt-Rasmussen, U., Skakkebæk, N.E., Hegedüs, L., Hilsted, L., Juul, A., Main, K.M., 2010. Childhood exposure to phthalates: associations with thyroid function, insulin-like growth factor I, and growth. Environ. Health Perspect. 118, 1458–1464. https://doi.org/10.1289/ehp.0901331.

Braun, J.M., 2017. Early life exposure to endocrine disrupting chemicals and childhood obesity and neurodevelopment. Nat. Rev. Endocrinol. 13, 161–173. https://doi.org/10.1038/nrendo.2016.186.

Buser, M.C., Murray, H.E., Scinicariello, F., 2014. Age and sex differences in childhood and adulthood obesity association with phthalates: analyses of NHANES 2007-2010. Int. J. Hyg Environ. Health 217, 687–694. https://doi.org/10.1016/j.ijheh.2014.02.005.

Carrico, C., Gennings, C., 2015. Characterization of a weighted quantile score approach for highly correlated data in risk analysis scenarios. J. Agric. Biol. Environ. Stat. 20, 100–120. https://doi.org/10.1007/s13253-014-0180-3.

Casals-Casas, C., Desvergne, B., 2011. Endocrine Disruptors: from endocrine to metabolic disruption. Annu. Rev. Physiol. 73, 135–162. https://doi.org/10.1146/annurev-physiol-012110-142200.

Casas, M., Basagaña, X., Sakhi, A.K., Haug, L.S., Philippat, C., Granum, B., Manzano-Salgado, C.B., Brochot, C., Zeman, F., de Bont, J., Andrusaityte, S., Chatzi, L., Donaire-Gonzalez, D., Giorgis-Allemand, L., Gonzalez, J.R., Gracia-Lavedan, E., Grazuleviciene, R., Kampouri, M., Lyon-Caen, S., Pañella, P., Petraviciene, I., Robinson, O., Urquiza, J., Vafeiadi, M., Vernet, C., Waiblinger, D., Wright, J., Thomsen, C., Slama, R., Vrijheid, M., 2018. Variability of urinary concentrations of non-persistent chemicals in pregnant women and school-aged children. Environ. Int. 121, 561–573. https://doi.org/10.1016/j.envint.2018.09.046.

CDC, 2009. Fourth National Report on Human Exposure to Environmental Chemicals. https://www.cdc.gov/exposurereport/pdf/FourthReport_UpdatedTables_Volume1_Jan2019-508.pdf. (Accessed 14 November 2020).

Correia-Sá, L., Kasper-Sonnenberg, M., Pälmke, C., Schütze, A., Norberto, S., Calhau, C., Domingues, V.F., Koch, H.M., 2018. Obesity or diet? Levels and determinants of phthalate body burden – a case study on Portuguese children. Int. J. Hyg Environ. Health 221, 519–530. https://doi.org/10.1016/j.ijheh.2018.02.001.

Deierlein, A., Wolff, M., Pajak, A., Pinney, S., Windham, G., Galvez, M.P., Silva, M.J., Calafat, A.M., Kushi, L.H., Biro, F.M., 2016. Longitudinal associations of phthalate exposures during childhood and body size measurements in young girls. Epidemiology 27, 492–499. https://doi.org/10.1097/EDE.0000000000000489.

Den Hond, E., Govarts, E., Willems, H., Smolders, R., Casteleyn, L., Kolossa-Gehring, M., Schwedler, G., Seiwert, M., Fiddicke, U., Castaño, A., Esteban, M., Angerer, J., Koch, H.M., Schindler, B.K., Sepai, O., Exley, K., Bloemen, L., Horvat, M., Knudsen, L.E., Joas, A., Joas, R., Biot, P., Aerts, D., Koppen, G., Katsonouri, A., Hadjipanayis, A., Krskova, A., Maly, M., Mørck, T.A., Rudnai, P., Kozepesy, S., Mulcahy, M., Mannion, R., Gutleb, A.C., Fischer, M.E., Ligocka, D., Jakubowski, M., Fátima Reis, M., Namorado, S., Gurzau, A.E., Lupsa, I.R., Halzlova, K., Jajcaj, M., Mazej, D., Tratnik, J.S., López, A., Lopez, E., Berglund, M., Larsson, K., Lehmann, A., Crettaz, P., Schoeters, G., 2015. First steps toward harmonized human biomonitoring in Europe: demonstration project to perform human biomonitoring on a European scale. Environ. Health Perspect. 123, 255–263. https://doi.org/10.1289/ehp.1408616.

Desvergne, B., Feige, J.N., Casals-Casas, C., 2009. PPAR-mediated activity of phthalates : a link to the obesity epidemic ? Mol. Cell. Endocrinol. 304, 43–48. https://doi.org/10.1016/j.mce.2009.02.017.

Engle, S., Wolff, M., 2013. Causal inference considerations for endocrine disruptor research in children's health. Annu. Rev. Publ. Health 34, 139–158. https://doi.org/10.1146/annurev-publhealth-031811-124556.

Environment and Climate Change Canada. Draft screening assessment phthalate substance grouping. http://ec.gc.ca/ese-ees/default.asp?lang=En&n=516A504A-1. (Accessed 14 November 2020).

Environment and Climate Change Canada, 2015. State of the science report phthalate substance grouping short-chain phthalate esters. http://ec.gc.ca/ese-ees/default.asp?lang=En&n=51624E94-1#Toc015. (Accessed 14 November 2020).

Fisher, M., Arbuckle, T.E., MacPherson, S., Braun, J.M., Feeley, M., Gaudreau, É., 2019. Phthalate and BPA exposure in women and newborns through personal care product use and food packaging. Environ. Sci. Technol. 53, 10813–10826. https://doi.org/10.1021/acs.est.9b02372.

Fisher, M., Arbuckle, T.E., Mallick, R., Leblanc, A., Hauser, R., Feeley, M., Koniecki, D., Ramsay, T., Provencher, G., Bérubé, R., Walker, M., 2015. Bisphenol A and phthalate metabolite urinary concentrations: daily and across pregnancy variability. J. Expo. Sci. Environ. Epidemiol. 25, 231–239. https://doi.org/10.1038/jes.2014.65.

Foster, P.M.D., Cattley, R.C., Mylchreest, E., 2000. Effects of Di-n-butyl phthalate (DBP) on male reproductive development in the rat: implications for human risk assessment. Food Chem. Toxicol. 38, S97–S99. https://doi.org/10.1016/s0278-6915(99)00128-3.

Garí, M., Koch, H.M., Pälmke, C., Jankowska, A., Wesołowska, E., Hanke, W., Nowak, D., Bose-O'Reilly, S., Polańska, K., 2019. Determinants of phthalate exposure and risk assessment in children from Poland. Environ. Int. 127, 742–753. https://doi.org/10.1016/j.envint.2019.04.011.

GBD 2013 Obesity Collaboration., 2014. Global, regional and national prevalence of overweight and obesity in children and adults 1980-2013: a systematic analysis. Lancet 384, 766–781. https://doi.org/10.1016/S0140-6736(14)60460-8.

Goodman, M., Lakind, J.S., Mattison, D.R., 2014. Do phthalates act as obesogens in humans? A systematic review of the epidemiological literature. Crit. Rev. Toxicol. 44, 151–175. https://doi.org/10.3109/10408444.2013.860076.

Gow, P.J., Ghabrial, H., Smallwood, R.A., Morgan, D.J., Ching, M.S., 2001. Neonatal hepatic drug elimination. Pharmacol. Toxicol. 88, 3–15. https://doi.org/10.1034/j.1600-0773.2001.088001003.x.

Grun, F., Blumberg, B., 2009. Endocrine disrupters as obesogens. Mol. Cell. Endocrinol. 304, 19–29. https://doi.org/10.1016/j.mce.2009.02.018.

Hao, C., Cheng, X., Xia, H., Ma, X., 2012. The endocrine disruptor mono-(2-ethylhexyl) phthalate promotes adipocyte differentiation and induces obesity in mice. Biosci. Rep. 32, 619–629. https://doi.org/10.1042/BSR20120042.

Harley, K., Berger, K., Rauch, S., Kogut, K., Henn, B., Calafat, A., Huen, K., Eskenazi, B., Holland, N., 2017. Association of prenatal urinary phthalate metabolite concentrations and childhood BMI and obesity. Pediatr. Res. 82, 405–415. https://doi.org/10.1038/pr.2017.112.

Hays, S.M., Aylward, L.L., Blount, B.C., 2015. Variation in urinary flow rates according to demographic characteristics and body mass index in NHANES: potential confounding of associations between health outcomes and urinary biomarker concentrations. Environ. Health Perspect. 123, 293–300. https://doi.org/10.1289/ehp.1408944.

Health Canada, 2019. Fifth report on human biomonitoring of environmental chemicals in Canada. Results of the Canadian Health Measures Survey Cycle 5.

Heindel, J.J., Newbold, R., Schug, T.T., 2015. Endocrine disruptors and obesity. Nat. Rev. Endocrinol. 11, 653–661. https://doi.org/10.1038/nrendo.2015.163.

Hill, C.E., Myers, J.P., Vandenberg, L.N., 2018. Nonmonotonic dose–response curves occur in dose ranges that are relevant to regulatory decision-making. Dose-Response 1–4. https://doi.org/10.1177/1559325818798282.

Hurst, C.H., Waxman, D.J., 2003. Activation of PPAR alpha and PPAR gamma by environmental phthalate monoesters. Toxicol. Sci. 308, 297–308. https://doi.org/10.1093/toxsci/kfg145.

Just, A.C., Adibi, J.J., Rundle, A.G., Calafat, A.M., Camann, D.E., Hauser, R., Silva, M.J., Whyatt, R.M., 2010. Urinary and air phthalate concentrations and self-reported use of personal care products among minority pregnant women in New York city. J. Expo. Sci. Environ. Epidemiol. 20, 625–633. https://doi.org/10.1038/jes.2010.13.

Keith, S.W., Redden, D.T., Katzmarzyk, P.T., Boggiano, M.M., Hanlon, E.C., Benca, R.M., Ruden, D., 2006. Putative Contributors to the Secular Increase in Obesity: Exploring the Roads Less Traveled, pp. 1585–1594. https://doi.org/10.1038/sj.ijo.0803326.

Lewin, A., Arbuckle, T.E., Fisher, M., Liang, C.L., Marro, L., Davis, K., Abdelouahab, N., Fraser, W.D., 2017. Univariate predictors of maternal concentrations of environmental chemicals: the MIREC study. Int. J. Hyg Environ. Health 220, 77–85. https://doi.org/10.1016/j.ijheh.2017.01.001.

Maresca, M.M., Hoepner, L.A., Hassoun, A., Oberfield, S.E., Mooney, S.J., Calafat, A.M., Ramirez, J., Freyer, G., Perera, F.P., Whyatt, R.M., Rundle, A.G., 2016. Prenatal exposure to phthalates and childhood body size in an urban cohort. Environ. Health Perspect. 124, 514–520. https://doi.org/10.1289/ehp.1408750.

Mylchreest, E., Sar, M., Cattley, R.C., Foster, P.M.D., 1999. Disruption of androgen-regulated male reproductive development by di(n- butyl) phthalate during late gestation in rats is different from flutamide. Toxicol. Appl. Pharmacol. 156, 81–95. https://doi.org/10.1006/taap.1999.8643.

Poti, J., Braga, B., Quin, B., 2017. Ultra-processed food intake and obesity: what really matters for health – processing or nutrient content ? Curr. Obes. Rep. 6, 420–431. https://doi.org/10.1007/s13679-017-0285-4.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://R-project.org.

Renzetti, S., Gennings, C., Curtin, P., 2019. gWQS: an R package for linear and generalized weighted quantile sum (WQS) regression. https://cran.asia/web/packages/gWQS/vignettes/gwqs-vignette.pdf. (Accessed 14 November 2020).

Romano, M.E., Savitz, D.A., Braun, J.M., 2014. Challenges and future directions to evaluating the association between prenatal exposure to endocrine-disrupting chemicals and childhood obesity. Curr. Epidemiol. Rep. 57–66. https://doi.org/10.1007/s40471-014-0007-3.

Schwedler, G., Rucic, E., Lange, R., Conrad, A., Koch, H.M., Pälmke, C., Brüning, T., Schulz, C., Schmied-Tobies, M.I.H., Daniels, A., Kolossa-Gehring, M., 2020. Phthalate metabolites in urine of children and adolescents in Germany. Human biomonitoring results of the German Environmental Survey GerES V, 2014–2017. Int. J. Hyg Environ. Health 225, 113444. https://doi.org/10.1016/j.ijheh.2019.113444.

Shoaff, J., Papandonatos, G.D., Calafat, A., Ye, Z., Chen, A., Lanphear, B.P., Yolton, K., Braun, J.M., 2017. Early-life phthalate exposure and adiposity at 8 years of age. Environ. Health Perspect. 125, 097008 https://doi.org/10.1289/ehp1022.

Tanner, E.M., Bornehag, C.G., Gennings, C., 2019. Repeated holdout validation for weighted quantile sum regression. Methods 6, 2855–2860. https://doi.org/10.1016/j.mex.2019.11.008.

Tanner, E.M., Hallerbäck, M.U., Wikström, S., Lindh, C., Kiviranta, H., Gennings, C., Bornehag, C.G., 2020. Early prenatal exposure to suspected endocrine disruptor mixtures is associated with lower IQ at age seven. Environ. Int. 134, 105185 https://doi.org/10.1016/j.envint.2019.105185.

Teitelbaum, S.L., Mervish, N., Moshier, E.L., Vangeepuram, N., Galvez, M.P., Calafat, A. M., Silva, M.J., Brenner, B.L., Wolff, M.S., 2012. Associations between phthalate metabolite urinary concentrations and body size measures in New York City children. Environ. Res. 112, 186–193. https://doi.org/10.1016/j.envres.2011.12.006.

Trasande, L., Attina, T.M., Sathyanarayana, S., Spanier, A.J., Blustein, J., 2013. Race/ethnicity-specific associations of urinary phthalates with childhood body mass in a nationally representative sample. Environ. Health Perspect. 121, 501–506. https://doi.org/10.1289/ehp.1205526.

Vafeiadi, M., Myridakis, A., Roumeliotaki, T., Margetaki, K., Chalkiadaki, G., Dermitzaki, E., Venihaki, M., Sarri, K., Vassilaki, M., Leventakou, V., Stephanou, E. G., Kogevinas, M., Chatzi, L., 2018. Association of early life exposure to phthalates with obesity and cardiometabolic traits in childhood: sex specific associations. Front Public Health 6, 1–11. https://doi.org/10.3389/fpubh.2018.00327.

van Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: multivariate imputation by chained equations in R. J. Stat. Software 45, 1–67.

Voerman, E., Santos, S., Golab, B.P., Amiano, P., Ballester, F., Barros, H., Bergström, A., Charles, M.A., Chatzi, L., Chevrier, C., Chrousos, G.P., Corpeleijn, E., Costet, N., Crozier, S., Devereux, G., Eggesbø, M., Ekström, S., Fantini, M.P., Farchi, S., Forastiere, F., Georgiu, V., Godfrey, K.M., Gori, D., Grote, V., Hanke, W., Hertz-Picciotto, I., Heude, B., Hryhorczuk, D., Huang, R.C., Inskip, H., Iszatt, N., Karvonen, A.M., Kenny, L.C., Koletzko, B., Küpers, L.K., Lagström, H., Lehmann, I., Magnus, P., Majewska, R., Mäkelä, J., Manios, Y., McAuliffe, F.M., McDonald, S.W., Mehegan, J., Mommers, M., Morgen, C.S., Mori, T.A., Moschonis, G., Murray, D., Chaoimh, C.N., Nohr, E.A., Andersen, A.M.N., Oken, E., Oostvogels, A.J.J.M., Pac, A., Papadopoulou, E., Pekkanen, J., Pizzi, C., Polanska, K., Porta, D., Richiardi, L., Rifas-Shiman, S.L., Ronfani, L., Santos, A.C., Standl, M., Stoltenberg, C., Thiering, E., Thijs, C., Torrent, M., Tough, S.C., Trnovec, T., Turner, S., van Rossem, L., von Berg, A., Vrijheid, M., Vrijkotte, T.G.M., West, J., Wijga, A., Wright, J., Zvinchuk, O., Sørensen, T.I.A., Lawlor, D.A., Gaillard, R., Jaddoe, V.W.V., 2019. Maternal body mass index, gestational weight gain, and the risk of overweight and obesity across childhood: an individual participant data meta-analysis. PLoS Med. 16, 1–22. https://doi.org/10.1371/journal.pmed.1002744.

Wang, H., Zhou, Y., Tang, C., He, Y., Wu, J., Chen, Y., Jiang, Q., 2013. Urinary phthalate metabolites are associated with body mass index and waist circumference in Chinese school children. PloS One 8, e56800. https://doi.org/10.1371/journal.pone.0056800.

Wang, Y., Zhu, H., Kannan, K., 2019. A review of biomonitoring of phthalate exposures. Toxics 7, 21. https://doi.org/10.3390/toxics7020021.

Weiss, J.M., Gustafsson, Å., Gerde, P., Bergman, Å., Lindh, C.H., Krais, A.M., 2018. Daily intake of phthalates, MEHP, and DINCH by ingestion and inhalation. Chemosphere 208, 40–49. https://doi.org/10.1016/j.chemosphere.2018.05.094.

WHO, 2015. The WHO Child Growth Standards. http://www.who.int/childgrowth/standards/en. (Accessed 14 November 2020).

Wittassek, M., Koch, H.M., Angerer, J., Brüning, T., 2011. Assessing exposure to phthalates - the human biomonitoring approach. Mol. Nutr. Food Res. 55, 7–31. https://doi.org/10.1002/mnfr.201000121.

Wormuth, M., Scheringer, M., Vollenweider, M., Hungerbühler, K., 2006. What are the sources of exposure to eight frequently used phthalic acid esters in Europeans? Risk Anal. 26, 803–824. https://doi.org/10.1111/j.1539-6924.2006.00770.x.

Yang, T.C., Peterson, K.E., Meeker, J.D., Sánchez, B.N., Zhang, Z., Cantoral, A., Solano, M., Tellez-Rojo, M.M., 2017. Bisphenol A and phthalates in utero and in childhood: association with child BMI z-score and adiposity. Environ. Res. 156, 326–333. https://doi.org/10.1016/j.envres.2017.03.038.

Zhang, Y., Meng, X., Chen, L., Li, D., Zhao, L., Zhao, Y., Li, L., Shi, H., 2014. Age and sex-specific relationships between phthalate exposures and obesity in Chinese children at puberty. PloS One 9. https://doi.org/10.1371/journal.pone.0104852.

Zota, A.R., Calafat, A.M., Woodruff, T.J., 2014. Temporal trends in phthalate exposures: findings from the national health and nutrition examination survey, 2001-2010. Environ. Health Perspect. 122, 235–241. https://doi.org/10.1289/ehp.1306681.

International Journal of Hygiene and Environmental Health 232 (2021) 113690

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

Short communication

Utilization of greenhouse effect for the treatment of COVID-19 contaminated disposable waste - A simple technology for developing countries

Osama Ali Maher [a,*], Sherif A. Kamal [b], Ahmed Newir [c], Kenneth M. Persson [a]

[a] Water Resources Engineering, Lund University, Lund, Sweden
[b] Building and Construction Department, Faculty of Engineering, October 6 University, Cairo, Egypt
[c] Faculty of Engineering, Mechatronics Department, October 6 University, Cairo, Egypt

ARTICLE INFO

Keywords:
IPC
Medical waste
Air temperature
COVID-19
Greenhouse effect

ABSTRACT

Countries with abundant solar radiation have the potential to invest in simple technologies for deactivation of many bacteria and viruses in medical solid waste. In addition to the traditional Infection and Prevention Control (IPC) measures, these simple technologies contribute to better protection of health care workers in countries with compromised solid management schemes. Monitoring of temperature, relative humidity and ultraviolet inside containers soundly designed to collect disposal infectious waste illustrated to deactivate several viruses and bacteria. Casanova et al., 2010, used some surrogate viruses to overcome the challenges of working with SARS-CoV, concluded that by temperature above 40 °C most of viruses become below levels of detection after 90 min. Here we are proposing a model of a simple transparent container almost 200 L in volume that allow solar energy to be accumulated inside. In summer conditions in the testing site, temperature inside the container reached above 50 °C when the ambient air temperature was around 30 °C. The container was built using epoxy glass to guarantee maximum heat penetration. Actual temperature measurement inside the container was measured in real time against ambient air temperature. We present a mathematical model for predication of maximum temperature at different positions inside the container and their relation to different ambient air temperature scenarios. The mathematical formulas used are based on the conservation laws and a good agreement of a full month of field measurements were obtained. Even in winter conditions in many of developing countries air temperature can maintain levels above 20 °C, which will produce temperature around 30 °C and viruses can reach levels below detection limit in maximum 3 h.

1. Introduction

Infectious solid waste management usually uses expensive techniques such as incineration, chemical treatment or autoclaving to insure minimizing or eliminating of the hazardous substances in these wastes. Personal Protective Equipment (PPE) for personnel managing patients suffering from infectious diseases can be either reusable or disposable. The reusable equipment needs to go through sterilization process to guarantee the elimination of any biological contamination. The disposable materials need to go through some processes before it is end of life disposal in a safe manner to guarantee low or no risk of infection to personnel responsible for its full life cycle from usage to final disposal. In many health care facilities in poor countries, financial capacities are limited to invest in modern techniques and protocols for how to deal with infectious waste is often lacking. There is not enough analysis on infection among health care waste handlers in low income countries or countries with no adequate medical solid waste management scheme. However, in such context it is important to explore all possible measures to protect workers under this category from getting infected as a result of exposure to biologically contaminated medical waste. In lack of enough Personal Protection Equipment (PPEs) and an integrated medical solid waste management schemes in health care facilities, it is important to investigate innovative ways to take advantage of environmental conditions to control the virus survivals on surfaces of medical waste. Little

* Corresponding author. Department for Water Resources Engineering, Lund University, Box 118, S-221 00, Lund, Sweden.
E-mail address: Osama.ali_maher@tvrl.lth.se (O.A. Maher).
URL: http://www.tvrl.lth.se, http://www.byggmiljo.lth.se (O.A. Maher).

https://doi.org/10.1016/j.ijheh.2021.113690
Received 17 September 2020; Received in revised form 27 December 2020; Accepted 3 January 2021
Available online 8 January 2021

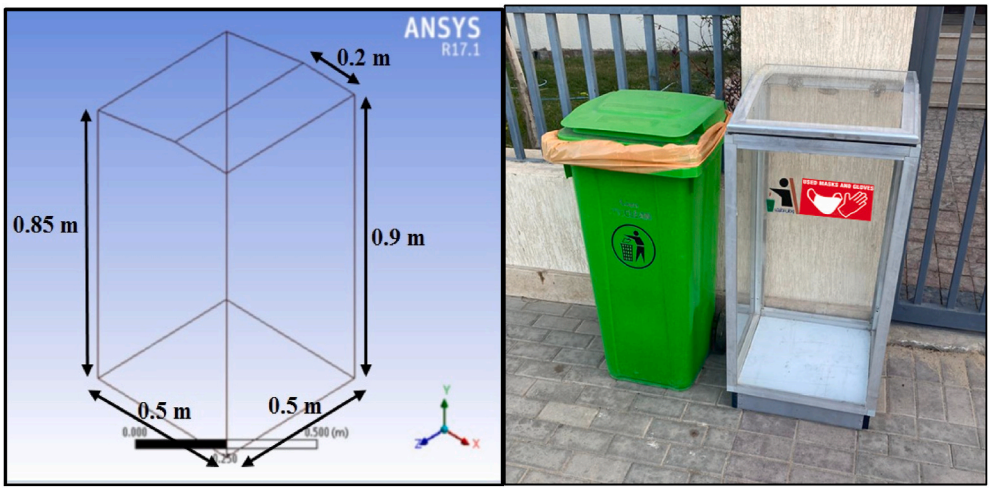


Fig. 1. The dimensions of the medical waste bin (to the left and the actual prototype is to the right). The dimensions are representing the clear dimensions of the acrylic sheets directly exposed to sun and connected using aluminum. The theoretical part for prediction of the temperature inside the bin was built on a three-dimensional model of greenhouse structure. The temperature environments for simulation calculation are in hot summer with no wind. Entire medical recycle bin model is divided into 47,887 elements. Iterative calculation is conducted using two CPU 3.07 GHz quad-core workstation in simulation.

attention was paid to advantages of some environmental conditions prevailing in some countries. However, it is reported that the COVID-19 can be transmitted largely by garbage collectors (Ouhsine et al., 2020). The survival of the COVID-19 virus on surfaces was subject to some research, however these researchers were not inclusive (WHO and UNICEF 2020). In this study we are examining three environmental aspects of affecting the survival of similar viruses as a proxy indicator for the COVID 19 survival on surfaces. These atmospheric indicators, Relative Humidity (RH), Air Temperature (AT) and Ultraviolet Irradiation (UV), aim to develop some guidance and operations procedures to an intermediate storage unit at health care facilities for COVID-19 PPEs in preparation for a final safe disposal among municipal waste.

2. Background

Several studies were directed examine the relation between environmental conditions such as Air Temperature (AT) on the survival of viruses and bacteria on different surfaces, for instance Hepatitis A Virus (Mbithi et al., 1991) and avian influenza viruses (Paek et al., 2010). Another study on the Coronavirus, Casanova et al., (2010), using two potential surrogates were evaluated in this study; transmissible gastroenteritis virus (TGEV) and mouse hepatitis virus (MHV) concluded that, these viruses can survive on different surfaces from hours to days and the optimum deactivation of the virus happened at around 50% RH. The same study confirmed that the virus infectivity with increased by temperatures higher than 40 °C and at this level of temperature, RH was not significant for the deactivation of the viruses. Following the SARS outbreak in 2003 some studies were conducted to examine the survival of the virus in relation to environmental conditions. In a study by Duan et al., (2003) on the SARS CoV–P9 on its deactivation when exposed to UV (260 nM) for about 60 min resulted in the destruction of virus infectivity. The same study pointed out that the virus was totally deactivated after 90, 60, 30 min at 56 °C, 67 °C and 75 °C respectively. In countries with abundance of solar radiation and long sunny hours, it is easy to achieve these favorable conditions for virus inactivation with simple technology. The proposed technology here depends on producing simple collection system based on minimum protection procedures and low costing.

With more than 15 million confirmed cases in Africa and Latin America alone, it is obvious that the amount of medical solid waste production has increased dramatically. The IPC measures usually not being optimal in theses settings, so complementary measures needs to be in place to minimize possible infections to these vulnerable group. The fate of the solid waste produced is usually that the medical waste will be mixed with municipal waste. According to our knowledge no detailed research concerning infection among solid waste collectors in low and middle income countries has been carried out.

3. Methodology

A model for a simple acrylic waste collection bin was constructed with clear shits dimensions of 0.9 m high and a square base 0.5 m in length. The bin had an open led 0.2 m length from one side of the top, Fig. 1. The bin was exposed to direct solar radiation at a site in Cairo, Egypt (longitude 29.9°, latitude 31.39°) and temperature was measured with 1 h interval between 11:00 a.m. and 4:00 p.m. for the full month of June in 2020. The bin was closed during the whole measurement period during the day. The temperature measured using TESTO 875i thermal imager with temperatures range up to 327 °C and the accuracy is ±2% of reading.

The storage volume of the bin is approximately 200 L (0.5 × 0.5*0.85 cm) and the different in head in the front 0.85 m and in the back 0.9 m is to make a slope to insure the direct sun light to enter the bin to give maximum temperature using greenhouse effect.

The bin is constructed in a practical way that it will not be well marked and easy to operate by the solid waste operators within the health care facility. The container should be placed outdoor in places directly exposed to the sun with clear marks showing that it is for the disposal of gloves and masks.

4. Results and discussions

4.1. Field measurements

The greenhouse effect is identified by the difference between the effective radiating temperature of the earth surface and its surface temperature. The difference between the energy emitted by the surface and can therefore be defined as the long wave energy trapped in the atmosphere (Berger and Tricot 1992). The origin of the term greenhouse was associated with providing a Controlled Environment Agriculture (CEA) through framed or inflated structure, covered by transparent or translucent material that permits optimum light and hence energy transmission. This closed structure will accumulate thermal energy if kept exposed to solar energy with little or no air movement (Jensen and Malter 1995). These conditions will generate a difference in temperature between inside the space and ambient air temperature. Egypt enjoys around 2900–3200 h of sunshine annually with annual direct normal energy density of 1970–3200 kWh/m^2 (Comsan 2010). The maximum day temperature, away from heat waves, usually reached during the month of August. The COVID -19 pandemic reached Egypt by mid-March and cases have been reported until September 2020. The experiment here was carried out during the month of June in the middle

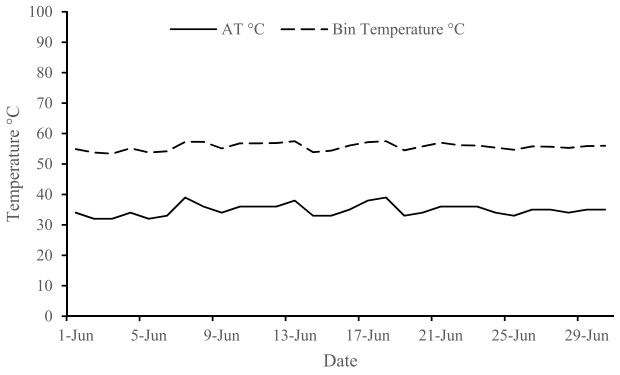


Fig. 2. Maximum recorded temperature outside and inside the bin during the month of June. During the whole the difference between the measured temperature exceeded 10 °C.

Table 1
Registered temperature on June 18, 2020, temperature was measured on 1 h interval and a difference in temperature was recorded already at the first measurement.

Time	AT (°C)	Bin Temp. (°C)
11:00	36	47.9
12:00	37	48.5
13:00	38	48.7
14:00	39	53.2
15:00	38	55.6
16:00	38	57.5

of the COVID-19 pandemic. Continuous measurements AT as well as the temperature inside the solid waste bin were carried throughout the month. Registered peaks of AT, corresponding temperature inside the bin, and times of these peaks are presented in Fig. 2.

The average peak air temperature during this month was above 30 °C and the maximum registered temperature difference between inside and outside temperature was more than 10 °C during the whole month. Table 1 shows the maximum registered daily air temperature of 39 °C around 2 p.m. on 18th of June.

4.2. Model iterations and validation

The governing equations of fluid flow and heat transfer can be considered as mathematical formulations of the conservation laws that govern all associated phenomena. These conservation laws describe the rate of change of a desired fluid property as a function of external forces in accordance with the continuity equation: the mass flows entering a fluid element must balance exactly with those leaving.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left[\rho \vec{V}\right] = S_m \tag{1}$$

where, ρ is the air density, t is the time, $\vec{v}$ is the velocity vector and S_m is the source term.

Conservation of momentum (Newton's second law): the sum of the external forces acting on the fluid particle is equal to its rate of change of linear momentum.

$$\frac{\partial}{\partial t}(\rho \vec{v}) + \nabla \cdot \left(\rho \vec{V} \vec{V}\right) = -\nabla p + \rho \vec{g} + \vec{F} \tag{2}$$

where, p is the static pressure and $\vec{g}$ and $\vec{F}$ are the gravitational body force and external body forces respectively.

Conservation of energy (the first law of thermodynamics): the rate of

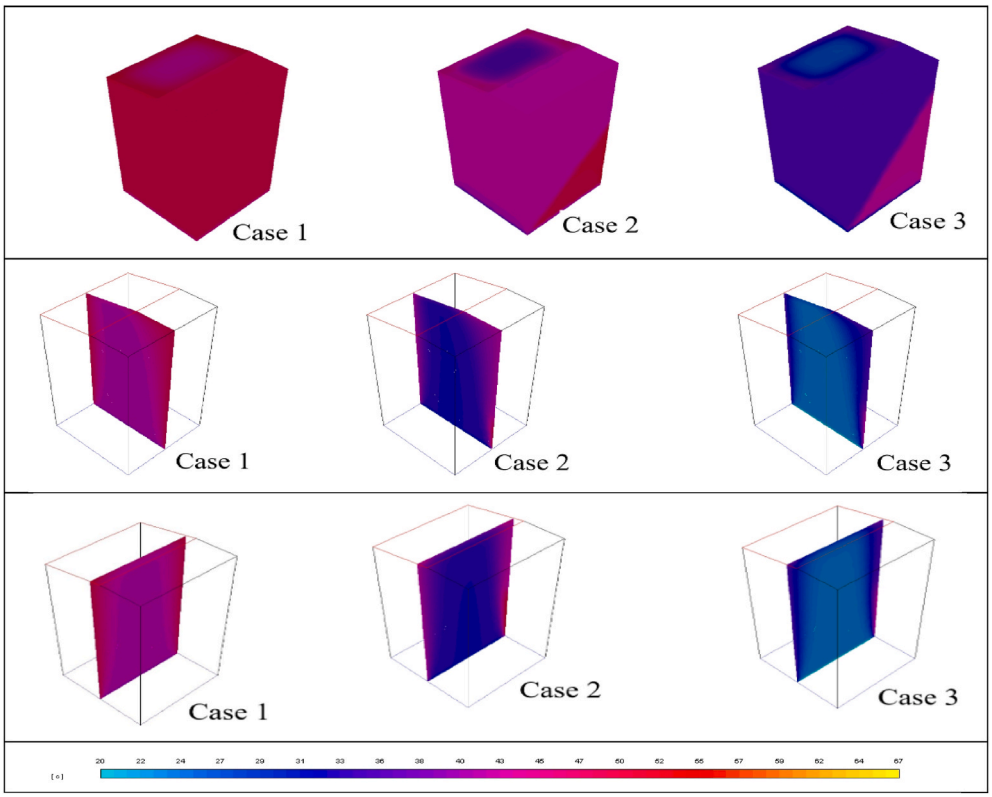


Fig. 3. Calculated temperature inside the bin with three scenarios, case 1 is using data registered at the day of maximum measured air temperature, case 2 and 3 for maximum air temperature of 20 and 30 °C respectively.

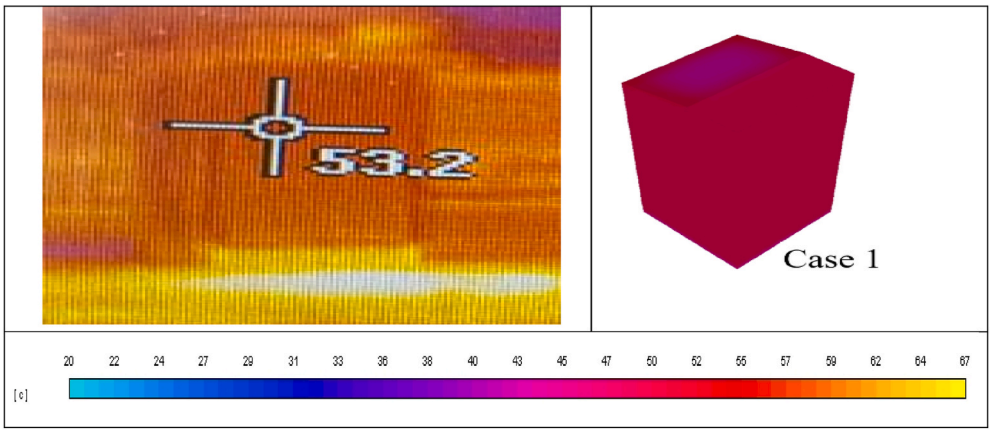


Fig. 4. Example of the maximum temperature registered and location inside the bin on June 18, 2020.

change of energy of a fluid particle is equal to the heat addition and the work done on the particle.

$$U_j\frac{\partial T}{\partial x_j}=\frac{\partial}{\partial x_j}\left(\alpha\frac{\partial T}{\partial x_j}-\overline{u_j t}\right) \quad (3)$$

The optimization of any structure to reach maximum utilization of solar thermal energy to be attained within a structure is highly dependent on the structure geometry. Here we used ANSYS © 17.1 DESIGN MODULER pre-processor software to create the simulated grid to the prototype. A good quality mesh can assist verifying model calculation and produces fast and accurate solution. Therefore, more than one mesh type is tested and compared with each other to attain a good computational fluid dynamics solution. The different meshes used are multizone and automatic, tetrahedral patch conforming. Mesh quality is depending on many parameters, but most important ones are meshed elements, mesh skewness ratio and orthogonal ratio.

These parameters are used to ensure that, mesh density should be high enough to capture all relevant flow features and the mesh adjacent to the wall should resolve the boundary layer flow.

The best method for the medical recycle bin geometry is tetrahedral patch conforming to 47,887 elements with average skewness 0.216 and the average orthogonal is 0.864, which falls into the "excellent" range according to the software standard. The acrylic walls in the model are 3 mm thickness and added in ANSYS the material properties, it is properties is 1180 kg/m^3 density, 1470 J/kg-k specific heat and 0.19 w/m-k thermal conductivity. Three cases were calculated in the mathematical model, case one is the experimental case which implemented in the practical experiment day rush hour at 2:00 p.m. (June 18, 2020), with AT of 40 °C. The second case calculated at the outside temperature was 30 °C and the third case at 20 °C. In addition, the longitude and latitude entered in the energy calculation to calculate the solar energy effect by solar calculator. Fig. 3 shows the calculated temperature distribution in outside walls at the top of the figure and the vertical and horizontal central contours in the rest of the figure.

In case 1 the temperature inside the medical recycle bin reaches more than 50 °C and the average temperature. Case 2 with AT of 30 °C resulted in temperature of 38 °C inside the bin. These calculated temperatures are expected to increase due to the greenhouse effect as it happened in the experimental study.

The obtained temperature from the actual measurement on 18th of June as in Fig. 4 shows 53.2° at an AT of 39 °C, which is in a good agreement with the model results for case 1 as mentioned earlier.

5. Conclusions

Inverse effects of elevated temperature on viruses and bacteria have been reported by several researchers. Most of the researcher reports 40 °C as the limits where most of viruses can be under detection limits. The utilization of the greenhouse effect in the area of treatment for biological contamination of disposal PPEs before the final disposal was rarely investigated. Our data showed that with a simple technology that guarantees enough penetration of solar radiation, air temperature of around 30 °C is enough to generate a temperature of 40 °C or higher inside an acrylic solid waste bin. The unit cost of the proposed bin can be 30 $ (current production cost in Egyptian market), which is fat cheaper than alternative high technologies in this area and can be afforded by health authorities in treatment facilities for patients with communicable diseases. The proposed technology is easy to handle and with some simple instruction regarding segregation of disposable waste and standard precautions while transporting it to the bin. Waste by the end of the day can be collected and dealt with as non-infectious waste and can be treated with municipal waste. There is still some limitation on using this method at air temperature lower than 20 °C. However, investing in these technologies can guarantee quick production and effective removal for viruses during the current pandemic and for other infectious diseases as well.

References

Berger, A., Tricot, C., 1992. The greenhouse effect. Surv. Geophys. 13 (6), 523–549.
Casanova, L.M., Jeon, S., Rutala, W.A., Weber, D.J., Sobsey, M.D., 2010. Effects of air temperature and relative humidity on coronavirus survival on surfaces. Appl. Environ. Microbiol. 76 (9), 2712–2717.
Comsan, M.N.H., 2010. Solar energy perspectives in Egypt. In: EPC'10: Environmental Physics Conference; Hurghada (Egypt); 10-14 Mar 2010, pp. 1–11.
Duan, S.M., Zhao, X.S., Wen, R.F., Huang, J.J., Pi, G.H., Zhang, S.X., Han, J., Bi, S.L., Ruan, L., Dong, X.P., 2003. Stability of SARS coronavirus in human specimens and environment and its sensitivity to heating and UV irradiation. Biomed. Environ. Sci. 16 (3), 246–255.
Jensen, M.H., Malter, A.J., 1995. Protected Agriculture: a Global Review, vol. 253. World Bank Publications.
Mbithi, J.N., Springthorpe, V.S., Sattar, S.A., 1991. Effect of relative humidity and air temperature on survival of hepatitis A virus on environmental surfaces. Appl. Environ. Microbiol. 57 (5), 1394–1399.
Ouhsine, O., Ouigmane, A., Layati, E., Aba, B., Isaifan, R.J., Berkani, M., 2020. Impact of COVID-19 on the qualitative and quantitative aspect of household solid waste. Global Journal of Environmental Science and Management 6 (4), 1–12.
Paek, M.R., Lee, Y.J., Yoon, H., Kang, H.M., Kim, M.C., Choi, J.G., Kwon, J.H., 2010. Survival rate of H5N1 highly pathogenic avian influenza viruses at different temperatures. Poultry Sci. 89 (8), 1647–1650.
WHO, UNICEF, 2020. Water, Sanitation, Hygiene and Waste Management for the COVID-19 Virus *Technical Brief*, pp. 2712–2717.