International Journal of Hygiene and Environmental Health 235 (2021) 113771

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

COVID-19: Hotspot hospital?- seroprevalence of SARS-CoV-2 antibodies in hospital employees in a secondary care hospital network in Germany: Intermediate results of a prospective surveillance study

Anke Hildebrandt [a,b,*], Oktay Hökelekli [a], Lutz Uflacker [a], Henrik Rudolf [c], Sören G. Gatermann [d]

[a] St. Vincenz Hospital Datteln, Department of Internal Medicine I, Germany
[b] Institute of Medical Microbiology, University Hospital Münster, Germany
[c] Department of Medical Informatics, Biometry and Epidemiology, Ruhr University Bochum, Germany
[d] National Reference Centre for Multidrug-resistant Gram-negative Bacteria, Ruhr University Bochum, Germany

ARTICLE INFO

Keywords:
SARS-CoV-2
Coronavirus
COVID-19
Antibodies
Healthcare workers

ABSTRACT

Purpose: The objective of the ongoing study was to investigate how SARS-CoV-2 infection spread within two hospitals in North Rhine-Westphalia, Germany by testing the employees working in high-risk, intermediate-risk and low-risk-areas for the presence of SARS-CoV-2 IgG antibodies. Presented intermediate results evaluate the first infection period until the end of September 2020.

Methods: The study "COVID-19: Hotspot hospital?- Seroprevalence of SARS-CoV-2 antibodies in hospital employees in a secondary care hospital network in Germany " is a prospective, single centre observational cohort study conducted at the St. Vincenz Hospital Datteln with 316 beds. The presented data include one other hospital: St. Laurentius Stift Waltrop, Germany with 172 beds.

Results: Between June 2020 and September 2020 we analyzed serum samples of 907 employees which represents 62.1% of all employees. Thirteen employees (1.4%), respectively 13/696 healthcare workers (HCWs) (1.9%) had detectable SARS-CoV-2 IgG antibodies. Among them, 4 (30.8%) were aware of COVID-19 exposure, and 5 (38.5%) reported clinical symptoms. HCWs working in high-risk areas had a seroprevalence rate of 1.6% (1/64), HCWs working in intermediate-risk area 1.7% (11/632) and 0.5% employees (1/211) in low-risk areas with no contact to patients were seropositive.

Conclusion: Even if we treated COVID-19 positive patients, we found no clear evidence that infection was transmitted to HCWs in contact to these patients. As knowledge about SARS-CoV-2 transmission evolves, the concept of infection prevention must be continuously reviewed and adapted as needed to keep hospitals a safe place.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel beta coronavirus that was first identified in December 2019 in Wuhan, China (Huang et al., 2020; Ralph et al., 2020). At the beginning of 2020 the virus spread and became pandemic (Abebe et al., 2020; Whitworth, 2020). The WHO declared a global health emergency on January 31, 2020; subsequently, on March 11, 2020, they declared it a pandemic situation (Dhama et al., 2020).

SARS-CoV-2 infection is presented clinically as corona virus disease 2019 (COVID-19) with a broad range of symptoms from asymptomatic and mild to critical courses (Guan et al., 2020; Pergolizzi et al., 2020). There are no specific symptoms that can suggest COVID-19 compared to symptoms of respiratory illnesses caused by other viruses, such as influenza and common cold (Abebe et al., 2020). The gold standard for diagnosing COVID-19 is the detection of SARS-CoV-2 viral nucleic acid using a quantitative real time-PCR (qRT-PCR) from respiratory tract samples (e.g, throat swabs) (Abebe et al., 2020). Rapid antigen tests provide a promising scheme for timely monitoring and eventual control of the global pandemic (Li et al., 2020). Antibody testing surveys can aid the investigation of an ongoing outbreak and retrospective assessment of the attack rate or extent of an outbreak. However, serological tests cannot be applied to early infection (Li et al., 2020).

The primary means of transmission is person to person through

* Corresponding author. St. Vincenz Hospital Datteln, Department of Internal Medicine I, Rottstr. 11, 45711, Datteln, Germany.
E-mail address: a.hildebrandt@vincenz-datteln.de (A. Hildebrandt).

https://doi.org/10.1016/j.ijheh.2021.113771
Received 15 February 2021; Received in revised form 8 May 2021; Accepted 10 May 2021
Available online 20 May 2021
1438-4639/

droplets that occurred during coughing or sneezing, through personal contact (shaking hands), or by touching contaminated objects (Abebe et al., 2020). Additionally, aerosols from infected persons may pose an inhalation threat even at considerable distances and in enclosed spaces, particularly if there is poor ventilation (Meselson, 2020).

As a consequence nosocomial transmission of inadequately protected health care workers (HCWs) can occur during aerosol generating procedures (Patel et al., 2020), but also in the regular contact to patients with delayed diagnosis of COVID-19 and in close contact to asymptomatic or presymptomatic virus carriers (patients or colleagues) which can also spread the virus (Chou et al., 2020a, b; Khonyongwa et al., 2020; Zhao et al., 2020).

In summer 2020, more than 1.3 Mio HCWs have been tested positive for SARS-CoV-2 worldwide (Fischer-Fels, 2020). Hence it is of great importance to implement infection prevention strategies in the health care sector and provide sufficient personal protection equipment (Chou et al., 2020a).

Data from German HCWs are scarce so far and mainly focussed on university hospitals (Bahrs et al., 2020; Brehm et al., 2021; Korth et al., 2020). The primary objective of this study was to investigate the SARS-CoV-2 infection spread within two hospitals of a secondary care hospital network in North Rhine-Westphalia, Germany by testing employees for the presence of SARS-CoV-2 IgG antibodies. Secondary objectives were to identify potential risk factors for infection and clinical symptoms of seropositive employees. Furthermore, we wanted to evaluate the results with regard to the number of treated COVID-19 positive patients and employees that were tested with PCR within the scope of contact tracking during the first period of SARS-CoV-2 infection.

2. Methods

2.1. Study design

The study "COVID-19: Hotspot hospital?- Seroprevalence of SARS-CoV-2 antibodies in hospital employees in a secondary care hospital network in Germany " is a prospective, single centre observational cohort study conducted at the St. Vincenz Hospital Datteln with 316 beds. The hospital belongs to the group Vestische Caritas Kliniken GmbH. Until end of September 2020 one other hospital of this group also took part in our study so far: St. Laurentius Stift Waltrop with 172 beds. The study is designed from June 2020 to June 2021. We want to publish intermediate results for the period June 2020 to September 2020 in order to look at the first surge of the new pandemic virus.

Research was conducted in accordance with the declaration of Helsinki and national standards. The study protocol was approved by the local ethics committee: Ärztekammer Westfalen-Lippe and Westfälische Wilhelms Universität Münster (approval no.2020-478-f-S). The study was registered at the German Clinical Trials Register (DRKS00022941).

2.2. Enrolment and data management

Participants were recruited since June 1, 2020. All employees of the St. Vincenz Hospital Datteln and St. Laurentius Stift Waltrop working with (HCWs) and without patient contact were addressed to take part. They received information about the study via intranet platform. Participation was voluntary and free of charge.

Employees were included if they put their laboratory number on a 3 paged document so that we could assign the sample to the person. This written informed consent included a questionnaire and agreement on providing a blood sample (not exceeding 9 ml of venous blood).

Pseudonymized blood samples were sent to our central laboratory for testing of antibodies against SARS-CoV-2. Data from pseudonymized questionnaires were collected and processed with MS Excel 2010.

2.3. Questionnaire

The questionnaire included information on personnel data like name, address, telephone number and working area for future contacting. Individual medical history contained clinical symptoms within the last two months such as fever, taste disturbances and smell disorders, dry cough, headache, growing pains, cold-like symptoms, exposure to confirmed COVID-19 cases, results of previous polymerase chain reaction (PCR) or previous serology.

2.4. SARS-CoV-2 antibody testing

Presence of SARS-CoV-2 antibodies were investigated with a chemiluminescence-based immunoassay Elecsys, Anti-SARS-CoV-2 (Roche, Basel, Switzerland). The immunoassay targets recombinant nucleocapsid protein and was carried out according to manufacturers' instructions. Sensitivity and specificity as provided by the manufacturer was high (≥99%). Volunteers with positive test results were regarded as SARS-CoV-2 seropositive. Re-testing was offered to all participants during the test period June 2020 to June 2021. We now only present results of the first point of testing.

2.5. Outcomes

The primary outcome of the study was to assess the seroprevalence of SARS-CoV-2 antibodies in hospital employees using an IgG detecting immunoassay. The study is still running. As we could see a clearly marked first pandemic infection period until end of September 2020, we decided to evaluate these data and publish intermediate results.

Secondary outcomes were: (i) differentiation between HCWs working in high-risk areas with contact to COVID-19 positive patients, HCWs working in intermediate-risk areas with contact to non-COVID-19 positive patients and employees working in low-risk areas with no contact to patients at all (personnel working in administration, kitchen, cleaning service, and others), (ii) potential risk factors and clinical symptoms for seropositive employees and (iii) evaluating the results with regard to infection risk of HCW according to the number of treated COVID-19 positive patients and employees that were tested with PCR within the scope of contact tracking.

2.6. Statistical analysis

In descriptive analysis participants demographics, professions, symptoms, and other attributes of COVID-19 exposure were determined and compared for the whole cohort and stratified by antibody test result using absolute and relative frequencies. Associations of characteristics with test results was statistically tested by Fisher exact test. Alterations of risks for a positive antibodies test result was estimated by univariable logistic regression, giving odds ratios and 95% confidence intervals versus the reference level for each main category of the characteristics or change per unit (for number of symptoms). We applied a significance level of 0.05. Analyses were done using the statistical programming software R (R Core Team (2020). R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria).

3. Results

3.1. Characteristics of the study participants

From June 2020 to the end of September 2020 nine hundred seven of 1460 (62.1%) employees of two hospitals (St. Vincenz Hospital Datteln, St. Laurentius Stift Waltrop) with together 488 beds took part in our study (Table 1).

Among the 907 participants 136 (15.0%) were males and 771 (85.0%) were females. We categorized three age groups: 16–25 years (n

Table 1

Basic information on the conditions in both hospitals in general and with regard to SARS-CoV-2 infections during the first infection period (until the end of September 2020).

Basic information	St. Vincenz Hospital Datteln	St. Laurentius Stift Waltrop	Total number
Beds	316	172	488
Employees	1085	375	1460
SARS-CoV-2 IgG antibodies tested employees	671 (61.8%)	236 (62.9%)	907 (62.1%)[a]
COVID-19 patients			
Hospitalized patients	29	0	29
Patients on intensive care unit	2	0	2
Patients died	4	0	4
Outpatients	24	0	24
SARS-CoV-2 infection in employees[b]			
IgG antibodies positive	5	8	13
PCR positive	4	0	4
PCR positive, IgG antibodies negative	2	0	2

[a] Total number of all tested employees is 907.

[b] Multiple answers possible.

= 124), 26–40 years (n = 254) and >40 years (n = 515). The background for this was the assumption that participants in these 3 categories might have different composition of their households (e.g. <25 years: less children, 26–40 years: young children, >40 years: older children) and consequently different risks for acquiring SARS-CoV-2 infection outside the hospital. However, we unfortunately did not collect these data (Table 2). The most common professions were nurses (n = 488; 53.8%), followed by medical doctors (n = 98; 10.8%), care workers (n = 78; 8.6%) administration staff (n = 81; 8.9%), therapists (n = 37; 4.1%), cleaning personnel (n = 31; 3.4%) and employees working in the kitchen service (n = 20; 2.2%). Six hundred ninety-six employees (76.7%) were HCWs with close contact to patients. Sixty-four HCWs (7.1%) were working in high-risk areas with regular contact to COVID-19 positive patients in the emergency department, COVID-19 ward or intensive care unit (ICU). Six hundred thirty-two HCWs (69.7%) worked in intermediate-risk areas with close contact to non-COVID-19 patients and two hundred eleven employees (23.3%) worked in low-risk areas without contact to patients. Fifty-three employees (5.8%) reported contact exposure to COVID-19 positive persons and 123 (13.6%) got a PCR test previously. Further details on characteristics of participants are provided in Table 2.

3.2. Seroprevalence from June 2020 to September 2020

Among the 907 participants, 894 (98.6%) were tested negative for SARS-CoV-2 IgG antibodies. Thirteen employees (1.4%) were tested positive for SARS-CoV-2 IgG antibodies.

When also considering previously reported PCR and serology results, cumulative SARS-CoV-2 infection rate was 1.7% as we identified two persons with previously positive PCR tests who had no detectable antibodies in our study (Table 1).

3.3. Risk factors and clinical symptoms of seropositive employees

Comparing both hospitals 0.7% of tested employees (5/671) working at the St. Vincenz Hospital Datteln and 3.4% employees (8/236) working at the St. Laurentius Stift Waltrop had SARS-CoV-2 IgG antibodies. The difference between both hospitals was statistical significant ($p = 0.007$) (Table 2).

A cluster of 7 seropositive HCWs (58.3%) worked on one psychiatric ward. One of these employees reported an exposure to a COVID-19 positive family member, but received an initially negative PCR result. Therefore this employee kept working with a surgical face mask until the second PCR test turned out to be positive. This happened at the beginning of April 2020, when we did not have established universal masking in the hospital, yet. Furthermore, we were not working generally with FFP-2 masks at that time. As the St. Laurentius Stift in Waltrop had no COVID-19 positive patient in the first infection period (Table 1), seropositive HCWs in this hospital presumably acquired their infection not in contact with patients, but more likely during break times or private contact.

Among the altogether 13 employees with detectable SARS-CoV-2 IgG antibodies we found the following characteristics: Twelve (92.3%) were HCWs, 1 (7.7%) was working with no direct contact to patients in the laboratory. The profession and risk group at work had no statistic significant influence on the risk of positive SARS-CoV-2 IgG antibody detection (Table 2).

Only 1 employee (7.7%) was working in contact with COVID-19 positive patients. Four (30.8%) reported known exposure to COVID-19 positive persons (at work or at home) and 10 (76.9%) had previously known positive PCR results. Statistical analysis revealed significance for the risk of SARS-CoV-2 positive antibody test in case of known contact to a COVID-19 positive person ($p = 0.005$) and previously performed PCR test ($p < 0.001$) (Table 2).

Three employees (23.1%) got to know about their previous infection only through the result of the SARS-CoV-2 antibody test. Only 5 employees (38.5%) reported clinical symptoms within the last two months: cold-like symptoms (2), fever (3), headache (3), cough (1) and taste or smell disorders (2). The number of clinical symptoms was not statistical significant. However, we found statistical significance for fever ($p = 0.036$) and taste or smell disorders ($p = 0.047$) as expected (Table 2).

3.4. Context to PCR positive patients and staff

From March 2020 to September 2020, we treated 53 SARS-CoV-2 positive patients in the St. Vincent Hospital Datteln: 29 patients were hospitalized in our COVID-19 ward, 2 of them were treated on the ICU, 4 patients died and 24 outpatients were seen at the emergency department. Surprisingly no COVID-19 positive patients were treated in the St. Laurentius Hospital Waltrop during the first period of infection (Table 1). Although we established general testing of patients on admission not until August 2020, all patients who were admitted from other hospitals to the hospital in Waltrop were tested for SARS-CoV-2 since April 2020 (Table 3).

On 27-Apr-2020 mandatory masking for all employees at the hospitals was implemented. Therewith we were nearly four weeks later than other regions in Germany, e.g. the University Hospital in Jena which implemented mandatory masking on 20-Mar-2020 (Bahrs et al., 2020).

We did not test employees routinely for SARS-CoV-2, but according to the recommendations of the Robert Koch Institute (RKI) at that time, employees were tested with PCR in case of cold-like symptoms of any severity, exposure to COVID-19 positive persons and returning from a region at risk (Robert Koch Institute, 2020). According to the RKI definitions we initiated 812 SARS-CoV-2 PCR tests in employees until the end of September 2020. Four employees working at the St. Vincenz Hospital Datteln were tested SARS-CoV-2 PCR positive, in 2 of them we found SARS-CoV-2 IgG antibodies in our study (Table 1).

Eight patients on regular wards turned out to be SARS-CoV-2 positive so that we did contact tracing and testing of contact patients and employees. At the time we did not have universal masking, one positive patient resulted in up to 54 contact persons that were not protected properly. Fortunately, none of the contact persons of these 8 patients was infected through the exposure.

We reduced the number of exposed persons in COVID-19 positive patients by implementation of universal masking of employees and patients in situation of close contact and training of the awareness of adequate protection.

Table 2
Characteristics of the study population - stratified by SARS-CoV-2 IgG antibody results.

Variable	Overall	SARS-CoV-2 IgG		Statistics	
		detectable	not detectable	OR [95% CI]	p-value
	n = 907	n = 13	n = 894		
Age					**0.722**
16–25 years	124 (13.9%)	1 (7.7%)	123 (14.0%)	ref.	
26–40 years	254 (28.4%)	5 (38.5%)	249 (28.3%)	2.22 [0.34; 59.2]	
>40 years	515 (57.7%)	7 (53.8%)	508 (57.7%)	1.51 [0.26; 38.8]	
Sex					**0.427**
Male	136 (15.0%)	3 (23.1%)	133 (14.9%)	1.78 [0.38; 6.01]	
Female	771 (85.0%)	10 (76.9%)	761 (85.1%)	ref.	
Hospital					**0.007**
St. Vincenz Hospital Datteln	671 (74.0%)	5 (38.5%)	666 (74.5%)	ref.	
St. Laurentius Stift Waltrop	236 (26.0%)	8 (61.5%)	228 (25.5%)	4.61 [1.49; 15.8]	
Profession					**0.721**
Nurse	488 (53.8%)	10 (76.9%)	478 (53.5%)	ref.	
Medical doctor	98 (10.8%)	0 (0.0%)	98 (11.0%)	.[.; .]	
Care worker	78 (8.6%)	2 (15.4%)	76 (8.5%)	1.26 [0.13; 6.07]	
Therapist	37 (4.1%)	0 (0.0%)	37 (4.1%)	.[.; .]	
Cleaning service	31 (3.4%)	0 (0.0%)	31 (3.5%)	.[.; .]	
Kitchen	20 (2.2%)	0 (0.0%)	20 (2.2%)	.[.; .]	
Administration staff	81 (8.9%)	0 (0.0%)	81 (9.1%)	.[.; .]	
Other profession	74 (8.2%)	1 (7.7%)	73 (8.2%)	0.66 [0.01; 4.73]	
Risk of COVID-19 infection					**0.430**
Low- risk group: working without patient contact	211 (23.3%)	1 (7.7%)	210 (23.5%)	ref.	
Intermediate-risk group HCWs	632 (69.7%)	11 (84.6%)	621 (69.5%)	3.32 [0.08; 130]	
High risk group HCWs	64 (7.1%)	1 (7.7%)	63 (7.1%)	3.29 [0.63; 81.3]	
Number of symptoms within the last 2 months[a]				**1.45 [0.87; 2.22]**	**0.106**
No symptoms	685 (75.5%)	8 (61.5%)	677 (75.7%)	.[.; .]	
1 symptom	113 (12.5%)	1 (7.7%)	112 (12.5%)	.[.; .]	
2 symptoms	53 (5.8%)	2 (15.4%)	51 (5.7%)	.[.; .]	
3 symptoms	47 (5.2%)	2 (15.4%)	45 (5.0%)	.[.; .]	
4 symptoms	9 (1.0%)	0 (0.0%)	9 (1.0%)	.[.; .]	
Clinical symptoms within the last 2 months[b]					
Cold-like symptoms	55 (6.1%)	2 (15.4%)	53 (5.9%)	3.05 [0.42; 12.0]	0.184
Headache	99 (10.9%)	3 (23.1%)	96 (10.7%)	2.58 [0.54; 8.76]	0.161
Fever	53 (5.8%)	3 (23.1%)	50 (5.6%)	5.22 [1.09; 18.0]	0.036
Cough	152 (16.8%)	1 (7.7%)	151 (16.9%)	0.46 [0.02; 2.40]	0.707
Hoarseness	12 (1.3%)	0 (0.0%)	12 (1.3%)	.[.; .]	1.000
Taste or smell disorders	25 (2.8%)	2 (15.4%)	23 (2.6%)	7.24 [0.98; 29.6]	0.047
Other reported information[b]					
Contact to COVID-19 positive person	53 (5.8%)	4 (30.8%)	49 (5.5%)	7.78 [1.98; 25.3]	0.005
Previous PCR testing[c]	123 (13.6%)	10 (76.9%)	113 (12.6%)	22.1 [6.53; 105]	<0.001

The number of participants (n) is provided. Furthermore results from univariable logistic regression are provided, giving odds ratios and 95% confidence intervals versus the reference level for each main category of the characteristics or as change per unit (number of symptoms). We applied a significance level of 0.05. Abbreviations: ref. - reference category,.[.; .] - not applicable.

[a] Clinical symptoms that were reported: cold-like symptoms, headache, fever, cough, hoarseness, taste or smell disorders.

[b] Multiple answers possible.

[c] Employees were previously tested with PCR according to the R.K.I. recommendations at the time of investigation: in case of i) cold-like symptoms of any severity, ii) exposure to COVID-19 positive persons and iii) returning from a region at risk. Previous positive PCR test: 2 employees with detectable SARS-CoV-2 IgG antibodies and 2 employees without detectable SARS-CoV-2 IgG antibodies.

As RKI recommendations were adapted continuously according to the knowledge of science we implemented infection control measures in both hospitals (Table 3).

4. Discussion

We found a low seroprevalence (1.4%) of SARS-CoV-2 IgG antibodies in the investigated employees of two hospitals belonging to a secondary care hospital network in North Rhine-Westphalia, Germany. Even if we also consider previously reported positive PCR results of seronegative employees we just reached an infection rate of 1.7%. Two other studies in Germany reported seroprevalence rates of hospital workers: 2.7% at the University hospital in Jena (Bahrs et al., 2020) and 1.8% at the University Medical Center Hamburg-Eppendorf (Brehm et al., 2021).

Regarding HCWs in our study, 1.7% of them (12/696) had detectable SARS-CoV-2 IgG antibodies. Similiar results were published from Korth et al. with 1.6% seropositive HCWs (5/316) at the University Hospital Essen, which is a closely related region in North Rhine-Westphalia, Germany (Korth et al., 2020). We detected the highest seroprevalence in intermediate-risk HCWs (1.7%), followed by high-risk HCWs (1.6%)

Table 3
Implementation of infection control measures in both hospitals during the first period of SARS-CoV-2 infection.

Date	Implementation of infection control measures
06-Mar-20	No visitors with cold-like symptoms are allowed
07-Mar-20	Restricted visitors (1 visitor/patient/day)
16-Mar-20	Corona Hotline for employees Cancellation of all elective operations and investigations Cancellation of all internal educational trainings
17-Mar-20	General prohibition of visitors
18-Mar-20	Closure of day hospitals in Waltrop (geriatrics and psychiatry)
24-Mar-20	Limited number of persons in lifts and rooms according to the size in m^2
25-Mar-20	Occupation of 3-bed room with only 2 patients
31-Mar-20	Facemasks for employees only after exposure to COVID-19 patients
02-Apr-20	Facemasks for employees only in direct contact to patients
03-Apr-20	Psychosocial Hotline for employees
07-Apr-20	PCR testing of patients before transfer to other hospitals or care institutions
13-Apr-20	Facemasks for hospitalized patients in investigations
22-Apr-20	Facemasks for outpatients
27-Apr-20	Universal masking for all employees
29-Apr-20	PCR testing of patients with ambulant nursing service before discharge
08-May-20	Risk-adapted PCR screening of all patients
19-May-20	Restricted visitors allowed: 1 visitor/patient/day
26-May-20	End of the psychosocial hotline for employees
10-Aug-20	General PCR screening for all patients
14-Oct-20	Restricted visitors allowed: 1 visitor/patient/every 5 days

and the lowest seroprevalence in low-risk employees (0.5%). Two other studies in Germany reported similiar results: a higher seroprevalence rate in intermediate-risk HCWs (Essen: 5.4%; Jena: 2.9%) compared to high-risk HCWs (Essen:1.2%; Jena: 1.5%) (Bahrs et al., 2020; Korth et al., 2020). Bahrs et al. even found the highest seroprevalence rate in employees working in low-risk areas (3.3%) (Bahrs et al., 2020).

Another study at the University hospital in Münster, Germany investigated HCWs with PCR soon after reported exposure to COVID-19 positive persons. In this setting they found 5.4% of tested HCWs infected. As HCWs with no known exposure were not tested in this setting, infection rate in HCWs in total was probably lower (Schwierzeck et al., 2020). According to a recent analysis of SARS-CoV-2 infections reported to the RKI, in Germany 273 720 laboratory confirmed infections were recorded until the end of September 2020, 15 946 (5.8%) in employees in medical institutions (Kramer et al., 2020).

Seroprevalence rates among HCWs outside Germany range from 4.0% to 11.9% (Garcia-Basteiro et al., 2020; Goenka et al., 2020; Iversen et al., 2020; Self et al., 2020). A New York City (NYC) hospital even reported a SARS-CoV-2 antibodies seroprevalence rate of 27% in HCWs (Venugopal et al., 2020).

Personal protective equipment was available in our hospitals all the time. As the RKI recommendations to prevent infections in healthcare facilities were adapted continuously, we started to screen all visitors for symptoms of COVID-19 infection with a questionnaire and visitors used facemasks. Additionally we implemented universal masking of employees and patients in close contact at 27-Apr-2020. Later on we extended the use of facemasks to HCWs all the time and implemented a risk adapted screening of all patients resulting in a PCR test of high-risk patients. Since 10-Aug-2020 all patients in both hospitals were screened with PCR on admission.

Even if we had no documented SARS-CoV-2 infection that clearly resulted from contact to a positive patient, we cannot rule out this scenario.

The strength of our study is the high percentage of employees that took part, representing 62.1% of all employees. Nevertheless, our study had limitations. As we did not test at defined points, we are not able to evaluate the effect of the described infection control interventions on prevention of nosocomial transmissions.

5. Conclusion

In our study we conclude that the two included hospitals were not hotspots for SARS-CoV-2 infection until the end of September 2020. The seroprevalence rate was low and we had no documented transmission of the infection that clearly resulted from contact to COVID-19 positive patients. Although we had numerically more HCWs with detectable SARS-CoV-2 IgG antibodies than employees with no contact to patients, community transmission might have played a larger role for COVID-19 infection than professional exposure during the first period of infection. However, this resulted from an overall low exposure of hospital employees to COVID-19 positive patients in the investigated hospitals at a time where the region was not a SARS-CoV-2 hotspot. As the knowledge about the way of transmission, symptoms and diagnosis of COVID-19 is increasing, it will be necessary to adapt the concept of infect prevention continuously to keep the hospital a safe place.

Funding

The study was financed by internal funding.

Availability of data and material

The datasets and materials used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Anke Hildebrandt and Oktay Hökelekli. Statistic analysis was performed by Henrik Rudolf. The first draft of the manuscript was written by Anke Hildebrandt and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Ethics approval

Research was conducted in accordance with the declaration of Helsinki and national standards. The study protocol was approved by the local ethics committee: Ärztekammer Westfalen-Lippe and Westfälische Wilhelms Universität Münster (approval no.2020-478-f-S). The study was registered at the German Clinical Trials Register (DRKS00022941).

Consent to participate

Informed consent was obtained from all included participants included in the study.

Consent for publication

Informed consent was obtained from all included participants that anonymized data will be published.

Declaration of competing interest

On behalf of all authors, the corresponding author states that there are no competing interests to declare.

Acknowledgements

We thank Wolfgang Mueller for the possibility of internal funding and Alide Vrielink for excellent technical support.

References

Abebe, E.C., Dejenie, T.A., Shiferaw, M.Y., Malik, T., 2020. The newly emerged COVID-19 disease: a systemic review. Virol. J. 17, 96.

Bahrs, C., Kimmig, A., Weis, S.J.A., Hagel, S., Maschmann, J., Stallmach, A., Steiner, S., Bauer, M., Behrigner, W., Baier, M., Kesselmeier, M., Richert, C., Zepf, F., Walter, M., Scherag, A., Kiehntopf, M., Loeffler, B., Pletz, M.W., 2020. Seroprevalence of SARS CoV-2 antibodies in healthcare workers and administration employees: a prospective surveillance study at a 1,400-bed university hospital in Germany. Transbound. Emerg. Dis. https://doi.org/10.1111/tbed.14041.

Brehm, T.T., Schwinge, D., Lampalzer, S., Schlicker, V., Kuchen, J., Thompson, M., Ullrich, F., Huber, S., Schmiedel, S., Addo, M.M., Lutgehetmann, M., Knobloch, J.K., Schulze Zur Wiesch, J., Lohse, A.W., 2021. Seroprevalence of SARS-CoV-2 antibodies among hospital workers in a German tertiary care center: a sequential follow-up study. Int. J. Hyg Environ. Health 232, 113671.

Chou, R., Dana, T., Buckley, D.I., Selph, S., Fu, R., Totten, A.M., 2020a. Epidemiology of and risk factors for coronavirus infection in health care workers: a living rapid review. Ann. Intern. Med. 173, 120–136.

Chou, R., Dana, T., Buckley, D.I., Selph, S., Fu, R., Totten, A.M., 2020b. Update alert 5: epidemiology of and risk factors for coronavirus infection in health care workers. Ann. Intern. Med. 173 (11) https://doi.org/10.7326/L20-1227. W154-W55, Epub 2020 Oct 20.

Dhama, K., Khan, S., Tiwari, R., Sircar, S., Bhat, S., Malik, Y.S., Singh, K.P., Chaicumpa, W., Bonilla-Aldana, D.K., Rodriguez-Morales, A.J., 2020. Coronavirus disease 2019-COVID-19. Clin. Microbiol. Rev. 33.

Fischer-Fels, J., 2020. Gesundheitspersonal und COVID-19: infektionszahlen nehmen zu. Dtsch. Ärztebl. 117. A-1484/B-1275.

Garcia-Basteiro, A.L., Moncunill, G., Tortajada, M., Vidal, M., Guinovart, C., Jimenez, A., Santano, R., Sanz, S., Mendez, S., Llupia, A., Aguilar, R., Alonso, S., Barrios, D., Carolis, C., Cistero, P., Choliz, E., Cruz, A., Fochs, S., Jairoce, C., Hecht, J., Lamoglia, M., Martinez, 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., Munoz, J., Trilla, A., Varela, P., Mayor, A., Dobano, C., 2020. Seroprevalence of antibodies against SARS-CoV-2 among health care workers in a large Spanish reference hospital. Nat. Commun. 11, 3500.

Goenka, M., Afzalpurkar, S., Goenka, U., Das, S.S., Mukherjee, M., Jajodia, S., Shah, B.B., Patil, V.U., Rodge, G., Khan, U., Bandyopadhyay, S., 2020. Seroprevalence of COVID-19 amongst health care workers in a tertiary care hospital of a metropolitan city from India. J. Assoc. Phys. India 68, 14–19.

Guan, W.J., Ni, Z.Y., Hu, Y., Liang, W.H., Ou, C.Q., He, J.X., Liu, L., Shan, H., Lei, C.L., Hui, D.S.C., Du, B., Li, L.J., Zeng, G., Yuen, K.Y., Chen, R.C., Tang, C.L., Wang, T., Chen, P.Y., Xiang, J., Li, S.Y., Wang, J.L., Liang, Z.J., Peng, Y.X., Wei, L., Liu, Y., Hu, Y.H., Peng, P., Wang, J.M., Liu, J.Y., Chen, Z., Li, G., Zheng, Z.J., Qiu, S.Q., Luo, J., Ye, C.J., Zhu, S.Y., Zhong, N.S., China Medical Treatment Expert Group for, C, 2020. Clinical characteristics of coronavirus disease 2019 in China. N. Engl. J. Med. 382, 1708–1720.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395, 497–506.

Iversen, K., Bundgaard, H., Hasselbalch, R.B., Kristensen, J.H., Nielsen, P.B., Pries-Heje, M., Knudsen, A.D., Christensen, C.E., Fogh, K., Norsk, J.B., Andersen, O., Fischer, T.K., Jensen, C.A.J., Larsen, M., Torp-Pedersen, C., Rungby, J., Ditlev, S.B., Hageman, I., Mogelvang, R., Hother, C.E., Gybel-Brask, M., Sorensen, E., Harritshoj, L., Folke, F., Sten, C., Benfield, T., Nielsen, S.D., Ullum, H., 2020. Risk of COVID-19 in health-care workers in Denmark: an observational cohort study. Lancet Infect. Dis. 20 (12), 1401–1408. https://doi.org/10.1016/S1473-3099(20)30589-2. Epub 2020 Aug 3.PMID: 32758438.

Khonyongwa, K., Taori, S.K., Soares, A., Desai, N., Sudhanva, M., Bernal, W., Schelenz, S., Curran, L.A., 2020. Incidence and outcomes of healthcare-associated COVID-19 infections: significance of delayed diagnosis and correlation with staff absence. J. Hosp. Infect. 106 (4), 663–672. https://doi.org/10.1016/j.jhin.2020.10.006. Epub 2020 Oct 13.

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. : Off. Publ. Pan Am. Soc. Clin. Virol. 128, 104437.

Kramer, R., Diercke, M., Klingeberg, A., Claus, H., Hecht, J., Sin, M.A., Eckmanns, T., Haller, S., 2020. COVID-19 – meldedatenanalyse zu Infektionen im Gesundheitswesen. Dtsch Arztebl Int. 117, 809–810.

Li, C., Zhao, C., Bao, J., Tang, B., Wang, Y., Gu, B., 2020. Laboratory diagnosis of coronavirus disease-2019 (COVID-19). Clinica Chimica Acta; Int. J. Clin. Chem. 510, 35–46.

Meselson, M., 2020. Droplets and aerosols in the transmission of SARS-CoV-2. N. Engl. J. Med. 382, 2063.

Patel, K.P., Vunnam, S.R., Patel, P.A., Krill, K.L., Korbitz, P.M., Gallagher, J.P., Suh, J.E., Vunnam, R.R., 2020. Transmission of SARS-CoV-2: an update of current literature. Eur. J. Clin. Microbiol. Infect. Dis. 39, 2005–2011.

Pergolizzi Jr., J.V., Magnusson, P., LeQuang, J.A., Breve, F., Paladini, A., Rekatsina, M., Yeam, C.T., Imani, F., Saltelli, G., Taylor Jr., R., Wollmuth, C., Varrassi, G., 2020. The current clinically relevant findings on COVID-19 pandemic. Anesthesiol. Pain Med. 10, e103819.

Robert Koch Institute, 2020. https://web.archive.org/web/20200424144252/https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Massnahmen_Verdachtsfall_Infografik_Tab.html.

Ralph, R., Lew, J., Zeng, T., Francis, M., Xue, B., Roux, M., Toloue Ostadgavahi, A., Rubino, S., Dawe, N.J., Al-Ahdal, M.N., Kelvin, D.J., Richardson, C.D., Kindrachuk, J., Falzarano, D., Kelvin, A.A., 2020. 2019-nCoV (Wuhan virus), a novel Coronavirus: human-to-human transmission, travel-related cases, and vaccine readiness. J. Infect. Develop. Countr. 14, 3–17.

Schwierzeck, V., Correa-Martinez, C.L., Schneider, K.N., Mellmann, A., Hennies, M.T., Hafezi, W., Czeschinski, P., Kampmeier, S., 2020. SARS-CoV-2 in the employees of a large university hospital. Dtsch Arztebl Int. 117, 344–345.

Self, W.H., Tenforde, M.W., Stubblefield, W.B., Feldstein, L.R., Steingrub, J.S., Shapiro, N.I., Ginde, A.A., Prekker, M.E., Brown, S.M., Peltan, I.D., Gong, M.N., Aboodi, M.S., Khan, A., Exline, M.C., Files, D.C., Gibbs, K.W., Lindsell, C.J., Rice, T. W., Jones, I.D., Halasa, N., Talbot, H.K., Grijalva, C.G., Casey, J.D., Hager, D.N., Qadir, N., Henning, D.J., Coughlin, M.M., Schiffer, J., Semenova, V., Li, H., Thornburg, N.J., Patel, M.M., Team, C.C.-R., Network, I.V.Y., 2020. Seroprevalence of SARS-CoV-2 among frontline health care personnel in a multistate hospital network - 13 academic medical centers, april-june 2020. MMWR. Morbid. Mortal. Week. Rep. 69, 1221–1226.

Venugopal, U., Jilani, N., Rabah, S., Shariff, M.A., Jawed, M., Batres, A.M., Abubacker, M., Menon, S., Pillai, A., Shabarek, N., Kasubhai, M., Dimitrov, V., Menon, V., 2020. SARS-CoV-2 seroprevalence among health care workers in a New York city hospital: a cross-sectional analysis during the COVID-19 pandemic. Int. J. Infect. Dis. : Int. J. Infect. Dis. 102, 63–69. https://doi.org/10.1016/j.ijid.2020.10.036 (official publication of the International Society for Infectious Diseases), Epub 2020 Oct 16.

Whitworth, J., 2020. COVID-19: a fast evolving pandemic. Trans. R. Soc. Trop. Med. Hyg. 114, 241–248.

Zhao, D., Wang, M., Wang, M., Zhao, Y., Zheng, Z., Li, X., Zhang, Y., Wang, T., Zeng, S., Hu, W., Yu, W., Hu, K., 2020. Asymptomatic infection by SARS-CoV-2 in healthcare workers: a study in a large teaching hospital in Wuhan, China. Int. J. Infect. Dis. : Int. J. Infect. Dis. 99, 219–225 official publication of the International Society for Infectious Diseases.

International Journal of Hygiene and Environmental Health

doi: 10.1016/S1438-4639(21)00101-2

International Journal of Hygiene and Environmental Health 235 (2021) 113766

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

Estimating ambient air pollutant levels in Suzhou through the SPDE approach with R-INLA

Neil Wright[a,*], Katherine Newell[a], Kin Bong Hubert Lam[a], Om Kurmi[b], Zhengming Chen[a], Christiana Kartsonaki[a,c]

[a] Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom
[b] Faculty Research Centre for Intelligent Healthcare, Coventry University, Coventry, United Kingdom
[c] MRC Population Health Research Unit, University of Oxford, Oxford, United Kingdom

ARTICLE INFO

Keywords:
Integrated nested Laplace approximation
Stochastic partial differential equation
Bayesian approach
Ambient air pollution
Covariate misalignment

ABSTRACT

Spatio–temporal models of ambient air pollution can be used to predict pollutant levels across a geographical region. These predictions may then be used as estimates of exposure for individuals in analyses of the health effects of air pollution. Integrated nested Laplace approximations is a method for Bayesian inference, and a fast alternative to Markov chain Monte Carlo methods. It also facilitates the SPDE approach to spatial modelling, which has been used for modelling of air pollutant levels, and is available in the R-INLA package for the R statistics software. Covariates such as meteorological variables may be useful predictors in such models, but covariate misalignment must be dealt with. This paper describes a flexible method used to estimate pollutant levels for six pollutants in Suzhou, a city in China with dispersed air pollutant monitors and weather stations. A two-stage approach is used to address misalignment of weather covariate data.

1. Introduction

Research into the health effects of ambient air pollution requires long-term measurements of pollutant exposure at the individual level. Studies on longer-term effects of air pollution exposure in China (and elsewhere) have used averaged concentrations from static ambient air pollution monitors (Cao et al., 2011; Chen et al., 2016; Dong et al., 2012; Li et al., 2014; Zhang et al., 2011; Zhou et al., 2014) or satellite data (Peng et al., 2017; Yin et al., 2017) for exposure information. For analyses that require individual exposure levels for study participants, spatio–temporal models of ambient air pollution may be used to predict pollutant levels across a geographical region. Predictions at individuals' residential or employment locations can then be used as estimates of ambient air pollution exposure.

Bayesian inference offers a practical method for applying such spatio–temporal models and producing predictions. Integrated nested Laplace approximations (INLA) (Rue et al., 2009) allow fast computation for Bayesian inference and enable the use of the SPDE approach for spatial modelling (Lindgren et al., 2011). These methods have been used in modelling of air pollutant levels in Italy (Cameletti et al., 2013; Fioravanti et al., 2021) and England (Blangiardo et al., 2016).

Meteorological variables can be useful predictors in models of ambient air pollution, but weather station locations may not coincide with the pollutant monitor locations or locations where predictions are sought. This is a case of the problem of covariate misalignment, where covariate data are not available at the same locations as observed dependent data. Joint modelling (Barber et al., 2016) or error models can be used to incorporate such covariates while accounting for uncertainty.

We obtained estimated pollutant exposure levels for participants in Suzhou in the China Kadoorie Biobank study. Data were available from static monitors for six pollutants: fine ($PM_{2.5}$), and coarse (PM_{10}) particulate matter, sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3). We used Bayesian spatial-temporal models to predict monthly levels of each pollutant at all clinic locations in the area. These predictions can be used as proxies for individual pollution exposure in analyses of health outcomes. This method exploits the spatial information from having monitors in different locations, providing localised exposure estimates that are not available by averaging pollution levels across the study area. Weather data were also

* Corresponding author. CTSU, Nuffield Department of Population Health, Big Data Institute, Old Road Campus, University of Oxford, Oxford, OX3 7LF, United Kingdom.
E-mail address: neil.wright@ndph.ox.ac.uk (N. Wright).

https://doi.org/10.1016/j.ijheh.2021.113766
Received 10 February 2021; Received in revised form 26 April 2021; Accepted 3 May 2021
Available online 24 May 2021

Fig. 1. Locations of clinics (black squares), pollutant monitors (triangles) and weather stations (circles).

available from stations in the city. However, the locations of weather monitors did not coincide with the pollutant monitors. Previous work (Blangiardo et al., 2016; Cameletti et al., 2013; Fioravanti et al., 2021) has used covariates that are fixed over space, or used the geographically closest available measurements. Given the limited number and placement of weather monitors, we used a two-stage approach to address misalignment of weather covariate data, and compare four models for including weather covariates in the pollutant models.

2. Background

2.1. China Kadoorie Biobank study

The China Kadoorie Biobank study (Chen et al., 2005) recruited 512,726 participants between 2004 and 2008, from ten diverse areas of China. Participants are followed up for a wide range of health outcomes via linkages with health insurance systems, established disease surveillance systems and death registries. Details of the study design and methods have been reported previously (Chen et al, 2005, 2011). In Suzhou, 53,269 study participants were recruited each of whom is linked to one of 77 local clinics. One clinic located outside the urban area of Suzhou was excluded from this analysis.

2.2. INLA and SPDE spatial models

Integrated nested Laplace approximations (INLA) (Rue et al., 2009; Wang et al., 2018) is a fast alternative to Markov chain Monte Carlo (MCMC) methods for Bayesian inference from latent Gaussian models. The method uses numerical integration and Laplace approximations for approximate Bayesian inference and is implemented in the R package R-INLA (R-INLA, 2020). The package includes many latent models, including SPDE spatial models (Lindgren et al., 2011), error models (Muff et al., 2013), and auto-regressive models. Posterior predictive distributions produced by fitting Bayesian models can be used to generate point or ranges of predictions.

The SPDE approach to spatial modelling, implemented in the R-INLA package, involves representing a continuously indexed Gaussian field with Matérn covariance as a discretely indexed Gaussian Markov

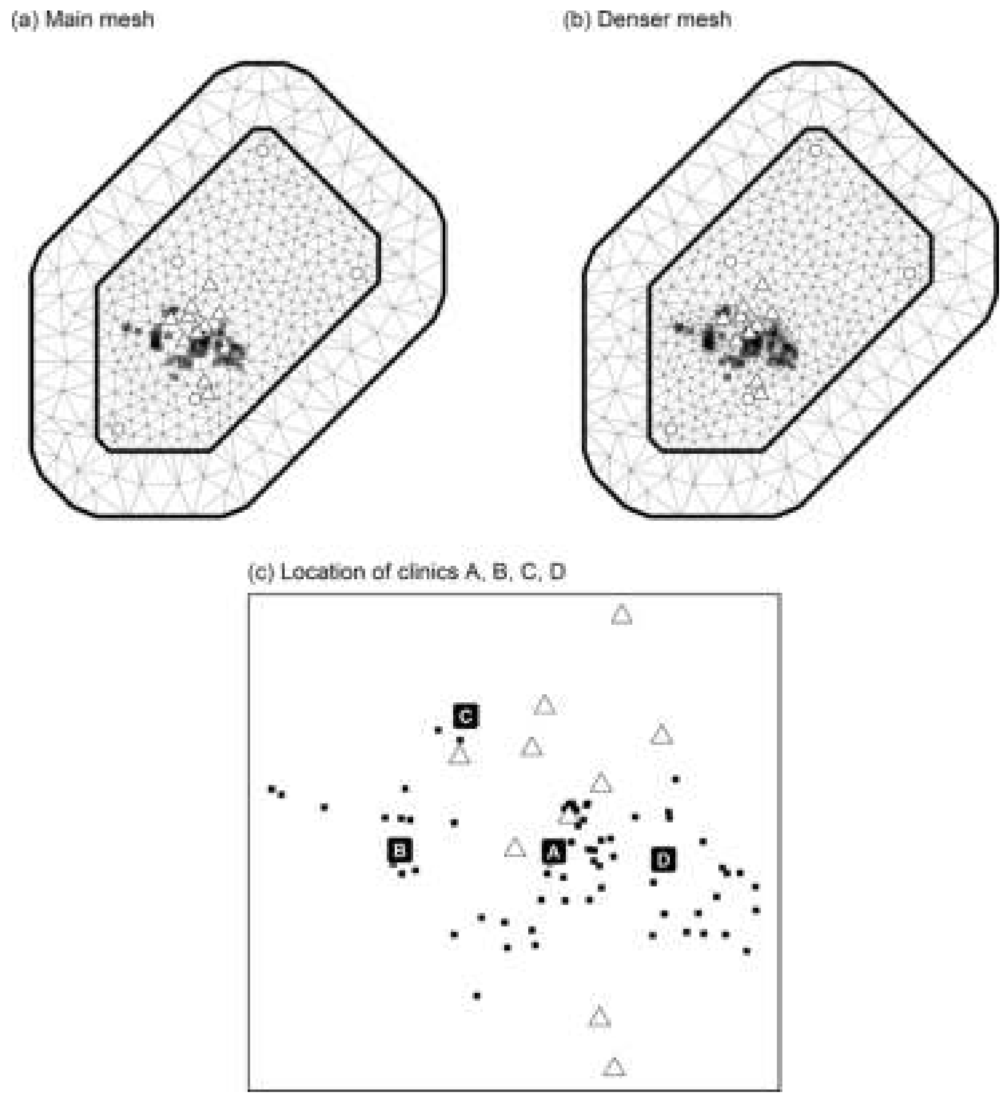


Fig. 2. Meshes and locations of clinics (black squares), pollutant monitors (triangles) and weather stations (circles).

random field (GMRF). This is achieved by means of a basis function representation defined on a triangulation of the domain. The GMRF has a sparse precision matrix and so computationally efficient methods for matrix factorisation, and INLA methods for Bayesian inference, can be used (Bakka et al., 2018; Bivand et al., 2015; Blangiardo et al., 2013; Blangiardo and Cameletti, 2015; Gomez-Rubio, 2020; Krainski et al., 2019; Lindgren et al., 2011; Lindgren and Rue, 2015; Moraga, 2019). The SPDE approach to spatial modelling and the Matérn covariance function and its parameters are well described in Chapter 6 of Blangiardo and Cameletti (2015). Separable space-time models, defined by the Kronecker product between the two precision matrices, can be constructed using the group feature in R-INLA. This allows spatial and temporal correlations to be jointly modelled. This form of spatio–temporal model is well described in Chapter 7 of Blangiardo and Cameletti (2015) and Chapter 10 of Moraga (2019). These methods have previously been used for spatio–temporal modelling of PM_{10} levels in Italy (Cameletti et al., 2013; Fioravanti et al., 2021) and NO_2 levels in England (Blangiardo et al., 2016).

2.3. Pollution and weather data

The data included daily average measurements of six pollutants: particulate matter with diameter of 2.5 μm or less ($PM_{2.5}$), particulate matter with a diameter between 2.5 and 10 μm (PM_{10}), sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3). Measurements were available between January 2013 and December 2015 from up to 10 pollution monitors situated in Suzhou (as shown in Fig. 1). Daily weather data, including temperature, pressure, precipitation and wind speed, were available from five monitors in the region from January 2013 to June 2016. The locations of the weather monitors are also shown in Fig. 1 and do not coincide with the locations of the pollution monitors or clinics. Five geographic covariates were available for all locations: elevation; distance to nearest major road; distance to nearest motorway; total length of major roads and motorways in a 1 km radius; and land use (a binary variable representing “urban” or “non-urban”). Elevation values were interpolated from the values of the four nearest raster cells.

3. Methods

In order to address the misalignment problem and use weather data variables as covariates in the pollutant models, two stages of models were used. Firstly, models for each weather variable were used to obtain predictions of each weather variable at pollution monitor and clinic locations. These predictions were then included as covariates in the models for the pollutants. Approximate Bayesian inference was performed using INLA with the R-INLA software package (R-INLA, 2020) for R.

Table 1
Summaries of observed data from five weather monitors (from January 2013 to June 2016) and up to 13 pollutant monitors (from January 2013 to December 2015).

	N	Mean	SD	Minimum	Median	Maximum
Daily values						
Weather variables						
Temperature (°C)	6385	16.8	8.8	−6.1	17.7	36.2
Wind speed (m/s)	6385	4.6	1.5	1.3	4.4	15.7
Humidity (%)	6385	73.8	13.5	29.0	75.0	100.0
Precipitation (mm)	5927	4.0	11.2	0.0	0.0	170.2
Pollutants (μg/m^3)						
PM_{10}	10,230	87.3	48.5	3.0	76.0	429.0
$PM_{2.5}$	10,209	63.1	39.1	3.0	55.0	405.0
SO_2	10,225	24.2	15.3	1.0	21.0	164.0
CO (mg/m^3)	10,237	0.9	0.4	0.1	0.8	3.5
NO_2	10,240	50.3	22.1	5.0	47.0	321.0
O_3	10,213	97.4	52.0	1.0	90.0	1251.0
Monthly values						
Weather variables						
Temperature (°C)	210	16.8	8.3	3.6	16.8	32.3
Wind speed (m/s)	210	4.6	0.7	3.5	4.6	6.7
Humidity (%)	210	73.8	6.7	56.4	74.0	91.9
Precipitation (mm)	210	4.1	3.5	0.2	3.3	25.0
Pollutants (μg/m^3)						
PM_{10}	348	87.2	27.2	43.2	80.2	194.3
$PM_{2.5}$	348	63.0	23.5	24.5	58.7	155.6
SO_2	348	24.1	10.2	8.5	21.9	63.4
CO (mg/m^3)	348	0.9	0.2	0.4	0.9	1.8
NO_2	348	50.7	15.1	24.3	48.5	109.3
O_3	348	95.9	34.5	18.6	104.5	182.2

This analysis used pollutant and weather variables aggregated to monthly means. This reduced the size of the data being used and number of values to be estimated, and made distributions approximately normal. For example, daily rainfall data are highly skewed with many zeroes, but the observed monthly average rainfall has a symmetric distribution. Observed daily pollutant levels of zero were set to missing, as these were believed to indicate errors in the data.

3.1. Meshes for SPDE spatial model

A mesh (triangulation) of the region was required to apply SPDE spatial models. The same mesh was used for all weather variable and pollutant models. Latitude and longitude coordinates of all locations were converted to Universal Transverse Mercator (UTM) coordinates. All clinic and monitor locations are in UTM zone 51. These coordinates were then re-scaled with centre equal to the midpoint of all monitor and clinic locations and so that 1 unit equals approximately 1 km. The R-INLA inla.mesh.2d function, which employs constrained refined Delaunay triangulation, was then used to construct a triangular mesh on the region. The domain was formed by the convex hull of all weather station, pollutant monitor and clinic locations, with an inner extension of 5 km and an outer extension of 15 km. The locations of all pollution monitors and weather monitors were used as initial triangulation nodes. The maximum edge ledge was set to 5 km (10 km in the outer extension), the minimum triangle angle was set to 28° (18° in the outer extension), and the minimum distance between points was set to 0.1 km. A denser mesh was also constructed using the locations of all clinics, as well as weather stations and pollutant monitors, as initial triangulation nodes. The weather and pollutant prediction models were additionally fit using this mesh, and point predictions (median of the posterior predictive distribution) of the pollutants were compared between using either mesh by Pearson correlation. The main mesh had 432 nodes and the denser mesh 711 nodes. The meshes are shown in Fig. 2.

3.2. Weather models

From the weather data, four variables were selected representing temperature (daily average temperature), humidity (daily average humidity), wind speed (daily 10 min maximum wind speed), and precipitation (total 24 h precipitation). The wind speed variable was log transformed. Each of the four weather variables was aggregated to monthly means and then re-scaled to have a mean of zero and a variance of one.

Each of the four weather variables was modelled as a Gaussian response. Model predictors with and without linear effects for space and time trends were compared using the Watanabe–Akaike (or "Widely Applicable") information criterion (WAIC) (Gelman et al., 2014; Watanabe, 2010). Calendar month was included as a factor. No level was dropped but the intercept term was dropped, so that prior distributions were exchangeable for levels of this factor. All models included a space-time model, using an SPDE spatial model (i.e. an approximation to a Gaussian field with Matérn covariance) for spatial correlations and a first order auto-regressive model for temporal correlations. Details and formulae for the models are provided in a supplementary file.

3.3. Pollutant models

After aggregation to monthly means, pollutant levels were log transformed and then modelled as Gaussian responses. The model predictors included spatial trends, a linear time trend, calendar month as a factor, five geographic covariates, and a space-time model with an SPDE spatial model (i.e. an approximation to a Gaussian field with Matérn covariance) and a first order auto-regressive model for temporal correlations. Continuous covariates were re-scaled to have mean zero and variance one for the pollution monitor and clinic locations. Spatial trends were included using terms for the x- and y-coordinates and the square of x- and y-coordinates. Allowing for a quadratic shape of trends (on the log scale for pollutants) prevented simple linear trends from being extrapolated in predictions for clinic locations far from the centre of the region. In particular, including only a simple linear trend led to extreme, implausible predictions for SO_2 levels at clinic locations in the far West of the region.

Four different approaches to include the standardised weather covariates in the pollutant models were compared:

1. Exclude weather covariates from the model predictor.
2. Include the mean of the values from each of the weather monitors, so that the same value is used for every location at the same time point.

Table 2
Medians and 95% HPD intervals of posterior distributions from models of monthly weather variables.

	Temperature (°C)	log (Wind speed m/s)	Humidity (%)	Precipitation (mm)
Month intercepts				
January	4.97 (−3.90, 13.72)	1.42 (0.64, 2.20)	73.93 (33.57, 114.84)	1.57 (−0.91, 4.06)
February	6.29 (−2.60, 15.04)	1.50 (0.72, 2.29)	76.31 (35.89, 117.27)	3.49 (1.01, 5.97)
March	11.01 (2.11, 19.77)	1.53 (0.75, 2.32)	71.92 (31.48, 112.91)	2.69 (0.21, 5.18)
April	16.11 (7.21, 24.87)	1.59 (0.81, 2.37)	70.57 (30.14, 111.57)	4.54 (2.05, 7.02)
May	21.14 (12.25, 29.89)	1.54 (0.75, 2.32)	73.82 (33.41, 114.80)	4.33 (1.85, 6.82)
June	24.01 (15.13, 32.74)	1.45 (0.66, 2.23)	83.38 (43.03, 124.30)	10.81 (8.33, 13.29)
July	28.53 (19.60, 37.31)	1.53 (0.75, 2.31)	77.39 (36.80, 118.57)	4.93 (2.08, 7.79)
August	28.15 (19.19, 36.97)	1.50 (0.71, 2.28)	78.71 (37.95, 120.05)	4.78 (1.91, 7.65)
September	24.09 (15.11, 32.92)	1.46 (0.67, 2.24)	78.91 (38.07, 120.33)	3.39 (0.52, 6.25)
October	19.20 (10.23, 28.03)	1.41 (0.62, 2.19)	74.32 (33.48, 115.75)	4.16 (1.30, 7.03)
November	13.20 (4.24, 22.02)	1.38 (0.60, 2.16)	77.13 (36.37, 118.47)	2.30 (−0.57, 5.17)
December	5.82 (−3.10, 14.61)	1.41 (0.63, 2.20)	70.13 (29.52, 111.29)	1.28 (−1.58, 4.13)
Hyperparameters				
SD for the Gaussian observations	0.08 (0.06, 0.09)	0.04 (0.03, 0.04)	1.15 (0.95, 1.38)	0.03 (0.01, 0.08)
Range of SPDE model (km)	378.63 (308.76, 461.48)	129.85 (99.81, 163.70)	247.64 (193.90, 310.69)	108.06 (91.61, 126.83)
Variance of SPDE model	0.01 (0.00, 0.03)	2.61 (0.81, 6.18)	1.04 (0.35, 2.44)	0.20 (0.16, 0.25)
Coefficient of AR model	0.96 (0.92, 0.99)	0.99 (0.97, 1.00)	0.96 (0.92, 0.99)	0.18 (0.03, 0.33)

Table 3
WAIC values for monthly pollutant models with different methods for using weather covariates.

Model	PM_{10}	$PM_{2.5}$	SO_2	CO	NO_2	O_3
1. Exclude weather covariates	−810.49	−932.64	−328.12	−467.15	−856.00	−1026.77
2. Mean values	−838.14	−1001.50	−350.41	−468.53	−2044.77	−2051.05
3. Means of posterior predictive distribution	−833.04	−998.91	−341.33	−467.04	−2027.96	−2057.84
4. Error models	**−1302.55**	**−1046.17**	**−389.43**	**−646.71**	**−2097.08**	**−2086.67**
5. Excluding SPDE model	−1885.11	−723.74	−319.38	−472.62	−1078.37	−2084.94
6. Excluding quadratic spatial terms	−2053.49	−1009.52	−378.29	−690.40	−2067.99	−2139.77

Table 4
Medians and 95% HPD intervals of posterior distributions from models of monthly means of particulate matter pollutants.

	PM_{10}	$PM_{2.5}$
Month intercepts (μg/m^3)		
January	65.69 (26.16, 136.98)	60.28 (20.52, 132.96)
February	52.57 (21.20, 106.33)	44.84 (15.98, 95.07)
March	63.38 (27.48, 116.50)	57.37 (23.88, 109.10)
April	66.83 (30.21, 116.49)	69.61 (31.51, 123.48)
May	83.60 (36.64, 145.90)	93.86 (40.65, 167.06)
June	88.18 (35.63, 158.04)	106.64 (40.72, 199.27)
July	90.04 (31.87, 168.05)	107.34 (33.93, 218.22)
August	91.20 (32.57, 170.06)	107.37 (34.28, 217.05)
September	75.56 (30.68, 135.92)	86.99 (33.60, 162.26)
October	79.34 (35.45, 138.56)	82.90 (36.95, 146.81)
November	77.35 (33.95, 139.55)	75.74 (32.37, 140.85)
December	66.15 (25.18, 142.37)	61.59 (19.71, 142.18)
Covariate coefficients (ratios)		
Time trend (per month)	1.00 (0.98, 1.02)	0.99 (0.97, 1.01)
Longitudinal trend (linear term)	1.13 (0.62, 1.76)	0.96 (0.51, 1.56)
Latitudinal trend (linear term)	1.00 (0.78, 1.25)	1.03 (0.79, 1.30)
Longitudinal trend (quadratic term)	1.55 (0.94, 2.25)	1.58 (0.93, 2.33)
Latitudinal trend (quadratic term)	1.07 (0.86, 1.31)	0.87 (0.69, 1.08)
Urban	0.85 (0.74, 0.98)	0.92 (0.80, 1.06)
Elevation (per 10m)	0.96 (0.73, 1.21)	0.85 (0.65, 1.10)
Distance from road (per 0.01)	0.77 (0.40, 1.28)	0.53 (0.27, 0.90)
Distance from motorway (per 0.01)	1.12 (0.98, 1.26)	1.02 (0.89, 1.16)
Length of roads and motorways in vicinity (per 1 km)	1.01 (0.98, 1.04)	1.00 (0.97, 1.03)
Coefficients in error models (ratios)		
Temperature (per 10C)	0.71 (0.50, 0.90)	0.59 (0.31, 0.90)
Wind speed (per 1 SD of log (wind speed))	0.94 (0.91, 0.97)	0.97 (0.94, 1.00)
Humidity (per 5%)	0.89 (0.87, 0.91)	0.97 (0.92, 1.01)
Precipitation (per 10 mm)	1.02 (0.90, 1.17)	0.97 (0.79, 1.16)
Hyperparameters		
SD of Gaussian observations (on log scale)	0.03 (0.02, 0.04)	0.04 (0.03, 0.05)
Range of the SPDE model (km)	42.36 (31.64, 51.60)	44.90 (28.53, 62.64)
SD of the SPDE model (on log scale)	0.21 (0.17, 0.25)	0.25 (0.18, 0.33)
Coefficient of AR model	0.80 (0.72, 0.86)	0.82 (0.68, 0.91)

Table 5
Medians and 95% HPD intervals of posterior distributions from models of monthly means of gaseous pollutants.

	SO_2	CO	NO_2	O_3
Month intercepts (μg/m³; mg/m³ for CO)				
January	26.26 (9.79, 52.99)	0.55 (0.01, 2.53)	64.50 (27.95, 117.11)	114.11 (57.02, 192.03)
February	18.24 (7.10, 36.05)	0.47 (0.01, 2.19)	47.24 (20.73, 85.19)	170.10 (86.03, 284.64)
March	25.36 (11.51, 45.60)	0.52 (0.01, 2.37)	66.30 (30.07, 117.38)	150.07 (79.16, 244.05)
April	32.02 (15.61, 54.86)	0.57 (0.02, 2.58)	67.36 (31.03, 117.93)	142.34 (76.50, 228.09)
May	35.70 (16.78, 61.54)	0.57 (0.02, 2.57)	58.56 (26.75, 102.62)	114.50 (60.87, 183.91)
June	42.10 (17.17, 76.31)	0.52 (0.01, 2.36)	56.44 (25.04, 100.51)	97.39 (49.84, 159.54)
July	52.03 (18.22, 101.37)	0.57 (0.01, 2.65)	53.89 (23.22, 98.06)	72.10 (35.11, 120.89)
August	51.60 (18.18, 100.35)	0.60 (0.02, 2.78)	54.95 (23.71, 99.87)	74.32 (36.25, 124.55)
September	45.19 (18.88, 82.12)	0.55 (0.01, 2.49)	56.42 (25.17, 100.18)	89.18 (45.84, 145.70)
October	39.89 (18.88, 68.25)	0.49 (0.01, 2.17)	62.15 (28.45, 108.73)	107.33 (57.24, 172.34)
November	40.12 (18.99, 69.65)	0.58 (0.02, 2.60)	67.05 (30.65, 117.96)	88.62 (47.34, 142.90)
December	31.55 (11.28, 64.90)	0.62 (0.02, 2.90)	66.10 (28.27, 120.83)	95.85 (47.34, 162.26)
Covariate coefficients (ratios)				
Time trend (per month)	0.98 (0.97, 0.99)	0.99 (0.97, 1.02)	1.00 (1.00, 1.01)	1.00 (1.00, 1.00)
Longitudinal trend (linear term)	0.42 (0.17, 0.78)	2.19 (0.03, 10.60)	0.71 (0.23, 1.46)	0.81 (0.34, 1.50)
Latitudinal trend (linear term)	1.02 (0.72, 1.38)	1.48 (0.35, 2.92)	0.96 (0.63, 1.36)	0.92 (0.65, 1.23)
Longitudinal trend (quadratic term)	0.85 (0.39, 1.49)	0.72 (0.02, 2.75)	1.61 (0.60, 3.16)	0.48 (0.21, 0.84)
Latitudinal trend (quadratic term)	1.15 (0.84, 1.51)	1.31 (0.33, 2.55)	0.77 (0.53, 1.04)	1.06 (0.79, 1.37)
Urban	1.53 (1.22, 1.88)	1.14 (0.57, 1.78)	0.85 (0.63, 1.09)	1.13 (0.88, 1.40)
Elevation (per 10m)	1.01 (0.67, 1.42)	1.16 (0.21, 2.43)	0.66 (0.40, 0.98)	1.19 (0.78, 1.68)
Distance from road (per 0.01)	1.33 (0.42, 2.80)	1.83 (0.01, 9.25)	0.23 (0.05, 0.58)	2.89 (0.83, 6.43)
Distance from motorway (per 0.01)	0.85 (0.70, 1.01)	1.18 (0.60, 1.83)	0.96 (0.76, 1.18)	0.91 (0.75, 1.09)
Length of roads and motorways in vicinity (per 1 km)	1.00 (0.96, 1.04)	0.98 (0.88, 1.09)	1.02 (0.98, 1.07)	1.01 (0.97, 1.05)
Coefficients in error models (ratios)				
Temperature (per 10C)	0.60 (0.35, 0.90)	0.90 (0.69, 1.15)	0.82 (0.63, 1.01)	1.96 (1.54, 2.45)
Wind speed (per 1 SD of log (wind speed))	0.99 (0.90, 1.08)	0.88 (0.84, 0.91)	0.90 (0.88, 0.93)	0.95 (0.89, 0.99)
Humidity (per 5%)	0.90 (0.85, 0.95)	0.97 (0.94, 0.99)	0.95 (0.92, 0.97)	0.97 (0.94, 0.99)
Precipitation (per 10 mm)	1.00 (0.78, 1.26)	1.11 (0.99, 1.24)	1.03 (0.95, 1.12)	0.90 (0.80, 1.00)
Hyperparameters				
SD of Gaussian observations (on log scale)	0.10 (0.07, 0.13)	0.07 (0.06, 0.09)	0.01 (0.00, 0.02)	0.01 (0.00, 0.03)
Range of the SPDE model (km)	14.48 (8.04, 22.75)	29.70 (18.49, 46.59)	7.61 (4.26, 11.05)	5.61 (3.66, 7.35)
SD of the SPDE model (on log scale)	0.20 (0.16, 0.25)	0.41 (0.19, 0.94)	0.17 (0.12, 0.21)	0.15 (0.13, 0.17)
Coefficient of AR model	0.60 (0.41, 0.75)	0.96 (0.88, 1.00)	0.79 (0.66, 0.87)	0.59 (0.49, 0.70)

3. Use the mean of the posterior predictive distribution at each location and time point, for each weather variable.
4. Use an error model. This was a Berkson error model, with observed values equal to the mean of the posterior predictive distribution at each location and time point, and precision fixed and equal to the precision of the posterior predictive distribution. This approach is similar to that investigated by Foster et al. (2012).

There is collinearity between the weather variables and calendar month, which complicates the interpretation of individual coefficients. However, the aim of these models is prediction of pollutant levels, rather than inference for individual coefficients, so this is not a concern and the prediction ability of the models is not affected (Shmueli, 2010).

Models were also fit for each pollutant excluding the SPDE model, but including the temporal first order auto-regressive random effects for stations. This allowed comparison between models which account for spatial correlation and models which ignore spatial correlation.

Details and formulae for the models are provided in a supplementary file.

3.4. Prediction models

Given a model with response variable Y and the predictor η,

$$Y \sim Normal\left(\eta, \sigma_e^2\right)$$

the posterior distribution of the predictor η is the posterior distribution of the mean response, not the posterior predictive distribution of the response itself. To obtain posterior predictive distributions of the response (including uncertainty due to all sources of error – modelled by σ_e^2) an adapted formulation of the model was used:

$$Y \sim Normal\left(\theta, e^{-20}\right)$$

The precision of the Gaussian response was fixed to be very large (e^{20}) so that the response Y is (effectively) equal to the value of the predictor θ. An independent and Gaussian distributed random effect was then added to the predictor, so $\theta = \eta + \varepsilon$ and $\varepsilon \sim Normal(0, \sigma_e^2)$. This strategy means that the posterior distribution of the predictor θ is the posterior predictive distribution of the response.

Posterior predictive distributions for pollutants were summarised by medians and 95% equal tailed intervals. Examples of the R code used for prediction models, for humidity and PM_{10}, are provided in a supplementary file.

3.5. Priors

The prior distributions for calendar month and other fixed effect parameters were Normal with mean 0 and precision 0.001. The priors for the precision of the responses were Gamma with a shape parameter of 1 and an inverse scale parameter of 5×10^{-5}. The priors for the coefficient, a, of the first order auto-regressive model were given by $log((1 + a)/(1 - a)) \sim Normal(0, 1/_{0.15})$. Normal priors were used for the coefficients of Berkson error models with mean 1 and precision 0.001. The mean and precision parameters for the error models were fixed values and therefore do not have prior or posterior distributions.

For the SPDE spatial models, penalised complexity (PC) priors were used with $P(r < 10) = 0.5$ and $P(\sigma > 1) = 0.5$, where r is the range and σ the standard deviation of the field.

3.6. Posteriors and model fit statistics

Posterior distributions for parameters and hyperparameters were

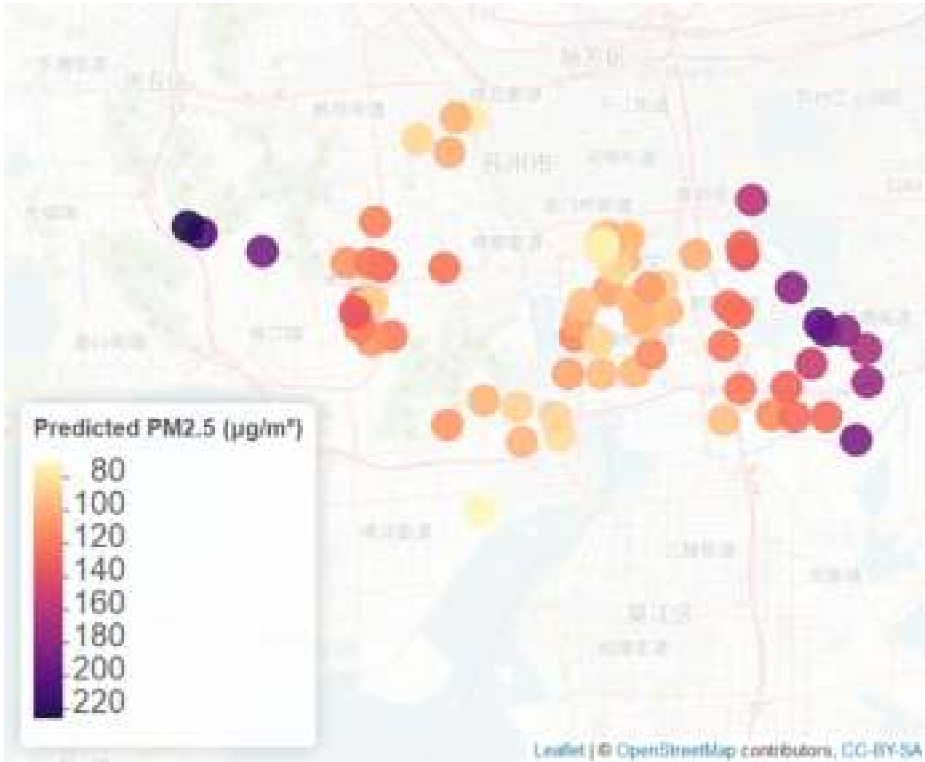


Fig. 3. Predicted levels (posterior medians) of $PM_{2.5}$ at clinic locations for January 2014.

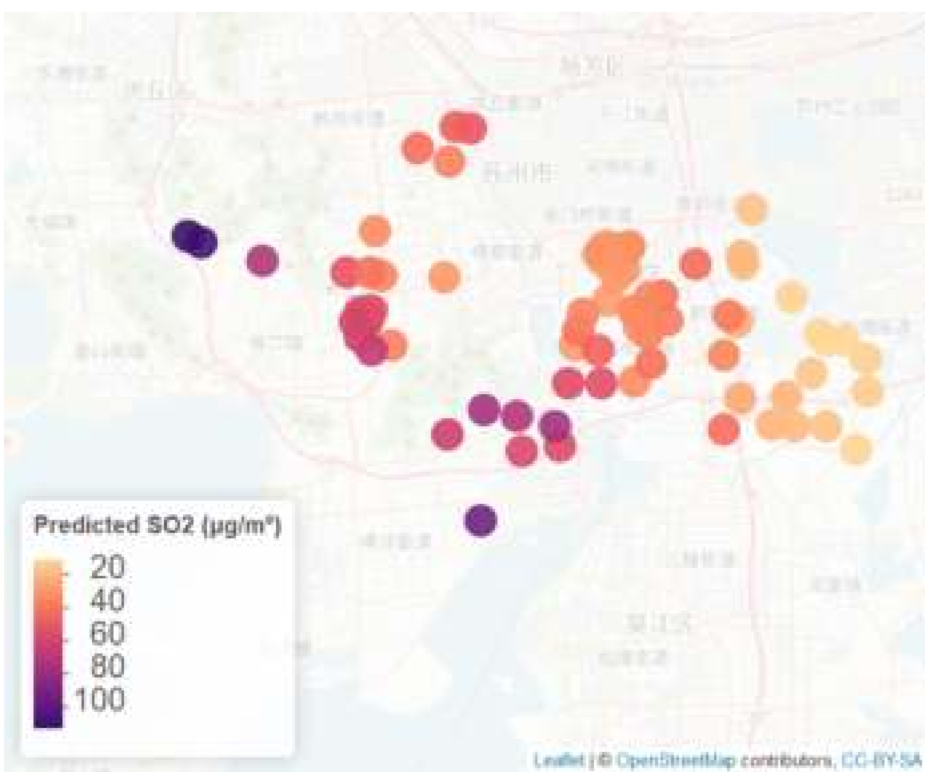


Fig. 4. Predicted levels (posterior medians) of SO_2 at clinic locations for January 2014.

transformed to the original scale of the dependent variable as applicable, and then summarised by medians and 95% highest posterior density (HPD) intervals. As pollutant levels were log transformed in the models, the exponentiated covariate coefficients are interpretable as ratios. For the SPDE models, hyperparameters were transformed to the range and variance.

Models were compared using the Watanabe–Akaike (or "Widely Applicable") Information Criterion (WAIC), a Bayesian approach for

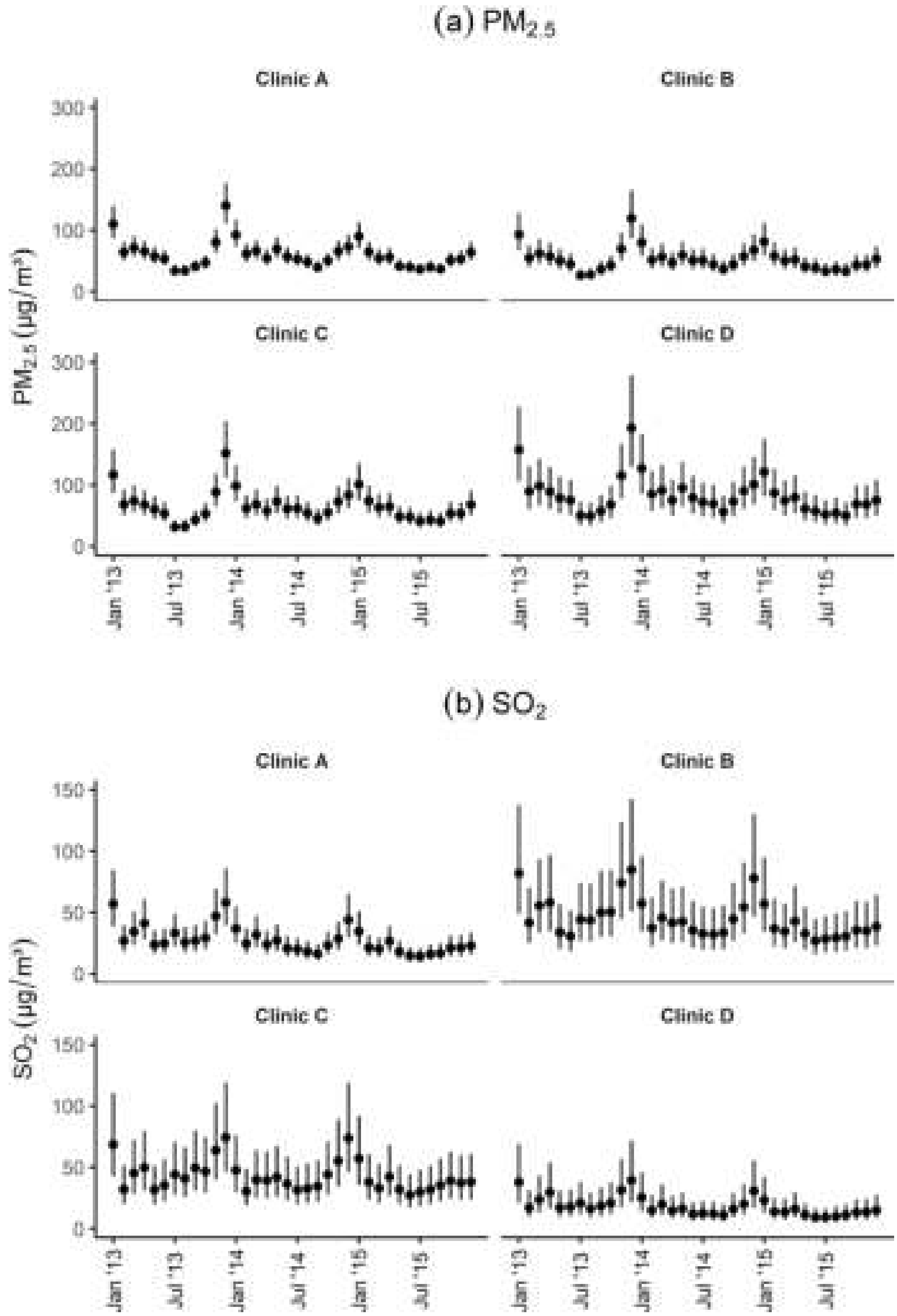


Fig. 5. Posterior medians and 95% predictive intervals for pollutant levels at four clinics.

estimating out-of-sample prediction error (Gelman et al., 2014; Watanabe, 2010).

To further asses the performance of the modelling approach, pollutant prediction models were also applied after excluding a sample of 50 pollutant observations. The sample was a simple random sample from all 348 combinations of monitor location and month (in which observed data were available). The RMSE and Pearson correlation between predicted (median of posterior predictive distribution) and observed values were then calculated.

4. Results

Observed weather and pollutant data are summarised in Table 1, using daily values and monthly means calculated for each monitor. Precipitation data are missing for 458 daily observations. There are at least 10,209 observations for each pollutant across the three year period.

4.1. Weather models

Including linear trends for time or space did not consistently decrease the WAIC values, so trends were not included in the weather prediction models. Medians and 95% HPD intervals of posterior distributions for the calendar month intercepts and hyperparameters of the four weather models are given in Table 2. Seasonal patterns are present for each of the weather variables. In particular, temperature and precipitation have much higher intercepts during the summer months (June to September) as in the observed data. The ranges of the SPDE models are large (posterior medians from 108 km to 379 km). Temperature,

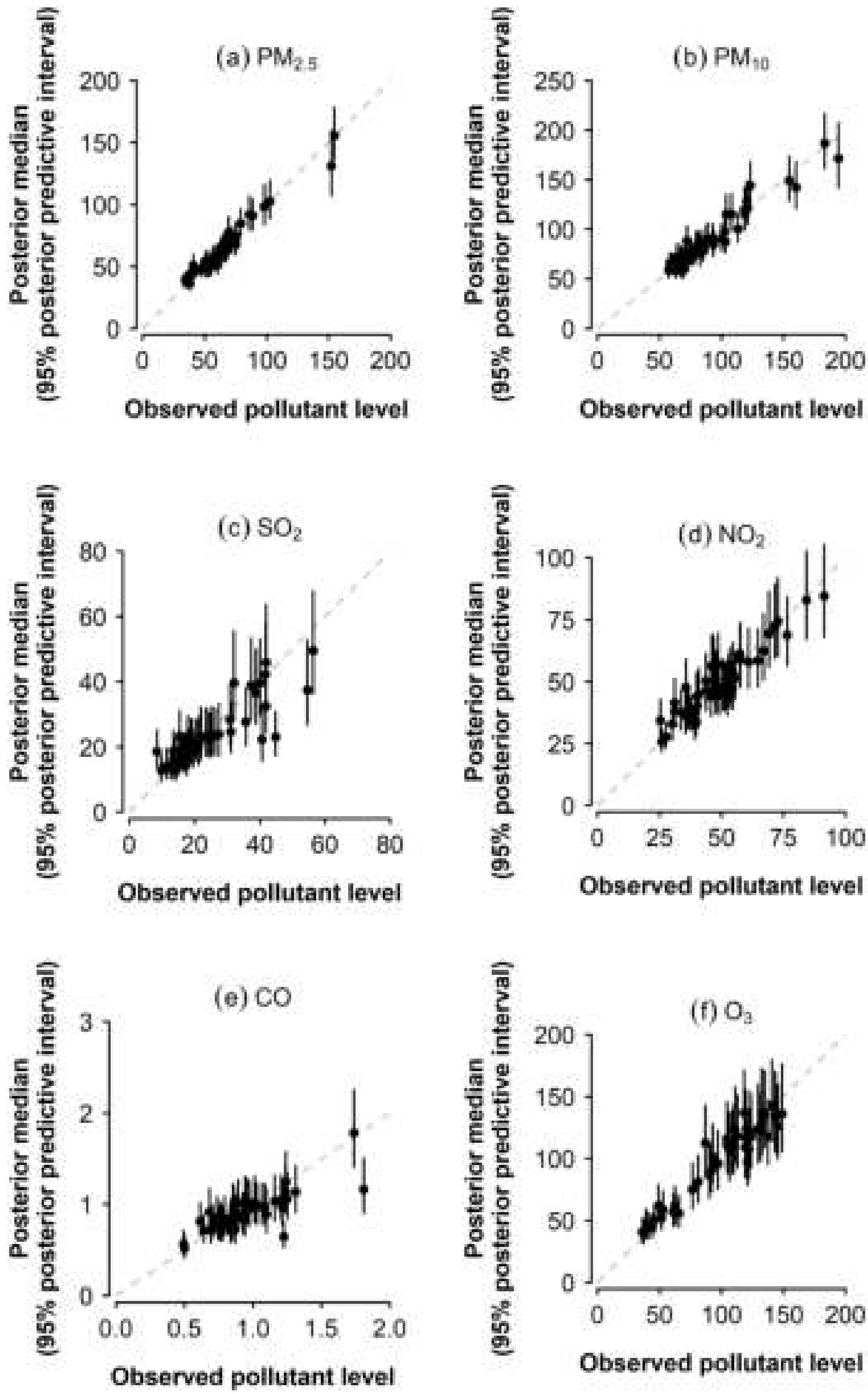


Fig. 6. Posterior medians and 95% predictive intervals for pollutant levels for random sample (having excluded observed data when fitting the model.

Table 6
RMSE and correlations between predicted values (posterior medians) and observed values for a random sample of fifty observations (excluded when fitting the models).

	PM_{10}	$PM_{2.5}$	SO_2	CO	NO_2	O_3
RMSE	8.20	5.03	5.90	0.16	5.05	9.14
Correlation	0.96	0.98	0.87	0.80	0.94	0.97

wind speed and humidity have high auto-correlation between months (posterior median of the AR coefficients of 0.96 or greater), whereas precipitation has weak auto-correlation (posterior median of the AR coefficient is 0.18).

Predicted temperature values (medians of posterior predictive distributions) at pollutant monitor and clinic locations vary between 4.02 and 32.18 °C. Wind speed predictions range from 3.38 to 5.83 m/s, and humidity from 56.84 to 84.97%. Predicted precipitation values range from 0.15 to 16.97 mm.

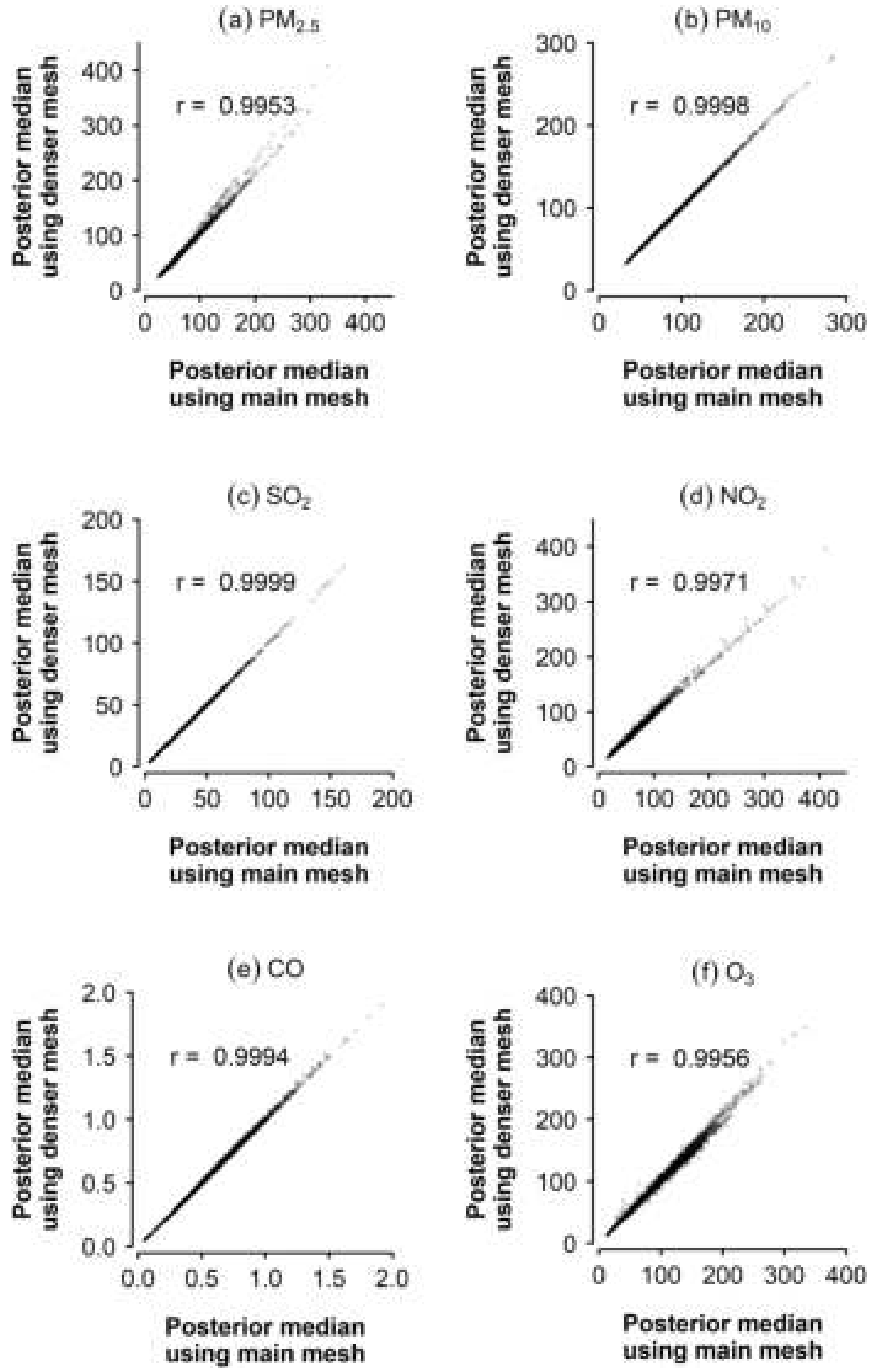


Fig. 7. Posterior medians for pollutant levels after using the main mesh and the denser mesh, and Pearson correlation coefficients.

4.2. Pollutant models

WAIC values for five models for each pollutant are given in Table 3. The models including error models for weather variables have lower WAIC values - indicating a better fit to the data - than models which use other methods to incorporate weather covariates, for all pollutants. For all pollutants, except PM_{10}, the WAIC for models excluding an SPDE spatial model is larger indicating that accounting for spatial correlations with an SPDE model improves the fit of the models. This also allows for individual predictions of pollutant levels at locations across the region. The following results and predictions use models which include error models for the weather covariates.

Medians and 95% HPD intervals of posterior distributions for the parameters and hyperparameters of the six pollutant models are given in Tables 4 and 5. Collinearity between calendar month intercepts and weather variables inhibits clear interpretation of these parameters. The particulate matter pollutants have the largest range for the spatial model (posterior medians 42 km and 45 km), followed by CO and SO_2 (30 km and 14 km), and then NO_2 and O_3 (8 km and 6 km).

Predicted levels (posterior medians) of $PM_{2.5}$ and SO_2 for January 2014 at all clinic locations are shown in Figs. 3 and 4. Posterior medians and 95% predictive intervals for $PM_{2.5}$ and SO_2 at four selected clinic locations (shown in Fig. 2) are given in Fig. 5. Predicted values (medians of posterior predictive distributions) at clinic locations have medians (inter-quartile range) of 66.12 (51.39–88.08) $\mu g/m^3$ for $PM_{2.5}$, 84.88 (58.65–102.24) $\mu g/m^3$ for PM_{10}, 25.90 (17.13–38.59) $\mu g/m^3$ for SO_2, 59.21 (42.65–81.67) $\mu g/m^3$ for NO_2, 0.61 (0.42–0.77) mg/m^3 for CO, and 91.09 (54.95–134.28) $\mu g/m^3$ for O_3.

After fitting the pollutant prediction models while excluding a random sample of fifty observations, posterior medians and 95% predictive intervals for pollutant levels are shown in Fig. 6. RMSE and correlations between predicted values (posterior medians) and observed values are given in Table 6. Correlations range from 0.80 for CO to 0.98 for $PM_{2.5}$.

Predicted pollutant levels are very similar (Pearson correlation coefficients greater than 0.99) when using the denser mesh for both weather variable and pollutant models. Posterior medians for pollutant levels after using either mesh are shown in Fig. 7.

5. Discussion

We have used Bayesian spatio–temporal models to predict levels of six pollutants at clinic locations in Suzhou, China. Inference was performed using the approximate INLA method and spatial models used the SPDE approach. The application of the SPDE approach for modelling pollutant levels has previously been reported by Cameletti et al. (2013) and Blangiardo et al. (2016). These analyses used covariates measured at or aligned to the same locations as the observed pollutant measurements. We extended this approach using a two-stage method to address misalignment of covariates. After using spatio–temporal models to produce predictions for four meteorological variables at all relevant locations, we used error models to add these as predictors in the models for pollutants. This ensured that the pollutant models incorporated the uncertainty in the predicted weather covariate values. To obtain predictions for pollutant levels at the set of clinic locations we extended the pollutant models so that posterior predictive distributions were obtained directly from R-INLA function calls.

The models and methods described in this paper provide a flexible approach to modelling ambient air pollutant levels in a region with dispersed monitors. The analysis incorporates fixed and time-varying covariate data from several sources, including misaligned covariates for which error models were used to ensure appropriate error propagation. This approach could be adapted for other scenarios, and models can be expanded with comparative ease.

These results are based on monthly pollutant levels, which were aggregated from daily data to monthly means before developing prediction models. However, the models could be adapted to use daily average values for meteorological variables and pollutant levels to enable more detailed time-series analyses. To capture dependencies over time, splines could be used with auto-regressive models for the values at knot locations.

We suggest pollutant levels at clinic locations could be used as proxies for individual exposure. It would be desirable to have individual participant residence, employment and other common locations to estimate exposure, however only clinic location (anticipated to be close to residence) is available in the given data. The methods described could be used to predict pollutant levels at any locations in the city, and if more extensive location data were available more specific estimates of exposure could be calculated.

The limited number of pollutant and weather monitors did not allow for detailed modelling of pollutants and weather variables across the city. It would be preferable to have data from more monitors throughout the city to allow better predictions of levels across the city. Given the available data, we have leveraged the geographic information available to predict pollutant levels at each clinic location. This is an alternative to ignoring the locations of pollutant monitors by using city-wide means in time-series analyses of health outcomes, or using pollutant levels are the nearest monitor as estimated levels at a clinic location.

The ranges of the SPDE models in the weather models are much larger than the extent of the area over which the models were applied. In such cases the model is usually indistinguishable from intrinsic random fields (Lindgren and Rue, 2015), but we do not expect that this affects the utility of predicted weather variables as covariates in the pollutant models.

The narrow locations (East to West) of the pollutant monitors caused a problem with including overall spatial trends. Extrapolating simple linear trends to out-of-sample x-coordinates caused predictions to be implausibly high (with small precision) in some models, but this was tempered by including quadratic terms for spatial trends. Ideally, observed pollutant data would be more geographically diverse. Alternatively, there may be better methods for ensuring reasonable out-of-sample predictions, and this potential problem should be considered when planning this type of analysis.

As an alternative to the error models used here misaligned covariates could be jointly modelled with the pollutant variables of interest. Further, health outcome data could be jointly modelled with pollutant levels. This would allow a single modelling framework for exposure, covariate, and outcome data, at the cost of more complex models and the time and resources for computation. However, the use of INLA as an efficient alternative to MCMC methods could make such an approach feasible.

Declaration of competing interest

None.

Acknowledgements

Professor Haidong Kan, School of Public Health Fudan University, Shanghai, for providing the fixed site monitoring data.

Dr Steve Hung Lam Yim, Chinese University of Hong Kong, for providing all meteorological variables.

Participants in the China Kadoorie Biobank study and the members of the survey teams in each of the 10 regional centres, as well as the project development and management teams based at Beijing, Oxford and the 10 regional centres.

The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been supported by Wellcome grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z) and grants from the National Key Research and Development Program of China (2016YFC0900500, 2016YFC0900501, 2016YFC0900504, 2016YFC1303904) and from the National Natural Science Foundation of China (91843302, 91846303, 81941018, 81390540). The UK Medical Research Council (MC_UU_00017/1, MC_UU_12026/2 MC_U137686851), Cancer Research UK (C16077/A29186, C500/A16896) and the British Heart Foundation (CH/1996001/9454), provide core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University for the project.

UK Medical Research Council Global Challenges Research Fund (Foundation Award MR/P025080/1).

Oxford-MRC Doctoral Training Partnership.

This research was funded in whole, or in part, by the Wellcome Trust [212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2021.113766.

References

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., Lindgren, F., 2018. Spatial modeling with R-INLA: a review. Wiley Interdisciplinary Reviews: Comput. Stat. 10, e1443.
Barber, X., Conesa, D., Lladosa, S., López-Quílez, A., 2016. Modelling the presence of disease under spatial misalignment using Bayesian latent Gaussian models. Geospatial health 11, 11–20.
Bivand, R., Gómez-Rubio, V., Rue, H., 2015. Spatial data analysis with R-INLA with some extensions. J. Stat. Software 63, 1–31.
Blangiardo, M., Cameletti, M., 2015. Spatial and Spatio-Temporal Bayesian Models with R - INLA. Wiley.
Blangiardo, M., Cameletti, M., Baio, G., Rue, H., 2013. Spatial and spatio-temporal models with R-INLA. Spatial and spatio-temporal epidemiology 4, 33–49.
Blangiardo, M., Finazzi, F., Cameletti, M., 2016. Two-stage Bayesian model to evaluate the effect of air pollution on chronic respiratory diseases using drug prescriptions. Spatial and spatio-temporal epidemiology 18, 1–12.
Cameletti, M., Lindgren, F., Simpson, D., Rue, H., 2013. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Advances in Statistical Analysis 97, 109–131.
Cao, J., Yang, C., Li, J., Chen, R., Chen, B., Gu, D., Kan, H., 2011. Association between long-term exposure to outdoor air pollution and mortality in China: a cohort study. J. Hazard Mater. 186, 1594–1600.
Chen, X., Zhang, L., Huang, J., Song, F., Zhang, L., Qian, Z., Trevathan, E., Mao, H., Han, B., Vaughn, M., Chen, K., Liu, Y., Chen, J., Zhao, B., Jiang, G., Gu, Q., Bai, Z., Dong, G., Tang, N., 2016. Long-term exposure to urban air pollution and lung cancer mortality: a 12-year cohort study in Northern China. Sci. Total Environ. 571, 855–861.
Chen, Z., Chen, J., Collins, R., Guo, Y., Peto, R., Wu, F., Li, L., Lancaster, G., Yang, X., Williams, A., Smith, M., Yang, L., Chang, Y., Guo, Y., Zhao, G., Bian, Z., Wu, L., Hou, C., Pang, Z., Wang, S., Zhang, Y., Zhang, K., Liu, S., Zhao, Z., Liu, S., Pang, Z., Feng, W., Wu, S., Yang, L., Han, H., He, H., Pan, X., Wang, S., Wang, H., Hao, X., Chen, C., Lin, S., Hu, X., Zhou, M., Wu, M., Wang, Y., Hu, Y., Ma, L., Zhou, R., Xu, G., Dong, B., Chen, N., Huang, Y., Li, M., Meng, J., Gan, Z., Xu, J., Liu, Y., Wu, X., Gao, Y., Zhang, N., Luo, G., Que, X., Chen, X., Ge, P., He, J., Ren, X., Zhang, H., Mao, E., Li, G., Li, Z., He, J., Liu, G., Zhu, B., Zhou, G., Feng, S., Gao, Y., He, T., Jiang, L., Qin, J., Sun, H., Liu, L., Yu, M., Chen, Y., Hu, Z., Hu, J., Qian, Y., Wu, Z., Chen, L., Liu, W., Li, G., Liu, H., Long, X., Xiong, Y., Tan, Z., Xie, X., Peng, Y., 2011. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int. J. Epidemiol. 40, 1652–1666.
Chen, Z., Lee, L., Chen, J., Collins, R., Wu, F., Guo, Y., Linksted, P., Peto, R., 2005. Cohort profile: the Kadoorie study of chronic disease in China (KSCDC). Int. J. Epidemiol. 34, 1243–1249.
Dong, G.-H., Zhang, P., Sun, B., Zhang, L., Chen, X., Ma, N., Yu, F., Guo, H., Huang, H., Lee, Y.L., Tang, N., Chen, J., 2012. Long-term exposure to ambient air pollution and respiratory disease mortality in Shenyang, China: a 12-year population-based retrospective cohort study. Respiration 84, 360–368.
Fioravanti, G., Martino, S., Cameletti, M., Cattani, G., 2021. Spatio-temporal modelling of PM10 daily concentrations in Italy using the SPDE approach. Atmos. Environ. 248, 118–192.
Foster, S.D., Shimadzu, H., Darnell, R., 2012. Uncertainty in spatially predicted covariates: is it ignorable? J. Roy. Stat. Soc.: Series C (Applied Statistics) 61, 637–652.
Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016.
Gomez-Rubio, V., 2020. Bayesian Inference with INLA. CRC Press.
Krainski, E.T., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., Rue, H., 2019. Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA. CRC Press/Taylor; Francis Group.
Li, G., Jiang, L., Zhang, Y., Cai, Y., Pan, X., Zhou, M., 2014. The impact of ambient particle pollution during extreme-temperature days in Guangzhou City, China. Asia Pac. J. Publ. Health 26, 614–621.
Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with r-INLA. J. Stat. Software 63, 1–25.
Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society. Ser. Bibliogr. 73, 423–498.
Moraga, P., 2019. Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny. Chapman; Hall/CRC.
Muff, S., Riebler, A., Rue, H., Saner, P., Held, L., 2013. Bayesian Analysis of Measurement Error Models Using INLA arXiv preprint arXiv:1302.3065.
Peng, Z., Liu, C., Xu, B., Kan, H., Wang, W., 2017. Long-term exposure to ambient air pollution and mortality in a Chinese tuberculosis cohort. Sci. Total Environ. 580, 1483–1488.
R-INLA, 2020. The R-INLA Project. www.r-inla.org.
Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. Roy. Stat. Soc. B Stat. Methodol. 71, 319–392.
Shmueli, G., 2010. To explain or to predict? Stat. Sci. 25, 289–310.
Wang, X., Yue, Y., Faraway, J.J., 2018. Bayesian Regression Modeling with INLA, first ed. Chapman; Hall/CRC.
Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11, 3571–3594.
Yin, P., Brauer, M., Cohen, A., Burnett, R.T., Liu, J., Liu, Y., Liang, R., Wang, W., Qi, J., Wang, L., Zhou, M., 2017. Long-term fine particulate matter exposure and nonaccidental and cause-specific mortality in a large national cohort of Chinese men. Environ. Health Perspect. 125, 117002.
Zhang, P., Dong, G., Sun, B., Zhang, L., Chen, X., Ma, N., Yu, F., Guo, H., Huang, H., Lee, Y.L., Tang, N., Chen, J., 2011. Long-term exposure to ambient air pollution and mortality due to cardiovascular disease and cerebrovascular disease in Shenyang, China. PloS One 6, e20827.
Zhou, M., Liu, Y., Wang, L., Kuang, X., Xu, X., Kan, H., 2014. Particulate air pollution and mortality in a cohort of Chinese men. Environ. Pollut. 186, 1–6.

International Journal of Hygiene and Environmental Health 235 (2021) 113773

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

Exposures to lead during urban combat training

Linda Schenk[a,*], Håkan Wingfors[b], Björn Skoog[c], Niklas Forsgard[d], Christina Nyberg[c], Fabian Taube[c,e]

[a] Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE 17177, Stockholm, Sweden
[b] Division of CBRN Defence and Security, Swedish Defence Research Agency, SE 901 82, Umeå, Sweden
[c] Swedish Armed Forces Centre for Defence Medicine (SWE CDM), SE 426 05, Västra Frölunda, Sweden
[d] Department of Clinical Chemistry, Sahlgrenska University Hospital, SE 413 45, Gothenburg, Sweden
[e] Institute of Medicine, School of Public Health and Community Medicine at Sahlgrenska Academy, Box 414, SE 405 30, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Military hygiene
Lead
Occupational health
Biological monitoring
Occupational exposure
Maximum allowable concentration

ABSTRACT

Lead exposure is still a major concern for occupations that regularly train or work with firearms, such as law enforcement and military personnel. Due to the increasing number of women of fertile age in such professions, there is a strong incentive to monitor lead exposures during firearms training. Personal air sampling was performed during two sessions of a nine-day urban combat training (UCT) course for cadets in the Swedish Armed Forces, one session employing leaded ammunition (*leaded scenario*) and one session employing unleaded ammunition (*unleaded scenario*). Blood lead levels (BLLs) were measured before and after the course for 42 cadets and five instructors. During the leaded scenario, the instructors' airborne exposure (geometric mean, GM, 72.0 μg/m^3) was higher than that of cadets (GM 42.9 μg/m^3). During the unleaded scenario, airborne concentrations were similar for instructors and cadets and considerably lower than during the leaded scenario (GM 2.9 μg/m^3). Despite comparably low external lead exposures during the course, we saw a statistically significant increase in systemic exposure for cadets (BLL GM increased from 1.09 to 1.71 μg/dL, $p < 0.001$). For the five instructors, notable differences were seen depending on task. The largest increase was seen for the two instructors performing close supervision during the leaded scenario (BLL GM increased from 2.41 μg/dL to 4.83 μg/dL). For the remaining three instructors the BLLs were unchanged (BLL GMs were 1.25 μg/dL before the course and 1.26 μg/dL after). None of the participants exceeded the applicable biological exposure limits, but extrapolating our findings shows that instructors in the leaded scenario may reach levels around 10 μg/dL after a year of repeated exposures. We conclude that comparably low airborne concentrations can contribute to the body burden of lead and that additional measures to reduce exposure are warranted, particularly for instructors.

1. Introduction

Shooting with firearms exposes the shooter to various air pollutants, often containing both metallic lead and soluble lead salts (Aurell et al., 2019; Grabinski et al., 2017; Laidlaw et al., 2017; Wingfors et al., 2014). Other metals, especially zinc, copper, chromium, barium, antimony, and iron, have also been identified in air samples related to weapon emissions, either in the dust or as vapor (Bergstrom et al., 2015; Mancuso et al., 2008; Vandebroek et al., 2019; Voie et al., 2014). The high temperatures caused by deflagration and friction generate these metals in the form of vapours, oxides, and fine particulates found in the breathing zone of shooters. Studies suggest that inhaled fine lead particles are more easily deposited and absorbed in the respiratory tract than coarse lead particles (Donguk and Namwon 2004; Lach et al., 2015; Park and Paik 2002).

Air monitoring of occupational groups that regularly train or work at firing ranges, such as law enforcement and military personnel, show that lead exposures vary considerably between settings, such as indoor and outdoor ranges, variable ventilation conditions, and type of weapons and ammunition (Novotny et al., 1987; Scott et al., 2012; Svensson et al., 1992; Valway et al., 1989). Furthermore, exercises vary in terms of duration, number of participants, and level of activity including number of rounds fired, making comparisons of exposure levels between studies challenging. Nevertheless, several studies report levels exceeding occupational exposure limits (OEL) for airborne lead (Bonanno et al., 2002; Council 2013; Laidlaw et al., 2017; Svensson et al., 1992).

* Corresponding author.
E-mail address: linda.schenk@ki.se (L. Schenk).

https://doi.org/10.1016/j.ijheh.2021.113773
Received 28 January 2021; Received in revised form 8 May 2021; Accepted 10 May 2021
Available online 23 May 2021

However, blood-lead levels (BLL) may be a more favourable method to assess lead exposure at firing ranges. Studies on the effect of weapons training on BLL in law enforcement and/or military staff mainly cover static shooting exercises, from a single shooting session at a covered outdoor (Tripathi et al., 1989) or indoor firing range (Vandebroek et al., 2019) to more extended periods of regular training at indoor (Fischbein et al., 1979; Rocha et al., 2014; Svensson et al., 1992; Valway et al., 1989; Vivante et al., 2008) or outdoor (Greenberg et al., 2016) firing ranges. Ample evidence shows that lead exposures at firing ranges increases the body burden both for recreational and occupational shooters, and long-term exposure increases BLLs to an extent where they are associated with adverse health outcomes (Laidlaw et al., 2017).

Lead is a highly toxic metal with a number of deleterious health effects, such as neurotoxicity, kidney toxicity, and cardiovascular outcomes. The developing foetus and child have been identified as particularly sensitive to the effects of lead, and there is general agreement that it is not possible to establish a safe level of exposure for this subpopulation (EFSA 2010; RAC 2020; NTP 2012). Preindustrial BLLs in humans have been estimated to be below 0.02 μg/dL (Flegal and Smith 1992).

A factor in the risk assessment of lead is its longevity in the human body. The half-life of lead in blood is around 30 days. In bone, the half-life is 10–30 years. Hence, directly after uptake, lead is primarily found in the erythrocytes and is distributed to organs such as the brain, kidneys, liver, and bones. Over time, a high proportion of absorbed lead is transferred to bone, where it accumulates over time. For adults, bone lead constitutes a very high proportion of the body burden of lead. If external exposure ceases (or is reduced), lead will be redistributed from bone to blood. During periods of degradation of bone, an increase in the endogenic exposure from lead stored in the bone could be a problem, especially among breastfeeding women or women at menopause with a history of high and/or long-term lead exposure. Excretion takes primarily place through urine and faeces. The interindividual variations in lead uptake and excretion are significant, as are differences in exposure by "hand to mouth" contact. Thus, it is difficult to establish a close correlation between airborne lead exposure and resulting BLL (RAC 2020).

According to the GESTIS database on occupational exposure limits (OELs, GESTIS n.d.), current OELs for lead vary from 3 μg/m^3 to 150 μg/m^3 (respirable, inhalable, and total dust fractions mixed) as time-weighted averages (TWA) over 8 h. The Swedish TWA OEL is 100 μg/m^3 for the inhalable fraction and 50 μg/m^3 for the respirable fraction (SWEA 2018). There are also a number of recommended biological exposure limits for the maximal allowable concentration of blood lead. ECHA's Committee for Risk Assessment (RAC, 2020) recently proposed a biological exposure limit value (BLV) of 15 μg lead/dL blood, based on neurotoxicity. ECHA RAC also recommended a biological exposure guidance value (BGV) of 4.5 μg/dL as an indicator of occupational exposure for groups at risk and a TWA OEL of 4 μg/m^3 (inhalable fraction) for lead and its inorganic compounds. The BGV is intended as guidance as neither the recommended BLV or OEL would be protective for women in fertile age (RAC 2020). Current Swedish BLL limits are listed in the ordinance on medical surveillance (recently updated in SWEA 2019). The ordinance states that no occupational exposure to lead for women under 50 years of age is allowed if the BLL exceeds 10 μg/dL (0.5 μmol/L). For other employees, the corresponding limit is set to 30 μg/dL (1.5 μmol/L). If BLLs are lower than 8 μg/dL (0.4 μmol/L) during three consecutive medical check-ups, no further BLL measurements are needed for men or for women over 50 years of age, as long as changes in work tasks do not lead to higher lead exposures.

Instructors and certain special task forces have been identified as more exposed due to frequent and more intensive exposures during complex shooting exercises, such as Urban Combat Training (UCT, Mancuso et al., 2008; Rocha et al., 2014; Vandebroek et al., 2019). In UCT, close-quarter combat is trained using small arms combined with explosive devices (e.g., stun grenades and smoke) in simulated urban zones and shoot houses (Weber et al., 2020). UCT reflects real combat scenarios but is rarely studied, despite being an essential part of the training, in addition to marksmanship at flat ranges.

During the last 20 years, women have become an increasingly natural part of the workforce within the military and law enforcement sectors. With respect to lead exposure, this positive progress may pose new challenges due to the potential for developmental toxicity, and consequently, require increased health surveillance and measures to control and reduce exposure.

The increasing concern for health effects at low lead exposures, reflected by the decreasing levels of OELs, warrant investigation of all kinds of lead exposure scenarios in the military occupation. In the present work, we analyse the findings from exposure measurements for lead during an introductory UCT course for cadets in the Swedish Armed Forces. Airborne exposures were measured during two sessions of the course, corresponding to a leaded and an unleaded exposure scenario in urban combat environments. BLL was assessed before and after the UCT, and data was further used to extrapolate to long-term exposures. The objective was to assess the effect of such training on systemic lead exposure for cadets and instructors that repeatedly conduct these exercises.

2. Materials and methods

2.1. Study population and settings

Study participants were military staff taking or leading an introductory UCT course. The instructors were 35–42 years old (n = 6, median 36 years, BLL data collected for 5), and the participating cadets were 21–39 years old (n = 43, median 24 years, BLL data collected for 41). Six of the participants were female (BLL data was collected for 5). Most of the participants had been employed by the Swedish Armed Forces for 1–5 years, five had been employed less than 1 year, and seven more than 5 years.

Forty of the participants reported firing more than 150 rounds a typical month. Two instructors and two cadets reported to be recreational shooters. Thirty-seven cadets reported to have participated in firearms-training in the preceding month, 31 also reported to have upcoming training in the coming months. Snus (Swedish/oral snuff, potentially increasing hand-to-mouth transfer) was used by 18 participants, of which two also reported to be smokers (less than 1 pack/day). No observations were performed to assess the extent of snus use during the UCT. A summary of demographic, lifestyle, and activity related factors of participants in blood monitoring is also presented in the Supplementary Material Table S1.

During the nine-day UCT, groups of cadets practiced different scenarios of urban warfare as well as regular outdoor flat range training (an overview of the course and sampling is presented in the supplementary material). During training, participants were given the opportunity to wash their hands prior to lunch. At the training sites, possibility for handwashing was limited. Two of the training sessions, one using leaded and one using unleaded ammunition, were selected for assessment of inhalable dust and lead in the breathing zone. In the session defined as the *leaded scenario,* live leaded rounds were used exclusively (5.56 diameter lead in bullet and primer) in a roofless UCT building. The average number of rounds per person was 35, reflecting the introductory nature of the training. In addition, lead-free explosive grenades (total use 63) and stun grenades (total use 113) were used. In the *unleaded scenario*, performed partially outdoors, unleaded ammunition (5.56 diameter, steel core with copper alloy jacket and lead-free primer, average use 96 per person) was used exclusively. Cadets performed one leaded exposure scenario, either on day 3 or 4 of the course. The same two instructors supervised both these sessions. During the remaining UCT, mainly lead-free ammunition (no lead in bullet or primer) was used. The average total use per person was 621 lead-free rounds and 12 lead-free bullets with leaded primer (7.62 diameter, used day 9). These unmonitored exercises are expected to have resulted in lower or similar lead exposures as the unleaded scenario.

The lead exposure assessments described herein were originally performed as part of the Armed Forces work environment assessment and not for research purposes. Participation in the exposure assessments during the UCT course was voluntary. This reanalysis of the data has been approved by the Swedish Ethical Review Authority (decision 2019–06339).

2.2. Exposure assessment

Personal airborne measurements were made for the two instructors supervising the leaded scenarios (day 3 and 4) and for 13 cadets (day 4). During the unleaded scenario (day 8), two instructors and 13 cadets were monitored, but for one cadet data were discarded due to a malfunctioning air sampling pump. Dust was collected from the participants' breathing zone using air pumps with a constant airflow (2.0 L/min) through preweighed cellulose filters (Millipore, diameter 25 mm, pore size 3 μm) mounted in an IOM-sampler. Lead content was analysed using inductively coupled plasma mass spectrometry (ICP-MS, Thermo Fisher Scientific iCAP-Q). Dust from the filter and the wall of the cassette was transferred to a 50 mL Falcon tube and digested using concentrated nitric acid with a small addition of hydrogen peroxide in a microwave oven (CEM MARS 5) at 115 °C. The solution was then diluted with Milli-Q water to a 1% nitric acid solution. The limit of detection (LOD) for lead was 0.5 μg/m^3.

Pre-exposure blood samples were drawn six days before (n = 33) or on the first day (n = 13) of the course. Post-exposure blood samples were collected the day after the course for all (41 male and 5 female participants). Venous blood was collected from the brachial vein using a 7 mL Vacutainer® trace element tube and stored in a refrigerator until analysis.

Lead content in the blood samples was analysed using ICP-MS (Agilent 7700x). Blood samples were diluted 20 times with an alkaline diluent containing 1-butanol (2% w/v), ethylenediaminetetraacetic acid (0.05%w/v), TritonX-100 (0.05%w/v), and ammonium hydroxide (1% w/v), prior to analysis. Measurements were performed in the helium collision mode. The LOD was 0.16 μg/L.

Data on demographics, lifestyle, and activity-related information, were collected through a questionnaire in connection to blood sampling.

2.3. Data analysis

We grouped participants into similarly exposed groups based on tasks during the exercise and measured exposures. For the leaded scenario's airborne concentrations and the BLL data the instructors and cadets were considered separate exposure groups. For the unleaded scenario's airborne concentrations, instructors and cadets were considered one mutual exposure group. Airborne concentrations were lognormally distributed (Shapiro-Wilk normality tests, $P > 0.05$). BLLs were lognormally distributed when separating cadets and instructors (Shapiro-Wilk normality tests, $P > 0.05$). Differences in BLLs for cadets were neither normally nor lognormally distributed (Shapiro-Wilk normality tests, $P < 0.05$). Pre- and post-BLL comparisons for cadets were performed with the paired *t*-test using log-transformed values. Instructors were divided into two exposure groups, "leaded scenario instructors" (n = 2) and "remaining instructors" (n = 3). Due to small sample size, pre- and post-BLL differences for instructors were not evaluated with any statistical test. Comparisons between groups mixing cadets and instructors were performed with the two sample Wilcoxon signed-ranks test. Correlations were tested with Kendall's rank correlation. Statistical analyses were performed with the R-4.0.1 software. Results were regarded as statistically significant at the $P < 0.05$ level. Statistical tests were two-tailed.

We performed two simple extrapolations to long-term BLL after repeated exposures for three different exposure scenarios: a worst-case instructor scenario, a high exposed cadet scenario, and a median cadet scenario. These extrapolation scenarios were based on our findings and are described further in results. For the first kind of extrapolation, we used the biokinetic slope factor (quasi-steady-state blood level) of 0.4 μg Pb/dL per 1 μg absorbed Pb/day defined by the US EPA adult lead model (US EPA, 2003). Lacking data on the height and bodyweight of cadets, we performed calculations assuming 85 kg and a blood volume of 70 mL/kg. It should be noted that calculations with a lower body weight and blood volume would yield lower BLLs and vice versa. For the second kind of extrapolation, we modelled the daily BLL over time based on a daily consumption of 0.083 μg Pb/kg bw/day (median for the Swedish population, SFA 2017) and 20% oral absorption, in addition to the exposure for each of three exposure scenarios, using a half-life in blood of 30 days. This latter model is not affected by differences in body weight assumptions. It does also not consider redistribution from other parts of the body back into blood (*i.e.*, it does not include a soft tissue compartment nor a bone compartment), nor any other lead exposure sources. For both kinds of extrapolation, we modelled the contribution from each repeated course as a bolus dose rather than a stepwise exposure across several days. We assumed that the corresponding pre-BLLs were background levels for each of the three exposure scenarios, driven by non-occupational exposure, which likely is an overestimation for instructors.

For risk characterisation of inhalation exposures, we compared the measurements, adjusted to an 8 h day, to the Swedish 8 h TWA OELs (SWEA 2018) and ECHA RAC's recommended TWA OEL (RAC 2020). For risk characterisation of systemic exposures, we compare the measured BLLs to the Swedish ordinance on medical surveillance (SWEA 2019) and ECHA RAC's recommended BLV (RAC 2020). As we are concerned about average exposures exceeding the OEL, we calculated one-tailed confidence intervals. Hence, for lognormally distributed data, we report the geometric mean (GM) and 95% upper confidence limit (UCL) of the GM. We also calculated the one-sided upper tolerance limit (UTL) to identify a value that we are 95% confident that 99% of the population would not exceed under the same conditions as the investigated UCT. For normally distributed data, the UTL is calculated as:

$$UTL = Mean + ((K_{\gamma,p,n})(SD))$$

Where $K_{\gamma,p,n}$ is a statistical coefficient whose value depends on the desired confidence level (γ), population proportion (p) and sample size (n). For a 95% confidence level that 99% of the population will be below the UTL the appropriate K-values are 7.042 (n = 4), 3.659 (n = 13), 3.585 (n = 14), 2.941 (n = 40) (Natrella, 1963). SD is standard deviation. For lognormally distributed data, the UTL was calculated using log-transformed values. It can be noted that our K-values are more conservative than the desired confidence levels (70%) and population proportion (95%) of the European Standards' strategy for testing compliance with occupational exposure limit values (CEN, 2019).

3. Results

The inhalable fraction of lead during the leaded scenario is shown in Fig. 1. Instructors were assessed during two days (3 and 4), geometric mean of the airborne concentrations was 72.0 μg/m^3 (UCL = 120.9 μg/m^3). Cadets' exposure assessments were from day 4, geometric mean 42.9 μg/m^3 (UCL = 48.5 μg/m^3).

After adjusting the exposure levels to an 8 h day, the airborne concentrations ranged from 14.6 to 42.8 μg/m^3 (GM 25.5 μg/m^3 UCL = 42.8 μg/m^3 UTL = 566 μg/m^3) for instructors and from 8.2 to 20.8 μg/m^3 for cadets (GM = 16.4 μg/m^3 UCL = 18.7 μg/m^3 UTL = 44.6 μg/m^3). Thus, on an 8 h basis, all measured lead exposures were below the Swedish TWA OEL of 100 μg/m^3, although all were also above the recent ECHA recommendation of 4 μg/m^3 (inhalable fraction). Moreover, we are 95% confident that 99% of cadets under the same circumstances will be exposed to no more than 44.6 μg/m^3 (the UTL), while the uncertainty

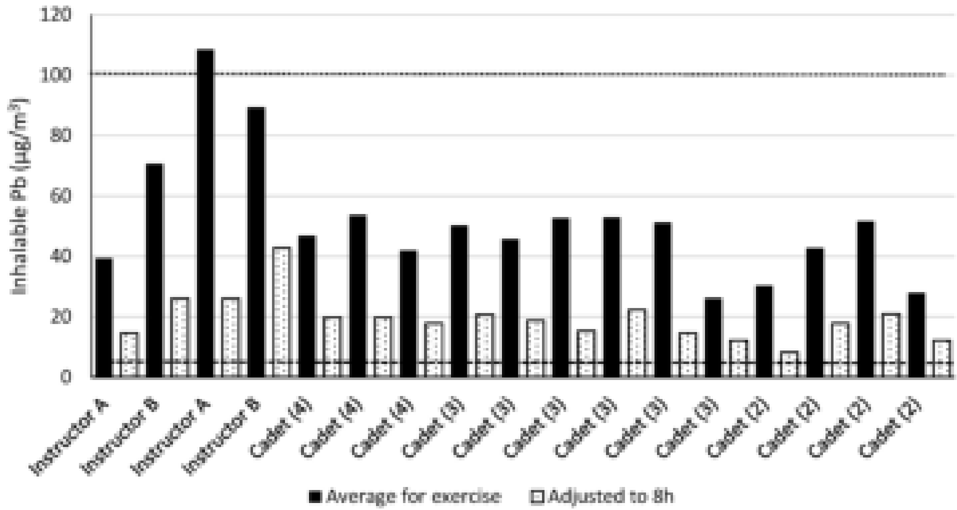


Fig. 1. Airborne concentrations of inhalable lead dust during the leaded exposure scenario, averaged over sessions (3–4 h) and over 8 h assuming no additional exposure that day. Cadets participated in 2–4 exercises per session (number in parentheses) on day 4. Instructors followed exercises continuously over the sessions (days 3 and 4, presented in that order). The dotted line indicates 100 μg/m^3 (Swedish OEL), dashed line indicates 4 μg/m^3 (ECHA RAC recommendation).

is very large regarding instructors.

The airborne concentrations of lead in the unleaded scenario are shown in Fig. 2. As the difference between instructors and cadets was less pronounced than in the leaded scenario, the two groups were combined. In the unleaded scenario, lead-free ammunition was used in combination with parts of the training being spent outdoors, leading to markedly lower lead exposures where a few were close to, but above, the LOD (GM = 2.9 μg/m^3 UCL = 4.1 μg/m^3).

After adjusting the exposure levels to an 8 h day, the airborne concentrations ranged from 0.6 to 3.4 μg/m^3 (GM = 1.8 μg/m^3 UCL = 2.7 μg/m^3 UTL = 13.5 μg/m^3, instructors and cadets combined). Thus, on an 8 h basis, all airborne concentrations were below the Swedish OELs for inhalable lead, and we are 95% certain that 99% of the population under the same conditions would be exposed to less than 13.5 μg/m^3. This UTL poses a considerable margin to the Swedish OEL but not to the recent ECHA recommendation of 4 μg/m^3.

Fig. 3 displays pre- and post-BLLs for instructors and cadets. For three instructors and one cadet, BLL did not increase over the course (Supplementary Material Fig. S1). The highest increase in BLL was seen for the two instructors supervising the leaded exposure scenario exercise (BLL GM increased from 2.41 μg/dL to 4.83 μg/dL, see also Supplementary Material Fig. S2), while the BLL geometric mean for the three other instructors was 1.25 μg/dL before the course and 1.26 μg/dL after (median 1.32 μg/dL and 1.33 μg/dL, respectively). For cadets, BLLs increased significantly (paired $t_{df = 40} = -19$, $p < 0.001$), from a geometric mean of 1.09 μg/dL (UCL = 1.18, median 1.06, UTL = 2.48 μg/dL) to a geometric mean of 1.71 μg/dL (UCL = 1.84, median 1.64, UTL = 3.82 μg/dL). The four recreational shooters had higher pre-BLL (median 1.73 μg/dL) than the median pre-BLL for both cadets and instructors, while the five females were below the median levels (Fig. 3). Across all participants, age and pre-BLL were positively correlated ($r_\tau = 0.27$, $p = 0.01$, see supplementary material, Fig. S3). Those with a higher pre-BLL were also more likely to have a higher absolute increase in BLL over the course ($r_\tau = 0.23$, $p = 0.027$, cadets and instructors combined). No statistically significant difference was seen between snus users and non-users, either in pre-BLL (W = 328, p = 0.09) or in the increase in BLL (W = 330, p = 0.08, see also the Supplementary Material Fig. S1). For eleven cadets, we had air monitoring data for the leaded scenario (i.e., the major source of lead exposure over the duration of the course) as well as blood monitoring data. However, the amount of lead on the filter (as proxy for total amount of inhaled lead) did not correlate with the increase in BLLs ($r_\tau = -0.02$, $p = 1$, see also the Supplementary Material Fig. S2 including also instructors).

Extrapolation to long-term quasi-steady-state BLL (EPA 2003) was

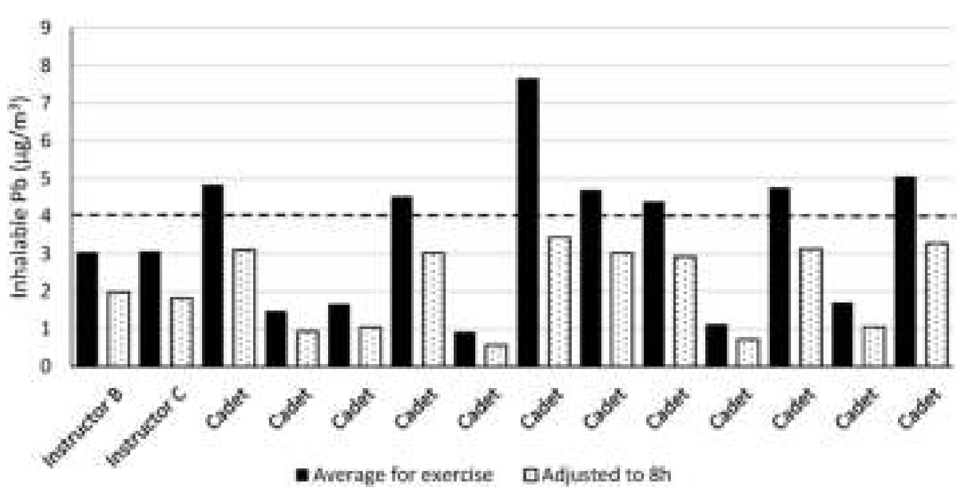


Fig. 2. Airborne concentrations of inhalable lead dust during unleaded exposure scenario, averaged over sessions (~5 h) and over 8 h assuming no additional exposure that day. The dashed line indicates 4 μg/m^3 (ECHA RAC recommendation).

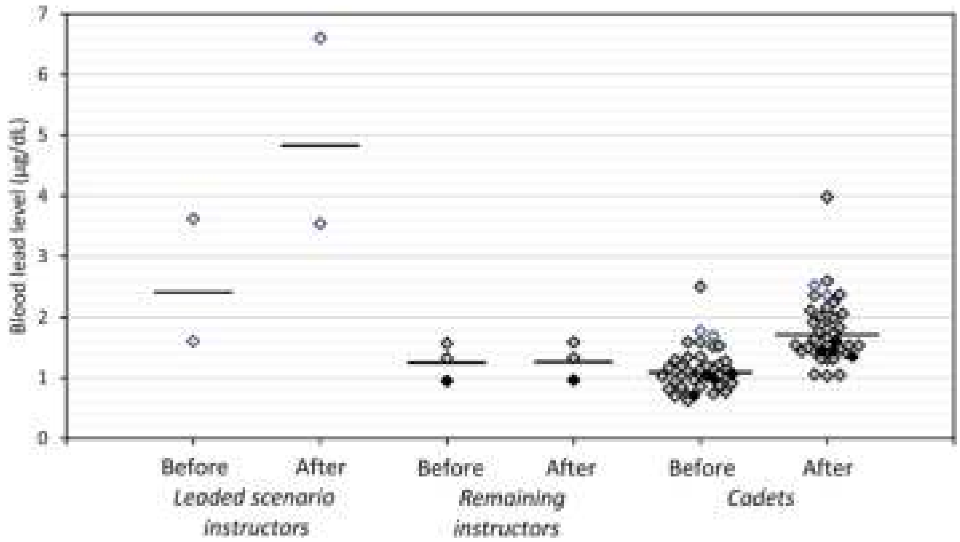


Fig. 3. Blood lead (μg/dL) for instructors of the leaded scenario (n = 2), remaining instructors (n = 3) and cadets (n = 41), before and after the 9-day course, horizontal lines indicate the geometric means, white markers indicate recreational shooters (n = 4, all male), black markers female participants (n = 5), grey markers male participants.

performed for three scenarios: A projected worst-case for instructors based on the highest instructor pre- and post-BLL values, a median cadet scenario, and a high uptake scenario for cadets based on the calculated UTLs for pre- and post-BLLs (Table 1, Fig. 4). Since some instructors teach continuously, we made a worst-case estimate of 24 exposures per year based on nine days per course and 220 working days per year. This includes a 5-week summer break and a 2-week winter break. Both cadet scenarios assumed cadets under education going through a similar course 11 times per year (monthly, with an exception for July). The instructor scenario yielded a contribution on a yearly basis of 4.7 μg/dL, meaning that if the background BLL of the exposed person is 3.6 μg/dL, BLL will reach a quasi-steady-state of 8.3 μg/dL. For the projected UTL cadet scenario, the quasi-steady-state reached 3.4 μg/dL, and for the median cadet scenario reached 1.5 μg/dL (Table 1). These levels are similar to the levels in Fig. 4, where we model the BLL over time using the same assumptions as in Table 1 and a dietary background contribution corresponding to the Swedish average.

All measured BLLs, as well as the UTLs for cadets, were below the action levels of the Swedish ordinance medical surveillance, including that for women under the age of 50 years (10 μg/dL) and the recently proposed BOELV of 15 μg/dL. However, as seen in Table 1 and Fig. 4, frequently exposed staff members could approach a BLL close to those levels in a year. Especially, if assignments also include activities leading to higher lead exposures than this introductory UCT course, or if there are additional non-occupational sources of exposure to lead, such as recreational shooting.

Table 1
Long-term extrapolation of blood lead level (BLL, μg/dL) for three different yearly scenarios using the biokinetic slope factor of the US EPA adult lead model (0.4 μg/dL per g/day, US EPA, 2003).

	Instructor	UTL cadet	Median cadet
Starting BLL (μg/dL)	3.6	2.5	1.1
Increase per course (μg/dL)	3.0	1.3	0.6
Times per year	24	11	11
Annual contribution (μg/dL)	4.7	0.9	0.4
Quasi-steady-state BLL (μg/dL)	8.3	3.4	1.5

4. Discussion

We have presented exposure assessments of inhalable lead for two UCT scenarios and measured BLL before and after participation in a 9-day UCT course. During the leaded scenario, airborne concentrations were about one-fourth of the Swedish OEL for inhalable lead (100 μg/m^3) when weighted over 8 h. During the unleaded scenario the GM of airborne concentrations was about 1/50 of the Swedish OEL or roughly half of ECHA RAC's recommended OEL (4 μg/m^3). Participation in this nine-day UCT-course increased participants' systemic exposure to lead, measured as BLL, significantly. However, the introductory character of the UCT course, (with a comparably low number of rounds fired - on average 100 per person and day), most likely resulted in lower airborne concentrations of lead than what would be expected during a regular UCT course. Due to differences in study design and sampling strategy, comparisons with previously reported air levels from firing ranges are complicated. Nevertheless, airborne concentrations of lead measured in this UCT-course are low compared to those reported for indoor ranges, and lower, or similar to, those for outdoor ranges (Greenberg et al., 2016) or UCT environments (Weber et al., 2020), corroborating that more intense exercises would result in higher exposure.

Although airborne concentrations were below the Swedish OEL for inhalable lead, ECHA RAC's recently recommended OEL was exceeded for several participants. However, there are limitations to using OELs for firearms training. OELs target repeated exposure during every day of a working life, while these activities are not a daily task, neither for cadets nor for instructors. Moreover, our adjustment from 4 h to 8 h, assumes that there is no additional exposure during the day in question, which may not be the case. In addition, the fact that we could detect airborne lead during the unleaded scenario indicates the presence of other sources of lead. Those include resuspension of lead-containing dust from training grounds or release from contaminated weapons and other equipment, which could also lead to exposure outside of firearms training. Taken together, BLL is a more informative measure of exposure, and associated risks, for this professional group.

The measured post-BLLs were below the Swedish BLV for adult men. In all but one case, post-BLLs were also below the recommended BGV of 4.5 μg l/dL (based on the background distribution in the EU population, RAC 2020). However, cadets' post-BLLs were significantly higher than their pre-BLLs (Fig. 3), despite the comparatively moderate exposure conditions of this course. We also saw a large variation between

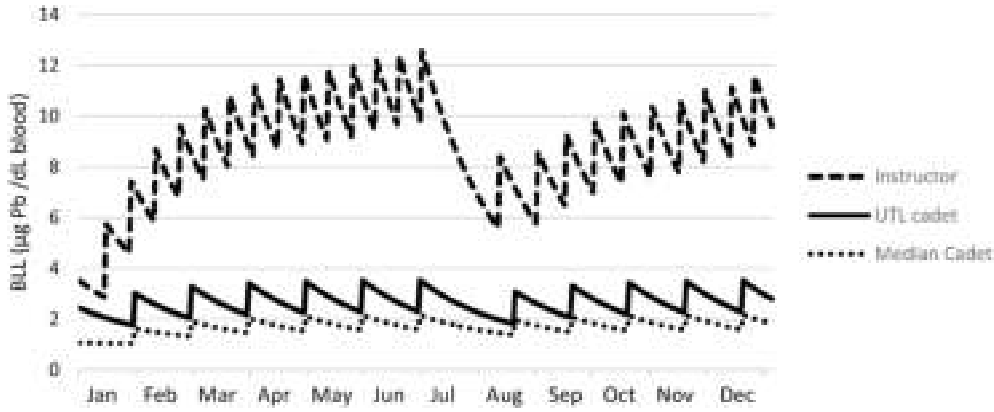


Fig. 4. A simple model of blood lead level (BLL) across one year of repeated exposures for three scenarios: instructor (worst-case (3 μg/dL increase 24 times per year), upper tolerance levels for cadets (1.3 μg/dL increase 11 times per year), and median cadet (0.6 μg/dL increase 11 times per year).

instructors, where instructors of the leaded scenario showed the largest BLL increases across all participants, while the BLLs of the three other instructors did not increase.

The most recent estimate of the average background BLL in Sweden is an investigation in Northern Sweden in 2014 (Wennberg et al., 2017). For 25–35 year-olds, the median BLL was 1.1 μg/dL for men and 0.97 μg/L for women, which is in line with the pre-BLL of cadets in the present study (median 1.06 μg/dL). In comparison to older studies on occupational firearm users, BLLs in the present study are comparably low (Fischbein et al., 1979; Goldberg et al., 1991; Rocha et al., 2014; Svensson et al., 1992; Tripathi et al., 1989; Valway et al., 1989; Vivante et al., 2008). To some extent, this reflects the age of the previous studies, since BLLs in the general population were much higher prior to the phase-out of leaded gasoline. Since lead accumulates in bone over our lifetime, age of participants is an important factor as well. A study following 18–19-year-old cadets' outdoor training noted an increase in geometric mean of BLL, from below the quantification limit of 0.29 μg/dL to 1.17 μg/dL after 15 weeks of basic training, which further increased to 3.92 μg/dL after an additional 15 weeks of advanced training (Greenberg et al., 2016). A one-day session in an indoor firing range among ten police officers (mean age 44.5 years) yielded a small but statistically significant increase from 1.41 to 1.47 μg/dL (Vandebroek et al., 2019). A 45-day UCT course increased the average BLL from 7.2 to 20.5 μg/dL among cadets. The implementation of exposure reduction measures and lead-free ammunition reduced the post-UCT average BLL to 10.5 μg/dL (Weber et al., 2020). As a comparison to studies on occupational firearm users, a recent study of seven Swedish brass foundry workers aged 22–57 years found levels up to 33 μg/dL (GM = 8 μg/dL) (Julander et al., 2020). Nevertheless, also the comparably low BLLs in the recently reported investigations of occupational firearm users are cause for concern, particularly if considering the potential accumulation over a military career (cf. Table 1 and Fig. 4).

The EU presently has a binding biological exposure limit of 70 μg/dL blood, but this limit is being reviewed. In 2002, the Scientific Committee on OELs (SCOEL) published a recommendation of 30 μg/dL, with the reservation that this "is not seen as being entirely protective of the offspring of working women". In 2007, the ACGIH adopted a biological exposure index (BEI) of 20 μg/dL based on neurological and reproductive toxicity. However, ACGIH recommends that any application of the BEI is complemented with additional information about developmental risks to women of child-bearing age. The recent recommendation of 15 μg/dL from ECHA RAC (2020) could result in a lowered binding limit in the EU. Regardless of its legal status, there is an increasing concern for the health effects of lead, also within a healthy population of male workers.

As a comparison to these occupational guidance values, the European Food Safety Authority (EFSA, 2010) concluded that there is no evidence of a threshold for the critical effects of lead exposure but that risk for effects on systolic blood pressure in the general population, is "low" if BLLs are below 3.6 μg/dL (the 1% benchmark dose lower confidence limit, $BMDL_{01}$). For renal effects, a $BMDL_{10}$ (the 10% benchmark dose lower confidence limit) of 1.5 μg/dL was identified (EFSA 2010). Also, National Toxicology Program (NTP) concluded that BLLs <5 μg/dL were associated with decreased kidney functions in adults (NTP 2012). It should be noted that these estimates are intended to cover the entire population, including more sensitive individuals that may not be part of the working population. Nevertheless, the most sensitive endpoint is developmental toxicity, relevant for children and pregnant women ($BMDL_{01}$ of 1.2 μg/dL identified by EFSA 2010), highlighting the importance of considering women of fertile age in risk management efforts.

The extrapolations in the present study are subject to several limitations. The UTL, as a measure of an exposure level that we are 95% certain 99% of the population will be below, only applies under conditions resembling those of the studied UCT. Since weather, and factors influencing particle behaviour such as wind and humidity varies considerably, repeated measurements across the year are recommended. As fumes from firearms mainly contain particulates from the smaller fractions, sampling specifically the respirable fraction may be more relevant than the inhalable fraction. We based the extrapolation of BLL to long-term conditions on the assumption of repeated low exposure conditions, an unlikely scenario as indoor and outdoor training will be mixed, and intensity increased with increasing experience. The median cadet scenario is thus a median of a relatively low exposure scenario. However, the UTL cadet and instructor scenarios are conservative. In particular, our assumptions for instructors targeted worst-case conditions. We selected the highest observed pre-BLL as the assumed background value and the highest BLL increase as the course contribution. Furthermore, we assumed instructors perform 24 similar courses per year, which would require all working days that year. As instructors indeed were below the extrapolated BLLs, these values should not be seen as predictions but illustrations of the need to regularly assess instructors' BLLs and plan their work with a long-term perspective.

Encouragingly, there are several measures to control and mitigate metal exposure in these settings. Besides the apparent accessibility of lead-free ammunition for training purposes (Weber et al., 2020; Voie et al., 2014; Wurster et al., 2006), improved assessment and control of ventilation systems (Grabinski et al., 2017) can be efficient to reduce exposures. Routines for ventilation through open windows and doors during and between exercises are necessary if a ventilation system is missing or out of order. Moreover, thorough and regular cleaning of ranges is advised (Scott et al., 2012), preferably using an EX-rated vacuum cleaner equipped with HEPA filters (or equivalent). If dry sweeping cannot be avoided, e.g., for the collection of empty sleeves, cleaning staff should wear suitable respiratory protection. Since inhalation may not be the only route for lead exposure, careful personal hygiene routines and the use of disposable gloves during firearm maintenance (Johnson-Arbor et al., 2020) can further be recommended to participants in these exercises. Hygiene routines include no snuffing or smoking during, and strict handwashing after, training. Furthermore, contaminated clothes should be washed at work, not at home. A final and effective measure, but not always feasible, is the use of respiratory protection.

5. Conclusions

We have shown that a 9-day training course with mostly low airborne lead concentrations due to low intensity, favourable outdoor conditions, and in part use of lead-free ammunition resulted in a measurable and statistically significant increase in BLLs. None of the instructors or cadets exceeded the Swedish biological exposure limits nor the recent recommendation of ECHA. However, because of the accumulating properties of lead and the uncertainty regarding whether any blood level is without adverse health effects lead exposure should always be minimised. The importance of minimising exposures is further underlined by the fact that the sensitive subpopulation of fertile women are increasingly represented in military training as the gender gap in recruitment closes. Since airborne exposure may be difficult to eliminate entirely in military training, duration and frequency of exposure, and other factors, need to be considered. Care should be taken to avoid all sources of lead as far as possible, especially during indoor training. Exposure reduction measures such as improved ventilation and hygiene routines should be implemented and evaluated with personal air-sampling during training sessions. Additionally, when planning instructors' tasks, the case of an instructor repeatedly trailing groups over several hours should be avoided.

Declaration of competing interest

The authors' affiliations are as stated on the title page. LS's participation was funded by the Swedish Armed Forces and Karolinska Institutet. The authors declare no other interests.

Acknowledgements

We thank the two anonymous reviewers for their insightful and constructive comments. Artur Podobas is acknowledged for helpful comments on the clarity of the manuscript. All remaining errors are our own.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2021.113773.

References

Aurell, J., Holder, A.L., Gullett, B.K., McNesby, K., Weinstein, J.P., 2019. Characterization of M4 carbine rifle emissions with three ammunition types. Environ. Pollut. 254 https://doi.org/10.1016/j.envpol.2019.112982.

Bergstrom, U., Ekstrand-Hammarstrom, B., Hagglund, L., Wingfors, H., 2015. Comparing acute toxicity of gunshot particles, from firing conventional and lead-free ammunition, in pulmonary epithelial cell cultures. J. Toxicol. Environ. Health 78 (10), 645–661. https://doi.org/10.1080/15287394.2015.1017682.

Bonanno, J., Robson, M.G., Buckley, B., Modica, M., 2002. Lead exposure at a covered outdoor firing range. Bull. Environ. Contam. Toxicol. 68 (3), 315–323. https://doi.org/10.1007/s00128-001-0256-y.

Council, N.R., 2013. Potential Health Risks to DOD Firing-Range Personnel from Recurrent Lead Exposure. National Academies Press, Washington DC.

Donguk, P., Namwon, P., 2004. Exposure to lead and its particles size distribution. J. Occup. Health 46 (3), 225–229. https://doi.org/10.1539/joh.46.225.

EFSA – European Food Safety Authority, 2010. Scientific opinion on lead in Food. EFSA panel on contaminants in the Food chain (CONTAM). EFSA Journal 8 (4). https://doi.org/10.2903/j.efsa.2010.1570.

Fischbein, A., Rice, C., Sarkozi, L., Kon, S., Petrocci, M., Selikoff, I., 1979. Exposure to lead in firing ranges. J. Am. Med. Assoc. 241 (11), 1141–1144. https://doi.org/10.1001/jama.1979.03290370045027.

Flegal, A.R., Smith, D.R., 1992. Lead levels in preindustrial humans. N. Engl. J. Med. 326 (19), 1293–1294. https://doi.org/10.1056/NEJM199205073261916.

GESTIS (n.d.) International limit values database. Available at: https://limitvalue.ifa.dguv.de/Last accessed 2021-01-04.

Goldberg, R.L., Hicks, A.M., O'Leary, L.M., London, S., 1991. Lead Exposure at Uncovered Outdoor Firing Ranges, 33, pp. 718–719. https://doi.org/10.1097/00043764-199106000-00013, 6.

Grabinski, C.M., Methner, M.M., Jackson, J.M., Moore, A.L., Flory, L.E., Tilly, T., Hussain, S.M., Ott, D.K., 2017. Characterization of exposure to byproducts from firing lead-free frangible ammunition in an enclosed, ventilated firing range. J. Occup. Environ. Hyg. 14 (6), 461–472. https://doi.org/10.1080/15459624.2017.1296232.

Greenberg, N., Frimer, R., Meyer, R., Derazne, E., Chodick, G., 2016. Lead exposure in military outdoor firing ranges. Mil. Med. 181 (9), 1121–1126. https://doi.org/10.7205/milmed-d-15-00454.

Johnson-Arbor, K., Soto, P., Liu, L., 2020. Adult lead exposure from ammunition reloading and indoor residential shooting. Am. J. Ind. Med. 63 (8), 733–737. https://doi.org/10.1002/ajim.23119.

Julander, A., Midander, K., Garcia-Garcia, S., Vihlborg, P., Graff, P., 2020. A case study of brass foundry workers' estimated lead (Pb) body burden from different exposure routes. Annals of Work Exposures and Health 64 (9), 970–981. https://doi.org/10.1093/annweh/wxaa061.

Lach, K., Steer, B., Gorbunov, B., Micka, V., Muir, R.B., 2015. Evaluation of exposure to airborne heavy metals at gun shooting ranges. Ann. Occup. Hyg. 59 (3), 307–323. https://doi.org/10.1093/annhyg/meu097.

Laidlaw, M.A.S., Filippelli, G., Mielke, H., Gulson, B., Ball, A.S., 2017. Lead exposure at firing ranges-a review. Environ. Health 16. https://doi.org/10.1186/s12940-017-0246-0.

Mancuso, J.D., McCoy, J., Pelka, B., Kahn, P.J., Gaydos, J.C., 2008. The challenge of controlling lead and silica exposures from firing ranges in a special operations force. Mil. Med. 173 (2), 182–186. https://doi.org/10.7205/MILMED.173.2.182.

Novotny, T., Cook, M., Hughes, J., Lee, S.A., 1987. LEAD-EXPOSURE IN a firing range. Am. J. Publ. Health 77 (9), 1225–1226. https://doi.org/10.2105/ajph.77.9.1225.

NTP – National Toxicology Program, 2012. NTP Monograph on Health Effects of Low-Level Lead. U.S. Department of Health and Human Services, Research Triangle Park, NC.

Park, D.U., Paik, N.W., 2002. Effect on blood lead of airborne lead particles characterized by size. Ann. Occup. Hyg. 46 (2), 237–243. https://doi.org/10.1093/annhyg/mef036.

RAC – Committee for Risk Assessment, 2020. Opinion on Scientific Evaluation of Occupational Exposure Limits for Lead and its Compounds. European Chemicals Agency, Helsinki.

Rocha, E., Sarkis, J., Carvalho, M., dos Santos, G., Canesso, C., 2014. Occupational exposure to airborne lead in Brazilian police officers. Int. J. Hyg Environ. Health 217 (6), 702–704. https://doi.org/10.1016/j.ijheh.2013.12.004.

Scott, E.E., Pavelchak, N., DePersis, R., 2012. Impact of housekeeping on lead exposure in indoor law enforcement shooting ranges. J. Occup. Environ. Hyg. 9 (3), D45–D51. https://doi.org/10.1080/15459624.2011.648571.

Svensson, B.G., Schutz, A., Nilsson, A., Skerfving, S., 1992. Lead-exposure in indoor firing ranges. Int. Arch. Occup. Environ. Health 64 (4), 219–221, 10.1007bf00378278.

SFA - Swedish Food Agency, 2017. Swedish Market Basket Survey 2015 – Per Capita-Based Analysis of Nutrients and Toxic Compounds in Market Baskets and Assessment of Benefit or Risk. Swedish Food Agency, Uppsala, p. 164.

SWEA - Swedish Work Envirnment Authority, 2018. Hygieniska Gränsvärden. AFS 2018:1. DanagårdLiTHO AB, Stockholm, p. 80.

SWEA – Swedish Work Envirnment Authority, 2019. Medicinska Kontroller I Arbetslivet. AFS 2019:3. Swedish Work Environment Authority, Stockholm.

Tripathi, R.K., Sherertz, P.C., Llewellyn, G.C., Armstrong, C.W., Ramsey, L., 1989. Overexposures to lead at a covered outdoor firing range. J. Am. Coll. Toxicol. 8 (6), 1189–1195. https://doi.org/10.3109/10915818909018076.

US EPA - United States Environmental Protection Agency, 2003. Recommendations of the Technical Review Workgroup for Lead for an Approach to Assessing Risks Associated with Adult Exposures to Lead in Soil, 2ed.

Valway, S.E., Martyny, J.W., Miller, J.R., Cook, M., Mangione, E.J., 1989. Lead absorption in indoor firing range users. Am. J. Publ. Health 79 (8), 1029–1032. https://doi.org/10.2105/ajph.79.8.1029.

Vandebroek, E., Haufroid, V., Smolders, E., Hons, L., Nemery, B., 2019. Occupational exposure to metals in shooting ranges: a biomonitoring study. Safety and Health at Work 10 (1), 87–94. https://doi.org/10.1016/j.shaw.2018.05.006.

Vivante, A., Hirshoren, N., Shochat, T., Merkel, D., 2008. Association between acute lead exposure in indoor firing ranges and iron metabolism. Isr. Med. Assoc. J. 10 (4), 292–295 (NA).

Voie, O., Borander, A.-K., Sikkeland, L.I.B., Grahnstedt, S., Johnsen, A., Danielsen, T.E., Longva, K., Kongerud, J., 2014. Health effects after firing small arms comparing leaded and unleaded ammunition. Inhal. Toxicol. 26 (14), 873–879. https://doi.org/10.3109/08958378.2014.970783.

Weber, A.K., Bannon, D.I., Abraham, J.H., Seymour, R.B., Passman, P.H., Lilley, P.H., Parks, K.K., Braybrooke, G., Cook, N.D., Belden, A.L., 2020. Reduction in lead exposures with lead-free ammunition in an advanced urban assault course. J. Occup. Environ. Hyg. https://doi.org/10.1080/15459624.2020.1836375.

Wennberg, M., Lundh, T., Sommar, J.N., Bergdahl, I.A., 2017. Time trends and exposure determinants of lead and cadmium in the adult population of northern Sweden 1990–2014. Environ. Res. 159, 111–117. https://doi.org/10.1016/j.envres.2017.07.029.

Wingfors, H., Svensson, K., Hagglund, L., Hedenstierna, S., Magnusson, R., 2014. Emission factors for gases and particle-bound substances produced by firing lead-free small-caliber ammunition. J. Occup. Environ. Hyg. 11 (5), 282–291. https://doi.org/10.1080/15459624.2013.858821.

Wurster, U., Ebert, H., Fleig, E., Ott, G., 2006. Reducing exposure to hazardous substances at indoor shooting ranges by using lead-free training ammunition. Gefahrst. Reinhalt. Luft 66 (7–8), 295–299 (NA).

International Journal of Hygiene and Environmental Health 235 (2021) 113756

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

Kenyan school book knowledge for water, sanitation, hygiene and health education interventions: Disconnect, integration or opportunities?

Carmen Anthonj[a,c,*], Sophie Githinji[b], Christoph Höser[c], Alfred Stein[a], Justine Blanford[a], Valentina Grossi[c]

[a] Faculty of Geo-Information Science and Earth Observation, ITC, University of Twente, Enschede, the Netherlands
[b] Independent Consultant, Nairobi, Kenya
[c] Institute for Hygiene and Public Health, Medical Faculty, University of Bonn, GeoHealth Centre, Bonn, Germany

ARTICLE INFO

Keywords:
Behaviour change
Curriculum
Education
Health promotion
Sub-Saharan Africa
WASH

ABSTRACT

Background: Schools, depending on their access to and quality of water, sanitation and hygiene (WASH) and the implementation of healthy behaviours, can be critical for the control and spread of many infectious diseases, including COVID-19. Schools provide opportunities for pupils to learn about the importance of hygiene and WASH-related practice, and build healthy habits and skills, with beneficial medium- and long-term consequences particularly in low- and middle-income countries: reducing pupils' absenteeism due to diseases, promoting physical, mental and social health, and improving learning outcomes. WASH services alone are often not sufficient and need to be combined with educational programmes. As pupils disseminate their acquired health-promoting knowledge to their (extended) families, improved WASH provisions and education in schools have beneficial effects also on the community. International organisations frequently roll out interventions in schools to improve WASH services and, in some cases, train pupils and teachers on safe WASH behaviours. How such interventions relate to local school education on WASH, health promotion and disease prevention knowledge, whether and how such knowledge and school books are integrated into WASH education interventions in schools, are knowledge gaps we fill.
Methods: We analyzed how Kenyan primary school science text book content supports WASH and health education by a book review including books used from class 1 through class 8, covering the age range from 6 to 13 years. We then conducted a rapid literature review of combined WASH interventions that included a behaviour change or educational component, and a rapid review of international policy guidance documents to contextualise the results and understand the relevance of books and school education for WASH interventions implemented by international organisations. We conducted a content analysis based on five identified thematic categories, including drinking water, sanitation, hygiene, environmental hygiene & health promotion and disease risks, and mapped over time the knowledge about WASH and disease prevention.
Results: The books comprehensively address drinking water issues, including sources, quality, treatment, safe storage and water conservation; risks and transmission pathways of various waterborne (Cholera, Typhoid fever), water-based (Bilharzia), vector-related (Malaria) and other communicable diseases (Tuberculosis); and the importance of environmental hygiene and health promotion. The content is broadly in line with internationally recommended WASH topics and learning objectives. Gaps remain on personal hygiene and handwashing, including menstrual hygiene, sanitation education, and related health risks and disease exposures. The depth of content varies greatly over time and across the different classes. Such locally available education materials already used in schools were considered by none of the WASH education interventions in the considered intervention studies.
Conclusions: The thematic gaps/under-representations in books that we identified, namely sanitation, hygiene and menstrual hygiene education, are all high on the international WASH agenda, and need to be filled especially now, in the context of the current COVID-19 pandemic. Disconnects exist between school book knowledge and

* Corresponding author. Faculty of Geo-Information Science and Earth Observation – ITC, University of Twente, Hengelosestraat 99, P.O. Box 217, 7500AE, Enschede, the Netherlands.
E-mail addresses: c.anthonj@utwente.nl (C. Anthonj), sophiegithinji@gmail.com (S. Githinji), christoph.hoeser@ukbonn.de (C. Höser), a.stein@utwente.nl (A. Stein), j.i.blanford@utwente.nl (J. Blanford), valentina.grossi@ukbonn.de (V. Grossi).

https://doi.org/10.1016/j.ijheh.2021.113756
Received 19 November 2020; Received in revised form 8 April 2021; Accepted 18 April 2021
Available online 15 May 2021

WASH education interventions, between policy and implementation, and between theory and practice, revealing missed opportunities for effective and sustainable behaviour change, and underlining the need for better integration. Considering existing local educational materials and knowledge may facilitate the buy-in and involvement of teachers and school managers in strengthening education and implementing improvements. We suggest opportunities for future research, behaviour change interventions and decision-making to improve WASH in schools.

1. Introduction

Many diseases are entirely or partially attributable to inadequate drinking water, sanitation and hygiene (WASH) (WHO, 2019). Therefore it is important to provide safe water and sanitation, and to implement healthy hygiene behaviours.

Schools, as critical settings for the development and the early life of all children as well as for the work of teachers and other school staff, play a vital role here. With many pupils spending a substantial amount of their time on the premises, schools are important for pupils' physical and cognitive developments, as well as for their well-being. Moreover, schools can be critical places for the control and spread of infectious diseases such as gastrointestinal diseases or respiratory infections, including COVID-19, depending on their access to and quality of WASH (Adams et al., 2009; Freeman et al., 2012; Greene et al., 2012; Jasper et al., 2012; Joshi and Amadi, 2013; McMichael, 2019; Munn et al., 2020; Onyango-Ouma et al., 2005; Patel et al., 2012).

Minimum requirements for safe WASH in schools, such as drinking water from an improved source, useable improved facilities and handwashing facilities with available water and soap (WHO and UNICEF, 2018a) are not provided in most rural schools in Kenya. According to a 2017 school-based study (Morgan et al., 2017), 25% of the 198 surveyed schools relied on unimproved drinking-water sources, 38% of the schools had unsafe sources, contaminated with *E. coli,* and 44% of the schools collected drinking-water off-premise. About 25% of the rural schools had unimproved sanitation facilities and insufficient provisions for menstrual hygiene, and overcrowding of sanitation facilities was common in 24% of schools. Handwashing facilities were largely lacking (40%), as were soap (87%) and hand drying materials (81%) (Morgan et al., 2017).

As concerns the WASH situation in households, Kenyan data from 2017 show that 32% of the households (15 million people) lack access to an improved drinking water source at home. WHO and UNICEF define safe drinking water source as one that by nature of its construction, protects the water from contamination and has the potential to deliver safe water. (WHO and UNICEF, 2018a). Nine percent (4.5 million Kenyans) spend 30 min (roundtrip, including queuing) or more to collect water from an improved source (WHO/UNICEF, n.d.). Access to sanitation remains a challenge as well, with 10% of Kenyan households continuing to practice open defecation, and with only 38% having access to an unimproved sanitation facility, such as pit latrines without a slab or platform, hanging latrines or bucket latrines (WHO and UNICEF, 2018a). The lack of adequate handwashing facilities with water and soap in 75% of the households poses additional significant health risks. Domestic WASH provisions are significantly worse in rural areas, despite the slow but steady improvements over the course of the past 7 years (WHO/UNICEF, n.d.).

At schools pupils can learn about the importance of hygiene and WASH-related practice, and build healthy habits, both of which might not always happen at home. Adequate WASH provision for healthy practices at school have beneficial medium- and long-term consequences: WASH interventions in schools, such as - the provision of soap, water treatment or improvements to sanitation - can reduce absenteeism due to diseases, improving pupils' mental and physical health, nutritional status, and learning outcomes (Jasper et al., 2012; Joshi and Amadi, 2013). Improved WASH services, including means for menstrual hygiene management, also improve enrolment and gender parity, increasing girls attendance and well-being (Garn et al., 2013; Sumpter and Torondel, 2013). Negative health practices such as toilet avoidance and poor hydration adversely impact on pupils' attention and cognitive performance in class and on their health and well-being (D'Anci et al., 2006; Lukacz et al., 2011; Merhej, 2019). Besides the environmental provisions for WASH services, WASH education is of importance as well: the risk of parasitic infections, for example, is lower in children with knowledge on hygiene and sanitation practices. Moreover, WASH interventions with an integrated education component are more efficient, and ensure commitment and adherence to healthy practices such as handwashing (Joshi and Amadi, 2013).

WASH provisions and WASH education in schools also have effects on the community, as pupils disseminate their acquired health-promoting knowledge to their (extended) families. Pupils have been described as hygiene and health change agents for their peers and parents/guardians in school and home environments in the past (Onyango-Ouma et al., 2005). The role that pupils can play in helping to break the transmission route of various WASH-related diseases is remarkable, and proven to create significant health improvements (Blanton et al., 2010; Dreibelbis et al., 2014; O'Reilly et al., 2008).

In summary, water, sanitation, hygiene and health education promotes and multiplies healthy behaviour (Freeman et al., 2015; Greene et al., 2012; Joshi and Amadi, 2013; Patel et al., 2012).

For this reason, international organisations roll out interventions in schools to train teachers and pupils on safe WASH behaviour. What we do not yet know is how such interventions relate to local school book knowledge in low- and middle-income country contexts. A knowledge gap is what pupils actually learn during their school curricula with respect to WASH, health promotion and disease prevention, and whether and how this knowledge is integrated into WASH education interventions in schools.

In this paper, we focus on how primary school science text books content supports WASH and health education to promote healthy behaviour by the means of a standard science school book review from Kenya across the 6–13 years age range. We assess whether there is an integration or a disconnect with WASH behaviour interventions in schools in Kenya and beyond, and elaborate opportunities that WASH education in schools holds for interventions.

The novelty of this paper to the scientific community, as well as to WASH- and education-related practitioners and decision-makers, includes:

1. First assessment of water, sanitation, hygiene, health and disease in school books.
2. Mapping over time what and in which level of detail pupils study about WASH and disease prevention in Kenyan primary schools.
3. Analysing the level of integration of school book knowledge into WASH-related education interventions and the potential for sustainable health programming.

2. Methods

Building on a study examining the risk perceptions and behaviours to water-related infectious disease exposure and WASH (Anthonj et al., 2016, 2019), we further assessed what water, sanitation, hygiene and health information was being taught at schools. To do so we used the water-related disease transmission classification by Bradley (1974).

According to this classification, water-related infectious disease transmission can be due to a water-related insect vector (e.g. Malaria), water-based (e.g. Schistosomiasis), waterborne (e.g. Typhoid fever) or water-washed (e.g. Trachoma) – and is overall closely linked to the provision of safe water and sanitation, and the implementation of healthy hygiene behaviours. Since water supply and storage, sanitation, personal as well as environmental hygiene and human behaviour are on the one hand risk factors for the contraction of infectious diseases, yet can function as well as health-promoting factors (Anthonj et al., 2018), we built our school book review on these exact themes. We first conducted a review of text books used to teach science in Kenyan primary schools under the 8-4-4 education system (Ministry of Education of the Republic of Kenya, 1984) to capture the representation of WASH issues. We then conducted a rapid literature review of behaviour change and health messaging interventions and global policy guidance documents to contextualise the results of the text book review and understand the relevance of books and school education for WASH interventions.

2.1. Review of Kenyan primary school science text books

We reviewed primary school science text books in order to answer two questions:

1. How are drinking water, sanitation and hygiene issues represented in primary school science text books in Kenya?
2. Which WASH dimensions are addressed in the books and is there a causal chain relating themes to each other?

2.1.1. Search and screening strategy

Data presented in this paper were gathered in 2016 as part of a research project on water-related infectious disease exposure and WASH in Rumuruti, Laikipia County, Kenya. Based on a rapid screening of teaching materials used in the study area (Anthonj et al., 2016, 2018, 2019), we identified science education as the single source of information related to WASH and disease risks in primary schools. We therefore collected the science text books in use at the time of this study. Eight primary school standard science books used from class 1 through class 8 covering the age range from 6 to 13 years (Table 1) were obtained and analyzed.

We conducted a content analysis and encoded the material into five categories in order to transform the raw data into analysable data. Building on the classification of water-related disease transmission by Bradley (1974), and the factors preventing the transmission of these diseases, we included drinking water; sanitation; hygiene; environmental hygiene & health promotion and disease risks as categories (corresponding to Anthonj et al., 2018) We manually screened the eight books for all information related to our identified categories.

2.1.2. Content analysis

Water supply and storage, sanitation, personal as well as environmental hygiene and human behaviour are on the one hand risk factors for the contraction of infectious diseases, yet can function as well as health-promoting factors. We therefore built this school book review on five themes, that included:

i. Drinking water (water source; water transport; water quality; water safety; household water treatment and safe storage; water conservation);
ii. Sanitation (sanitation facility; toilet; sanitary hygiene; sanitation behaviour);
iii. Hygiene (personal hygiene; handwashing, hygiene item; food hygiene; food storage);
iv. Environmental hygiene with disease prevention and health promotion; and
v. WASH-related disease risks and transmission pathways.

2.2. Rapid literature review of behaviour change and health messaging interventions and global policy guidance documents

We conducted a rapid review of peer-reviewed and grey literature to identify publications on WASH education and how this may affect WASH behaviour change interventions in schools to answer two additional questions:

3. In what ways is school book information integrated in WASH interventions in schools?
4. Which potential does WASH education in schools hold for interventions?

2.2.1. Search and screening strategy

Our a rapid review aimed at identifying publications from (i) grey literature that address WASH education in schools and (ii) peer-reviewed journal articles that address the use of WASH in schools (WinS) interventions. Two comprehensive literature reviews on WASH in schools (Jasper et al., 2012; Joshi and Amadi, 2013) served as a starting point for approaching the topic. We screened them for interventional studies addressing components related to education, health promotion and behaviour change.

Additionally, we deployed the electronic literature database MEDLINE for computer-based searches, to retrieve studies published after the two reviews. The search within titles, abstracts and keywords included the keyword combination of school AND water OR sanitation OR

Table 1
Kenyan primary school science textbooks for classes 1–8 used for book review.

Class	Code used	Name of book	Authors	Year published	Publisher	Place published	Taught at age
1	C1	Science Matters 1. A Science Course for Primary Schools	Ojwang & K'opiyo	2003	East African Publishers Ltd.	Nairobi	6 years
2	C2	Science Matters 2. A Science Course for Primary Schools	Embeywa & Ndungi	2003	Oxford University Press, East Africa Ltd.	Nairobi	7 years
3	C3	Science in Action 3	Embeywa et al.	2004	Oxford University Press, East Africa Ltd.	Nairobi	8 years
4	C4	Understanding Science 4. Pupil's Book 4	Karaka et al.	2005	Longhorn Publishers Ltd.	Nairobi	9 years
5	C5	Primary Science 5. Pupils' Book for Standard Five	Gichuki	2005	Kenya Literature Bureau	Nairobi	10 years
6	C6	Primary Science 6. Pupils' Book for Standard Six	Gichuki	2005	Kenya Literature Bureau	Nairobi	11 years
7	C7	Primary Science 7. Pupils' Book for Standard Seven	Gichuki	2005	Kenya Literature Bureau	Nairobi	12 years
8	C8	Primary Science 8. Pupils' Book for Standard Eight	Gichuki	2005	Kenya Literature Bureau	Nairobi	13 years

hygiene AND intervention.

The results of the search were sorted by “best match”, an option on MEDLINE to identify the most relevant publications, and the first identified 100 papers were non-systematically screened in order to identify potentially relevant studies for abstract review. Of those identified, only publications on intervention studies on WASH in school settings that addressed our five categories of the school book review, namely (i) drinking water; (ii) sanitation; (iii) hygiene; (iv) environmental hygiene with disease prevention and health promotion; and (v) WASH-related disease risks and transmission pathways, were included. In case multiple articles were retrieved for the same intervention, only one was considered. For every included study, the list of related articles suggested by the database as well as the list of articles citing the study were screened to possibly find additional studies. Moreover, included studies were hand-searched for additional bibliographical references. Publications in languages other than English were excluded.

Grey literature on WASH education for schools was searched on Google using the following keywords combination: school, water, sanitation, hygiene, and education (manual). The first 50 results were screened to identify potentially relevant materials. Of those identified, only publications on educational programs and objectives on WASH in school settings that addressed our five categories of the school book review were considered.

2.2.2. Content analysis

Of the n = 100 papers screened during the rapid review, 23 matched the search criteria of being an intervention study in a school settings addressing WASH aspects, considering an educational or behavioural change component. This final set of eligible texts was subjected to analysis and synthesis. Corresponding to the analysis of the school book data, the same five categories were used in order to compare the included studies to what we found in the school book data. We extracted and tabulated relevant information, including type of educational intervention, means of provision of education or health promotion, topics addressed, educational materials considered and location of intervention. We interpreted, contextualised and synthesised the results. From the grey literature, n = 5 publications were included and we analyzed suggestions on main WASH topics to be included in WASH education interventions, learning objectives and WASH curricula.

3. Results

3.1. Overview mapping of WASH themes in Kenyan primary school science text books

3.1.1. Drinking water, water safety, household water treatment and safe storage

Drinking water issues are addressed in four of the eight primary school science text books reviewed. These were C2–C4 and C7 (Table 2), a total of 22 text book pages (Table 3) and numerous illustrations (Table 4).

Learning about drinking water starts in class 2 with a reminder to “*always drink clean water*” and information on different ways of removing substances from water. Step-by-step activities navigate the pupils through using different cloth materials as filters (C2).

C3 teaches the pupils how to make water safe for drinking, and introduces step-by-step guides to filtering and boiling drinking water because “*dirty water can make us sick*”. The book recommends filtered water to be boiled prior to drinking. The need to safely store water in clean, covered containers is described, with storage options including bottle, tank, jerry can and pot. Moreover, different ways of transporting water are addressed, such as carrying water in a pot, using oxen or a cart, carrying a jerry can on the head, or on the back, using a canal or pipes.

C4 entails more details and specifications on the ways of storing water in tanks, drums, pots, buckets and dams. Different ways of water use are described, including using water at home, in the farm, for recreation, industry or transport. Water use at home includes drinking and health promotion, cooking, personal hygiene, environmental hygiene and sanitation, and others.

C7 highlights the “*proper care and use of water and water sources*” in

Table 2
Overview mapping of WASH themes in Kenyan primary school science textbooks for classes 1–8 by themes.

Class book / Theme covered	1	2	3	4	5	6	7	8
Drinking water		Water treatment, removal of substances from water	Water transport, household water storage, water safety	Water use, household water storage			Water conservation	
Sanitation		Sanitation facilities and use	Sanitary hygiene					
Hygiene	Personal hygiene, hygiene items, handwashing	Hygiene items, cleaning personal hygiene items				Food hygiene, preservation and storage		
Environmental hygiene & health promotion		Cleaning compound, waste disposal	Medical care, clean surroundings of house		Health education, use and storage of medicine	Prevention of malaria, tuberculosis, waterborne diseases	Control of water pollution, livestock parasites and intestinal worms	
Disease risks						Communicable diseases, malaria, tuberculosis, waterborne diseases	Effects of water pollution, livestock diseases, parasites	Livestock diseases, effects of livestock diseases

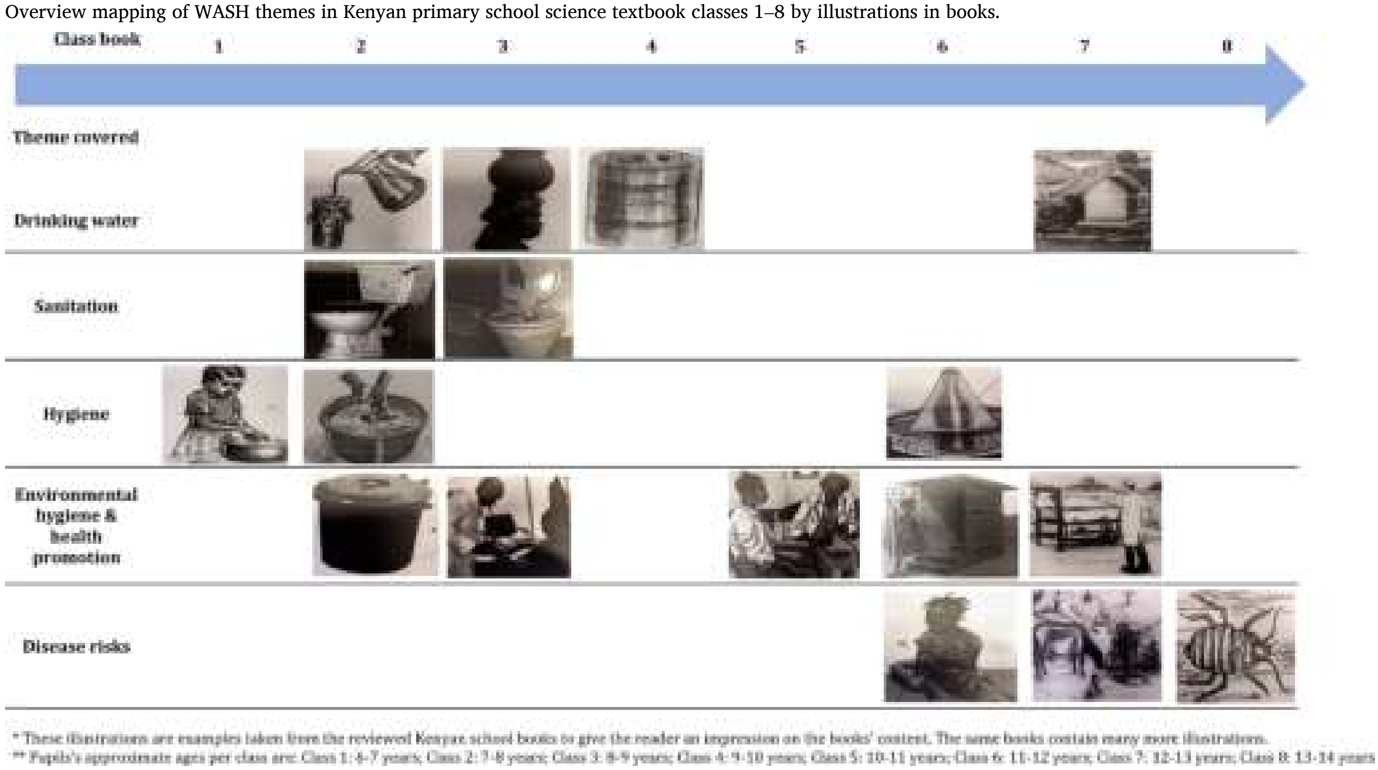
* Pupils's approximate ages per class are: Class 1: 6-7 years; Class 2: 7-8 years; Class 3: 8-9 years; Class 4: 9-10 years; Class 5: 10-11 years; Class 6: 11-12 years; Class 7: 12-13 years; Class 8: 13-14 years

Table 3
Overview mapping of WASH themes in Kenyan primary school science textbook classes 1–8 by number of pages in books.

Class book / Theme covered	1	2	3	4	5	6	7	8
Drinking water		3	4	11			4	
Sanitation		1.5	2					
Hygiene	2	2				6		
Environmental hygiene & health promotion		0.5	1		6	3.5	5	
Disease risks						8	9.5	2

* The darker the shades in the table, the higher the number of pages in school books addressing WASH themes in school books.
** Pupils's approximate ages per class are: Class 1: 6-7 years; Class 2: 7-8 years; Class 3: 8-9 years; Class 4: 9-10 years; Class 5: 10-11 years; Class 6: 11-12 years; Class 7: 12-13 years; Class 8: 13-14 years.

Table 4
Overview mapping of WASH themes in Kenyan primary school science textbook classes 1–8 by illustrations in books.



order to "*ensure that water is spared for future use,*". The book also teaches ways of water conservation including rainwater harvesting, recycling water, reusing water, constructing dams and others.

3.1.2. Sanitation facilities, sanitary hygiene and sanitation behaviour

Sanitation issues are addressed in C2–C3 (Table 2), with 3.5 text book pages covering sanitation in total (Table 3).

C2 describes different types of sanitation facilities, including pit latrines and flush toilets. Moreover, a simple manual guides the pupils on how to use toilets: (i) enter the toilet and close the door; (ii) prepare yourself and sit or squat on the toilet seat; (iii) after you finish, clean yourself; (iv) flush the toilet; (v) wash your hands. Guides on the right use of latrines and urinals are provided as well.

C3 teaches the pupils how to clean toilets, urinals and latrines, and reminds the pupils that "*you will need soap, water and a hard brush*". It also states that "*good health requires clean surrounding"* and that *"toilets, urinals and latrines must be kept clean all the time"*.

3.1.3. Personal hygiene and handwashing

Hygiene issues are addressed in C1–C2 and C6 (Table 2), with a total of ten text book pages covering the topic (Table 3).

Learning about hygiene starts in class 1 with a photo story on how to clean and wash hands with soap and clean water. Different items for personal hygiene are presented, including soap, towel, sponge and water. Moreover, the book calls for washing hands at key times, including (i) before and after eating; (ii) before preparing food; (iii) after visiting the toilet (C1).

C2 focuses on different hygiene items such as toothbrush, face towel and handkerchief. It includes a step-by-step guide on how to clean a handkerchief, because "*using dirty handkerchiefs can make you sick*". Moreover, it reminds the pupils "*not share your handkerchief with other people*".

C6 addresses the topic of food hygiene and proper food storage, and describes methods of food preservation.

3.1.4. Environmental hygiene, disease prevention and health promotion

Environmental hygiene and health promotion issues are addressed in five science text books: in C2–C3 and C5–C7 (Table 2), 16 text book pages were dedicated to this issue (Table 3).

C2 briefly touches upon the need of cleaning the domestic compound and to "*put all the rubbish in the rubbish pit*".

C3 refers to the need to keep the house and classroom clean as well. Besides, the book underlines the value of good health, as "*people feel well when their bodies are in good health*", "*good health means the body feeling*" and "*the mind is thinking well*". The book calls for pupils feeling unwell to notify their teacher or parents to take them to a doctor for assessment and cure.

In C5, pupils learn about the proper use and storage of medicines by the means of checklists they can follow.

C6 provides details on how to prevent the following diseases: (i) **Malaria** (i.e. draining stagnant water and covering any water storage containers to prevent breeding places for mosquitoes; sleeping under a mosquito net; protecting windows with mosquito gauze; taking malarial preventative drugs); (ii) **Tuberculosis** (i.e. immunisation; avoiding exposure to a lot of dust through environmental hygiene; well-ventilated rooms; avoiding overcrowded places; high standards of cleanliness); (iii) **Cholera** (i.e. proper disposal of faeces; use of sanitation facilities; sanitary hygiene; handwashing at key times; food hygiene; treatment of drinking water; prevention of water pollution; treatment of infected persons; cleanliness); (iv) **Typhoid fever** (i.e. drinking treated water; food hygiene; handwashing at key times; proper disposal of faeces; sanitary hygiene); and (v) **Bilharzia** (i.e. use of sanitation facility; draining stagnant water; cleanliness; prevention of direct water contact; killing water snails using chemicals.

In C7, different ways to control water pollution are presented, including practicing proper hygiene, using sanitation facilities, and avoiding surface water sources. Moreover, the textbook recommends "*drawing water for animals instead of taking them to water sources. This will prevent them from releasing waste into the water*" and "*controlling the dumping of industrial waste into water sources. All waste should be treated to make it harmless to the environment*". C7 also suggests different methods of controlling human intestinal worms. According to the book, the main ways of breaking the transmission route from one person to another or from an animal to a person is through proper sanitation practices and "*keeping the sources of water such as wells and rivers clean. Animals should not be watered in water sources where people get their drinking water. This ensures that animals do not drop their faeces in the sources of water. This will help to control the spread of the parasites. Faeces and other garbage should not be dumped in rivers or dams where people draw their water*." Moreover, "*burning or burying garbage to destroy the eggs of intestinal worms*" is recommended, as well as "*using latrines always so that pigs and other animals do not get into contact with human waste. This can help to control the worms*". Besides, food hygiene and regular deworming are advised.

3.1.5. WASH-related disease risks and transmission pathways

Disease risks and transmission pathways related to inadequate drinking water, sanitation, hygiene and environmental health are addressed in C6–C8 (Table 2). A total of 19.5 textbook pages deal with WASH-related diseases (Table 3).

Learning about specific WASH-related disease risks and transmission pathways starts in C6 with very detailed information on communicable diseases. The signs, symptoms and causes, along with explanatory illustrations, are detailed for malaria and tuberculosis as follows:

- Malaria, *"caused by some small parasites which feed on the red blood cells of the human body", with "parasites passed from one person to another through mosquito bites. When a mosquito bites a person who is infected with malaria, the parasites enter the body of the mosquito. When the mosquito carrying malaria-causing parasites bites a healthy person, he or she gets infected with malaria. In this way, malaria is spread from one person to another"*;
- Tuberculosis, *"caused by some germs called bacteria" that "mainly attack the lungs. When a person who has TB coughs, these bacteria get into the surrounding air. If a healthy person breathes in the air carrying tuberculosis-causing bacteria, he or she may contract the disease. In this way, TB is spread from one person to another"*.

Moreover, waterborne diseases are discussed in detail. According to C6, "*there can be many disease-causing germs in the water around us. These disease-causing germs are very small; we cannot see them with the naked eye. A person can get a disease by drinking, walking in or swimming in water containing germs. Diseases that are transmitted through water are called waterborne diseases. Water which contains germs is said to be contaminated.*" The signs, symptoms and causes, along with explanatory illustrations, are detailed for the following waterborne diseases:

- Cholera, *"caused by germs found in contaminated water. It is especially common when there are floods and water gets contaminated with human faeces. Cholera germs are passed on through drinking contaminated water. Flies also carry cholera germs to food. Cholera spreads rapidly from one person to another"*;
- Typhoid, *"a very serious waterborne disease" affecting the intestines and "caused by germs found in contaminated water. A person may get typhoid fever after eating food and drinking water contaminated with faeces or urine from a typhoid carrier. Water which is contaminated with sewage can spread typhoid very quickly*";
- Bilharzia, *"caused by bilharzia worms", "first found in human beings". "They enter into fresh water snails and are later passed through contaminated fresh water. A person suffering from bilharzia normally passes out the eggs of bilharzia worms in his/her urine or faeces. If the eggs are passed into water through the urine or faeces, they hatch into larvae which enter the body of a freshwater snail and get into the water, ready to attack human beings. If one walks in, swims in or drinks contaminated water, the larvae enter his or her body through the skin, particularly the legs"*.

C7 teaches pupils that "*polluted water is not safe for drinking by human beings and animals or for watering crops*", "*because it contains harmful substances that can cause diseases*". It aims at making pupils understand that water pollution, caused by floods, human and animal waste and uncontrolled use of pesticides and fertilisers, could be easily prevented. Moreover, pupils learn about different types of parasites that can affect livestock and that, "*most parasites, both external and internal, that affect livestock, are also a nuisance to human health. Internal parasites are spread* *from animals to human beings through food and water*". Safe disposal of sanitary waste and maintenance of environmental hygiene help to control the spread of these parasites.

In C8, the effects of livestock diseases on livestock are described. Some of these include stunted growth, loss of weight, reduced yields, rough coat, coughing and blood in stool. According to the book, "*when livestock are infected with diseases, they affect the farmer and the community*".

3.2. Overview mapping of WASH in schools interventions; rapid literature review

3.2.1. Content of grey literature on WASH education

Several guidelines and manuals have been developed by national and international organisations such as UNICEF, WHO, governmental and development organisations, specifying curricular content and learning objectives for WASH education, including tips and examples of activities. We identified the top five relevant guidance documents in our rapid review (Kenya Water for Health Organization, 2018; UNICEF, 2012; UNICEF Sri Lanka, n.d.; USAID/Splash, 2015; WHO Regional Office for Europe, 2019) (Table 5).

Corresponding to what we found in Kenyan school books, WASH education generally addressed one or more of the following themes (Table 2 vs. 5):

- Drinking water treatment, handling and storage (UNICEF, 2012; UNICEF Sri Lanka, n.y.; Kenya Water for Health Organization, 2018; USAID/Splash, 2015), and hydration (Kenya Water for Health Organization, 2018; WHO Regional Office for Europe, 2019)
- Safe faecal disposal and safe use of toilets and urinals (UNICEF, 2012; WHO Regional Office for Europe, 2019; UNICEF Sri Lanka, n.y.; Kenya Water for Health Organization, 2018; USAID/Splash, 2015)
- Personal hygiene and handwashing with soap (UNICEF, 2012; WHO Regional Office for Europe, 2019; UNICEF Sri Lanka, n.y.; USAID/Splash, 2015), food hygiene (UNICEF, 2012; WHO Regional Office for Europe, 2019; UNICEF Sri Lanka, n.y.; Kenya Water for Health Organization, 2018; USAID/Splash, 2015), male and female hygiene and menstrual hygiene management (UNICEF, 2012; WHO Regional Office for Europe, 2019; USAID/Splash, 2015)
- Environmental hygiene, waste management and water drainage (UNICEF, 2012; WHO Regional Office for Europe, 2019; UNICEF Sri Lanka, n.y.; Kenya Water for Health Organization, 2018)
- WASH-related diseases (UNICEF, 2012; Kenya Water for Health Organization, 2018; USAID/Splash, 2015)

Content on faecal disposal includes a lesson on safe disposal of faeces, and use of toilets. It covers the importance of sanitation for the prevention of diseases such as diarrhoea and worm infections and the F-diagram, healthy voiding/defecation, the use of toilets and sanitary hygiene. For waste management, the content often includes safe handling of waste and importance of keeping the environment clean, or detail about reducing, reusing and recycling waste. For older pupils it covers aspects such as managing infectious waste or dangerous waste.

3.2.2. Content of publications on WASH education interventions

We identified 23 intervention studies addressing WASH or hygiene education in schools in our rapid review. Twelve of the studies were conducted in Kenya and eleven were done elsewhere (Table 6).

WASH education interventions typically targeted entire schools and several **classes**, grades and a large **age range** at a time. Only one study from China targeted only one grade, class 1 (Bowen et al., 2007).

The focus of WASH education interventions was mainly on hand hygiene (n = 21). Half of the interventions targeted drinking water safety (n = 12). One third of interventions addressed WASH-related diseases such as diarrhoea, cholera, influenza (n = 7). Few of the identified intervention studies included education on waste management (n = 4), sanitation (n = 3), cleanliness of the environment (n = 3), and food hygiene (n = 3). Of the WASH in schools intervention studies identified through our rapid review, n = 13 included an **educational component** such as ad-hoc lessons or a specific curriculum.

Less than half of interventions included a **curricular component** (n = 9) (Bowen et al., 2007; Karon et al., 2016; Muckelbauer et al., 2009;

Table 5
Overview mapping of WASH in schools interventions. Content of grey literature on WASH education.

Grey literature	UNICEF (2012)	WHO Regional Office for Europe (2019)	UNICEF Sri Lanka (n.y.)	Kenya Water for Health Organization (2018)	USAID/Splash (2015)
Theme covered					
Drinking water	Water testing and treatment, removal of faecal or chemical contamination	Hydration and voiding matters.	Water treatment, handling and storage, water safety, water mapping, water quality testing	Water use, water treatment methods	Water transport, household water storage, water safety, water treatment measures, hydrological cycle
Sanitation	Toilet and urinal use	Toilet use, sanitary hygiene, healthy voiding/defecation, link to social education and respect	Toilet use, mapping sanitation, de-rehydration	Sanitation practices to keep the home, school and community surroundings clean	Safe disposal of faeces
Hygiene	Personal hygiene and handwashing at key times, genital and menstrual hygiene, importance of food hygiene	Key personal hygiene behaviours, handwashing tips and instructions, importance of genital hygiene, menstrual hygiene management, importance of food hygiene	Importance of handwashing, personal hygiene, importance of food hygiene and safe handling	Hygiene practices and importance of personal hygiene	Personal hygiene, hygiene items, cleaning personal hygiene items, handwashing with soap/ash, building a handwashing station, menstrual hygiene management, food hygiene
Environmental hygiene, waste management, health promotion	Solid waste and water management to reduce mosquito breeding, reduction of diarrhoeal and acute respiratory diseases through WASH, healthy nutrition	Waste management (reducing, reusing and recycling), environmental protection and resulting health benefits, management of infectious waste, WASH-related health benefits	Solid waste and water drainage mapping, sorting solid waste	Environmental hygiene practices, waste management	Environmental hygiene: animals, disinfection, disposal of waste, wastewater, food hygiene practices to prevent food-borne contamination
Disease risks	Water-related diseases, e.g. malaria, risks related to contaminated food	Health risks related to poor WASH conditions and practices, food-borne disease risks	F-diagram and disease risks related to poor sanitation, risks related to unsafe food handling	Water pollution and waterborne diseases, difference between clean and safe water, food-borne disease risks	Faecal-oral transmission of germs, risks related to faecal contamination and open defecation, effects of pollution

Table 6
Description of included studies on WASH education interventions in schools based on n = 23 publications identified by rapid review.

#	Study	Study setting	Class and age range	Intervention(education/ behaviour component)	education topics							curricular				book	
					water	sanitation	hand hygiene	food hygiene	cleanliness	waste management	disease	clubs	curricular	extracurricular	unclear	ad-hoc material	not considered
Studies conducted in Kenya																	
1	Blanton et al.	Nyanza Province	Classes not specified, ages 8-19	Training on hygiene as well as instructional books.	x		x				x	x		x		x	
2	Caruso et al. (2014)	Nyando, Kisumu, Rachuonyo Districts, Nyanza Province	Classes and ages not specified	Head teacher and health patron training behaviour change component: (i) Making & using soapy water, (ii) handwashing techniques at critical wash times, (iii) latrine cleaning and monitoring instruction.	x		x		x					x			x
3	Freeman et al. (2012)	Nyando, Kisumu & Rachuonyo, Suba Districts, Nyanza Province	Classes not specified, average age 13	Training of teachers on health promotion, behaviour change and water treatment methods and regular follow-up visits throughout the school year.	x										x		x
4	Garn et al. (2013)	Rachuonyo, Suba, Nyando, Kisumu Districts, Nyanza Province	Classes 1–7, ages not specified	Hygiene promotion (no details).	x	x	x								x		x
5	Greene et al. (2012)	Nyanza Province	Classees 4 through 8 (typically ages 6–16)	Teacher training on maintenance of drinking and handwashing facilities and behaviour change promotion lessons with pupils through health clubs or other venues.			x					x		x			x
6	La Con et al. (2017)	Western Kenya	Classes 1 through 8, ages not specified	Teacher training and provision of educational materials.	x		x					x		x		x	
7	Onyango-Ouma et al. (2005)	Bondo District, Western Kenya	Classes 2,3 and 5, ages range 9–15	Action-oriented and participatory health education and follow-up phase with pupils working as health communicators in school, community and families.	x	x	x	x	x	x	x		x			x	
8	O'Reilly et al. (2008)	Nyanza Province	Classes 4–8, ages not specified	Teacher training on SWS use and proper six step handwashing practices, provision of training materials for classroom, form safe water clubs with pupils, teach SWS and hygiene.	x		x					x	x	x		x	
9	Pasewaldt et al. (2019)	Kenya & Uganda	Classes 3–8, ages not specified	Hand hygiene curriculum consisting of various lessons and activities to educate pupils about handwashing and healthy handwashing behaviour.			x				x		x			x	
10	Patel et al. (2012)	Nyando Division, Nyanza Province	Classes 4–8, ages not specified	Teacher training on handwashing and water treatment, provision of instructional materials for pupils.	x		x								x	x	

(continued on next page)

Table 6 (*continued*)

#	Study	Study setting	Class and age range	Intervention(education/ behaviour component)	education topics							curricular				book	
					water	sanitation	hand hygiene	food hygiene	cleanliness	waste management	disease	clubs	curricular	extracurricular	unclear	ad-hoc material	not considered
11	Saboori et al. (2013)	Nyanza Province	Classes 4–7, ages not specified	Teacher training on handwashing prior to implementation.			x					x		x			x
12	Schlegelmilch et al. (2016)	District of Kinango, Coast Province	Classes and ages not specified	Health and hygiene promotion education to communities and schools, employing Community Led Total Sanitation (CLTS) methods.		x	x				x			x			x
Studies conducted outside Kenya																	
13	Bowen et al. (2007)	China	Classes 1, ages not specified	Handwashing programme.			x				x		x			x	
14	Chard et al. (2019)	Laos	Classes 3–5, ages not specified	Behaviour change component: hygiene action led by pupils in schools implemented after the installation of WASH hardware components.	x		x		x	x				x			x
15	Freeman and Clasen (2011)	Southern India	Classes and ages not specified	Teacher and pupil training, and provision of basic hygiene and water treatment information.	x		x							x			x
16	Karon et al. (2016)	Indonesia	Classes 4–5, ages not specified	Hygiene promotion with activities including capacity development of children, teachers, headmasters, parents and community members for improved hygiene behaviour among school children.	x		x	x		x	x		x	x		x	
17	Lau et al. (2012)	Chigaco	Classes pre-kindergarten to grade 8, ages 4–14	Short repetitive instruction in hand hygiene every two months.			x							x			x
18	Muckelbauer et al. (2009)	Germany	Classes 2–3, average age 8	Teachers presenting prepared classroom lessons to promote water consumption.	x								x			x	
19	Schulte et al. (2012)	North Texas	Classes and ages not specified	Sustained instruction in handwashing and monitoring of hygiene practices among pupils.			x							x			x
20	Shrestha et al. (2020)	Nepal	Classes not specified, ages 8-17	Health promotion activities.			x	x		x			x	x		x	
21	Talaat et al. (2011)	Egypt	Classes not specified, average age 8	Intensive campaign to promote hand hygiene, requiring pupils to wash their hands at least twice during the school day for 45 s, followed by proper rinsing and drying with a clean towel.			x							x			x
22	Tidwell et al. (2020)	India	Classes not specified, age 8-13	Teacher training and curriculum for students on handwashing.			x						x			x	
23	Vally et al. (2019)	Philippines	kindergarten to grade 6, ages 5–12	Children Hygiene and Sanitation Transformation (CHAST) methodology, using exercises and educational games to teach children about the links between personal hygiene and health.			x				x		x			x	

Onyango-Ouma et al., 2005; O'Reilly et al., 2008; Pasewaldt et al., 2019; Shrestha et al., 2020; Tidwell et al., 2020; Vally et al., 2019). Interventions mainly promoted **extracurricular educational activities** (n = 14) (Blanton et al., 2010; Caruso et al., 2014; Chard et al., 2019; Freeman and Clasen, 2011; Greene et al., 2012; Karon et al., 2016; La Con et al., 2017; Lau et al., 2012; O'Reilly et al., 2008; Saboori et al., 2013; Schlegelmilch et al., 2016; Schulte et al., 2012; Shrestha et al., 2020; Talaat et al., 2011) and **health clubs** (n = 5) (Blanton et al., 2010; Greene et al., 2012; La Con et al., 2017; O'Reilly et al., 2008; Saboori et al., 2013). Such extracurricular educational activities were often health promotion programs or campaigns or behaviour change interventions rather than a structured education program with written curriculum/lessons, frontal lessons, and specific content/learning objectives.

The interventions with an **educational component** were usually limited in time – ranging from 1 day at the minimum to 2 months at the maximum - and included ad-hoc developed materials for teachers and children and lesson programmes (n = 12) (Blanton et al., 2010; Bowen et al., 2007; Karon et al., 2016; La Con et al., 2017; Muckelbauer et al., 2009; Onyango-Ouma et al., 2005; O'Reilly et al., 2008; Pasewaldt et al., 2019; Patel et al., 2012; Shrestha et al., 2020; Tidwell et al., 2020; Vally et al., 2019).

Three interventions did **not provide enough information** on the type of intervention (whether curricular or extracurricular) (Freeman et al., 2012; Garn et al., 2013; Patel et al., 2012).

Notably, none of our identified WASH education interventions (n = 23) explicitly considered evaluating or using the existing school curriculum or integrating the intervention lessons into the school books in use (Table 6).

4. Discussion

4.1. Why what is in the books matters: primary school teaching material from Kenya supports WASH and health education and promotes healthy behaviour

WASH education is not only important to improve knowledge about healthy behaviours, but it can also improve attitudes and generate commitment to healthy conditions which eventually drive change, making students improve their hygiene practices through health-protective actions (Jasper et al., 2012; Onyango-Ouma et al., 2005; UNICEF, 2012; WHO, 2019).

The reviewed Kenyan science textbooks address water, sanitation, hygiene, environmental hygiene, health promotion and disease risks in a total of 71 book pages over the course of eight years of primary school education (Table 3). The topic addressed in most detail is drinking water (22 pages), with pupils learning about different ways of treating, safely storing and using drinking water. The books discuss the need for and value of water conservation strategies, such as rainwater harvesting. Altogether, the information provided by the school textbooks present a comprehensive understanding on the value and benefits of drinking safe water, while contextualising drinking water in the bigger hydrological water cycle and various, competing uses (Table 2).

WASH-related disease risk factors and transmission pathways are well covered across school years (19.5 pages) and suitably developed in the Kenyan school books. Specifically, various waterborne (Cholera, Typhoid fever), water-based (Bilharzia), vector-related (Malaria) and other communicable diseases (Tuberculosis) are presented. The books well explain several aspects of the causal chain and interconnection of (the lack of) WASH and exposure to these diseases such as the adverse effects of water pollution on human health and risks related to lack of hygiene when interacting with livestock. This link is also discussed in the context of environmental hygiene (16 pages), which in the books covers the need to keep the domestic compound clean and dispose waste properly, or control water pollution in order to prevent disease.

Drinking water-, WASH-related disease risk- and environmental health-related topics are therefore comprehensively covered in the Kenyan 8-4-4 primary school science textbooks. However, given that **WASH** includes the three dimensions of water AND sanitation AND hygiene, outstanding gaps were identified in lack of detailed educational content on sanitation (3.5 pages) and hygiene (10 pages). Hygiene and handwashing, a key measure to prevent water-related and other infectious diseases, are only taught briefly in C1-2 (4 pages). The content solely covers the importance and implementation of personal hygiene, handwashing and cleaning of hygiene items. Vital hygiene issues such as food hygiene could be added in earlier classes, as well as biological and epidemiological concepts on the need for hygiene and disease prevention. To favour long-lasting behaviour change and healthy practices, it is recommended to provide sustained hygiene promotion, as negative habits re-emerge over time even after initially successful interventions (Neal et al., 2015). Sanitation education only covers the different types and use of sanitation facilities, as well as the importance of sanitary hygiene, with a brief reference to WASH-related disease risks. Further details on risks related to poor sanitation are not addressed. Open defecation, a major risk factor to WASH-related infectious disease, is not addressed at all even though 12% of the Kenya's population still practice it. Menstrual hygiene management, a key topic in WASH in schools, is not addressed at all in any of the eight science school textbooks reviewed either. This is a key gap in the core teaching material given the importance of menstrual hygiene to prevent infections and reduce girls' absenteeism in school.

Another challenge in fully understanding the causal chain of (lack of) WASH and health risk – or health promotion – results from the fact that the different themes are not consistently covered through the primary school classes. While hygiene, for example, is addressed in class 1 and 2, there is a gap until class 6, when pupils continue to learn about hygiene. Likewise, sanitation is only covered in C2-3, and not at all before or thereafter. Finally, pupils do not learn about disease risks until C6.

In summary, the depth of content of educational materials varies greatly across the primary school classes. While Kenyan primary school teaching material greatly supports WASH, health education and promotes healthy behaviours, there remains room for improvement for them to serve as entry point to strengthen WASH.

A comparison of our assessment of WASH, health and disease knowledge and the mapping over time on what and in which level of detail pupils study about WASH and disease prevention in Kenyan primary schools (Tables 2–4) with the WASH-related education intervention studies (Table 6) showed the following: Similar to school book education on WASH, that mainly focuses on drinking water aspects, most school-based WASH interventions in Kenya also include drinking water components, covering water treatment (Freeman et al., 2012; Patel et al., 2012), teacher training on maintenance of drinking water facilities (Greene et al., 2012), and student water clubs (O'Reilly et al., 2008). Sanitation interventions in schools in Kenya according to our review are limited, just as sanitation issues were also only discussed in school books (Tables 2–4) in a limited way, and include sanitation behaviour components (Onyango-Ouma et al., 2005) as well as community-led total sanitation methods (Schlegelmilch et al., 2016). Hygiene and related behaviour education, on the other hand, which come up short in our reviewed school books, are well represented in the school-based WASH interventions, covering hand hygiene training (Blanton et al., 2010; Garn et al., 2013; Greene et al., 2012; Patel et al., 2012; Schlegelmilch et al., 2016) including the preparation and use of soapy water, handwashing techniques and handwashing at critical times (Caruso et al., 2014; Pasewaldt et al., 2019; O'Reilly et al., 2008). While vastly addressed among older primary school students (grades 5–7, Tables 2–4), environmental hygiene and waste management for health promotion are included in few intervention studies in Kenya only (Caruso et al., 2014; Onyango-Ouma et al., 2005). Likewise, disease risks that were addressed in detail in Kenyan school books for older pupils (grades 6–8, Tables 2–4), are considered only in very few WASH in school education interventions in Kenya (Onyango-Ouma et al., 2005).

Similar to the school books that cover drinking water-, sanitation-, hygiene-, environmental hygiene-, and disease risk-related themes not in every grade in a similar depth and detail (Tables 2–4), WASH in schools education interventions in Kenya do not either. Such interventions typically target entire schools and several classes, grades and large age ranges at a time. Of our identified WASH in school education intervention studies from Kenya (n = 12), half do not specify grades or age ranges of targeted pupils (Caruso et al., 2014; Freeman et al., 2012; Garn et al., 2013; La Con et al., 2017; Schlegelmilch et al., 2016). The remaining studies cover multiple age ranges and classes, e.g. from 4 through 8 (Greene et al., 2012; O'Reilly et al., 2008; Patel et al., 2012).

4.2. Local school book knowledge versus global WASH education interventions: disconnect or integration?

The WASH in school education intervention studies that were included in our rapid review (Table 6), did not consider using locally available education materials prior to or during the intervention. Only few studies (e.g. Blanton et al., 2010; Karon et al., 2016; Pasewaldt et al., 2019; Vally et al., 2019) evaluated the pupils' knowledge prior to an intervention as a baseline value. They included limited general questions on hand hygiene or water treatment, and only one study included questions specifically asking about the education received in class (Karon et al., 2016). Similarly, none of the intervention studies conducted in Kenya (Table 6) made use of the content available in the primary school science textbooks. A possible reason for this could be deficient WASH infrastructure in areas where the studies were conducted. The interventions thus centred mainly around infrastructural improvements instead of educational dimension. However, where educational materials and traditional lessons-style interventions were used, ad-hoc materials were generally developed and used instead of existing school books.

This observation shows a disconnect and lack of integration of WASH interventions with the local educational systems and characteristics. This may affect the success and the sustainability of any short-term behaviour change interventions. It is important to consider that international organisations rolling out interventions in schools to improve WASH services and train pupils and teachers on safe WASH behaviours are usually responding to a funder-determined need. Such projects are usually implemented in response to requests by, and with involvement of, ministries in the target countries (Garn et al., 2013).

Integrated interventions, addressing education as well as infrastructure needs, are known to be more efficacious than purely infrastructure-focused interventions. Integrated interventions have the potential to last in the long-term, especially when they include approaches that stimulate and/or facilitate the development of attitude and ownership among school users, beyond the mere gain in knowledge or access (WHO Regional Office for Europe, 2016). An integration of WASH education into standardised school curricula ensures that capacity on all topics of relevance is built among all teachers in the system and that all pupils receive the same evidence-based information and lessons developed specifically to favour the development of habits. WHO and UNICEF have long recommended countries and schools to integrate hygiene and WASH education in the national curriculum to ensure sustained and standardised hygiene promotion across all schools. Integration of WASH content in the school curriculum, and training teachers in the course of their standardised college education, is considered most sustainable and efficient. Teaching materials must also be available and accessible for all teachers and pupils at reasonable prices.

4.3. Implications for WASH-related interventions, health programming and messaging

Our analysis suggests that integrating local school book knowledge into WASH education interventions is not common. Considering, on the other hand, that the inclusion of school book content into interventions is no indicator for its effective uptake in practice, and that many WASH interventions are designed as extra-curricular programs to either address gaps in implementation or augment existing curricula, integrating WASH education interventions into existing school dynamics and educational programs holds an evident potential for future projects. This will enable interventions to complement and strengthen the health messaging around WASH and health outcomes.

According to our literature review of WASH in schools education interventions, such interventions typically target entire schools and several classes, grades and large age ranges at a time (Table 6). Considering the breadth and depth – but also gaps - of WASH themes covered in primary school textbooks in Kenya, interventions complementing the school book knowledge with skills-based and participatory education could further support the development of healthy habits among pupils.

In many schools, health or hygiene education are an integral part of the curriculum, although in some countries and regions, the number and breadth of WASH-related topics in schools is limited only to some classes or is not comprehensive (WHO Regional Office for Europe, 2016). Considering and building on such curricula, existing local educational materials and knowledge may facilitate the buy-in and involvement of teachers and school managers in strengthening education and the implementation of improvements. As school books represent in part the knowledge that exists and is provided to children as well as the national priorities, their long-term consideration is key.

To facilitate translation of the information in the school books into life-long skills that empower children to take healthy choices, an interactive, child-centred, participatory approach that engages the pupils in the learning process is of great importance (WHO Regional Office for Europe, 2019). Complementary hygiene or WASH clubs and extracurricular activities which are often considered in interventions (Table 6), have proven effective, although integration of such educational messages and behaviour change measures into school curricular activities is the most sustainable option (WHO Regional Office for Europe, 2019).

New approaches should best be developed around existing knowledge and educational programs to ensure the relevance and uptake at the local or the national level, particularly when implemented in collaboration with educational authorities. Finally, to link WASH education and real hygiene needs and priorities of the communities in which the pupils live, dialogue and participation of local actors is vital.

The thematic gaps/under-representations in books that we identified in our work, namely sanitation, hygiene and menstrual hygiene education, are all high on the development and WASH agenda, and altogether challenging topics which require breaking cultural taboos. Menstrual hygiene has long been recommended by UNICEF and WHO to be better integrated into education to ensure the equity, well-being and dignity of girls in schools (Alexander et al., 2018; Sommer et al., 2016; WHO Regional Office for Europe, 2019). Likewise, healthy hydration as well as proper voiding and defecation and learning about the health of bowel and bladder need to find their space in the books (WHO Regional Office for Europe, 2019). Finally, it is important that teachers' materials stress the importance and suggest approaches of building on child centred and skills-based education to facilitate the development of healthy attitudes and habits (WHO Regional Office for Europe, 2019).

4.4. Contextualisation of our findings with the reality in the local settings, with the unprecedented COVID-19 pandemic and potential of WASH in school education for future public health emergencies

The Kenya Environmental Sanitation and Hygiene Policy mandates the Ministry of Health to promote health and hygiene education as part of the curriculum at nursery, primary and secondary school levels (GoK, 2016). Results of our school book review, however, showed critical gaps in health and hygiene topics in primary school books. Hand hygiene, for example, was only covered in the lower C1 and C2 books and was not developed in the subsequent classes. In the context of the current unprecedented COVID-19 pandemic that the woeld is facing, the lack of sufficient and continuous education on basic hand hygiene - one of the key preventive measures for COVID-19 – is particularly problematic. According to Ritchie et al. (2020), there has been a steady increase in COVID-19-related mortality overall in Kenya, and as schools are reopening across the country, many have reported positive coronavirus cases (Yusuf, 2020).

Besides, there is a theory versus practice gap:

As previously mentioned, combined measures addressing both infrastructure and education are most effective in building healthy WASH habits. Kenya's National School Health Policy states that:

- adequate, safe drinking water points shall be available in each school;
- hand-washing facilities including soap shall be provided in each school within the vicinity of the toilet/latrine (Republic of Kenya, 2009).

The reality in Kenyan schools however, looks different: In 2014, only 43% of public primary schools had access to reliable water from taps or boreholes, 48% of the schools relied on rain or river water and 9.5% did not have access to any water source (Ministry of Education of the Republic of Kenya, 2014). Moreover, 40% of rural schools in Kenya do not have handwashing facilities and more than 80% of those that have them lack soap (Morgan et al., 2017). Thus, even though the pupils gain a comprehensive understanding of water-related health risks and healthy behaviours through the text books, putting this gained knowledge into practice is impeded by inaccessible WASH infrastructure. WASH, economic and social conditions are highly heterogenous in the country, depending on the local realities. Various geographical factors, including environmental conditions like climate, weather, water availability, may play a role in the level of depth that pupils are familiar with WASH infrastructure and healthy practices and in the level of access of these services in schools. WASH conditions, perceptions and cultural preferences or taboos may differ depending on the region and local settings and influence perceptions, practices and knowledge (Anthonj et al., 2016, 2019).

The Kenyan Ministry of Education is responsible for implementing WASH in school programmes and ensuring that all schools are provided with adequate sanitation and hygiene facilities and services, while taking into account the special needs of girls and children with disabilities (GOK, 2016). Although the policy clearly recognised these special needs, none of the science books covered topics related to menstrual hygiene, uncovering also a science education versus WASH policy gap.

In the face of the ongoing COVID-19 pandemic particularly, there is an urgent need to improve and standardise approaches for WASH education and WASH infrastructure in schools. Kenya's Ministry of Education health and safety protocols for reopening of basic education institutions in the context of the pandemic require provision of adequate handwashing points at strategic locations, sufficient clean running water (at least 5 L per person per day), and adequate liquid soap at all hand washing points (Republic of Kenya, 2020). In order to meet these requirements, stakeholders in the education sector need to upgrade WASH facilities, ensure continuous supply of hand washing soap and promote hand hygiene as part of COVID-19 control measures in all schools. Financial aids and efforts to control and reduce the spread of COVID-19 could become the drive to review hygiene curricula and conditions in schools to provide more comprehensive and participatory education programmes, including hands-on practice experience, and ensure access to basic WASH provisions. This would turn in long-term improvements to protect the health of students and teachers and prevent future public health emergencies.

4.5. Limitations

Our school book review was limited to Kenya. Within Kenya, only one region was sampled to access a set of school books which were commonly used in Laikipia County, according to head teachers consulted prior to the data collection and analysis,. These science education books represent what was being taught in the study area at the time of the study (2015–2016). These books do not represent the entirety of school books approved for primary schools in Kenya, many of which are dated well after 2005 (the latest publication date of the books used in the area) (Table 1) (Ministry of Education of the Republic of Kenya, 2015). What is being taught on WASH and environmental health in schools may have changed since the time of the study (Table 1). We acknowledge that this dataset is not be comprehensive – neither for Kenya, nor in general but it does capture topics covered during primary school. While in our research area WASH and disease risk was only taught in science class, elsewhere, it may be covered also by arts or social science text books. Nevertheless, it enabled us to gain a good understanding on what level of WASH- and water-related disease information pupils at primary schools may have access to; how this relates to WASH education interventions; and how this knowledge can be better used for WASH in school programming. Our rapid review of peer-reviewed and grey literature of behaviour change and health messaging interventions on WASH in schools and global policy guidance documents was not aligned with the PRISMA criteria for systematic literature reviews (Moher et al., 2009). Instead, we identified our included publications strategically via "best match", an option on MEDLINE to identify the most relevant peer-reviewed publications, and grey literature via first and most relevant results on Google. This review of literature was restricted to publications that were available in English, which may have affected the studies identified and included.

The intervention studies included from Kenya underlie a reporting bias, with eight out of twelve studies conducted in Nyanza province (Table 6), where a large-scale research and intervention project designed to develop, test and promote improved water treatment and hygiene through schools and communities was rolled out in response to a high diarrhoeal disease burden and poor drinking water access by an international non-governmental organisation over the past years (Caruso et al., 2014; Freeman et al., 2012; Garn et al., 2013; Greene et al., 2012; O'Reilly et al., 2008). Interventions in the larger part of the country is therefore not represented in the review.

Finally, the Kenyan 8-4-4 educational system is currently transitioning towards a competence-based curriculum (CBC). The roll-out of the new curriculum started in 2018 with children who started their primary school in that year, while the other classes continue with the 8-4-4 system as the new curriculum is phased in (Kenya Institute of Curriculum Development, 2017). This CBC will also use different teaching materials. While this could be considered a limitation of our analyses of 8-4-4 system teaching materials, it can also serve as an entry point for continuous improvement of the existing and newly developed teaching materials.

5. Conclusions

Kenyan school books provide pupils with a comprehensive understanding of drinking water sources and management - including quality, treatment, safe storage and water conservation -, water-related disease risks and transmission pathways, and the importance of environmental hygiene, disease prevention and health promotion. While this serves as an entry point for strengthening WASH, there remains room for improvement. Topics in particular needed for improvement with respect to thematic and school years coverage are: personal hygiene and handwashing, menstrual hygiene management, and sanitation education, as well as related health risks and disease exposures. These themes are all high on the development and WASH agenda, which urges for integration in school books, making an effort to breaking cultural taboos (WHO Regional Office for Europe, 2019).

In the Kenyan context, WASH improvements still require efforts addressing both accessibility of services and behaviour change to ensure healthy practices, as though education is included in school books, services do not meet national requirements in many schools. This analysis revealed a gap between policies and implementation as well as a disconnect between local education curricular programmes and WASH education interventions, a missed opportunity for effective and sustainable behaviour change. There is still need for scaling-up improvement interventions and integrating WASH education interventions into existing school dynamics and educational programs holds an evident potential for future projects. Considering and building on existing local educational materials and knowledge may facilitate the buy-in and involvement of teachers and school managers in strengthen education and lasting implementation of improvements with sustained beneficial effects.

Efforts are needed now more than ever, due to the challenges of schools and communities in the context of the current COVID-19 pandemic.

To our best knowledge, this is the first study addressing the coverage of WASH in school books and the consideration of curricular materials for WASH in interventions.

6. Recommendations and future research

Thus we recommend researchers, practicioners and educational authorities to

- integrate WASH in school books and educational curricula, besides WASH infrastructure;
- investigate how the WASH content of schools books is used and taught to the pupils, and to what extent knowledge is retained and multiplies in their households and communities;
- promote implementation research with integrated interventions, that better consider school book knowledge.

We recommend international organisations working on WASH in schools behaviour change interventions to

- work even closer with ministries of education to jointly address gaps in implementation and augment the existing curricula;
- integrate text book knowledge for specific classes and age ranges into their programs, adapted to thegeographical and cultural contexts that pupils live in.

We recommend political decision-makers in WASH and education to consider

- working closer with local communities to help provide suitable and sustainable WASH programming;
- working closer with international organisations providing WASH programming in their country to coordinate and scale-up improvement and monitor integration of interventions with existing materials and local dynamics;
- regularly reviewing evidence-based guidance and recommendations on WASH education and update curricular programs and learning objectives - the current revision of the school system provides a good entry point;
- implementing approaches for child-centred participatory teaching and skills based education to integrate in regular curriculum and school materials;
- reviewing school books and teaching methods to cover the gaps between policy and implementation and ensure comprehensive coverage of topics such as sanitation and hygiene;
- adapting school learning materials to be comprehensive and applicable in all the realities of different regional and environmental conditions (e.g. climate, weather, water availability), settings (urban vs. peri-urban vs. rural vs. informal) and necessities, all of which are important drivers for WASH practices, behaviours, perceptions, and knowledge.

Our future work will continue to build on these analyses, extend them to other countries and contextualise them with the COVID-19 pandemic and the implications, challenge and solutions of WASH in schools.

Acknowledgements

This work was part of the project GlobE Wetlands in East Africa– Reconciling future food production with environmental protection, funded by the German Federal Ministry of Education and Research (grant number FKZ 031A250 D).

References

Adams, J., Bartram, J., Chartier, Y., Sims, J., 2009. Water, Sanitation and Hygiene Standards for Schools in Low-Cost Settings. WHO.

Alexander, K.T., Zulaika, G., Nyothach, E., Oduor, C., Mason, L., Obor, D., Eleveld, A., Laserson, K.F., Phillips-Howard, P.A., 2018. Do water, sanitation and hygiene conditions in primary schools consistently support schoolgirls' menstrual needs? A longitudinal study in rural Western Kenya. Int. J. Environ. Res. Publ. Health 15 (8). https://doi.org/10.3390/ijerph15081682.

Anthonj, C., Rechenburg, A., Kistemann, T., 2016. Water, sanitation and hygiene in Wetlands. The case of Ewaso Narok swamp, Kenya. Int. J. Hyg Environ. Health 219 (7), 606–616. https://doi.org/10.1016/j.ijheh.2016.06.006.

Anthonj, C., Githinji, S., Kistemann, T., 2018. The impact of water on health and ill-health in a sub- Saharan African wetland: Exploring both sides of the coin. Sci. Total Environ. 624, 1411–1420. https://doi.org/10.1016/j.scitotenv.2017.12.232.

Anthonj, C., Diekkrüger, B., Borgemeister, C., Kistemann, T., 2019. Health risk perceptions and local knowledge of water-related infectious disease exposure among Kenyan wetland communities. Int. J. Hyg Environ. Health 222 (1), 34–48. https://doi.org/10.1016/j.ijheh.2018.08.003.

Blanton, E., Ombeki, S., Oluoch, G.O., Mwaki, A., Wannemuehler, K., Quick, R., 2010. Evaluation of the role of school children in the promotion of point-of-use water treatment and handwashing in schools and households—Nyanza province, Western Kenya, 2007. Am. J. Trop. Med. Hyg. 82 (4), 664–671. https://doi.org/10.4269/ajtmh.2010.09-0422.

Bowen, A., Ma, H., Ou, J., Billhimer, W., Long, T., Mintz, E., Hoekstra, R.M., Luby, S., 2007. A cluster-randomized controlled trial evaluating the effect of a handwashing-promotion program in Chinese primary schools. Am. J. Trop. Med. Hyg. 76 (6), 1166–1173.

Bradley, D., 1974. Water supplies: the consequences of change. In: Elliott, K., Knight, J. (Eds.), Human Rights in Health. Ciba Foundation Symposium 23. Associated Publishers, Amsterdam, London, New York, pp. 81–98.

Caruso, B.A., Freeman, M.C., Garn, J.V., Dreibelbis, R., Saboori, S., Muga, R., Rheingans, R., 2014. Assessing the impact of a school-based latrine cleaning and handwashing program on pupil absence in Nyanza Province, Kenya: a cluster-randomized trial. Trop. Med. Int. Health: TM & IH 19 (10), 1185–1197. https://doi.org/10.1111/tmi.12360.

Chard, A.N., Garn, J.V., Chang, H.H., Clasen, T., Freeman, M.C., 2019. Impact of a school-based water, sanitation, and hygiene intervention on school absence, diarrhea, respiratory infection, and soil-transmitted helminths: results from the WASH HELPS cluster-randomized trial. Journal of Global Health 9 (2). https://doi.org/10.7189/jogh.09.020402.

Dreibelbis, R., Freeman, M.C., Greene, L.E., Saboori, S., Rheingans, R., 2014. The impact of school water, sanitation, and hygiene interventions on the health of younger siblings of pupils: a cluster-randomized trial in Kenya. Am. J. Publ. Health 104 (1), e91–e97. https://doi.org/10.2105/AJPH.2013.301412.

D'Anci, K.E., Constant, F., Rosenberg, I.H., 2006. Hydration and cognitive function in children. Nutr. Rev. 64 (10 Pt 1), 457–464. https://doi.org/10.1111/j.1753-4887.2006.tb00176.x.

Embeywa, H.E., Ndungi, B., 2003. Science Matters 2. A Science Course for Primary Schools. Oxford University Press, East Africa Ltd, Nairobi.

Embeywa, H.E., Mbiyu, R.W., Mwakio, C.B., 2004. Science in Action 3. Oxford University Press, East Africa Ltd, Nairobi.

Freeman, M.C., Clasen, T., 2011. Assessing the impact of a school-based safe water intervention on household adoption of point-of-use water treatment practices in southern India. Am. J. Trop. Med. Hyg. 84 (3), 370–378. https://doi.org/10.4269/ajtmh.2011.10-0361.

Freeman, M.C., Greene, L.E., Dreibelbis, R., Saboori, S., Muga, R., Brumback, B., Rheingans, R., 2012. Assessing the impact of a school-based water treatment, hygiene and sanitation programme on pupil absence in Nyanza Province, Kenya: a cluster-randomized trial. Trop. Med. Int. Health 17 (3), 380–391. https://doi.org/10.1111/j.1365-3156.2011.02927.x.

Freeman, M.C., Chard, A.N., Nikolay, B., Garn, J.V., Okoyo, C., Kihara, J., Njenga, S.M., Pullan, R.L., Brooker, S.J., Mwandawiro, C.S., 2015. Associations between school-and household-level water, sanitation and hygiene conditions and soil-transmitted helminth infection among Kenyan school children. Parasites Vectors 8 (412). https://doi.org/10.1186/s13071-015-1024-x.

Garn, Joshua V., Greene, Leslie E., Robert Dreibelbis, Saboori, Shadi, Rheingans, Richard D., Freeman, M.C., 2013. A cluster-randomized trial assessing the impact of school water, sanitation, and hygiene improvements on pupil enrollment and gender parity in enrollment. J. Water, Sanit. Hyg. Dev. 3. https://doi.org/10.2166/washdev.013.217.
Gichuki, 2005a. Primary Science 5. Pupils' Book for Standard Five. Kenya Literature Bureau, Nairobi.
Gichuki, 2005b. Primary Science 6. Pupils' Book for Standard Six. Kenya Literature Bureau, Nairobi.
Gichuki, 2005c. Primary Science 7. Pupils' Book for Standard Seven. Kenya Literature Bureau, Nairobi.
Gichuki, 2005d. Primary Science 8. Pupils' Book for Standard Eight. Kenya Literature Bureau, Nairobi.
Government of the Republic of Kenya GoK, 2016. Kenya Environmental Sanitation and Hygiene Policy 2016–2030. Ministry of Health.
Greene, L.E., Freeman, M.C., Akoko, D., Saboori, S., Moe, C., Rheingans, R., 2012. Impact of a school-based hygiene promotion and sanitation intervention on pupil hand contamination in Western Kenya: a cluster randomized trial. Am. J. Trop. Med. Hyg. 87 (3), 385–393. https://doi.org/10.4269/ajtmh.2012.11-0633.
Jasper, C., Le, T.-T., Bartram, J., 2012. Water and sanitation in schools: a systematic review of the health and educational outcomes. Int. J. Environ. Res. Publ. Health 9 (8), 2772–2787. https://doi.org/10.3390/ijerph9082772.
Joshi, A., Amadi, C., 2013. Impact of water, sanitation, and hygiene interventions on improving health outcomes among school children. Journal of Environmental and Public Health 984626. https://doi.org/10.1155/2013/984626, 2013.
Karaka, J., Nyangasi, L., Githui, M.K., 2005. Understanding Science 4. Pupil's Book 4. Longhorn Publishers Ltd, Nairobi.
Karon, A.J., Cronin, A.A., Cronk, R., Hendrawan, R., 2016. Improving water, sanitation, and hygiene in schools in Indonesia: a cross-sectional assessment on sustaining infrastructural and behavioral interventions. Int. J. Hyg Environ. Health 220 (3), 539–550. https://doi.org/10.1016/j.ijheh.2017.02.001.
Kenya Institute of Curriculum Development, 2017. Basic Education Curriculum Framework.
Kenya Water for Health Organization KWHO, 2018. Safe Water, Better Health! Hygiene Promotion Manual for Primary School. Siemens Stiftung, Munich, p. 2018.
La Con, G., Schilling, K., Harris, J., Person, B., Owuor, M., Ogange, L., Faith, S., Quick, R., 2017. Evaluation of student handwashing practices during a school-based hygiene program in rural Western Kenya, 2007. Int. Q Community Health Educ. 37 (2), 121–128. https://doi.org/10.1177/0272684X17701263.
Lau, C.H., Springston, E.E., Sohn, Min-Woong, Mason, I., Gadola, E., Damitz, M., Gupta, R.S., 2012. Hand hygiene instruction decreases illness-related absenteeism in elementary schools: a prospective cohort study. BMC Pediatr. 12, 52. Doi: 1471-2431/12/52.
Lukacz, E.S., Sampselle, C., Gray, M., MacDiarmid, S., Rosenberg, M., Ellsworth, P., Palmer, M.H., 2011. A healthy bladder: a consensus statement. Int. J. Clin. Pract. 65 (10), 1026–1036. https://doi.org/10.1111/j.1742-1241.2011.02763.x.
McMichael, C., 2019. Water, sanitation and hygiene (WASH) in schools in low-income countries: a review of evidence of impact. Int. J. Environ. Res. Publ. Health 16 (3), 359. https://doi.org/10.3390/ijerph16030359.
Merhej, R., 2019. Dehydration and cognition: an understated relation. International Journal of Health Governance 24 (1), 19–30. https://doi.org/10.1108/IJHG-10-2018-0056.
Ministry of Education, Science and Technology of the Republic of Kenya, 2014. Basic Education Statistical Booklet.
Ministry of Education, Science and Technology of the Republic of Kenya, 2015. Approved list of school textbooks and other instructional materials for ECDE, primary schools and teach training colleges. Revised 14th Edition Volume One. Available at. https://icta.go.ke/pdf/approvered-list-of-books-primary-2015.pdf.
Ministry of Education of the Republic of Kenya, 1984. 8-4-4 System of Education. Government Printer, Nairobi.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151 (4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.
UNICEF Sri Lanka, n.y. Manual on Hygiene Promotion in Schools - Teachers Guidebook. UNICEF.
Morgan, C., Bowling, M., Bartram, J., Kayser, G.L., 2017. Water, sanitation, and hygiene in schools: status and implications of low coverage in Ethiopia, Kenya, Mozambique, Rwanda, Uganda, and Zambia. Int. J. Hyg Environ. Health 220 (6), 950–959. https://doi.org/10.1016/j.ijheh.2017.03.015.
Muckelbauer, R., Libuda, L., Clausen, K., Toschke, A.M., Reinehr, T., Kersting, M., 2009. Promotion and provision of drinking water in schools for overweight prevention: randomized, controlled cluster trial. Pediatrics 123 (4), e661–667. https://doi.org/10.1542/peds.2008-2186.
Munn, Z., Tufanaru, C., Lockwood, C., Stern, C., McAneney, H., Barker, T.H., 2020. Rinse-free hand wash for reducing absenteeism among preschool and school children. Cochrane Database Syst. Rev. 4 (4), CD012566. https://doi.org/10.1002/14651858.CD012566.pub2.
Neal, D., Vujcic, J., Hernandez, O., Wood, W., 2015. The Science of Habit: Creating Disruptive and Sticky Behavior Change in Handwashing Behavior. United States Agency for International Development/WASHplus, Washington DC.
Ojwang, A., K'opiyo, F., 2003. Science Matters 1. A Science Course for Primary Schools. East African Publishers Ltd, Nairobi.
Onyango-Ouma, W., Aagaard-Hansen, J., Jensen, B.B., 2005. The potential of schoolchildren as health change agents in rural Western Kenya. Soc. Sci. Med. 61 (8), 1711–1722. https://doi.org/10.1016/j.socscimed.2005.03.041.
O'Reilly, C.E., Freeman, M.C., Ravani, M., Migele, J., Mwaki, A., Ayalo, M., Ombeki, S., Hoekstra, R.M., Quick, R., 2008. The impact of a school-based safe water and hygiene programme on knowledge and practices of students and their parents: Nyanza province, Western Kenya, 2006. Epidemiol. Infect. 136 (1), 80–91. https://doi.org/10.1017/S0950268807008060.
Pasewaldt, S.E., Baller, S.L., Blackstone, S.R., Malenke, L.B., 2019. Impact of a hand hygiene curriculum and group handwashing station at two primary schools in East africa. Int. Q Community Health Educ. 39 (3), 175–187. https://doi.org/10.1177/0272684X18819968.
Patel, M.K., Harris, J.R., Juliao, P., Nygren, B., Were, V., Kola, S., Sadumah, I., Faith, S. H., Otieno, R., Obure, A., Hoekstra, R.M., Quick, R., 2012. Impact of a hygiene curriculum and the installation of simple handwashing and drinking water stations in rural Kenyan primary schools on student health and hygiene practices. Am. J. Trop. Med. Hyg. 87 (4), 594–601. https://doi.org/10.4269/ajtmh.2012.11-0494.
Republic of Kenya, 2009. National School Health Policy. Ministry of Public Health & Sanitation and Ministry of Education. Ministry of Health, Nairobi.
Republic of Kenya, 2020. Kenya Basic Education Covid-19 Emergency Response Plan. May 2020.
Ritchie, H., Ortiz-Ospina, E., Beltekian, E., Mathieu, E., Hasell, J., Macdonald, B., Giattino, C., Roser, M., 2020. Kenya: coronavirus pandemic country profile. Available at. https://ourworldindata.org/coronavirus/country/kenya?country=~KEN. accessed on 17.11.2020.
Saboori, S., Greene, L.E., Moe, C.L., Freeman, M.C., Caruso, B.A., Akoko, D., Rheingans, R.D., 2013. Impact of regular soap provision to primary schools on hand washing and E. coli hand contamination among pupils in Nyanza Province, Kenya: a cluster-randomized trial. Am. J. Trop. Med. Hyg. 89 (4), 698–708. https://doi.org/10.4269/ajtmh.12-0387.
Schlegelmilch, M.P., Lakhani, A., Saunders, L.D., Jhangri, G.S., 2016. Evaluation of water, sanitation and hygiene program outcomes shows knowledge-behavior gaps in Coast Province, Kenya. The Pan African Medical Journal 23, 145. https://doi.org/10.11604/pamj.2016.23.145.7546.
Schulte, J.M., Williams, L., Asghar, J., Dang, Thi, Bedwell, S., Guerrero, K., Hamaker, D., Stonecipher, S., Zoretic, J., Chow, C., 2012. How we didn't clean up until we washed our hands: shigellosis in an elementary and middle school in North Texas. South. Med. J. 105 (1), 1–4. https://doi.org/10.1097/SMJ.0b013e31823c411e.
Shrestha, A., Schindler, C., Odermatt, P., et al., 2020. Nutritional and health status of children 15 months after integrated school garden, nutrition, and water, sanitation and hygiene interventions: a cluster-randomised controlled trial in Nepal. BMC Publ. Health 20 (158). https://doi.org/10.1186/s12889-019-8027-z.
Sommer, M., Caruso, B.A., Sahin, M., Calderon, T., Cavill, S., Mahon, T., Phillips-Howard, P.A., 2016. A time for global action: addressing girls' menstrual hygiene management needs in schools. PLoS Med. 13 (2), e1001962 https://doi.org/10.1371/journal.pmed.1001962.
Sumpter, C., Torondel, B., 2013. A systematic review of the health and social effects of menstrual hygiene management. PloS One 8 (4), e62004. https://doi.org/10.1371/journal.pone.0062004.
Talaat, M., Afifi, S., Dueger, E., El-Ashry, N., Marfin, A., Kandeel, A., Mohareb, E., El-Sayed, N., 2011. Effects of hand hygiene campaigns on incidence of laboratory-confirmed influenza and absenteeism in schoolchildren, Cairo, Egypt. Emerg. Infect. Dis. 17 (4), 619–625. https://doi.org/10.3201/eid1704.101353.
Tidwell, J.B., Gopalakrishnan, A., Unni, A., Sheth, E., Daryanani, A., Singh, S., Sidibe, M., 2020. Impact of a teacher-led school handwashing program on children's handwashing with soap at school and home in Bihar, India. PloS One 15 (2). https://doi.org/10.1371/journal.pone.0229655.
UNICEF, 2012. Water, Sanitation and Hygiene (WASH) in Schools: a Companion to the Child Friendly Schools Manual.
Vally, H., McMichael, C., Doherty, C., Li, X., Guevarra, G., Tobias, P., 2019. The impact of a school-based water, sanitation and hygiene intervention on knowledge, practices, and diarrhoea rates in the Philippines. Int. J. Environ. Res. Publ. Health 16 (21). https://doi.org/10.3390/ijerph16214056.
WASHplus, 2015. A Teacher's Guide to Integrating WASH in Schools. USAID/WASHplus Project, Washington DC.
WHO, 2019. Safer Water, Better Health. World Health Organization, Geneva. Available at. http://www.who.int/water_sanitation_health/publications/safer-water-better-health/en/. accessed on 17.11.2020.
WHO and UNICEF, n.y. Household data – Kenya – service levels. Available at: www.washdata.org [accessed on 17.11.2020] .
WHO, UNICEF, 2018a. Core Questions on Drinking Water, Sanitation and Hygiene for Household Surveys: 2018 Update.
WHO, UNICEF, 2018b. Core Questions and Indicators for Monitoring WASH in Schools in the Sustainable Development Goals.
WHO and UNICEF b, n.y. School data - SDG regions - Sub-Saharan Africa - 2019 - Service Levels. Available at: www.washdata.org [accessed on 17.11.2020] .
WHO Regional Office for Europe WHO/Europe, 2016. The Situation of Water, Sanitation and Hygiene in Schools in the Pan-European Region.
WHO Regional Office for Europe WHO/Europe, 2019. Improving Health and Learning through Better Water, Sanitation and Hygiene in Schools. An Information Package for School Staff.
Yusuf, M., 2020. Kenyan union says COVID-19 in schools putting teachers at risk. Published on VOA news on 2 November 2020. Available at. https://www.voanews.com/africa/kenyan-union-says-covid-19-schools-putting-teachers-risk, 2020. accessed on 17.11.2020.

International Journal of Hygiene and Environmental Health 235 (2021) 113755

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

Prenatal exposure to pyrethroid and organophosphate insecticides and language development at age 20–36 months among children in the Odense Child Cohort

Helle Raun Andersen [a,*], Louise Dalsager [a], Inge Kjær Jensen [a], Clara Amalie Gade Timmermann [a], Trine Staak Olesen [a], Fabio Trecca [b], Flemming Nielsen [a], Greet Schoeters [a,c], Henriette Boye Kyhl [d,e], Philippe Grandjean [a,h], Niels Bilenberg [f,g], Dorthe Bleses [b], Tina Kold Jensen [a,d]

[a] Department of Public Health, University of Southern Denmark, Odense, J.B. Winsløws Vej 17A, 5000, Odense, Denmark
[b] School of Communication and Culture – Trygfondens Centre for Child Research, Aarhus University, Denmark
[c] Environmental Risk and Health, Flemish Institute for Technological Research (VITO), University of Antwerp, Belgium
[d] Hans Christian Andersen Children's Hospital, Odense University Hospital, Sdr. Boulevard 29, 5000, Odense C, Denmark
[e] Odense Patient Data Explorative Network (OPEN), Odense, Denmark
[f] Department of Child and Adolescent Mental Health Odense, Mental Health Services in the Region of Southern Denmark, Odense, Denmark
[g] Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark
[h] Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

ARTICLE INFO

Keywords:
Pyrethroids
Organophosphates
Insecticides
Language development
Prenatal exposure
Birth cohort

ABSTRACT

Background: Prenatal exposure to organophosphate and pyrethroid insecticides has been associated with impaired neurodevelopment. Few longitudinal studies have investigated associations with early language development in populations with mainly low dietary exposure.

Objective: To investigate associations between biomarkers of maternal gestational exposure to organophosphate and pyrethroid insecticides and the child's language development at age 20–36 months in the prospective Odense Child Cohort.

Methods: Metabolites of organophosphate and pyrethroid insecticides were measured in maternal urine samples collected at gestational week 28. Language development was assessed among 755 singletons at age 20–36 months using the Vocabulary and Complexity scores of the MacArthur-Bates Communicative Development Inventories, standardized into age and sex specific percentile scores according to a Danish reference study. Multiple logistic regression models were used to estimate the odds of scoring below the 15th percentile scores in relation to maternal urinary insecticide metabolite concentrations after adjustment for confounders.

Results: The generic pyrethroid metabolite 3-phenoxybenzoic acid (3-PBA) and the chlorpyrifos metabolite 3,5,6-trichloro-2-pyridinol (TCPY) were detectable in more than 90% of the urine samples analyzed. Likewise, 82.2% had detectable concentrations of diethyl phosphates (DE) and 58.4% of dimethyl phosphates (DM), both of which are common metabolites of organophosphate insecticides. None of the metabolites was associated with higher odds of delayed results below the 15th percentile language scores. In contrast, reduced probability for scoring below the 15th percentile Vocabulary score was seen for the highest tertile of 3-PBA in boys and for the upper tertile of TCPY and DE in girls.

Conclusion: In this prospective cohort, with predominantly dietary insecticide exposure, we found no evidence that gestational exposure to organophosphate or pyrethroid insecticides adversely affected early language development in the children. The observed indication of a positive effect of insecticides on language development may be explained by residual and unmeasured confounding from socioeconomic factors and dietary habits. Follow-up of these children should include assessment of more complex cognitive functions in later childhood, as well as associations with their own postnatal insecticide exposure.

* Corresponding author.
E-mail address: hrandersen@health.sdu.dk (H.R. Andersen).

https://doi.org/10.1016/j.ijheh.2021.113755
Received 29 October 2020; Received in revised form 8 April 2021; Accepted 16 April 2021
Available online 5 May 2021

1. Introduction

Pyrethroids and organophosphates are two major classes of insecticides. Due to their acute toxicity, organophosphates were prohibited for indoor use as biocides in both the EU and the U.S. more than a decade ago, but they are still in use in agriculture, although not in Denmark. The organophosphate chlorpyrifos has for decades been one of the most widely used insecticides in agriculture worldwide (Eaton et al., 2008). However, from February 2020 chlorpyrifos and chlorpyrifos-methyl are no longer approved for use in the EU because of concern for developmental neurotoxicity and genotoxicity (EFSA, 2019), although exposure will still occur from produce originating outside the EU. While the use of organophosphates is decreasing, the use of pyrethroids is increasing as substitutes for organophosphates in biocide products and also, to some degree, in agriculture. Human biomonitoring studies suggest widespread exposure to both organophosphates and pyrethroids, also among pregnant women, based on high detection rates of organophosphate and pyrethroid metabolites in urine from pregnant women (Dalsager et al., 2019; Dereumeaux et al., 2018; Sokoloff et al., 2016; Yolton et al., 2013). The insecticides are able to cross the human placenta, thus causing fetal exposures as well, as illustrated by detection of these substances in umbilical cord blood and meconium (Berton et al., 2014; Rauh et al., 2011; Silver et al., 2016).

Most insecticides target the nervous system of insects. Due to similarities in neurochemistry they also have neurotoxic properties in mammals (Abreu-Villaca and Levin, 2017). The developing brain is particularly vulnerable to neurotoxicants, and major windows of vulnerability occur in utero and during early postnatal life (Grandjean and Landrigan, 2014; Rice and Barone, 2000). Several birth cohort studies have reported associations between prenatal exposure to organophosphate and pyrethroid insecticides and impaired childhood neurodevelopment, i.e., cognitive and/or behavioural deficits, although mainly in populations with elevated exposure from occupational or indoor use or from residence in proximity to pesticide-treated agricultural areas (Andersen et al., 2015; Bouchard et al., 2011; Eskenazi et al., 2018; Gunier et al., 2017; Liu et al., 2016; Llop et al., 2013; Rauh et al., 2006; Shelton et al., 2014). A few studies have investigated neurodevelopmental outcomes associated with prenatal organophosphate and pyrethroid exposure in populations considered to be mainly exposed through ingestion of residues in foods (Cartier et al., 2016; Dalsager et al., 2019; Donauer et al., 2016; Jusko et al., 2019; Nkinsa et al., 2020; Tanner et al., 2020; Viel et al., 2015, 2017; Yolton et al., 2013), and the results have been equivocal perhaps in part due to differences in study design. Some of these studies reported on language development as part of IQ assessment in relation to prenatal pyrethroid (Viel et al., 2015) or organophosphate exposure (Cartier et al., 2016; Donauer et al., 2016; Nkinsa et al., 2020). Only one study assessed language skills as a separate outcome and found no relationships between prenatal organophosphate exposure and language development at age 4 years (Donauer et al., 2016). Since early language development can be seen as an important indicator of early neurodevelopment and is correlated to subsequent intelligence and later academic achievement (Bleses et al., 2016; Flensborg-Madsen and Mortensen, 2018; Liao et al., 2015), we aimed to investigate associations between maternal urinary concentrations of organophosphate and pyrethroid metabolites in pregnancy and language development in the children up to age 36 months among mother-child pairs from the Odense Child Cohort (OCC). From this cohort, mainly exposed to organophosphate and pyrethroid insecticides from dietary sources, we have previously reported associations between maternal pyrethroid exposure in pregnancy and more attention deficit hyperactivity disorder (ADHD) symptoms among the children at age 2–4 years (Dalsager et al., 2019).

2. Methods

2.1. Study population

We used data from the prospective Odense Child Cohort (OCC). From January 2010 to December 2012, all pregnant women living in the Municipality of Odense in Denmark were invited to participate in the cohort either at a voluntary meeting about ultrasound, at the first antenatal visit or at the first ultrasound examination at Odense University Hospital at gestational week (GW) 8–16 weeks. A total of 2874 women agreed to participate. Of these women, 52 dropped out before giving birth, 112 had a miscarriages or stillbirth, and 210 withdrew from the cohort after giving birth, leaving 2500 families (2549 children) in the cohort. Participating women tended to be older, more often non-smokers, better educated, had fewer children, and were more often of Danish origin than non-participants (Kyhl et al., 2015). At both inclusion and GW 28 the women filled in questionnaires about health, lifestyle and social factors. All children were examined at birth, and information on birth weight (g), gestational age of the child as well as maternal pre pregnancy BMI (kg/m^2), smoking status, parity and age at delivery was obtained from obstetric records. Maternal country of birth was obtained from a register in the Municipality of Odense. For 228 women, information on education level was missing from the questionnaires. In these cases, information on occupation obtained from obstetric records was used to estimate educational level whenever possible. Information on duration of breastfeeding was obtained from questionnaires answered by the parents when the child was 18 months and supplemented with data from a sub-project on breastfeeding reported via text messages (Bruun et al., 2016). For this study, we excluded twins and children of mothers not born in Denmark (to ensure that all mothers were native Danish speakers).

2.2. Pyrethroid and organophosphate insecticide exposure measurements

Spot urine samples were collected in GW 28 (median 28.7, range 26.4–34.0) after overnight fasting and before 9.30 a.m. The samples were stored at −80 °C until analyses. Out of 1603 women who provided a urine sample, a subset of 1207 samples (including mothers of twins and mothers not born in Denmark) was analyzed for the specific metabolite of chlorpyrifos/chlorpyrifos-methyl, TCPY (3,5,6-trichloro-2-pyridinol), and the generic pyrethroid metabolite, 3-PBA (3-phenoxybenzoic acid), representing exposure to most pyrethroids. In addition, we measured some more specific pyrethroid metabolites: *cis*- and *trans*-DCCA (3-(2,2-dichlorovinyl)-2,2-dimethylcyclopropane-1-carboxylic acid) representing exposure to the *cis*- and *trans*-isomers of permethrin, cypermethrin, and cyfluthrin; 4-F-3PBA (4-fluoro-3-phenoxybenzoic acid) which is a specific metabolite of cyfluthrin; and *cis*-DBCA (*cis*-3-(2,2-dibromovinyl)-2,2-dimethylcyclopropane-1-carboxylic acid) being a specific metabolite of deltamethrin. The analyses were performed by high performance liquid chromatography and tandem mass spectrometry (LC-MS/MS) as previously described (Dalsager et al., 2019). Calibration curves, solvent blanks, and quality control samples were included in each batch of samples. In-house made quality control (QC) samples (low and high level) with all compounds were made in diluted urine (1:3). The accuracy of the analysis was also controlled by participation in the German External Quality Assessment Scheme (G-EQUAS), (for 3-PBA, trans-DCCA, and cis-DBCA). Excess sample material from this program was also used as QC samples. The accuracy of the analysis for all the compounds ranged from 92.7 to 103.3%. The between batch variation (CV%) ranged from 3.5 to 17.6%. The limit of detection (LOD) for the compounds were: 0.3 μg/L for TCPY; 0.03 μg/L for 3-PBA, 0.2 μg/L for 4-F-3PBA; 0.4 μg/L for *trans*-DCCA, and 0.5 μg/L for *cis*-DCCA and *cis*-DBCA. Spectrophotometric determination of creatinine concentrations was conducted on a Konelab 20 Clinical Chemistry Analyzer, using a commercial kit (Thermo, Vantaa, Finland).

Out of the 1207 urine samples, a subset of 564 samples were also analyzed for concentrations of six non-specific dialkyl phosphate (DAP) metabolites of organophosphate insecticides at the Flemish Institute for Technological Research NV (VITO), Belgium, using solid phase extraction (SPE) followed by Ultra Performance Liquid Chromatography-tandem mass spectrometry (UPLC-MS/MS) as previously described (Dalsager et al., 2018). The DAP metabolites comprise three dimethyl (DM) phosphate metabolites (dimethyl phosphate (DMP), dimethyl thiophosphate (DMTP), dimethyl dithiophosphate (DMDTP)) and three diethyl (DE) phosphate metabolites (diethyl phosphate (DEP), diethyl thiophosphate (DETP), and diethyl dithiophosphate (DEDTP)). LODs were 1.49 μg/L for DMP and 0.30 μg/L for DMTP, DMDTP, DEP, DEDTP, and DEDTP. The metabolite concentrations were converted to their molar concentrations and summed to yield total DM, total DE, and total DAP concentrations (nmol/L).

The first 200 urine samples analyzed for insecticides were selected randomly from the 1603 available urine samples, whereas the remaining samples were selected based upon availability of information from baseline questionnaires, birth records and clinical examinations of the child.

2.3. Language development assessment

Language development was assessed using the validated Danish adaption of the MacArthur-Bates Communicative Development Inventories (MB-CDI) (Bleses et al., 2008; Fenson et al., 2000) which is a parent report instrument used for assessing children's early language development (e.g., vocabulary and grammar) at different age steps. Every third month from the age of 20–36 months, the parents completed an electronic version of the MB-CDI: *Words and Sentence* questionnaire regarding their child's current language skills. The questionnaire contained seven subscales, of which two (Vocabulary and Complexity) were included in the present study to represent important domains of language development during the age span of interest (Fenson et al., 2000). The Vocabulary subscale focuses on the child's use of commonly used words for toddlers (lexical development), whereas the Complexity subscale focuses on use of grammar and syntactic complex sentences (morphosyntactic development). In the questionnaire, parents marked all words and sentences their child produced prior to or at the time of reply.

Although parents repeatedly filled in questionnaires during the 16-month period, we included only one questionnaire per category per child because some parents only filled in few questionaries while others filled in many, and at different ages during the 16 month period. For Vocabulary, we used the first questionnaire filled out to create a productive Vocabular summary score (number of words reported to be expressed), whereas for Complexity, we used the first questionnaire after the age of 26 months in girls and 30 months in boys for the purpose of creating a summary score (number of mastered Complexity items). These age limits were applied as more than 15% of children in the reference population had a score of zero before the ages chosen (Bleses et al., 2008). For each child, the summary scores were converted into an age and sex specific percentile score by comparing the obtained score with the sex and age specific score for the Danish reference population (Bleses et al., 2008). A percentile-score ≤ 15 was considered as delayed language development.

2.4. Ethics

The project was performed in accordance with the Helsinki Declaration II, with written informed consent, and approved by the Regional Scientific Ethical Committees for Southern Denmark (Project-ID S-20090130) and the Danish Data Protection Agency (j.no. 2016-231-0188).

2.5. Statistical analysis

Insecticide metabolites with detection frequencies below 20% were dichotomized (≥LOD vs < LOD). For all other metabolites, values below the limit of detection (LOD) were substituted by the metabolite specific LOD/$\sqrt{2}$. Urinary concentrations were expressed as volume-based (μg or nmol/L) or creatinine-based (μg or nmol/g creatinine) and reported as medians and 25–75 percentiles. To investigate possible effects of combined exposure to chlorpyrifos and pyrethroids, we also created a variable indicating whether the combined exposure was low (both metabolites below the median), medium (one above the median) or high (both above the median). Correlations between urinary pesticide metabolite concentrations and between Vocabulary and Complexity percentile scores were evaluated by Spearman correlation. The language score percentiles were reported as medians and 25–75 percentiles, and the language scores were further dichotomized (below/above the 15th percentile) to indicate potentially delayed language development (Bleses et al., 2008).

Associations between maternal insecticide metabolite concentrations and the age and sex specific percentile language scores were estimated by logistic regression models. For these analyses, the volume-based metabolite concentrations were divided into tertiles and the continuous concentrations were $\ln_2$ transformed to obtain a normal distribution and ln-transformed creatinine (g/L) was included as a covariate. The odds ratio (OR) for having a language score below the 15th percentile across increasing tertiles of maternal urinary insecticide metabolites concentrations as well as for doubling the concentrations using continuous $\ln_2$-transformed variables were calculated. Potential confounders were identified as variables shown in the literature to be associated with either pesticide exposure or child language development. Of the covariates considered *a priori*, we had information on maternal age, pre-pregnancy body mass index (BMI), educational level, smoking during pregnancy, parity, preterm birth (child born before GW 37), child sex and birth weight and duration of breastfeeding. Breastfeeding information was missing for 11% of the women and a category consisting of these women was added to the breastfeeding variable to avoid exclusion of these data. Differences in characteristics between participants and non-participants as well as differences in exposure and outcome variables in relation to maternal and child characteristics were tested using non-parametric Kruskal Wallis/Mann Whitney tests and chi-square tests. Covariates included in the adjusted analyses were maternal education (as the best available estimate of socio-economic status), breastfeeding and child sex. We did not adjust for child age, as language scores were categorized into age specific percentile scores according to age of the reference population. Although the language scores were also sex standardized, sex was included to account for potential sex-differences in vulnerability towards pesticides, and interaction between exposure and child sex was investigated by including an interaction term in the adjusted analyses. Further, logistic regression analyses were performed for boys and girls separately. In a sensitivity analysis, we extended the adjusted model to include also maternal pre-pregnancy BMI, as this characteristic was associated with both exposure and outcome in our data set. The results are presented as OR with 95% confidence intervals, and p-values < 0.05 were considered statistically significant.

3. Results

Of 2217 singletons of mothers born in Denmark, MB-CDI questionnaires were completed for 1,360, and 1109 had insecticides measured in the urine (Fig. 1). A total of 755 mother-child pairs, with data on both vocabulary outcome and exposure, were included in this study. The participants were more often non-smokers, had fewer preterm deliveries, longer duration of breastfeeding, and the children had higher birth weight compared to non-participating singletons of mothers born in Denmark (N = 1462). There was no difference in maternal age, BMI,

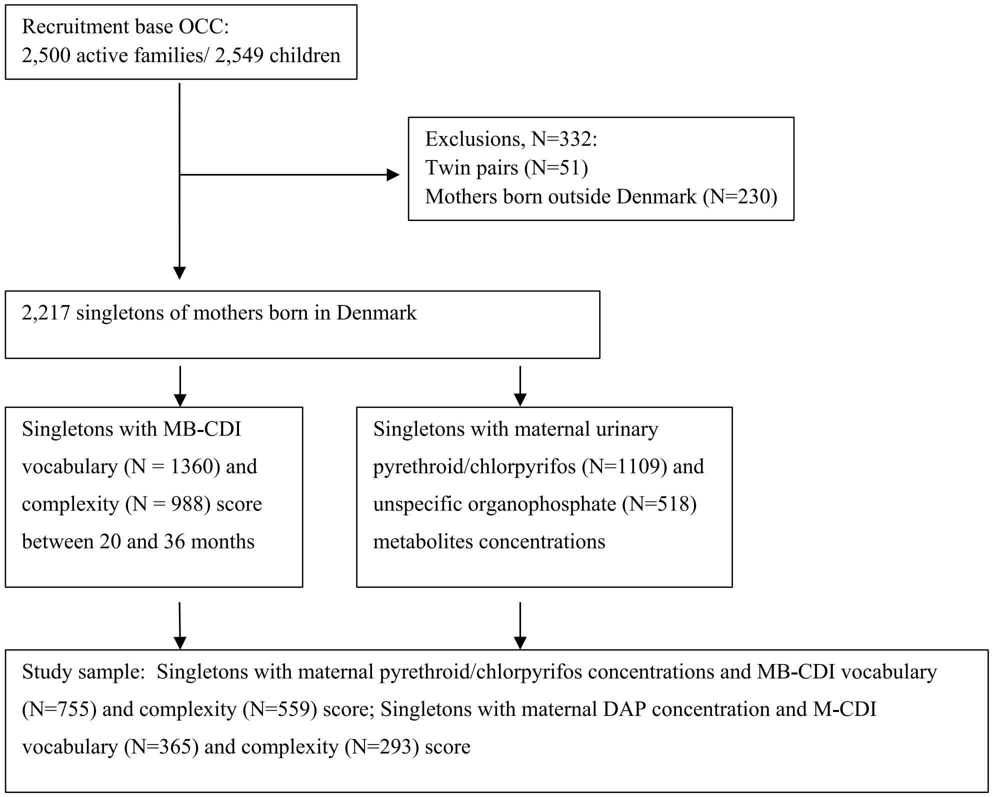


Fig. 1. Flow diagram of study population based on mother-child pairs from the Odense Child Cohort (OCC) with maternal insecticide metabolite concentration and MacArthur-Bates Communicative Development Inventories (MB-CDI) language assessment in the offspring.

level of education or parity (data not shown).

Most of the women had detectable concentrations of TCPY (92.3%) and 3-PBA (94.3%). The median concentrations were 1.73 µg/g creatinine and 0.24 µg/g creatinine, respectively (Table 1). The specific pyrethroid metabolites were detectable only in 12.2% (*trans*-DCCA), 2.6% (*cis*-DCCA), 3.2% (*cis*-DBCA) while 4-F-3PBA was detected only in a single urine sample (data not shown). For the DAPs, 82.2% of the 365 samples analyzed had detectable concentrations of at least one DE metabolite and 58.4% of at least one DM metabolite. The median molar sum of the six metabolites (DAPs) was 59.6 nmol/g creatinine (Table 1) and 23.7 and 30.9 nmol/g creatinine for DE and DM, respectively (data not shown). Further, 3-PBA was weakly correlated with TCPY (Spearman $r_s = 0.17$). Also, as expected, TCPY and DAP were correlated ($r_s = 0.34$). The TCPY concentration was higher among women with longer education and women with low pre-pregnancy BMI and among women who did not breastfeed (Table 1). None of the selected participant characteristics were statistically significantly associated with 3-PBA or DAP concentrations.

Due to incomplete responses, the Complexity percentile score could be calculated only for 559 of the 755 children with maternal 3-PBA and TCPY concentrations, and for 293 of the 365 with maternal DAPs (Fig. 1). The MB-CDI questionnaires on Vocabulary were completed at a median age of 21 months (range 20–36) for both boys and girls. For the Complexity subscale, the median age was 31 months (range 30–36) for boys and 27 months (range 26–36) for girls. The Vocabulary percentile score median was 50 (55 for boys and 50 for girls), and the median Complexity percentile score was 45 (40 for boys and 45 for girls) (Table 1). Vocabulary and Complexity percentile scores were highly correlated ($r_{s\ Boys} = 0.69$ and $r_{s\ Girls} = 0.67$, $p < 0.0001$). The language percentile scores were higher among children who were breastfed the longest, and the Complexity percentile score was higher among children of women with longer education and women with lower pre-pregnancy BMI (Table 1).

The ORs for scoring below the age and sex standardized 15th percentile on the MB-CDI Vocabulary and Complexity subscales according to maternal urinary insecticide metabolite concentrations are shown in Tables 2 and 3. Generally, none of the insecticide metabolites were associated with higher odds of scoring below the 15th percentile language scores in either the crude or the adjusted analyses. For the highest tertile of 3-PBA, a reduced OR for having a Vocabulary score below the 15th percentile was seen in the adjusted analyses (OR = 0.51 (95% CI: 0.27; 0.94), p = 0.03). This association in a direction opposite to the hypothesis was driven by boys (OR = 0.30 (0.12; 0.72), p = 0.007) while no association was observed for girls. The p-value for sex-exposure interaction was 0.05. For the upper tertile of TCPY and DE, again a reduced OR for scoring below the 15th percentile was found for girls (OR for TCPY = 0.38 (0.15, 0.96), p = 0.04 and OR for DE = 0.21 (0.05, 0.86), p = 0.03) but not for boys or both sexes combined. However, the associations described were not apparent when 3-PBA, TCPY and DE were included as continuous variables. No statistically significant associations were seen between maternal insecticide metabolites and odds of scoring below the 15th percentile on the MB-DI Complexity subscale. Combined maternal 3-PBA and TCPY concentrations were not associated with altered odds of scoring below the 15th percentile language scores, but a tendency of a reduced OR among boys was seen for the two categories with 3-PBA above the median (Table 2). Further adjustment for maternal pre-pregnancy BMI did not affect these results (data not shown).

4. Discussion

In this prospective study of mother-child pairs from the Odense Child Cohort, we did not observe any associations between low-level gestational exposure to pyrethroid and organophosphate insecticides and impaired language development in the children at age 20–36 months. The observed suggestion of a possibly protective effect of pyrethroids

Table 1
Maternal urinary concentrations (Median (M) and 25th −75th percentiles (25–75)) of 3-PBA, TCPY, and DAP in gestational week 28 and M (25–75) of MB-CDI Vocabulary and Complexity score at age 1.5–3 years according to maternal and child characteristics.

		3-PBA µg/g creatinine	TCPY µg/g creatinine	DAP nmol/g creatinine		Vocabulary percentile score		Complexity percentile score	
Maternal characteristics	N	M (25–75)	M (25–75)	N	M (25–75)	N	M (25–75)	N	M (25–75)
All	755	0.24 (0.14–0.45)	1.73 (1.05–3.02)	365	59.6 (37.0–104.8)	755	50 (25–75)	559	45 (15–70)
Maternal age at delivery (years)									
<26	81	0.22 (0.12–0.47)	1.64 (1.06–2.72)	40	69.0 (43.8–117.9)	81	50 (25–75)	61	45 (20–65)
26–34	512	0.24 (0.14–0.45)	1.72 (1.02–3.09)	242	58.4 (34.3–105.6)	512	55 (30–75)	376	45 (20–75)
>34	162	0.23 (0.12–0.45)	1.80 (1.11–3.05)	83	59.8 (44.0–89.8)	162	45 (19–75)	122	35 (15–66)
Pre-pregnancy BMI (kg/m^2)									
<18.5	21	0.16 (0.09–0.37)	2.13 (1.56–3.23)*	12	57.2 (46.2–121.6)	21	60 (40–85)	17	55 (40–75)*
18.5–24.9	450	0.23 (0.14–0.43)	1.78 (1.10–3.31)*	217	64.4 (41.8–107.7)	450	55 (25–75)	319	50 (20–75)*
≥25	284	0.25 (0.14–0.50)	1.63 (0.91–2.73)*	136	55.6 (29.9–89.8)	284	50 (20–75)	223	40 (15–70)*
Smoking									
No	730	0.24 (0.14–0.45)	1.73 (1.05–3.01)	352	59.8 (37.4–104.8)	730	55 (25–75)	539	45 (15–75)
Yes	25	0.29 (0.16–0.48)	1.61 (0.90–3.50)	13	43.1 (20.4–103.7)	25	55 (28–78)	20	35 (15–56)
Education level[a]									
Low	199	0.24 (0.13–0.42)	1.52 (0.93–2.51)*	111	59.8 (37.7–104.4)	199	45 (20–70)	151	35 (20–60)*
Intermediate	389	0.23 (0.14–0.48)	1.78 (1.08–3.26)*	174	59.1 (35.9–105.6)	389	55 (25–75)	280	45 (15–70)*
High	160	0.24 (0.13–0.47)	1.88 (1.17–3.32)*	75	65.8 (38.8–103.7)	160	55 (31–85)	123	55 (25–85)*
missing	7	0.37 (0.21–0.41)	0.78 (0.53–1.34)	5	26.6 (12.9–108.1)	7	85 (55–98)	5	70 (35–98)
Parity									
Nulliparous	427	0.24 (0.13–0.46)	1.73 (1.02–3.03)	197	59.6 (37.6–104.1)	427	55 (25–75)	312	45 (20–70)
Multiparous	328	0.23 (0.14–0.44)	1.69 (1.05–3.02)	168	59.6 (36.0–105.3)	328	53 (25–75)	248	40 (15–75)
p-value		0.46	0.99		0.93		0.45		0.93
Child characteristics									
Sex									
Boy	390	0.24 (0.14–0.46)	1.72 (1.00–3.01)	188	60.6 (38.0–102.9)	390	55 (25–75)	262	40 (15–75)
Girl	365	0.24 (0.14–0.44)	1.73 (1.09–3.05)	177	59.5 (35.1–106.0)	365	50 (25–75)	297	45 (20–70)
Preterm (<37 weeks)									
no	743	0.24 (0.14–0.45)	1.70 (1.02–3.00)	360	59.4 (37.0–105.3)	743	55 (25–75)	549	45 (15–75)
yes	12	0.31 (0.21–0.44)	2.74 (1.89–3.19)	5	62.7 (59.6–75.9)	12	40 (25–84)	10	30 (14–51)
Birth weight (g)									
<3000	84	0.24 (0.12–0.48)	1.68 (0.89–2.74)	39	59.8 (26.6–86.6)	84	45 (16–70)	67	35 (15–65)
3000-4000	521	0.24 (0.14–0.45)	1.77 (1.08–3.23)	253	60.8 (39.7–105.6)	521	55 (25–75)	379	45 (20–75)
≥4000	150	0.24 (0.14–0.43)	1.64 (1.00–2.60)	73	55.6 (34.4–112.7)	150	50 (25–75)	113	40 (15–73)
Total breastfeeding (weeks)									
0	19	0.29 (0.14–0.90)	2.05 (1.25–3.27)*	9	73.63 (35.6–123.2)	19	55 (15–70)*	14	20 (10–36)*
1–25	229	0.24 (0.13–0.42)	1.55 (0.83–2.72)*	100	57.3 (30.9–98.6)	229	45 (20–70)*	161	35 (15–60)*
>25	426	0.24 (0.14–0.48)	1.79 (1.09–3.18)*	199	60.0 (38.3–103.2)	426	60 (30–85)*	330	50 (20–80)*
Missing	81	0.21 (0.13–0.37)	1.91 (1.23–3.03)	57	57.8 (40.6–122.6)	81	40 (20–65)	54	50 (25–65)

*p < 0.05 in Kruskal Wallis test.
[a] Low: High school and vocational education or less, Intermediate: High school + 1–4 years, High: High school + 5 years or more.

and chlorpyrifos on language development is considered biologically implausible and may be explained by residual and unmeasured confounding from socioeconomic factors and dietary habits. Fruit, vegetables, and cereals often contain pesticide residues (Lu et al., 2008; Morgan, 2012; Vanacker et al., 2020) but at the same time a high intake of these commodities may be beneficial, thus potentially outweighing detrimental effects of low organophosphate and pyrethroid exposure, as has also been suggested elsewhere (Donauer et al., 2016; Yolton et al., 2013). Unfortunately, we did not have any information on maternal food habits. Further, due to multiple testing, the few statistically significant associations may well be chance findings.

Regarding organophosphates, our findings are in accordance with two other studies with low maternal DAP-concentrations comparable to our study. In an urban/suburban birth cohort (HOME) in Ohio, no associations between maternal DAP concentrations (geometric mean (GM): 73.7 nmol/g creatinine) and child cognition, including language development, at 1–5 years of age were found (Donauer et al., 2016). Further, a study based on a French birth cohort (PELAGIE) reported no consistent pattern of associations between maternal DAP (median:11.3 nmol/L) and cognitive scores, inclusive of verbal comprehension scores, in the offspring at age 6 years (Cartier et al., 2016). As in our study, a few positive associations were reported (Cartier et al., 2016; Donauer et al., 2016), likely due to similar residual confounding. In contrast, a recent study from the Canadian MIREC cohort, with a slightly higher median maternal DAP concentration (86.1 nmol/L), found higher concentrations of DE to be associated with poorer verbal IQ scores at 3–4 years of age among boys, but not girls (Nkinsa et al., 2020). Interestingly, we found a reduced probability for scoring below the 15th percentile Vocabulary score related to high maternal DE and TCPY concentrations among girls only. Assuming this association is due to residual negative confounding, it may suggest that boys' benefit from these confounding factors is hampered by prenatal exposure to chlorpyrifos and other diethyl organophosphates. Higher vulnerability among boys was found in two studies from China, in which maternal DE concentrations were associated with neurodevelopmental delay at 24 months of age in boys only, but mainly in the adaptive (Liu et al., 2016) or social (Wang et al., 2017) domains, while language development was not significantly affected. The exposure levels were considerably higher than in our study, with median DAP concentrations of 295.8 nmol/L and 352.7 nmol/g creatinine, respectively, but similar to the median of 310 nmol/g creatinine reported in the Generation R study from the Netherlands (Jusko et al., 2019). In the Generation R study, lower nonverbal IQ was associated with maternal DM and total DAP concentrations in late pregnancy (above GW 25), though the results were inconsistent across repeated urine sampling periods (<18, 18–25 >25 GW) in pregnancy. Language development was not included in this study, and the associations were not modified by child sex. Maternal DAP concentrations in the Generation R study were markedly higher than reported from other urban/suburban studies in the EU or US and also higher than in the CHAMACOS study (GM: 114.9 nmol/L) from an agricultural community in California (Bouchard et al., 2011). From the CHAMACOS study, adverse associations with neurodevelopment at 2 years of age (Eskenazi

Table 2
Crude and adjusted odds ratio (OR) and 95% confidence intervals (95% CI) for scoring below the age and sex standardized 15th percentile on the MB-CDI Vocabulary and Complexity subscales according to maternal urinary concentrations of pyrethroid and chlorpyrifos metabolites (categorical and continuous after ln_2-transformation) among children from the Odense Child Cohort.

	MB-CDI Vocabulary score below 15th percentile				MB-CDI Complexity score below 15th percentile			
	All (N = 755)	All (N = 748)	Boys (N = 389)	Girls (N = 359)	All (N = 559)	All (N = 554)	Boys (N = 261)	Girls (N = 293)
	Crude	Adjusted[a]	Adjusted[a]	Adjusted[a]	Crude	Adjusted[a]	Adjusted[a]	Adjusted[a]
3-PBA								
1the tertile (<0.13 μg/L)	Reference				Reference			
2nd tertile (0.13–0.36 μg/L)	1.21 (0.74; 1.98)	0.99 (0.59; 1.66)	0.71 (0.34; 1.48)	1.38 (0.64; 2.95)	1.17 (0.70; 1.97)	1.06 (0.60; 1.83)	0.53 (0.23; 1.18)	2.04 (0.93; 4.49)
3rd tertile (>0.36 μg/L)	0.71 (0.41; 1.23)	0.51 (0.27; 0.94)	0.30 (0.12; 0.72)	0.88 (0.36; 2.17)	0.78 (0.45; 1.34)	0.66 (0.35; 1.23)	0.49 (0.20; 1.19)	0.95 (0.40; 2.38)
p-trend	0.24	0.03	0.007	0.80	0.37	0.19	0.12	0.89
Continuous[b]	0.94 (0.83; 1.05)	0.87 (0.76; 1.00)	0.81 (0.67; 0.98)	0.93 (0.75; 1.15)	0.92 (0.82; 1.05)	0.89 (0.77; 1.02)	0.91 (0.74; 1.10)	0.87 (0.70; 1.07)
Trans-DCCA								
<LOD (0.4 μg/L)	Reference				Reference			
≥LOD	1.12 (0.64; 2.19)	1.09 (0.58; 2.04)	1.22 (0.53; 2.81)	0.96 (0.35; 2.66)	0.80 (0.39; 1.63)	0.74 (0.35; 1.53)	0.83 (0.31; 2.23)	0.63 (0.21; 1.93)
TCPy								
1the tertile (<1.13 μg/L)	Reference				Reference			
2nd tertile (1.13–2.50 μg/L)	1.20 (0.72; 2.00)	1.06 (0.61; 1.82)	1.37 (0.61; 3.10)	0.88 (0.42; 1.85)	0.98 (0.58; 1.64)	0.87 (0.50; 1.53)	1.06 (0.48; 2.35)	0.78 (0.36; 1.72)
3rd tertile (>2.50 μg/L)	1.00 (0.60; 1.70)	0.79 (0.44; 1.43)	1.49 (0.66; 3.40)	0.38 (0.15; 0.96)	0.72 (0.42; 1.24)	0.61 (0.33; 1.12)	0.65 (0.27; 1.58)	0.59 (0.24; 1.38)
p-trend	0.99	0.41	0.35	0.04	0.23	0.10	0.33	0.22
Continuous[b]	1.00 (0.87; 1.15)	0.93 (0.79; 1.10)	1.05 (0.84; 1.31)	0.83 (0.65; 1.05)	0.90 (0.78; 1.04)	0.84 (0.71; 1.00)	0.87 (0.68; 1.12)	0.82 (0.63; 1.06)
Combined 3-PBA and TCPy								
Both≤p50	Reference				Reference			
3-PBA > p50 & TCPy≤p50	1.15 (0.63; 2.09)	0.94 (0.50; 1.79)	0.42 (0.14; 1.21)	1.69 (0.72; 3.96)	1.06 (0.56; 1.98)	0.97 (0.50; 1.90)	0.88 (0.32; 2.40)	1.12 (0.44; 2.83)
3-PBA≤p50 & TCPy > p50	1.11 (0.60; 2.04)	0.91 (0.48; 1.73)	1.23 (0.51; 2.99)	0.67 (0.25; 1.79)	1.14 (0.62; 2.10)	1.03 (0.54; 1.98)	0.94 (0.34; 2.61)	1.06 (0.43; 2.58)
Both > p50	1.15 (0.63; 2.09)	0.94 (0.50; 1.79)	0.42 (0.14; 1.21)	1.69 (0.72; 3.96)	1.06 (0.56; 1.98)	0.97 (0.50; 1.90)	0.88 (0.32; 2.40)	1.12 (0.44; 2.83)

P50: median (0.21 μg/L for 3-PBA and 1.69 μg/L for TCPy).
[a] Adjusted for creatinine (ln-transformed), maternal education, weeks of breastfeeding, and child sex (for all children).
[b] For doubling the concentration.

et al., 2007) and with full-scale IQ, including poorer scores for processing speed and verbal comprehension, at 7 years of age were reported (Bouchard et al., 2011).

For chlorpyrifos, dose-response associations have been reported between the concentration in umbilical cord blood and delayed psychomotor development in 3-year-old children (Rauh et al., 2006) and cognitive development at age 7 years (Rauh et al., 2011) in a New York inner-city birth cohort established before the ban of chlorpyrifos for residential use. However, only two previous studies have included TCPY in maternal urine as biomarker for chlorpyrifos exposure (Eskenazi et al., 2007; Guo et al., 2019). Both studies reported higher TCPY concentrations than in our study, with median TCPY concentrations of 3.54 and 5.39 μg/L, respectively. Neither of these studies found adverse associations with cognitive development in the children, although associations with maternal DAP concentrations were seen. Thus, adverse neurodevelopment seems more strongly correlated to prenatal exposure to organophosphates overall than to chlorpyrifos specifically, which is reasonable assuming additive effects of organophosphates on neurodevelopment (EFSA et al., 2019).

Regarding pyrethroids, our findings are in accordance with some other studies with low maternal pyrethroid exposure. In the French birth cohort, PELAGIE, verbal comprehension and working memory scores at age 6 years were not associated with maternal pyrethroid metabolites (Viel et al., 2015), but some associations with behavioural problems were reported (Viel et al., 2017) as also seen in the OCC (Dalsager et al., 2019) and in a study from New York (Furlong et al., 2017). However, a recent study from Sweden found associations between maternal 3-PBA and reduced IQ at age 7 years (Tanner et al., 2020), and a study from Mexico City reported association between maternal 3-PBA and delayed mental development at 2 years of age (Watkins et al., 2016). A direct comparison of the urinary pyrethroid metabolite concentrations is hampered by differences in analytical methods and LODs but reported maternal 3-PBA concentrations were below 0.25 μg/L in all the studies. Cognitive deficits have also been associated with higher maternal exposure levels from residential proximity to agricultural pyrethroid use (Gunier et al., 2017; Xue et al., 2013) or high indoor use for malaria control (Eskenazi et al., 2018). Maternal median 3-PBA concentrations reported from these studies were between 0.70 μg/L (Eskenazi et al., 2018) and 2.24 μg/L (Xue et al., 2013).

The neurotoxic properties of organophosphates and pyrethroids are well-known, and several animal studies have demonstrated long-lasting alterations in brain function, including deficits in learning and memory, related to prenatal or early postnatal exposure to organophosphates and pyrethroids (Abreu-Villaca and Levin, 2017; Aldridge et al., 2005; Laugeray et al., 2017). Nonetheless, the present study failed to identify adverse associations between maternal exposure to these insecticides and early language development. Differences in instruments used to assess cognitive function as well as child age at examination might explain some of the inconsistency across the studies. Furthermore, variability and differences in maternal exposure levels, heterogeneity in exposure routes, exposure patterns and individual susceptibility are other likely explanations for the disparate findings. Dermal and inhalation exposures might lead to higher fetal exposure to the parent compounds because maternal hepatic first-pass metabolism is avoided. Interestingly, in the PELAGIE cohort, no association between maternal metabolite concentrations and impaired cognitive development was seen, despite the children's own urinary concentration of DE being associated with lower working memory scores (Cartier et al., 2016) and

Table 3
Crude and adjusted odds ratio (OR) and 95% confidence intervals (95% CI) for scoring below the age and sex standardized 15th percentile on the MB-CDI Vocabulary and Complexity subscales according to maternal urinary concentrations of dialkyl phosphate metabolites of organophosphate insecticides (categorical and continuous after ln_2-transformation) among children from the Odense Child Cohort.

	MB-CDI Vocabulary score below 15th percentile				MB-CDI Complexity score below 15th percentile			
	All (N = 365)	All (N = 360)	Boys (N = 187)	Girls (N = 173)	All (N = 293)	All (N = 290)	Boys (N = 140)	Girls (N = 150)
	Crude	Adjusted[a]	Adjusted[a]	Adjusted[a]	Crude	Adjusted[a]	Adjusted[a]	Adjusted[a]
Diethyl phosphates (DE)								
1the tertile (<13.6 nmol/L)	Reference				Reference			
2nd tertile (13.6–35.9 nmol/L)	1.17 (0.61; 2.23)	1.04 (0.51; 2.11)	1.39 (0.47; 4.09)	0.68 (0.24; 1.90)	0.94 (0.47; 1.88)	0.78 (0.36; 1.67)	1.39 (0.45; 4.26)	0.37 (0.12; 1.15)
3rd tertile (>35.9 nmol/L)	0.62 (0.30; 1.28)	0.52 (0.23; 1.20)	0.82 (0.24; 2.78)	0.21 (0.05; 0.86)	0.71 (0.34; 1.46)	0.57 (0.25; 1.33)	0.93 (0.26; 3.33)	0.33 (0.10; 1.10)
p-trend	0.21	0.12	0.56	0.03	0.35	0.19	0.93	0.07
Continuous[b]	0.89 (0.74; 1.07)	0.83 (0.67; 1.03)	0.87 (0.64; 1.18)	0.76 (0.55; 1.05)	0.99 (0.82; 1.20)	0.96 (0.76; 1.20)	1.14 (0.82; 1.59)	0.80 (0.57; 1.12)
Dimethyl phosphates (DM)								
1the tertile (<14.3 nmol/L)	Reference				Reference			
2nd tertile (14.3–41.1 nmol/L)	0.50 (0.25; 1.01)	0.46 (0.22; 0.96)	0.39 (0.14; 1.07)	0.73 (0.23; 2.30)	1.25 (0.59; 2.65)	1.21 (0.55; 2.68)	1.38 (0.46; 4.14)	1.26 (0.37; 4.27)
3rd tertile (>41.1 nmol/L)	0.75 (0.39; 1.44)	0.68 (0.34; 1.37)	0.52 (0.20; 1.35)	0.94 (0.31; 2.89)	1.53 (0.75; 3.15)	1.44 (0.67; 3.09)	1.21 (0.41; 3.62)	2.11 (0.67; 6.71)
p-trend	0.37	0.28	0.20	0.91	0.24	0.31	0.78	0.19
Continuous[b]	1.00 (0.84; 1.22)	1.00 (0.82; 1.22)	0.96 (0.74; 1.25)	1.02 (0.75; 1.40)	1.10 (0.91; 1.33)	1.09 (0.88; 1.34)	1.10 (0.83; 1.48)	1.08 (0.79; 1.47)
Dialkyl phosphates (DAP)								
1the tertile (<34.5 nmol/L)	Reference				Reference			
2nd tertile (34.5–89.0 nmol/L)	0.61 (0.30; 1.23)	0.57 (0.27; 1.21)	0.78 (0.27; 2.29)	0.36 (0.11; 1.13)	1.56 (0.74; 3.29)	1.57 (0.71; 3.48)	2.11 (0.65; 6.89)	1.24 (0.41; 3.78)
3rd tertile (>89.0 nmol/L)	0.90 (0.47; 1.72)	0.81 (0.39; 1.70)	0.98 (0.34; 2.86)	0.55 (0.18; 1.75)	1.45 (0.69; 3.04)	1.45 (0.63; 3.37)	1.67 (0.48; 5.76)	1.43 (0.42; 4.89)
p-trend	0.73	0.63	0.95	0.24	0.36	0.43	0.53	0.57
Continuous[b]	0.96 (0.78; 1,19)	0.93 (0.73; 1.19)	0.97 (0.70; 1.36)	0.84 (0.58; 1.23)	1.09 (0.88; 1.36)	1.07 (0.84; 1.40)	1.25 (0.86; 1.83)	0.96 (0.66; 1.39)

[a] Adjusted for creatinine (ln-transformed), maternal education, weeks of breastfeeding, and child sex (for all children).
[b] For doubling the concentration.

their 3-PBA concentration being negatively associated with verbal comprehension scores (Viel et al., 2015). Similar associations between child urinary concentrations of metabolites from pyrethroids and/or organophosphates and neurodevelopment have been reported in several other recent studies (Guo et al., 2019; Oulhote and Bouchard, 2013; Wagner-Schuman et al., 2015; Wang et al., 2016). These findings could indicate that the child's own dietary exposure to the parent compounds may be more hazardous for neurodevelopment than maternal exposure in pregnancy. Unfortunately, we did not have information on postnatal exposure at this young age, but it will be interesting to include the children's own metabolite concentrations when more complex cognitive functions can be assessed later in childhood.

In our study, language development was reported by parents, which can be seen as both a strength and a limitation, since it is independent of the current well-being of the child and test setting, but at the same time dependent on the subjective perception of the parents. Thus, some measurement error is likely, but since the women were unaware of their pesticide exposure status when responding to the questionnaire, this is expected to be non-differential and may contribute to our null findings. The major strengths of the present study are the longitudinal design, the high number of mother-child pairs enrolled, and the access to a large reference population of more than 3500 Danish children for language development (Bleses et al., 2008). The MB-CDI instrument applied at preschool age has shown good predictive validity in regard to language development (Bleses et al., 2008; Fenson et al., 2000). However, the young age of the children when assessing their language skills could limit the ability to detect subtle effects that might manifest later in childhood because of cascading developmental processes (Rice and Barone, 2000) and better opportunity for examination of more complex language skills and other cognitive functions. Among other limitations was that only 43% of the eligible pregnant women (N = 2874) were recruited and included in the OCC, and these women were better educated and more likely to be non-smokers, older and nulliparous than non-participants (Kyhl 2015). We cannot rule out the possibility that this enrollment bias might attenuate potential neurotoxic impacts of the insecticides due to more stimulating environments and healthier lifestyles. Furthermore, it was only possible to analyze urine samples and obtain information about child language development from around a third of the eligible women in the OCC. Participating families may differ from non-participating families leading to selection bias. They were probably more likely to have a healthier lifestyle (e.g., higher intake of organic food and thereby lower pesticide exposure) and exposure contrast may therefore be smaller reducing the possibility to find an association. In addition, families of children with poor language development may have been more or less likely to complete the MB-CDI questionnaire. However, participants were unaware of their urinary pesticide concentration when they responded to the questionnaire and it is therefore unlikely to have affected our findings. An additional concern is the fact that the insecticide metabolite concentrations were determined in a single spot urine sample. Pyrethroids and organophosphates are rapidly metabolized and excreted from the body within few days. When combining this with within-subject variability, misclassification is likely and may bias again the findings toward the null (Spaan et al., 2015). Under the assumption of stable dietary habits, the exposure variation in our study might be lower than in studies with additional exposure from agricultural or indoor use of insecticides. Furthermore, the urinary metabolites may not entirely reflect exposure of the parent compounds but also intake of metabolites preformed in food items. Thus, food samples have been found to contain TCPY and DAP residues (Morgan et al., 2011; Zhang et al., 2008), while the content of pyrethroid degradates was reported to be very low (Morgan et al., 2018). Finally, we cannot exclude the possibility of residual and unmeasured confounding by socioeconomic, dietary, lifestyle or behavioural factors. Accordingly, the women with high education had highest urinary concentrations of TCPY, and the same tendency was seen for DAPs. Most likely organophosphate and pyrethroid exposure was primarily through intake of fruit, vegetables, and cereals (Lu et al., 2008; Morgan, 2012; Vanacker et al., 2020), which also contain essential vitamins and antioxidants with beneficial effects that may diminish potential harmful effects of insecticides as has been reported for other environmental exposures (Choi et al., 2014; Kim et al., 2011). Finally, we did not have information on parental IQ, although we did take into account maternal education. Given these considerations, the study findings suggest that no detectable pesticide-associated deficits in language development occurred below age 3 years, but extended follow-up will be needed to obtain more conclusive data.

5. Conclusion

In this relatively large prospective study, more than 90% of the pregnant women had measurable, though low, concentrations of organophosphate and pyrethroid metabolites in urine. We found no adverse associations between maternal insecticide exposure and early language development in their children at age 20–36 months among these rather privileged families with mainly dietary exposure to insecticides. The unexpected suggestion of a positive effect of pyrethroids and chlorpyrifos on language development could likely be explained by residual and unmeasured confounding from socioeconomic factors and dietary habits that may mask early adverse effects of insecticides. Follow-up of these children should include assessment of more complex language skills and cognitive functions in later childhood as well as associations with their own postnatal insecticide exposure.

Funding

The work was funded by Odense University Hospital, the Region of Southern Denmark, The Municipality of Odense, The University of Southern Denmark, Odense Patient data Exploratory Network (OPEN), The Danish Environmental Protection Agency's Pesticide Research Programme (MST-667-00164), Novo Nordisk Foundation (grant nr. NNF17OC0029404 and NNF10OC0058266), the Danish Council for Independent Research (4004-00352B_FSS), and the EU Horizon 2020 Research and Innovation Programme, HBM4EU (grant no. 733032).

Declaration of competing interest

None.

Acknowledgements

The families in OCC are acknowledged for their participation and commitment to the study. The technicians at Hans Christian Andersen's Children's Hospital and at the Department of Environmental Medicine are acknowledged for their careful examination of the children and analyses of pesticide metabolites in urine samples, respectively.

References

Abreu-Villaca, Y., Levin, E.D., 2017. Developmental neurotoxicity of succeeding generations of insecticides. Environ. Int. 99, 55–77.
Aldridge, J.E., Levin, E.D., Seidler, F.J., Slotkin, T.A., 2005. Developmental exposure of rats to chlorpyrifos leads to behavioral alterations in adulthood, involving serotonergic mechanisms and resembling animal models of depression. Environ. Health Perspect. 113, 527–531.
Andersen, H.R., Debes, F., Wohlfahrt-Veje, C., Murata, K., Grandjean, P., 2015. Occupational pesticide exposure in early pregnancy associated with sex-specific neurobehavioral deficits in the children at school age. Neurotoxicol. Teratol. 47, 1–9.
Berton, T., Mayhoub, F., Chardon, K., Duca, R.C., Lestremau, F., Bach, V., Tack, K., 2014. Development of an analytical strategy based on LC-MS/MS for the measurement of different classes of pesticides and theirs metabolites in meconium: application and characterisation of foetal exposure in France. Environ. Res. 132, 311–320.
Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T.O., Basboll, H., 2008. The Danish Communicative Developmental Inventories: validity and main developmental trends. J. Child Lang. 35, 651–669.
Bleses, D., Makransky, G., Dale, P.S., Hojen, A., Ari, B.A., 2016. Early productive vocabulary predicts academic achievement 10 years later. Appl. Psycholinguist. 37, 1461–1476.
Bouchard, M.F., Chevrier, J., Harley, K.G., Kogut, K., Vedar, M., Calderon, N., Trujillo, C., Johnson, C., Bradman, A., Barr, D.B., Eskenazi, B., 2011. Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. Environ. Health Perspect. 119, 1189–1195.
Bruun, S., Wedderkopp, N., Molgaard, C., Kyhl, H.B., Zachariassen, G., Husby, S., 2016. Using text messaging to obtain weekly data on infant feeding in a Danish birth cohort resulted in high participation rates. Acta Paediatr. 105, 648–654.
Cartier, C., Warembourg, C., Le Maner-Idrissi, G., Lacroix, A., Rouget, F., Monfort, C., Limon, G., Durand, G., Saint-Amour, D., Cordier, S., Chevrier, C., 2016. Organophosphate insecticide metabolites in prenatal and childhood urine samples and intelligence scores at 6 Years of age: results from the mother-child PELAGIE cohort (France). Environ. Health Perspect. 124, 674–680.
Choi, A.L., Mogensen, U.B., Bjerve, K.S., Debes, F., Weihe, P., Grandjean, P., Budtz-Jorgensen, E., 2014. Negative confounding by essential fatty acids in methylmercury neurotoxicity associations. Neurotoxicol. Teratol. 42, 85–92.
Dalsager, L., Christensen, L.E., Kongsholm, M.G., Kyhl, H.B., Nielsen, F., Schoeters, G., Jensen, T.K., Andersen, H.R., 2018. Associations of maternal exposure to organophosphate and pyrethroid insecticides and the herbicide 2,4-D with birth outcomes and anogenital distance at 3 months in the Odense Child Cohort. Reprod. Toxicol. 76, 53–62.
Dalsager, L., Fage-Larsen, B., Bilenberg, N., Jensen, T.K., Nielsen, F., Kyhl, H.B., Grandjean, P., Andersen, H.R., 2019. Maternal urinary concentrations of pyrethroid and chlorpyrifos metabolites and attention deficit hyperactivity disorder (ADHD) symptoms in 2-4-year-old children from the Odense Child Cohort. Environ. Res. 176, 108533.
Dereumeaux, C., Saoudi, A., Goria, S., Wagner, V., De Crouy-Chanel, P., Pecheux, M., Berat, B., Zaros, C., Guldner, L., 2018. Urinary levels of pyrethroid pesticides and determinants in pregnant French women from the Elfe cohort. Environ. Int. 119, 89–99.
Donauer, S., Altaye, M., Xu, Y., Sucharew, H., Succop, P., Calafat, A.M., Khoury, J.C., Lanphear, B., Yolton, K., 2016. An observational study to evaluate associations between low-level gestational exposure to organophosphate pesticides and cognition during early childhood. Am. J. Epidemiol. 184, 410–418.
Eaton, D.L., Daroff, R.B., Autrup, H., Bridges, J., Buffler, P., Costa, L.G., Coyle, J., McKhann, G., Mobley, W.C., Nadel, L., Neubert, D., Schulte-Hermann, R., Spencer, P. S., 2008. Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment. Crit. Rev. Toxicol. 38 (Suppl. 2), 1–125.
EFSA, 2019. Statement on the available outcomes of the human health assessment in the context of the pesticides peer review of the active substance chlorpyrifos. EFSA J. 17 e05809.
EFSA, Crivellente, F., Hart, A., Hernandez-Jerez, A.F., Hougaard Bennekou, S., Pedersen, R., Terron, A., Wolterink, G., Mohimont, L., 2019. Establishment of cumulative assessment groups of pesticides for their effects on the nervous system. EFSA J. 17 e05800.
Eskenazi, B., Marks, A.R., Bradman, A., Harley, K., Barr, D.B., Johnson, C., Morga, N., Jewell, N.P., 2007. Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. Environ. Health Perspect. 115, 792–798.
Eskenazi, B., An, S., Rauch, S.A., Coker, E.S., Maphula, A., Obida, M., Crause, M., Kogut, K.R., Bornman, R., Chevrier, J., 2018. Prenatal exposure to DDT and pyrethroids for malaria control and child neurodevelopment: the VHEMBE cohort, South Africa. Environ. Health Perspect. 126, 047004.
Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J.S., Thal, D., 2000. Measuring variability in early child language: don't shoot the messenger. Child Dev. 71, 323–328.
Flensborg-Madsen, T., Mortensen, E.L., 2018. Associations of early developmental milestones with adult intelligence. Child Dev. 89, 638–648.
Furlong, M.A., Barr, D.B., Wolff, M.S., Engel, S.M., 2017. Prenatal exposure to pyrethroid pesticides and childhood behavior and executive functioning. Neurotoxicology 62, 231–238.
Grandjean, P., Landrigan, P.J., 2014. Neurobehavioural effects of developmental toxicity. Lancet Neurol. 13, 330–338.
Gunier, R.B., Bradman, A., Harley, K.G., Kogut, K., Eskenazi, B., 2017. Prenatal residential proximity to agricultural pesticide use and IQ in 7-year-old children. Environ. Health Perspect. 125, 057002.
Guo, J., Zhang, J., Wu, C., Lv, S., Lu, D., Qi, X., Jiang, S., Feng, C., Yu, H., Liang, W., Chang, X., Zhang, Y., Xu, H., Cao, Y., Wang, G., Zhou, Z., 2019. Associations of prenatal and childhood chlorpyrifos exposure with Neurodevelopment of 3-year-old children. Environ. Pollut. 251, 538–546.
Jusko, T.A., van den Dries, M.A., Pronk, A., Shaw, P.A., Guxens, M., Spaan, S., Jaddoe, V. W., Tiemeier, H., Longnecker, M.P., 2019. Organophosphate pesticide metabolite concentrations in urine during pregnancy and offspring nonverbal IQ at age 6 years. Environ. Health Perspect. 127, 17007.
Kim, H., Hwang, J.Y., Ha, E.H., Park, H., Ha, M., Lee, S.H., Hong, Y.C., Chang, N., 2011. Fruit and vegetable intake influences the association between exposure to polycyclic aromatic hydrocarbons and a marker of oxidative stress in pregnant women. Eur. J. Clin. Nutr. 65, 1118–1125.

Kyhl, H.B., Jensen, T.K., Barington, T., Buhl, S., Norberg, L.A., Jorgensen, J.S., Jensen, D. F., Christesen, H.T., Lamont, R.F., Husby, S., 2015. The odense child cohort: aims, design, and cohort profile. Paediatr. Perinat. Epidemiol. 29, 250–258.

Laugeray, A., Herzine, A., Perche, O., Richard, O., Montecot-Dubourg, C., Menuet, A., Mazaud-Guittot, S., Lesne, L., Jegou, B., Mortaud, S., 2017. In utero and lactational exposure to low-doses of the pyrethroid insecticide cypermethrin leads to neurodevelopmental defects in male mice-An ethological and transcriptomic study. PloS One 12 e0184475.

Liao, S.F., Liu, J.C., Hsu, C.L., Chang, M.Y., Chang, T.M., Cheng, H., 2015. Cognitive development in children with language impairment, and correlation between language and intelligence development in kindergarten children with developmental delay. J. Child Neurol. 30, 42–47.

Liu, P., Wu, C., Chang, X., Qi, X., Zheng, M., Zhou, Z., 2016. Adverse associations of both prenatal and postnatal exposure to organophosphorous pesticides with infant neurodevelopment in an agricultural area of Jiangsu province, China. Environ. Health Perspect. 124, 1637–1643.

Llop, S., Julvez, J., Fernandez-Somoano, A., Santa Marina, L., Vizcaino, E., Iniguez, C., Lertxundi, N., Gascon, M., Rebagliato, M., Ballester, F., 2013. Prenatal and postnatal insecticide use and infant neuropsychological development in a multicenter birth cohort study. Environ. Int. 59, 175–182.

Lu, C., Barr, D.B., Pearson, M.A., Waller, L.A., 2008. Dietary intake and its contribution to longitudinal organophosphorus pesticide exposure in urban/suburban children. Environ. Health Perspect. 116, 537–542.

Morgan, M.K., 2012. Children's exposures to pyrethroid insecticides at home: a review of data collected in published exposure measurement studies conducted in the United States. Int. J. Environ. Res. Publ. Health 9, 2964–2985.

Morgan, M.K., Sheldon, L.S., Jones, P.A., Croghan, C.W., Chuang, J.C., Wilson, N.K., 2011. The reliability of using urinary biomarkers to estimate children's exposures to chlorpyrifos and diazinon. J. Expo. Sci. Environ. Epidemiol. 21, 280–290.

Morgan, M.K., MacMillan, D.K., Zehr, D., Sobus, J.R., 2018. Pyrethroid insecticides and their environmental degradates in repeated duplicate-diet solid food samples of 50 adults. J. Expo. Sci. Environ. Epidemiol. 28, 40–45.

Nkinsa, P.N., Muckle, G., Ayotte, P., Lanphear, B.P., Arbuckle, T.E., Fraser, W.D., Bouchard, M.F., 2020. Organophosphate pesticides exposure during fetal development and IQ scores in 3 and 4-year old Canadian children. Environ. Res. 190, 110023.

Oulhote, Y., Bouchard, M.F., 2013. Urinary metabolites of organophosphate and pyrethroid pesticides and behavioral problems in Canadian children. Environ. Health Perspect. 121, 1378–1384.

Rauh, V.A., Garfinkel, R., Perera, F.P., Andrews, H.F., Hoepner, L., Barr, D.B., Whitehead, R., Tang, D., Whyatt, R.W., 2006. Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. Pediatrics 118, E1845–E1859.

Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D.B., Whyatt, R., 2011. Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ. Health Perspect. 119, 1196–1201.

Rice, D., Barone Jr., S., 2000. Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. Environ. Health Perspect. 108 (Suppl. 3), 511–533.

Shelton, J.F., Geraghty, E.M., Tancredi, D.J., Delwiche, L.D., Schmidt, R.J., Ritz, B., Hansen, R.L., Hertz-Picciotto, I., 2014. Neurodevelopmental disorders and prenatal residential proximity to agricultural pesticides: the CHARGE study. Environ. Health Perspect. 122, 1103–1109.

Silver, M.K., Shao, J., Chen, M., Xia, Y., Lozoff, B., Meeker, J.D., 2016. Distribution and predictors of pesticides in the umbilical cord blood of Chinese newborns. Int. J. Environ. Res. Publ. Health 13.

Sokoloff, K., Fraser, W., Arbuckle, T.E., Fisher, M., Gaudreau, E., LeBlanc, A., Morisset, A. S., Bouchard, M.F., 2016. Determinants of urinary concentrations of dialkyl phosphates among pregnant women in Canada - results from the MIREC study. Environ. Int. 94, 133–140.

Spaan, S., Pronk, A., Koch, H.M., Jusko, T.A., Jaddoe, V.W., Shaw, P.A., Tiemeier, H.M., Hofman, A., Pierik, F.H., Longnecker, M.P., 2015. Reliability of concentrations of organophosphate pesticide metabolites in serial urine specimens from pregnancy in the Generation R Study. J. Expo. Sci. Environ. Epidemiol. 25, 286–294.

Tanner, E.M., Hallerback, M.U., Wikstrom, 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.

Vanacker, M., Quindroit, P., Angeli, K., Mandin, C., Glorennec, P., Brochot, C., Crepet, A., 2020. Aggregate and cumulative chronic risk assessment for pyrethroids in the French adult population. Food Chem. Toxicol. 143, 111519.

Viel, J.F., Warembourg, C., Le Maner-Idrissi, G., Lacroix, A., Limon, G., Rouget, F., Monfort, C., Durand, G., Cordier, S., Chevrier, C., 2015. Pyrethroid insecticide exposure and cognitive developmental disabilities in children: the PELAGIE mother-child cohort. Environ. Int. 82, 69–75.

Viel, J.F., Rouget, F., Warembourg, C., Monfort, C., Limon, G., Cordier, S., Chevrier, C., 2017. Behavioural disorders in 6-year-old children and pyrethroid insecticide exposure: the PELAGIE mother-child cohort. Occup. Environ. Med. 74, 275–281.

Wagner-Schuman, M., Richardson, J.R., Auinger, P., Braun, J.M., Lanphear, B.P., Epstein, J.N., Yolton, K., Froehlich, T.E., 2015. Association of pyrethroid pesticide exposure with attention-deficit/hyperactivity disorder in a nationally representative sample of U.S. children. Environ. Health 14, 44.

Wang, N., Huang, M., Guo, X., Lin, P., 2016. Urinary metabolites of organophosphate and pyrethroid pesticides and neurobehavioral effects in Chinese children. Environ. Sci. Technol. 50, 9627–9635.

Wang, Y., Zhang, Y., Ji, L., Hu, Y., Zhang, J., Wang, C., Ding, G., Chen, L., Kamijima, M., Ueyama, J., Gao, Y., Tian, Y., 2017. Prenatal and postnatal exposure to organophosphate pesticides and childhood neurodevelopment in Shandong, China. Environ. Int. 108, 119–126.

Watkins, D.J., Fortenberry, G.Z., Sanchez, B.N., Barr, D.B., Panuwet, P., Schnaas, L., Osorio-Valencia, E., Solano-Gonzalez, M., Ettinger, A.S., Hernandez-Avila, M., Hu, H., Tellez-Rojo, M.M., Meeker, J.D., 2016. Urinary 3-phenoxybenzoic acid (3-PBA) levels among pregnant women in Mexico City: distribution and relationships with child neurodevelopment. Environ. Res. 147, 307–313.

Xue, Z., Li, X., Su, Q., Xu, L., Zhang, P., Kong, Z., Xu, J., Teng, J., 2013. Effect of synthetic pyrethroid pesticide exposure during pregnancy on the growth and development of infants. Asia Pac. J. Publ. Health 25, 72S, 9S.

Yolton, K., Xu, Y., Sucharew, H., Succop, P., Altaye, M., Popelar, A., Montesano, M.A., Calafat, A.M., Khoury, J.C., 2013. Impact of low-level gestational exposure to organophosphate pesticides on neurobehavior in early infancy: a prospective study. Environ. Health 12, 79.

Zhang, X., Driver, J.H., Li, Y., Ross, J.H., Krieger, R.I., 2008. Dialkylphosphates (DAPs) in fruits and vegetables may confound biomonitoring in organophosphorus insecticide exposure and risk assessment. J. Agric. Food Chem. 56, 10638–10645.

International Journal of Hygiene and Environmental Health 235 (2021) 113767

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

Review

The association between natural environments and childhood mental health and development: A systematic review and assessment of different exposure measurements

Zoë Davis[a,*], Martin Guhn[b], Ingrid Jarvis[a], Michael Jerrett[c,d], Lorien Nesbitt[e], Tim Oberlander[f], Hind Sbihi[f,g], Jason Su[h], Matilda van den Bosch[a,i,j,k,l]

[a] *Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada*
[b] *Human Early Learning Partnership, School of Population and Public Health, University of British Columbia, 2206 East Mall, Vancouver, BC, V6T 1Z3, Canada*
[c] *Department of Environmental Health Sciences, Fielding School of Public Health, University of California, Los Angeles, 650 Charles E. Young Drive S, Los Angeles, CA, 90095, United States*
[d] *Center for Occupational and Environmental Health, Fielding School of Public Health, University of California, Los Angeles, 650 Charles E. Young Drive S, Los Angeles, CA, 90095, United States*
[e] *Department of Forest and Resource Management, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada*
[f] *Department of Pediatrics, Faculty of Medicine, University of British Columbia, 4480 Oak St., Vancouver, BC, V6H 3V4, Canada*
[g] *BC Children's Hospital Research Institute, 950 W 28th Avenue, Vancouver, BC, V5Z 4H4, Canada*
[h] *Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, 2121 Berkeley Way #5302, Berkeley, CA, 94720, United States*
[i] *School of Population and Public Health, Faculty of Medicine, University of British Columbia, 2206 East Mall, Vancouver, BC, V6T 1Z3, Canada*
[j] *ISGlobal, Parc de Recerca Biomèdica de Barcelona, Carrer de Dr. Aiguader, 88, 08003, Barcelona, Spain*
[k] *Universitat Pompeu Fabra, Plaça de La Mercè, 10-12, 08002, Barcelona, Spain*
[l] *Centro de Investigación Biomédica en Red Instituto de Salud Carlos III, Calle de Melchor Fernández Almagro, 3, 28029, Madrid, Spain*

A R T I C L E   I N F O

Keywords:
Green spaces
Urban nature
Public health
Exposure
GIS
Epidemiology

A B S T R A C T

Background: Several studies have assessed the relationship between exposure to natural environments (NEs) and childhood mental health and development. In most cases, a positive association has been found, but results are inconsistent, and the strength of association is unclear. This inconsistency may reflect the heterogeneity in measurements used to assess NE.
Objectives: This systematic review aims to identify the most common NE metrics used in childhood mental health and development research. Our second aim is to identify the metrics that are most consistently associated with health and assess the relative strength of association depending on type of NE exposure measurement, in terms of metric used (i.e., measurement technique, such as remote sensing), but also rate (i.e., spatial and temporal exposure).
Methods: We used the PRISMA protocol to identify eligible studies, following a set of pre-defined inclusion criteria based on the PECOS strategy. A number of keywords were used for retrieving relevant articles from Medline, Embase, PsychINFO, and Web of Science databases between January 2000–November 2020. From these, we extracted data on type of NE measurement and relative association to a number of indicators of childhood mental health and development. We conducted a systematic assessment of quality and risk of bias in the included articles to evaluate the level of evidence. Case studies and qualitative studies were excluded.
Results: After screening of title (283 studies included), abstract, and full article, 45 studies were included in our review. A majority of which were conducted in North America and Europe (n = 36; 80%). The majority of studies used land use or land covers (LULC, n = 24; 35%) to determine exposures to NEs. Other metrics included the normalized difference vegetation index (NDVI), expert measures (e.g., surveys of data collection done by experts), surveys (e.g., self-reported assessments), and use of NE (e.g., measures of a participant's use of NE such as through GPS tracts or parent reports). Rate was most commonly determined by buffer zones around residential addresses or postal codes. The most consistent association to health outcomes was found for buffers of 100 m, 250 m, 500 m, and within polygons boundaries (e.g., census tracts). Six health categories, academic achievement, prevalence of doctor diagnosed disorders, emotional and behavioral functioning, well-being, social

* Corresponding author.
E-mail address: zoedavis@student.ubc.ca (Z. Davis).

https://doi.org/10.1016/j.ijheh.2021.113767
Received 22 September 2020; Received in revised form 30 April 2021; Accepted 3 May 2021
Available online 11 May 2021

functioning, and cognitive skills, were created post hoc. We found sufficient evidence between NDVI (Landsat) and emotional and behavioral well-being. Additionally, we found limited evidence between LULC datasets and academic achievement; use of NE, parent/guardian reported greenness, and expert measures of greenness and emotional and behavioral functioning; and use of NE and social functioning.

Discussion: This review demonstrates that several NE measurements must be evaluated further before sufficient evidence for a potential association between distinct NE exposure metrics and childhood mental health and development can be established. Further, we suggest increased coordination between research efforts, for example, by replication of studies and comparing different NE measurements systematically, so that effect sizes can be confirmed for various health outcomes. Finally, we recommend implementing research designs that assess underlying pathways of nature-health relations and utilize measurement techniques that adequately assess exposure, access, use, and perception of NEs in order to contribute to a better understanding of health impacts of surrounding natural environments.

1. Introduction

Childhood mental health and development are strong predictors of health, well-being, and social ability throughout life (Gluckman et al., 2016). From the life course perspective, it is important to understand how the complex set of early life environmental exposures, sometimes referred to as the exposome, may influence the origin and development of disease (Tamayo-Uria et al., 2019). Harmful exposures have been studied most often, but an increasing number of studies have analyzed natural environments (NEs) as potential health promoting factors on childhood development. NEs consist of green spaces like parks, street trees, and gardens, and blue spaces, like beaches and waterfronts. A majority of studies have shown exposure to NEs improves childhood health (Mygind et al., 2021; Vanaken and Danckaerts, 2018), however, a stronger evidence base is needed to provide recommendations of what constitutes healthy environments in early life, particularly for supporting physical, social, and emotional development (Poore et al., 2017).

Many studies have investigated the relationship between NE exposure and various health outcomes. Positive effects of NEs have been observed for multiple childhood health outcomes including decreased emotional and behavioral difficulties (Vanaken and Danckaerts, 2018), improved birth outcomes (Dzhambov et al., 2014), higher academic achievement (Browning and Rigolon, 2019) and improved overall mental well-being, and cognitive development (McCormick, 2017). Multiple pathways have been hypothesized to explain the association between NE exposure and health, such as through a reduction in heat, air, and noise pollution (Christian et al., 2015; Hartig et al., 2014), increased physical activity (Gray et al., 2015), restoration of attention (Markevych et al., 2017), and decreased stress (Hartig et al., 2014; Vanaken and Danckaerts, 2018). In order to measure these pathways, various NE metrics have been used to capture different forms of exposure. Childhood mental health and development encompasses socio-emotional functioning, cognitive development, academic achievement, overall development and well-being, and mental disorders such as Attention Deficit Hyperactivity Disorder (ADHD).

While previous reviews have investigated associations between childhood health and NEs (e.g., Browning and Lee, 2017; Labib et al., 2019; Smith et al., 2021), the role of different NE metrics, such as remote sensing, land cover, use of, or perceived NEs, have not been specifically evaluated in childhood health studies. NEs are dynamic and may differentially capture pathways linking NE to health outcomes, which creates challenges when assessing exposure to NEs (Labib et al., 2019). These challenges are reflected in the different and numerous techniques for measuring NEs that have been used in research to date. The high heterogeneity in NE measurements (e.g., differences in data collection, data attributes, and definitions of exposure) and different types of study designs may result in inconsistent findings across studies (Labib et al., 2019). For instance, some studies found an association between physical health and well-being, social competence, and emotional well-being (Christian et al., 2017), while others using similar study designs did not (Bell et al., 2020). By investigating if some metrics are more related to health outcomes than others, we may be able to determine the relative importance and potential differences in association depending on NE metric or rate.

Common NE measures used in health research relate to surrounding greenness, presence of street trees, and distance to nearest park areas (Labib et al., 2019; Smith et al., 2021). NE measurements are typically composed of ways to quantify NE (metric) and the amount of exposure (spatial and/or temporal rate). Techniques used to derive NE metrics include remotely sensed indices, land use/land cover (LULC) datasets, on-site evaluations, and self-reports (Table 1). Depending on the metric used, rate can be determined as spatial (e.g., different sizes of buffer zones or polygons around residential address) or temporal (e.g., reported time spend in NEs). NE measurements can also be accessed via the mode of exposure, such as active interaction with green spaces (e.g., use through physical activity) or passive exposure (e.g., surrounding greenness) (Smith et al., 2021).

Among the most common metrics used for NE exposure assessments are vegetation indices, such as the normalized difference vegetation index (NDVI) and soil-adjusted vegetation index (SAVI) (Gascon et al., 2016; Labib et al., 2019). In principle, both NDVI and SAVI are indicators of vegetation health but use different areas of the electromagnetic spectrum to determine vegetation health (Hay, 2000). The quality of the remote sensing derived metrics (e.g., NDVI, SAVI) depends on various sensor characteristics, such as spatial and temporal resolutions, and are essential for determining how many details an image holds and the frequency of image sampling. As such, the choice of sensor usually incurs a trade-off between these features (Hay, 2000; Pettorelli et al., 2018). The Moderate Resolution Imaging Spectroradiometer (MODIS, 250 m spatial resolution), Landsat (30 m spatial resolution), and RapidEye (5 m spatial resolution) are the most frequently used sensors, due to their availability both geographically and temporally (Table 1).

LULC based metrics incorporate data from various sources, usually including high-resolution remote sensing imagery, but also input from, for example, on-site assessments, lidar-derived datasets, zoning, and municipal datasets of urban features, to create discrete representations (classes) of features on the ground (Foody, 2002) (Table 1). Such classes can represent urban NEs (possibly distinguishing between types of green spaces, such as trees, shrubs, or grass), built-up surfaces, and water. The quality of a LULC dataset is usually assessed by the number of pixels that are accurately classified in the dataset and by the resolution of the classification (Finegold et al., 2016).

On-site evaluations can be conducted by experts or by lay people through surveys (Table 1). These evaluations may be potentially less objective than remote sensing indices or LULC datasets, but can provide more detailed data on features and qualities of the landscape (Mazumdar et al., 2020). Although similarly subjective, self-reports can indicate people's perception and experience of NEs, adding another dimension to the assessment (Macintyre et al., 2008; Mazumdar et al., 2020).

With a few exceptions, metrics indicating use of NE have rarely been applied. Use of NE can be measured by, for example, mobile devices and global navigation satellite systems (GNSS, e.g., Almanza et al., 2012), tracking people's mobility patterns and location (Helbich, 2018) or through surveys.

For metrics derived from remotely sensed imagery or LULC datasets, spatial rate is often determined by buffers or polygons of varying sizes around a location of interest (e.g., residential address), wherein distance-to-address or proportion of NE within the area can be determined. Buffers are used as proxies for estimating amount or type of exposure – in the near neighborhood or in a larger area of activity space. These proxies may be indicative of different pathways between NE exposure and health; for instance, smaller buffers may represent immediate NE exposure possibly contributing to stress reduction, whereas larger buffers may be more representative of opportunities for physical activity (Cusack et al., 2017; Jarvis et al., 2020). For metrics like surveys or expert assigned walks, the temporal rate can be assessed by the observed or reported time spent in an NE.

In this review, we aim to complete a systematic evaluation of relative associations between different NE measurements and childhood mental health and development indicators such as socioemotional functioning, cognitive development, academic achievement, mental disorders, and overall development and well-being. Our focus is to assess the metric component of the NE measurement, but rate will be considered when relevant. Our objectives are two-fold: firstly, we aim to identify the most common NE measurements used in childhood mental health and development research. Secondly, we aim to identify which NE measurements are most consistently related to childhood mental health and development and determine the level of evidence for an association between different NE measurements and respective health outcomes.

2. Materials and methods

2.1. Search strategy and methods

We developed a review protocol, defining our design, strategy, and methods a priori (Supplementary Materials, Section 2). We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol (Moher et al., 2009) and in accordance with the PECOS framework, our search referred to: children (Participants), exposure to natural environments (Exposure), any other type of environment in case of controlled studies (Comparator), childhood mental health and development (Outcomes) (Morgan et al., 2018). The initial search for relevant articles was conducted on July 31, 2018 in four databases: EMBASE, MEDLINE (National Library of Medicine), Web of Science, and PsycINFO (American Psychological Association) using Medical Subject Headings (MeSH) and/or keywords (Supplemental Materials, Table S1). For full reproducible queries, see Supplementary Materials, Table S2. The keywords for children's health and development were derived from a reproducible search filter curated by the University of Alberta (Desmeules, 2014) and modified for queries in the Web of Science and PsycINFO. The original search was limited to studies published in English between January 1, 2000 and July 31, 2018 in order to capture recent studies focused on child outcomes. Given the ambitious and time-consuming scope of this review, a complementary search was conducted on November 2, 2020 in MEDLINE using the same search strategy previously used (Supplementary Materials, Table S2b), to update the results for studies published after July 2018. The initial screening for both search events considered title and abstract. Studies deemed potentially valid were retrieved for full article screening and review.

2.2. Study eligibility

The selection criteria were as follows: (a) the study was an original research article published in a peer-review journal in English; (b) the population consisted of children below teenage years. Studies which analyzed population which exceeded the age cut-off were included only if the mean age of the included children was below 13 and at least 80% of the population was below 15. Where possible, only children below the age of 12 were included; (c) the study used quantitative methods of reporting mental health and development (e.g., disease classification, validated scales, or school achievement); (d) environmental exposure included an NE measurement, including self-reports of neighborhood greenness; and (e) the study was either of experimental, longitudinal, cross-sectional, or ecological design. Case studies and qualitative studies were excluded. The selection of studies to be included was determined by two authors (ZD and MvdB) and disagreements were resolved by discussion between the authors until consensus. For details on inclusion and exclusion criteria, see Supplementary Materials, Section 2.

2.3. Data extraction

We reviewed and evaluated the study characteristics of each included article, following similar approaches as in previous reviews (Gascon et al, 2015, 2017). We extracted a large set of data from each study, including author(s), title, publishing journal, year of publication, study design, study country, population, sample size, NE measurement (metric and rate), mental health and development outcome (measure and collection method), main results (including effect sizes), statistical methods, confounding factors/covariates, reported sources of bias,

Table 1
Characteristics and examples of different NE metrics including spatial and temporal resolution and the relationship of the metric to the participant. Exposure type is broken into passive and active, where passive refers to metrics of simple exposure to surrounding NEs and active denotes actual use by participants.

Metric	Example	Common source	Spatial resolution	Temporal resolution	Exposure type
Vegetation indices	NDVI, SAVI	Landsat	30 m	• Imagery available since 1984 (Landsat 5) • 16-day revisit time	Passive
		MODIS	250 m	• Imagery available since 2000 • 16-day composite imagery	Passive
		RapidEye	5 m	• Imagery available from 2009 • 5.5-day revisit time	Passive
		Sentinel-2	10 m	• Imagery available since 2015 • 5-day revisit time	Passive
LULC	Proportion of NE, distance to nearest NE	High resolution imagery, orthophotos, lidar	Depends on source imagery (<10 m most common)	• Availability depends on the agency (e.g., municipality) responsible for production	Passive
Survey	Self-reported greenness or quality of neighborhood	Self-reported surveys			Passive/ Active
Expert measures	Survey conducted by professional (e.g., sky-view index, objective quality measures)			• Greenness measures are sometimes collected at the same time as health assessment	Passive/ Active
Use	Self-reported survey or GPS			• Availability depends on the method of collection, for instance, GPS measurements could be for a week, and surveys could cover a year	Passive/ Active

reported strengths and limitations, and other pertinent information. A modified version of our data extraction sheet is provided in Supplementary Materials, Table S4. Initially, one reviewer (ZD) worked independently to extract this information. A second reviewer (MG, IJ, MJ, LN, TO, HS, JS, or MvdB) then reread the articles and reviewed the extracted data for control and accuracy. We only reported statistically significant effect sizes for specific subscales and subcategories of outcomes (positive or negative association) in the data extraction sheet. We also reported if no statistically significant association was found for the main outcome of the study.

2.4. Quality assessment and classification of evidence

Many different approaches have been suggested for reaching a composite score of study quality in systematic reviews (Russell et al., 2009). In this study, the quality assessment criteria to derive a numerical score per study were based on an established methodology previously published (Gascon et al, 2015, 2017) and included, for instance, quality of remote sensing product (e.g., spatial and temporal resolution), expert or lay-person NE evaluation, quality of health outcome measure (e.g., self-reported or objective test or diagnose), study design, and risk of bias. As an adaptation of the Gascon et al. (2017, 2015) method, we also included separate assessments of the NE metric used due to the specific objective of this review. For more information on the decision making for each category, please refer to the protocol (Supplementary Materials, Section 2). Similar, to Gascon et al. (2015, 2017), we did not consider criteria that were not applicable for the study as to not double count or penalize studies for different study designs and use of NE measurements. For criteria that were not applicable (NA), we gave no points for the category and removed the points that the category was worth from the overall total of possible points. This meant that each study had a different number of possible maximum points, ranging from 21 to 25 points. This method resulted in a final three-grade quality assessment of the study – poor (≤40%), moderate (41–60%), or good quality (≥61%), where each category was determined by the relative percentage. The quality assessment was conducted independently by ZD and a second reviewer (either MG, IJ, MJ, LN, TO, HS, JS, or MvdB). For each study, the reviewers provided a numerical score for each criterion item (Supplementary Materials, Table S5). A final decision of the quality score was made after iterative reviews and discussion between the reviewers until consensus. The quality scores were used to establish a degree of evidence for relationships between separate measurements of NE and childhood mental health and development. Similar to previous reviews Gascon et al. (2017, 2015), we used an adapted version of the International Agency for Research on Cancer (IARC) definitions to describe causal relationships (International Agency for Research on Cancer, 2019) to create categories for the level of evidence for an association between the selected health outcomes and NE exposure. While acknowledging that most of the studies included in this review would not allow for establishing causal relationships, we included the following evidence categories for commonly observed association between exposure and outcome – sufficient, limited, insufficient, and lack of evidence. Sufficient evidence indicates significant relationships were observed between NE and improved health outcome in studies rated as predominantly moderate or good quality. Limited evidence indicates that significant relationships were observed in studies of varying quality, but more studies are required to rule out bias and confounding. Insufficient evidence suggests that there is a lack of good quality studies, lack of replication, and/or lack of statistical power, which prevents a conclusion of evidence to be made. Finally, lack of evidence is when there are several good quality studies which indicate no association between outcome and exposure.

3. Results

We are presenting the full search results from both search dates. We identified 2050 articles in Embase, 3547 articles in Medline, 6808 articles in Web of Science and 1364 articles in PsycINFO. 3045 duplicates were identified and removed. After screening titles, 283 articles were selected for an abstract screening. Additionally, we identified 56 articles from other sources (e.g., known citations) that passed the title screen. The bibliography search yielded eight additional papers. We evaluated the abstract of each paper and selected 95 papers for full-text screening. During the full-text screening, we eliminated 50 studies. Thus, 45 articles were included in the final review (Fig. 1). Of full articles selected, 35 were identified during the initial search and 10 were identified during the complementary search conducted in 2020. Individual PRISMA diagrams for each search are presented in the Supplementary Materials, Figures S1 and S2. For the extracted data of each paper, including study design, population, sample size, NE metric and rate, health outcome, and main results and effect sizes, see Table S4 in Supplementary Materials.

A majority of the studies had a cross-sectional design (n = 33; 73%), of which twelve (26%) had an ecological design. Six (13%) studies were longitudinal, two (4%) had a pseudo-experimental design, and one (2%) had a case-control design (Supplementary Materials, Table S4).

The 45 studies were conducted in ten different countries. A majority of the studies were conducted in the United States (n = 19; 42%). The other studies were conducted in the United Kingdom (n = 6; 13%), Australia (n = 3; 6%), China (n = 3; 6%), New Zealand (n = 3; 6%), Spain (n = 3; 6%), Canada (n = 2; 4%), Germany (n = 2; 4%), Lithuania (n = 2; 4%), and Sweden (n = 2; 4%). Most of the studies were published within the last five years (2015–2020) (n = 31; 69%) and only five studies were published prior to 2004. The sample sizes ranged from 17 (Faber Taylor and Kuo, 2009; Wells, 2000) to 66,823 (Markevych et al., 2018). The sample sizes in studies of ecological design ranged between eleven schools (Mårtensson et al., 2009) and 1772 schools (MacNaughton et al., 2017). Thirteen (28%) studies received a composite score of poor quality, 27 (60%) of moderate quality, and five (11%) of good quality (Supplementary Materials, Table S5).

3.1. Natural environment measurements

Multiple types of measurements were used to evaluate NEs, including remote sensing products such as NDVI and LULC datasets, information collected via survey or by experts, exposures assigned to participants by experts, and evaluations of participants' use of NEs (Table 2). In many cases, multiple rates were used for the same NE metric (Table 2; Supplementary Materials, Table S4). As such, we considered them separately in our analysis. LULC datasets were the most commonly used metric (n = 24; 35%). Thirteen of these datasets had high accuracy (>80%) and were derived from moderate-to high-resolution imagery (5- to 30-m imagery). Seven studies had moderate accuracy (60–80%) and four studies had low accuracy (<60% accurate or no information on dataset development). When accuracy level of the LULC data was not reported in the article, we tracked the original source. NDVI was the second most common metric used across all the studies (n = 21; 31%). In particular, NDVI derived from Landsat was used in 10 (15%) studies and NDVI derived from MODIS (250 m resolution) was used in six (9%) studies. One (1%) study used NDVI from RapidEye (5 m resolution) and one (1%) study used NDVI from the US National Agricultural Imagery Program (NAIP, 1 m resolution). In addition, NDVI derived from RapidEye and Landsat were used to estimate cumulative exposure during home-school commutes (n = 2; 3%).

3.2. Natural environment measures and relation to health outcomes

Based on the data extraction of health outcomes from the included studies, we created six health outcome categories: academic achievement (n = 9; 20%), prevalence of doctor diagnosed disorders (n = 3; 6%), emotional and behavioral functioning (n = 22; 49%), well-being (n = 9; 20%), social functioning (n = 11; 24%), and cognitive skills (n = 5,

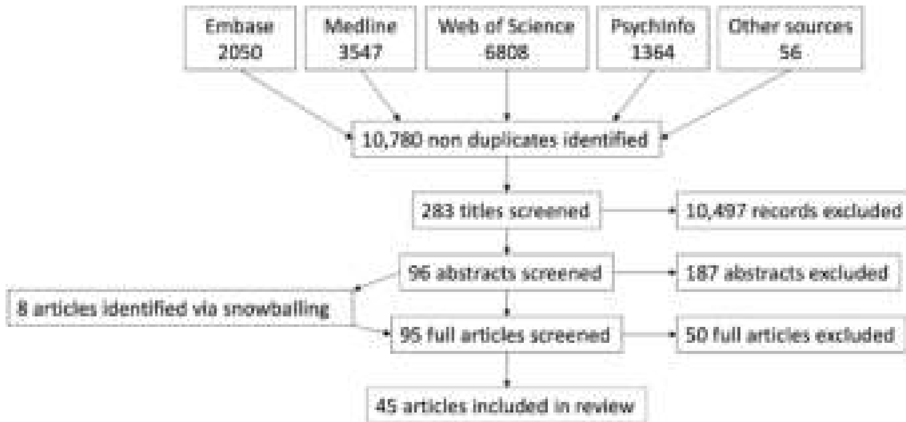


Fig. 1. Combined PRISMA diagram of the procedure for article selection. Snowballing refers to studies that were identified during citation screenings on the studies put forward for full text screening (i.e., bibliography search). For separate PRISMA diagrams for each search date, please see Supplemental Materials, Section 1 Figure S1 for 2018 and Figure S2 for 2020.

Table 2
NE metric categories and rates included in the studies. Count and percentage refer to the number of studies (68 significant findings) that employed the metric and rate. See Table 1 for details of metrics and rates.

Metric	Rate
NDVI (n = 21; 31%)	• NDVI summarized (e.g., mean, median) within specified area around residence (e.g., buffer or polygon) (n = 19; 28%). Buffer sizes varying between 50 m and 2000 m • NDVI as a proxy for exposure during commute (n = 2; 3%)
LULC (n = 24; 35%)	• Proportion of classes considered NE within a specified area around a residence (e.g., buffer or polygon) (n = 16; 24%) • Distance to nearest NE from residence (continuous or discrete) (n = 7; 10%) • Landscape distribution of NE patches (n = 1; 1%)
Survey (n = 5; 7%)	• Parent or guardian reported amount of "greenness" of surroundings (n = 4; 6%) • Parent or guardian reported "quality" of NE around residence (n = 1; 1%)
Expert Measures (n = 7; 10%)	• On-site measurements of NE around residence (e.g., surveys conducted by professionals, OPEC[a], sky-view index) (n = 5; 7%) • Expert assignments of participant exposure to pre-selected environment (e.g., a walk through an urban or natural environment pre- and post- health assessment) (n = 2; 3%)
Use (n = 10; 15%)	• Survey on child's use of NE completed by parents (n = 9; 13%) • Time child spent in NE measured by GPS tracts (n = 1; 1%)

[a] OPEC: Outdoor Play Environment Categories. Percentages were rounded up to the nearest whole.

11%), where multiple outcomes could be assessed within the same study. Academic achievement describes a child's performance in school either through scores on standardized tests or through attendance. Prevalence of doctor diagnosed disorders refers to the rate of disease diagnosis within populations. The emotional and behavioral function describes various disabilities related to, for example, attention, interpersonal relationships, and also includes some of the common symptoms associated with ADHD. The well-being category refers to measures of overall development and well-being. The social functioning category refers to prosocial behavior and positive functioning. Finally, the cognitive skills category refers to tests that evaluate memory performance and cognitive performance. Descriptions of the outcome measures used in each health category are provided in Table 3.

Across the health outcomes, different NE metrics were favored in different health outcome categories. LULC and NDVI metrics were used across all health categories (Table 4). Specifically, academic achievement and prevalence of doctor diagnosed mental disorders only used metrics derived from remote sensing data (e.g., vegetation indices and LULC datasets), whereas other health outcomes, such as emotional and behavioral functioning, well-being, and social functioning, used the most diverse array of NE metrics. We found positive associations between NEs and childhood mental health and development in 56 cases (82%), an inverse association in nine cases (13%), and no statistically significant association in three cases (4%) (Table 5). The counts do not sum up to the number of studies included (n = 45), because outcomes were evaluated separately, and some studies analyzed more than one outcome. When available, effect sizes for the respective health outcomes are reported in Table S4 in the Supplementary Materials. In the following we provide a synthesis of the results from the studies included in the review as well as the level of evidence per each outcome.

3.2.1. Academic achievement

Nine studies (20%) evaluated academic achievement, including success of students (ages 5–12) on standardized tests or school attendance (Table 5; Supplementary Materials, Table S4). Most students were in the third grade (ages 8–9) (n = 5; 56%). All studies assessed NEs at the school, either by estimating the spatial rate of NEs within school property or within the neighborhood that the school serves. Results for academic achievement were mixed (Table 5) – six studies showed positive associations between NE and measures of academic achievement via standardized test scores (Hodson and Sander, 2017; Kuo et al., 2018; Kweon et al., 2017; Sivarajah et al., 2018; Tallis et al., 2018; Wu et al., 2014) and two studies showed inverse associations between NE and the outcomes (Beere and Kingham, 2017; Browning et al., 2018). One study analyzed the association between chronic absenteeism and school surrounding NE (MacNaughton et al., 2017) and found a 2.6% decrease in absenteeism with each 1-IQR increase of NDVI (250 m MODIS).

Level of evidence: Given the ecological study design of these studies and the mixed results, we classified the evidence between NE, as measured by LULC datasets, and standardized test scores as limited. Due to differences in findings, we classified the association between standardized test scores and NDVI as insufficient. We classified the level of evidence for an association between NE and absenteeism as insufficient (Table 6).

3.2.2. Prevalence of doctor diagnosed disorders

Three studies (6%) analyzed prevalence of doctor-diagnosed mental disorders; two studies of the studies evaluated ADHD (Donovan et al., 2019; Markevych et al., 2018) and one study evaluated autism (Wu and Jackson, 2017) (Table 5; Supplementary Materials, Table S4). All three studies found that NEs was correlated with lower disease prevalence,

Table 3
Specific tests and subscales used in each health outcome categories.

Category	Health outcome (Tool)
Academic achievement	• Standardized test scores • Rates of absenteeism
Prevalence of doctor diagnosed disorders	• Rate of disease diagnosis
Emotional and behavioral functioning	• SDQ total difficulties (SDQ) • Hyperactivity (SDQ) • Internalizing difficulties (SDQ) • Externalizing difficulties (SDQ) • School functioning (HRQOL) • Emotional functioning (HRQOL) • Aggressive behavior • Anxiety (parent reported questionnaire) • Conduct problems (SDQ) • Behavioral problems (CBCL) • Depression (CBCL, Behavioral Assessment System for Children) • Inattention (ECADDES, ADDES, ADHD/DM-IV, SDQ) • Attention (Digit Span Backwards, Stroop Color-word test, Symbol Digit Modalities, ANT, K-CPT) • Parent reported symptoms
Well-being	• Physical health and well-being (AEDC) • Total health-related quality of life (HRQOL) • Life satisfaction (Life Satisfaction Scale) • Happiness • Well-being index • Global self-worth • Self-esteem (Kid KINDL) • Subjective health reports
Social functioning	• Social competence (AEDC) • Peer relationship problem (SDQ) • Pro-social behavior (SDQ) • Social functioning (HRQOL) • Psychosocial functioning (HRQOL) • Psychosocial health (Lewis Stressful Life Events Scale) • Friends (Kid KINDL) • Social problems (CBCL)
Cognitive skills	• Spatial working memory • Working memory (n back test) • Superior working memory (n back test) • Cambridge Neuropsychological Test Automated Batter • Neurodevelopment (BSID) • Cognitive ability (WPPSI-R)

Abbreviations.
ADDES = Attention Deficit Disorder Evaluation Scale.
ADHD/DSM IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition.
AEDC = Australian Early Development Census.
ANT = Attentional Network Test.
BSID = Bayley Scales of Infant Development.
CBCL = Child Behavior Checklist.
ECADDES = Early Childhood Attention Deficit Disorder Evaluation Scale.
HRQOL = health-related quality of life.
HRT-SE = hit-reaction time – standard error.
K-CPT = Conners' Kiddie Continuous Performance test.
Kid KINDL = Kid KINDL Health-related quality of life.
SDQ = strengths and difficulties questionnaire.
WPPSI-R = Wechsler Preschool and Primary Scales of Intelligence - Revised.

while using different NE measurements. Donovan et al. (2019) and Markevych et al. (2018) used NDVI, derived from 30 m Landsat and 250 m MODIS. In New Zealand, Donovan et al. (2019) found that rural meshblocks protective against ADHD for children who had always lived in them, however, early life exposure to green space had no significant results. For every 0.1 increase in NDVI, Markevych et al. (2018) found a decrease in the rate of ADHD diagnosis in Germany. Wu and Jackson (2017) utilized a LULC dataset and found a decrease in autism prevalence with increase in proportion of NE within school district areas in California, USA.

Level of evidence: Due to the limited number of studies, we classified the evidence for an association between NE exposure and each of the mental health and developmental disorders as insufficient (Table 6).

3.2.3. *Emotional and behavioral functioning*

Twenty-two studies (49%) assessed emotional and behavioral functioning. NDVI from Landsat was used to evaluate exposure to NEs in eight studies (Amoly et al., 2014; Balseviciene et al., 2014; Dadvand et al., 2017; Liao et al., 2020; Madzia et al., 2019; McEachan et al., 2018; Yang et al., 2019). In addition, one study, Dadvand et al. (2015), utilized NDVI from RapidEye (5 m spatial resolution) and one study, Yang et al. (2019), used the soil-adjusted vegetation index (SAVI) derived from Landsat to estimate NE exposure. Eight buffer sizes were used, ranging from 50 m to 1000 m, the most common being the 100 m buffer (n = 6; 32%) and the 500 m buffer (n = 4; 21%). All eight of these studies found decreased emotional and behavioral problems with increased exposure to NEs as measured through vegetation indices.

Six studies utilized LULC datasets for evaluating exposure to NEs (Balseviciene et al., 2014; Bell et al., 2020; Christian et al., 2017; Feng and Astell-Burt, 2017; Larson et al., 2018; Markevych et al., 2014b). Two used the relative distance between participant's home address and the nearest NE (Balseviciene et al., 2014; Christian et al., 2017) and found mixed results. Markevych et al. (2014) evaluated the presence of NEs within 500 m of home addresses and found an inverse relationship between park presence and emotional and behavioral functioning. Three studies evaluated the proportion of NEs within polygons surrounding home addresses (Bell et al., 2020; Feng and Astell-Burt, 2017; Larson et al., 2018). Bell et al. (2020) and Feng and Astell-Burt (2017) conducted similar studies in the Perth, Australia metropolitan area, and found opposing results; however, Bell et al. (2020) evaluated home yard space, whereas Feng and Astell-Burt (2017) estimated park space within census tracts. Finally, Larson et al. (2018) evaluated the association between the proportion of tree canopy and anxiety in a population diagnosed with autism in California.

Five studies evaluated exposure to NE via parent-reported survey (Faber Taylor et al, 2001, 2002; Faber Taylor and Kuo, 2009, 2011; Kuo and Faber Taylor, 2004). Of these, four studies (Faber Taylor et al., 2001; Faber Taylor and Kuo, 2009, 2011; Kuo and Faber Taylor, 2004) stated that participants had a formal ADHD diagnosis. Faber Taylor et al. (2002) found that girls who had more views of NE from their residence had better concentration, impulse inhibition, and self-discipline compared to those with less view. Faber Taylor et al. (2001) found that children who played in areas with higher amounts of NEs had fewer symptoms than children who played indoors or in built environments (Table 5). Similarly, children had more severe ADHD symptoms when playing "deep indoors" compared to "open grass" or "big trees and grass" (Faber Taylor and Kuo, 2011) (Table 5). Kuo and Faber Taylor (2004) used parent-reported survey data to assess ADHD symptoms in after-school or weekend activities and found no significant association to NE (Table 5). One study (Faber Taylor and Kuo, 2009) evaluated NE exposure through expert measures. Faber Taylor and Kuo (2009) found participants performed better on a puzzle task after they had gone for a 20-min walk in a park compared to after they had taken a walk in a downtown area (Table 5; Supplementary Materials, Table S4).

Two studies used expert measures to evaluate exposure to NEs (Mårtensson et al., 2009; Wells, 2000). Mårtensson et al. (2009) found a decrease in hyperactivity of children whose play areas had a higher outdoor play environment categories (OPEC) score in Sweden. Wells (2000) followed children who moved from a neighborhood with less NE to one with more and found the children's directed attention capacity increased after the move.

Finally, three studies assessed use of NE through surveys to identify the amount of time (e.g., hours/year, days/week) children spent in NEs (Amoly et al., 2014; Flouri et al., 2014; Richardson et al., 2017) and all three found decreased behavioral problems with increased NE exposure.

Level of evidence: Given the positive findings and similar study designs, the evidence of association between NDVI (Landsat), particularly

Table 4
Number of NE metrics used for each health outcome. Note: For studies that used multiple NE metrics, each significant finding was counted separately, therefore, the total n exceeds the total number of studies evaluated (n = 45).

Health outcome	NDVI	SAVI	LULC	Survey	Expert measures	Use of NE	Total
Academic achievement	3	0	6	0	0	0	9
Prevalence of doctor diagnosed disorders	2	0	1	0	0	0	3
Emotional and behavioral functioning	9	1	6	5	3	3	27
Well-being	3	0	5	0	2	3	11
Social functioning	3	0	5	0	1	4	13
Cognitive skills	1	0	1	0	1	0	5
Total	21	1	24	5	7	10	68

within in 100 and 300 m buffers, and emotional and behavioral function was classified as sufficient (Table 6). The evidence of association between LULC datasets and emotional and behavioral functioning was classified as limited due to the differences in study design, quality of studies, and inconsistent results. The evidence of association for surveys of NEs and expert measures was classified as limited due to the quality and heterogeneity of the studies. Finally, due to the lack of replicated studies, the evidence of association between use of NE was classified and emotional and behavioral development was classified as insufficient.

3.2.4. *Well-being*

Nine studies (20%) assessed the association between NEs and childhood well-being (Andrusaityte et al., 2020; Bell et al., 2020; Christian et al., 2017; Kim et al., 2016; McCracken et al., 2016; Söderström et al., 2013; Tillmann et al., 2018; Ward et al., 2016; Wells and Evans, 2003) (Table 5; Supplementary Materials, Table S4). Use of NE, as defined by parent survey or GPS tracking, was used in three studies (Andrusaityte et al., 2020; McCracken et al., 2016; Ward et al., 2016). Four studies utilized LULC datasets for assessing the relative distance to a park and for determining the proportion of NE within areas surrounding a residence (Bell et al., 2020; Christian et al., 2017; Kim et al., 2016; Tillmann et al., 2018). One study used average NDVI from Landsat (30 m) within 100 m, 300 m, 500 m buffers to evaluate NE exposure (Andrusaityte et al., 2020). Finally, two studies used expert evaluations of the amount of vegetation visible within school yards or from inside participants' homes (Söderström et al., 2013; Wells and Evans, 2003). Eight studies found some positive associations between NE exposure and outcome, while two studies, both using LULC datasets to quantify the proportion of land cover types, found inverse associations (Larson et al., 2018; Tillmann et al., 2018) and one study found no significant results (Söderström et al., 2013).

Level of evidence: Due to the varying quality of studies and the heterogeneity in NE measurements used for NDVI, LULC datasets, expert measures, and use of NE, we classified the evidence of association as insufficient (Table 6).

3.2.5. *Social functioning*

Eleven studies (24%) analyzed various aspects of social functioning (Amoly et al., 2014; Andrusaityte et al., 2020; Balseviciene et al., 2014; Bell et al., 2020; Christian et al., 2017; Flouri et al., 2014; Liao et al., 2020; McCracken et al., 2016; Richardson et al., 2017; Tillmann et al., 2018; Wells and Evans, 2003) (Table 5; Supplementary Materials, Table S4). Three studies utilized average NDVI from Landsat within 100 m, 250 m, and 300 m buffers to estimate green space exposure (Amoly et al., 2014; Balseviciene et al., 2014; Liao et al., 2020).

Five studies used LULC datasets to evaluate NE exposure. Two studies found a positive association between social functioning and estimated exposure to NE through the distance to the nearest NE from a residential address (Balseviciene et al., 2014; Christian et al., 2017). Three studies used the proportion of NE within either a polygon boundary or a 500 m buffer from a residential address (Bell et al., 2020; Richardson et al., 2017; Tillmann et al., 2018). Wells and Evans (2003) quantified vegetation by assessing views from windows using surveys conducted by professionals. Finally, parent reported use of NE (e.g., hours/year, days/week) was included in four studies (Amoly et al., 2014; Andrusaityte et al., 2020; Flouri et al., 2014; McCracken et al., 2016).

Level of evidence: We classified the level of evidence of association between NE exposure, as measured by use of NE, and social functioning as limited following the overall positive results and quality of the studies (Table 6). Evidence of association for NE exposure determined by NDVI (Landsat), LULC, and expert measures was classified as insufficient due to the quality and small number of studies (Table 6).

3.2.6. *Cognitive skills*

Five studies (11%) analyzed associations between memory function, neurodevelopment, or cognitive ability and NE exposure (Dadvand et al., 2015; Flouri et al., 2018; Liao et al., 2019; Reuben et al., 2019; Schutte et al., 2017) (Supplementary Materials, Table S4). Dadvand et al. (2015) used NDVI (30 m Landsat) to quantify exposure over a year and observed an increase in working memory and superior working within the schools with more NEs compared with schools with less NEs. Flouri et al. (2019) estimated NE exposure through determining the proportion of green space within each child's neighborhood of residence. They found that a decile increase in green space was associated with higher spatial working memory. Schutte et al. (2017) assigned 20-min walks in different environments prior to and after completing a memory test and found children had more accurate spatial working memory after a walk in a natural environment compared to a walk in an urban environment (Table 5; Supplementary Materials, Table S4). Liao et al. (2019) and Reuben et al. (2019) both utilized average NDVI (250 m MODIS) within 300 m or 1-mile buffers, and both found increased cognitive performance with increased greenness levels.

Level of evidence: Due to the differences in NE measurements and limited number of studies, there is insufficient evidence in regard to the association between NE exposure and memory (Table 6).

4. Discussion

To our knowledge, this is the first review examining the evidence for an association between NE and childhood mental health and development that includes a systematic assessment of relative influence on outcome depending on NE measurement. While the number of studies representing various metrics differ, our results suggest that the statistical association to childhood health, to some extent, depends on the metric used. We found sufficient evidence of association between NDVI (derived from Landsat) and emotional and behavioral outcomes. There was also evidence, although limited, for an association between academic achievement and LULC based metrics; emotional and behavioral functioning and use of NE, parent/guardian measure of NE, and expert measures of NE; and social functioning and use of NE (Table 6). For the remaining metrics, there was insufficient evidence to confirm an association to health outcomes, mostly due to small number of studies and high variability of metrics (Table 6).

LULC dataset and NDVI were the most common metrics used across all health outcomes (Table 4) and may be the most common choice due to the ease of assessing such data and for assigned individual NE exposures in large cohorts. In studies using NDVI, exposure was typically

Table 5
Associations (adjusted results) between NE and outcome for each study. ↑ indicates positive associations between NE exposure and health outcome, i.e., NEs improves health; ↓indicates an inverse association between NE exposure and health outcome, i.e., NEs worsens health; x̂ indicates no significant association between NE exposure and health outcome. The score indicates the quality of each study: P indicates poor; M, moderate; G, good.

Health Outcome	Article	Metric	Sensor/Source	Rate	Direction of association	Score
Academic achievement						
Standardized tests	Browning et al. (2018)	NDVI	MODIS	250 m, 500 m, 1000 m, 2000 m	↓	P
	Wu et al. (2014)	NDVI	MODIS	250 m, 500 m, 1000 m, 2000 m	↑	M
	Beere and Kingham (2017)	LULC	Proportion of NE	School ground, school zone	↓	P
	Hodson and Sander (2017)	LULC	Proportion of NE	Polygon, attendance zones	↑	P
	Kuo et al. (2018)	LULC	Proportion of NE	Polygon, school zone	↑	M
	Kweon et al. (2017)	LULC	Proportion of NE	Polygon, school boundary	↑	M
	Sivarajah et al. (2018)	LULC	Proportion of NE	Polygon, school boundary	↑	M
	Tallis et al. (2018)	LULC	Proportion of NE	750 m, 100 m	↑	M
Absenteeism	MacNaughton et al. (2017)	NDVI	MODIS	250 m	↑	P
Prevalence of Doctor Diagnosed Disorders						
ADHD prevalence	Markevych et al. (2018)	NDVI	MODIS	Polygon	↑	M
	Donovan et al. (2019)	NDVI	Landsat	Polygon	x̂	M
Autism	Wu and Jackson (2017)	LULC	Proportion of NE	Polygon, school boundary	↑	M
Emotional and Behavioral Functioning						
	Amoly et al. (2014) [a]	NDVI	Landsat	100 m, 250 m, 500 m	↑	M
	Balseviciene et al. (2014)	NDVI	Landsat	300 m	↑	P
	Dadvand et al. (2017)	NDVI	Landsat	100 m, 300 m, 500 m	↑	G
	Liao et al. (2020)	NDVI	Landsat	100 m around residence	↑	M
	Liao et al. (2020)	NDVI	Landsat	100 m around school	↑	M
	Madzia et al. (2019)	NDVI	Landsat	400, 800 m	↑	M
	McEachan et al. (2018)	NDVI	Landsat	100 m, 300 m, 500 m	↑	M
	Yang et al. (2019)	NDVI	Landsat	100 m, 500 m, 1000 m	↑	M
	Dadvand et al. (2015)	NDVI	RapidEye	50 m, 100 m	↑	G
	Yang et al. (2019)	SAVI	Landsat	100 m, 500 m, 1000 m	↑	M
	Balseviciene et al. (2014)	LULC	Distance to NE	Address-to-NE-distance	↑	P
	Christian et al. (2017)	LULC	Distance to NE	Address-to-NE-distance	↑	P
	Markevych et al. (2014b)	LULC	Distance to NE	500 m	↓	G
	Bell et al. (2020)	LULC	Proportion of NE	Polygon	↓	M
	Feng and Astell-Burt (2017)	LULC	Proportion of NE	Polygon	↑	M
	Larson et al. (2018)	LULC	Proportion of NE	Polygon	↓	M
	Faber Taylor et al. (2002)	Survey	Parent/guardian survey	NA	↑	P
	Faber Taylor and Kuo (2011) [a]	Survey	Parent/guardian survey	NA	↑	P
	Faber Taylor et al. (2001)	Survey	Parent/guardian survey	NA	↑	P
	Kuo and Faber Taylor (2004)	Survey	Parent/guardian survey	NA	x̂	P
	Feng and Astell-Burt (2017)	Survey	Quality of NE	NA	↑	M
	Faber Taylor and Kuo (2009) [a]	Expert measure	Assigned walk	20 min walk	↑	M
	Mårtensson et al. (2009)	Expert Measure	On-site	NA	↑	G
	Wells (2000)	Expert Measure	On-site	NA	↑	M
	Amoly et al. (2014) [a]	Use of NE	Parent/guardian survey	Hours/year	↑	M
	Flouri et al. (2014) [a]	Use of NE	Parent/guardian survey	Day/week	↑	M
	Richardson et al. (2017) [a]	Use of NE	Parent/guardian survey	Private garden	↑	M
Well-being						
	Andrusaityte et al. (2020) [a]	NDVI	Landsat	100 m, 500 m	↑	M
	Christian et al. (2017)	LULC	Distance to NE	Address-to-NE-distance	↑	P
	Kim et al. (2016)	LULC	Distance to NE	400 m, 800 m	↑	P
	Bell et al. (2020)	LULC	Proportion of NE	Polygon	↓	M
	Tillmann et al. (2018)	LULC	Proportion of NE (park)	500 m	↑	M
	Tillmann et al. (2018)	LULC	Proportion of NE (water, grass)	500 m	↓	M
	Söderström et al. (2013)	Expert Measure	On-site	NA	x̂	M
	Wells and Evans (2003)	Expert Measure	On-site	NA	↑	P
	Andrusaityte et al. (2020) [a]	Use of NE	Survey	Hour/week	↑	M
	McCracken et al. (2016) [a]	Use of NE	Parent/guardian survey	Day/week	↑	M
	Ward et al. (2016) [a]	Use of NE	GPS	For 7 days	↑	M
Social Functioning						
	Amoly et al. (2014) [a]	NDVI	Landsat	250 m	↑	M
	Balseviciene et al. (2014)	NDVI	Landsat	300 m	↓	P
	Liao et al. (2020)	NDVI	Landsat	100 m around residence	↑	M
	Balseviciene et al. (2014)	LULC	Distance to NE	Address-to-NE-distance	↑	P
	Christian et al. (2017)	LULC	Distance to NE	Address-to-NE-distance	↑	P
	Bell et al. (2020)	LULC	Proportion of NE	Polygon	↓	M
	Richardson et al. (2017) [a]	LULC	Proportion of NE	500 m	↑	M
	Tillmann et al. (2018)	LULC	Proportion of NE	500 m	↑	M
	Wells and Evans (2003)	Expert Measure	On-site	NA	↑	P
	Amoly et al. (2014) [a]	Use of NE	Survey	Hours/year	↑	M
	Andrusaityte et al. (2020) [a]	Use of NE	Parent/guardian survey	Hours/week	↑	M
	Flouri et al. (2014) [a]	Use of NE	Parent/guardian survey	Days/week	↑	M
	McCracken et al. (2016) [a]	Use of NE	Parent/guardian survey	Days/week	↑	M
Cognitive Skills						
	Dadvand et al. (2015)	NDVI	RapidEye	50 m, 100 m	↑	G
	Liao et al. (2019)	NDVI	MODIS	300 m	↑	P
	Reuben et al. (2019)	NDVI	MODIS	1 mile	↑	M

(continued on next page)

Table 5 (*continued*)

Health Outcome	Article	Metric	Sensor/Source	Rate	Direction of association	Score
	Flouri et al. (2019)	LULC	Proportion of NE	Polygon	↑	P
	Schutte et al. (2017)	Expert measure	Assigned walk	20 min walk	↑	M

The spatial resolutions of the sensors are as follows: Landsat, 30 m; MODIS, 250 m; and RapidEye, 5 m.
Abbreviations: NDVI = normalized difference vegetation index; MODIS = the moderate resolution imaging spectroradiometer; m = meter; LULC = land use/land cover; NE = natural environment; NA = not applicable.

[a] Indicates that the study assessed temporal rate, or the amount of time spent in NEs. For more information, see Supplementary Materials, Table S4. For detail assessment of the Quality Scores, see Supplementary Materials, Table S5.

Table 6
Summarized results of the associations between health outcomes and NE metrics.

Outcome		NE Metric	Evidence of Association
Academic achievement	Standardized test scores	LULC	Limited
		NDVI (MODIS)	Insufficient
	Absenteeism	NDVI (MODIS)	Insufficient
Prevalence of mental doctor diagnosed mental disorders	Autism	LULC	Insufficient
	Autism	LULC	Insufficient
Emotional and Behavioral Functioning		NDVI (Landsat)	Sufficient
		LULC, survey, expert measures	Limited
		Use of NE	Insufficient
Well-being		NDVI (Landsat, MODIS), LULC, expert measures, use of NE	Insufficient
Social functioning		Use of NE	Limited
		NDVI (Landsat), LULC, expert measures	Insufficient
Cognitive skills		NDVI (Landsat), LULC, expert measures	Insufficient

assessed via circular buffers around residential addresses. Studies that most commonly employed LULC datasets for evaluating NE exposure applied polygon boundaries as spatial exposure unit. Rate was most commonly determined by buffer zones around residential address or postal code and the most consistent association to health outcomes was found within 100 m, 250 m, 500 m, and within polygons boundaries (e.g., census tracts). Only a few studies (n = 2; 3%), namely the ones that evaluated NE exposures during commutes, utilized network buffer, or buffers that follow road paths.

Of note, NDVI derived from Landsat was the only measure to have a sufficient level of evidence with a health outcome (emotional and behavioral functioning). This was due to the high number of the studies of relatively high quality and overall positive associations, which may suggest that NDVI is an adequate way of estimating exposure to NEs. It is possible, however, that this finding is caused simply by the common use of NDVI. While NDVI has been one of the most common methods for evaluating exposure, and has shown to closely resemble self-reported greenness estimates by adults (Rhew et al., 2011), it does not offer street-level estimates of greenness or other greenness qualities, which may be important for determining specific pathways linking health and NEs (Labib et al., 2020).

Use of NE, particularly when assessed via parent-reported survey, was the third most common NE metric and was utilized in three health categories (emotional and developmental function, well-being, and social functioning). Time periods was assessed in days or hours per week (Andrusaityte et al., 2020; Flouri et al., 2014; McCracken et al., 2016; McEachan et al., 2018; Richardson et al., 2017), hours per year (Amoly et al., 2014), or by GPS tracking over the course of a week (Ward et al., 2016). Other reviews have evaluated moderate to vigorous physical activity of children within NEs (e.g., McCrorie et al., 2014), however, more research is needed to connect use of NEs to mental health and well-being. Due to the limited number of studies and lack of replication, it was not possible to determine an optimal time for children to spend in NEs. Previous research has estimated optimal rates for adult populations. For instance, White et al. (2019) found that a minimum of 120 min in NE per week was required for positive association to self-reported health and subjective well-being among adults, however, similar study designs are needed in order to determine optimal rates of NE exposure for children.

NE measurements based on surveys to assess surrounding NEs (n = 5; 7%) and expert measures of NEs (n = 6; 9%) were less commonly used and, therefore, we were unable to determine consistency in the results. This does not necessarily mean that they would be less useful or appropriate; it rather calls for replication of studies and efforts to validate and standardize such measurements for increased application. The assessment of an optimal NE measurement was challenged by the large number of different scales or indicators for the various health outcomes included in this review. In the majority of cases, even though most scales were of high validity and reliability, heterogeneous measures were applied across studies.

It is plausible that there may be a stronger association between certain health outcomes and NE metrics based on the pathway examined. For instance, social functioning may be more related to NE as measured by metrics of perception or experience through surveys or on-site evaluations. Comparatively, emotional and behavioral functioning may relate more to ability to play and exercise, such as through measures of use. Before a conclusion about a universal or optimal NE metric for studying associations to childhood mental health and development can be drawn, study replication across different geographical and social contexts and comparisons between different metrics for one defined outcome are required.

In general, our findings of an association between NE exposure and childhood development and mental health are consistent with previous reviews (Browning and Rigolon, 2019; McCormick, 2017; Mygind et al., 2021). For instance, Browning and Rigolon (2019) noted a positive association between NE surrounding schools and academic achievement, for which we found some supporting evidence. Additionally, our findings of increased mental health and development with NE exposure are supported by a review by Mygind et al. (2021). Finally, a recent review by Labib et al. (2019) also concluded that the most common metrics used in studies around NE and health were LULC, NDVI, or a combination of the two.

4.1. Strengths and limitations of included studies

In general, studies of predominantly high quality made efforts to include more than one NE measurement, such as NDVI in combination with high accuracy LULC datasets (e.g., Cusack et al., 2017b; Dadvand et al., 2012a; Markevych et al., 2014a, 2014b). In urbanizing areas with rapid development of commercial and residential areas, substantial land cover changes can occur over short periods of time (Aguilera et al., 2011), therefore, temporal alignment of spatial and health data is of importance to avoid exposure misclassification (Helbich, 2019). This aspect was taken into consideration in the majority of the studies. Additionally, due to variation of naturalness across study areas, it may be possible that some effects go undetected, even in cases of high-resolution exposure metrics.

Only a small number of studies considered quality of NE in their assessment (Feng and Astell-Burt, 2017; Kim et al., 2016; Mårtensson et al., 2009; McEachan et al., 2018; Söderström et al., 2013). The Outdoor Play and Environment Categories (OPEC) is an example of quality assessment that can be applied to assess children's recreational and psychological benefits from NEs (Mårtensson et al., 2009). Another common characteristic for studies of high quality was the accuracy of exposure assignment, in other words, that the NE exposure was linked to the child per accurate home address, rather than on aggregated neighborhood scale. Related to this is whether the exposure is assigned per the child's activity space and use of NE. This was considered in one of the high-quality studies (Dadvand et al., 2015), though it was accounted for in several studies that got an overall score of moderate quality (Amoly et al., 2014; Flouri et al., 2014; Mårtensson et al., 2009; McEachan et al., 2018; Söderström et al., 2013). Other strengths that characterized some of the high-quality studies were appropriate control for a number of adequate confounders (Cusack et al., 2017; Dadvand et al, 2012, 2017; Markevych et al., 2014b), and application of longitudinal or case-control study designs (Dadvand et al, 2015, 2017; Richardson et al., 2017).

A common limitation of the included studies was the risk of bias, for example, residential selection bias effects, mostly a consequence of cross-sectional study design. The self-selection bias means that positive associations may be a result of residential-choice processes so that neighborhoods with an abundance of NEs tend to be more expensive, thus attracting residents of higher income and education who already have healthy habits and lifestyles (Yu and Zhu, 2015). Other limitations include non-validated outcome variables, limited control of confounders, lack of effect estimates, suboptimal statistical approaches (e.g., failure to control for spatial autocorrelation), and failure to account for the actual time period a child had resided in the environment.

4.2. Strengths and limitations of this review

The main strengths of this review are the systematic assessment of various NE metrics and rates (spatial and temporal) and the analysis of relative associative consistency to a wide variety of childhood mental health and developmental outcomes. The inclusion of a large number of studies provided an opportunity to assess a broad set of NE metrics and rates, contributing to fulfill our objective to evaluate association and evidence per different measurements. We used an established method (see Gascon et al., 2017, 2015 for details) for assessing the quality of each study and by compiling the results of the quality assessment, we were able to conclude the evidence level per separate health outcomes. We made substantial efforts to make valid evaluations of the quality of NE metrics in the included studies and, for example, tracked the original sources for LULC datasets to determine accuracy. Article selection, data extraction, and quality assessment were conducted by at least two authors independently.

While we invested in a careful and broad selection of keywords, partly derived from an established and previously used search filter for childhood mental health and development (Desmeules, 2014), and conducted our search across several databases, we did not assess grey literature outside of peer-reviewed journals. This may have resulted in exclusion of some studies, especially for geographic regions not represented in this review, and economically disadvantaged areas which may not have access to the peer-reviewed literature needed to support and justify their own publications. We focused on an inclusive search approach, directed at population rather than outcome, creating a large search yield. This may have potentially resulted in low specificity, although we made substantial efforts to carefully go through the search results, aiming for high sensitivity. We were limited in our search to articles published in English and may thus involuntarily failed to assess studies in other languages. Due to the heterogeneity in study designs, the large variation in types of NE measurements, and many different indicators of health outcomes, we were not able to pool the effect sizes or conduct a meta-analysis. This also prevented us from creating a funnel plot or use other methods to control for publication bias (Mavridis and Salanti, 2014). This may skew the results towards more positive associations compared to if more negative results are published.

The quality assessment and evidence scoring used in this review may inherently have an element of subjectivity. We tried to reduce this subjectivity by replicating methods from previous reviews and quality assessment tools (Gascon et al., 2017; 2015; Higgins and Green, 2011; International Agency for Research on Cancer, 2019) and having each article evaluated by two different authors independently, aiming for as high reliability as possible in our approach. We aimed for following the previously published quality assessment method, although some aspects may potentially be improved in future reviews, such as including aspects of whether the NE metric was appropriate for the study area, whether masking of water bodies or level of cloud free images was appropriately conducted in NDVI calculations, and increase the granularity and specificity in the scaling of, for example, health outcome measurement quality, confounder control, and bias assessment. Nevertheless, in spite of these limitations, the large number of studies included and numerous attempts to avoid assessment bias contribute to the accuracy of our overall assessment of the level of evidence. It is unlikely that this conclusion would change significantly by a quality reclassification of some studies. It should be noted that some of the included studies used overlapping datasets. Amoly et al. (2014) and Dadvand et al. (2015) used children who participated in the BREATHE study; Andrusaityte et al. (2016) and Balseviciene et al. (2014) utilized the Kaunas birth cohort; Bell et al. (2020) and Christian et al. (2017) both used the 2012 Australian Early Development Census (AEDC); and finally, the Millennium Cohort study was used in both Flouri et al. (2014) and Flouri et al. (2018). While utilizing overlapping datasets may contribute to inflated findings, these studies analyzed either different outcome metrics (e.g., Flouri et al. (2014) analyzed emotional and behavioral outcomes whereas Flouri et al. (2018) analyzed spatial working memory) or targeted different populations (e.g., Bell et al. (2020) analyzed a random sample of children in a localized area compared with Christian et al. (2017) who analyzed a regional dataset).

4.3. Recommendations for future work and conclusions

In this systematic review, we found sufficient or limited evidence for an association between NE and several childhood mental health and developmental outcomes, but for a number of outcomes the evidence was insufficient and conclusions for decision making are thus difficult to draw. We have several recommendations for future studies. Firstly, to improve the level of evidence, studies should prioritize elements of high-quality studies in terms of study design and control of bias. For example, the common issue of residential self-selection can, at least partly, be addressed via statistical methods, such as propensity score matching (Crouse et al., 2018). Likewise, studies can harness data on individuals who change their NE exposure over time due to, for example, moving (van den Bosch et al., 2015) or natural changes in the environment, such as tree loss due to pest infestation (Donovan et al., 2015).

Secondly, we also suggest replication of studies and consistency in both NE and health measurements to provide options for comparability and identification of effect sizes across populations and regions. Future studies should also aim for considering length of exposure by including data of residential history of the participants to identify cumulative effects (Helbich, 2018). To date, most studies have used residential address as a proxy for exposure and we recommend that future studies aim for accessing and incorporating a child's daily exposures at, for example, daycare facilities and schools, to better capture the full extent of children's living environments.

Thirdly, to address causality, longitudinal study designs, randomized controlled trials, or natural experiments should be prioritized. In addition, biological mechanisms for effects of NE on childhood mental health and development needs to be investigated. This would need to include collection of biomarkers, such as cortisol to measure stress, or brain imaging techniques to evaluate neurological responses to environmental exposures (Tost et al., 2019). However, these kinds of studies are often costly, thus raising challenges for individual researchers. Additionally, innovative techniques for ambulatory assessments, including mobile devices for monitoring of momentary psychological and physiological reactions, in combination with environmental parameters, including GPS and accelerometers for monitoring physical behavior and activity, should be further refined and developed for managing and analysis of big data. This would provide longitudinal data and opportunities for "real-life", "real-time" studies, increasing the ecological validity and thereby study quality and evidence level (Reichert et al., 2020).

While aiming for consistency among measurements, it is important to further improve the knowledge by novel measurements and inclusion of different user perspectives of the environment by making use of, for example, volunteered geographic information, including the collection of environmental perception data from social media or participatory GIS (Brown et al., 2017; Reichert et al., 2020; Schwartz et al., 2019). These kinds of techniques are still in their infancy, but findings from existing studies are promising (Almanza et al., 2012; Epstein et al., 2014). A recently developed NE measurement reflects human perception at a vertical level by creating green view indices from Google Street View (GSV) that (Li et al., 2015; Nutsford et al., 2015), likely capturing a different component of NE exposure compared to, for example, remote sensing products (Larkin and Hystad, 2018). GSV datasets are becoming increasingly available, but to our knowledge, no studies on childhood health have harnessed on this technique. In parallel with development and incorporation of new, possibly more accurate, NE exposure measurements, we would suggest to regularly include NDVI in forthcoming studies, at least for sensitivity analyses, to provide comparable material and opportunities for assessing effect sizes across studies.

Finally, findings from this review reflect the unequal distribution of research resources by demonstrating a dominance of studies from the Minority World (i.e., high income countries, mainly North America, Western Europe, and Australasia) while research from the Majority World (i.e., low- and middle-income countries), where the most rapidly urbanizing countries exist, is lacking (Punch and Tisdall, 2012). Ultimately, many of the suggested improvements are dependent on resources both for the individual researcher and for institutions. Therefore, it is imperative that funders and health agencies recognize the importance of investing in knowledge production around health promotion, by for example healthy environments, in research grants and investments, especially in the majority world. It is particularly important to understand what factors contribute to improved health of children, given the life course impacts (Gluckman et al, 2007, 2016). If we are to fully realize the importance of NEs in urban areas in different regions of the world, this type of research needs to be strongly prioritized, and only then can we provide sufficient evidence for decisions around healthy urban planning where every child can live, play, and thrive.

Funding

This project received funding from the Canadian Institutes of Health Research [grant number 156152]. The funders played no role in study design, collection, analysis, interpretation of data, writing of the report, or in the decision to submit the paper for publication. They accept no responsibility for the contents. The authors have no conflicts of interest.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Colleen Pawliuk, librarian at the BC Children's Hospital Research Institute (BCCHRI), for her advice and support during the search process. We also thank the two anonymous reviewers for their thorough review of our manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2021.113767.

References

Aguilera, F., Valenzuela, L.M., Botequilha-Leitão, A., 2011. Landscape metrics in the analysis of urban land use patterns: a case study in a Spanish metropolitan area. Landsc. Urban Plann. 99, 226–238. https://doi.org/10.1016/j.landurbplan.2010.10.004.

Almanza, E., Jerrett, M., Dunton, G., Seto, E., Ann Pentz, M., 2012. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. Health Place 18, 46–54. https://doi.org/10.1016/j.healthplace.2011.09.003.

Amoly, E., Dadvand, P., Forns, J., Lopez-Vicente, M., Basagana, X., Julvez, J., Alvarez-Pedrerol, M., Nieuwenhuijsen, M.J., Sunyer, J., 2014. Green and blue spaces and behavioral development in Barcelona schoolchildren: the BREATHE project. Environ. Health Perspect. 122, 1351–1358. https://doi.org/10.1289/ehp.1408215.

Andrusaityte, S., Grazuleviciene, R., Dedele, A., Balseviciene, B., 2020. The effect of residential greenness and city park visiting habits on preschool Children's mental and general health in Lithuania: a cross-sectional study. Int. J. Hyg Environ. Health 223, 142–150. https://doi.org/10.1016/j.ijheh.2019.09.009.

Andrusaityte, S., Grazuleviciene, R., Kudzyte, J., Bernotiene, A., Dedele, A., Nieuwenhuijsen, M.J., 2016. Associations between neighbourhood greenness and asthma in preschool children in Kaunas, Lithuania: a case-control study. Bmj Open 6, e010341. https://doi.org/10.1136/bmjopen-2015-010341.

Balseviciene, B., Sinkariova, L., Grazuleviciene, R., Andrusaityte, S., Uzdanaviciute, I., Dedele, A., Nieuwenhuijsen, M.J., 2014. Impact of residential greenness on preschool children's emotional and behavioral problems. Int. J. Environ. Res. Publ. Health 11, 6757–6770. https://doi.org/10.3390/ijerph110706757.

Beere, P., Kingham, S., 2017. Assessing the relationship between greenspace and academic achievement in urban New Zealand primary schools. N. Z. Geogr. 73, 155–165. https://doi.org/10.1111/nzg.12155.

Bell, M.F., Turrell, G., Beesley, B., Boruff, B., Trapp, G., Zubrick, S.R., Christian, H.E., 2020. Children's neighbourhood physical environment and early development: an individual child level linked data study. J. Epidemiol. Community Health 74, 321–329. https://doi.org/10.1136/jech-2019-212686.

Brown, G., Strickland-Munro, J., Kobryn, H., Moore, S.A., 2017. Mixed methods participatory GIS: an evaluation of the validity of qualitative and quantitative mapping methods. Appl. Geogr. 79, 153–166. https://doi.org/10.1016/j.apgeog.2016.12.015.

Browning, M., Lee, K., 2017. Within what distance does "greenness" best predict physical health? A systematic review of articles with GIS buffer analyses across the lifespan. Int. J. Environ. Res. Publ. Health 14, 675. https://doi.org/10.3390/ijerph14070675.

Browning, M.H.E.M., Kuo, M., Sachdeva, S., Lee, K., Westphal, L., 2018. Greenness and school-wide test scores are not always positively associated – a replication of "linking student performance in Massachusetts elementary schools with the 'greenness' of school surroundings using remote sensing.". Landsc. Urban Plann. 178, 69–72. https://doi.org/10.1016/j.landurbplan.2018.05.007.

Browning, M.H.E.M., Rigolon, A., 2019. School green space and its impact on academic performance: a systematic literature review. Int. J. Environ. Res. Publ. Health 16, 429. https://doi.org/10.3390/ijerph16030429.

Christian, H., Ball, S.J., Zubrick, S.R., Brinkman, S., Turrell, G., Boruff, B., Foster, S., 2017. Relationship between the neighbourhood built environment and early child development. Health Place 48, 90–101. https://doi.org/10.1016/j.healthplace.2017.08.010.

Christian, H., Zubrick, S.R., Foster, S., Giles-Corti, B., Bull, F., Wood, L., Knuiman, M., Brinkman, S., Houghton, S., Boruff, B., 2015. The influence of the neighborhood physical environment on early child health and development: a review and call for research. Health Place 33, 25–36. https://doi.org/10.1016/j.healthplace.2015.01.005.

Crouse, D.L., Balram, A., Hystad, P., Pinault, L., van den Bosch, M., Chen, H., Rainham, D., Thomson, E.M., Close, C.H., van Donkelaar, A., Martin, R.V., Ménard, R., Robichaud, A., Villeneuve, P.J., 2018. Associations between living near water and risk of mortality among urban Canadians. Environ. Health Perspect. 126, 077008 https://doi.org/10.1289/EHP3397.

Cusack, L., Larkin, A., Carozza, S.E., Hystad, P., 2017. Associations between multiple green space measures and birth weight across two US cities. Health Place 47, 36–43. https://doi.org/10.1016/j.healthplace.2017.07.002.

Dadvand, P., de Nazelle, A., Figueras, F., Basagana, X., Su, J., Amoly, E., Jerrett, M., Vrijheid, M., Sunyer, J., Nieuwenhuijsen, M.J., 2012. Green space, health inequality and pregnancy. Environ. Int. 1, 110–115. https://doi.org/10.1016/j.envint.2011.07.004.

Dadvand, P., Nieuwenhuijsen, M.J., Esnaola, M., Forns, J., Basagana, X., Alvarez-Pedrerol, M., Rivas, I., Lopez-Vicente, M., De Castro Pascual, M., Su, J., Jerrett, M., Querol, X., Sunyer, J., 2015. Green spaces and cognitive development in primary schoolchildren. Proc. Natl. Acad. Sci. U. S. A. 112, 7937–7942. https://doi.org/10.1073/pnas.1503402112.

Dadvand, P., Tischer, C., Estarlich, M., Llop, S., Dalmau-Bueno, A., Lopez-Vicente, M., Valentin, A., de Keijzer, C., Fernandez-Somoano, A., Lertxundi, N., Rodriguez-Dehli, C., Gascon, M., Guxens, M., Zugna, D., Basagana, X., Nieuwenhuijsen, M.J., Ibarluzea, J., Ballester, F., Sunyer, J., 2017. Lifelong residential exposure to green space and attention: a population-based prospective study. Environ. Health Perspect. 125, 097016 https://doi.org/10.1289/EHP694.

Desmeules, R., 2014. Filter to Retrieve Pediatric Articles in the OVIDMedline Database. John W. Scott Health Sciences Library, University of Alberta.

Donovan, G.H., Michael, Y.L., Gatziolis, D., Mannetje, A. 't, Douwes, J., 2019. Association between exposure to the natural environment, rurality, and attention-deficit hyperactivity disorder in children in New Zealand: a linkage study. The Lancet Planetary Health 3, e226–e234. https://doi.org/10.1016/S2542-5196(19)30070-1.

Donovan, G.H., Michael, Y.L., Gatziolis, D., Prestemon, J.P., Whitsel, E.A., 2015. Is tree loss associated with cardiovascular-disease risk in the Women's Health Initiative? A natural experiment. Health Place 36, 1–7. https://doi.org/10.1016/j.healthplace.2015.08.007.

Dzhambov, A.M., Dimitrova, D.D., Dimitrakova, E.D., 2014. Association between residential greenness and birth weight: systematic review and meta-analysis. Urban For. Urban Green. 13, 621–629. https://doi.org/10.1016/j.ufug.2014.09.004.

Epstein, D.H., Tyburski, M., Craig, I.M., Phillips, K.A., Jobes, M.L., Vahabzadeh, M., Mezghanni, M., Lin, J.-L., Furr-Holden, C.D.M., Preston, K.L., 2014. Real-time tracking of neighborhood surroundings and mood in urban drug misusers: application of a new method to study behavior in its geographical context. Drug Alcohol Depend. 134, 22–29. https://doi.org/10.1016/j.drugalcdep.2013.09.007.

Faber Taylor, A., Kuo, F.E., 2009. Children with attention deficits concentrate better after walk in the park. J. Atten. Disord. 12, 402–409. https://doi.org/10.1177/1087054708323000.

Faber Taylor, A., Kuo, F.E., Sullivan, W.C., 2002. Views of nature and self-discipline: evidence from inner city children. J. Environ. Psychol. 22, 49–63. https://doi.org/10.1006/jevp.2001.0241.

Faber Taylor, A., Kuo, F.E., Sullivan, W.C., 2001. Coping with ADD: the surprising connection to green play settings. Environ. Behav. 33, 54–77. https://doi.org/10.1177/00139160121972864.

Faber Taylor, A., Kuo, F.E.M., 2011. Could exposure to everyday green spaces help treat ADHD? Evidence from children's play settings: everyday green space and ADHD symptoms. Appl. Psychol.: Health and Well-Being 3, 281–303. https://doi.org/10.1111/j.1758-0854.2011.01052.x.

Feng, X., Astell-Burt, T., 2017. Residential green space quantity and quality and child well-being: a longitudinal study. Am. J. Prev. Med. 53, 616–624. https://doi.org/10.1016/j.amepre.2017.06.035.

Finegold, Y., Ortmann, A., Lindquist, E., d'Annunzio, R., Sandker, M., 2016. Map Accuracy Assessment and Area Estimation: A Practical Guide (No. Working Paper No. 46/E), National Forest Mointoring Assessment. Food and Agriculture Organization of the United Nations, Rome, Italy.

Flouri, E., Midouhas, E., Joshi, H., 2014. The role of urban neighbourhood green space in children's emotional and behavioural resilience. J. Environ. Psychol. 40, 179–186. https://doi.org/10.1016/j.jenvp.2014.06.007.

Flouri, E., Papachristou, E., Midouhas, E., 2018. The role of neighbourhood greenspace in children's spatial working memory. Br. J. Educ. Psychol. https://doi.org/10.1111/bjep.12243.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Rem. Sens. Environ. 80, 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4.

Gascon, M., Cirach, M., Martinez, D., Dadvand, P., Valentin, A., Plasencia, A., Nieuwenhuijsen, M., 2016. Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: the case of Barcelona City. Urban For. Urban Green. 19, 88–94. https://doi.org/10.1016/j.ufug.2016.07.001.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., Nieuwenhuijsen, M., 2015. Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. Int. J. Environ. Res. Publ. Health 12, 4354–4379. https://doi.org/10.3390/ijerph120404354.

Gascon, M., Zijlema, W., Vert, C., White, M.P., Nieuwenhuijsen, M.J., 2017. Outdoor blue spaces, human health and well-being: a systematic review of quantitative studies. Int. J. Hyg Environ. Health 220, 1207–1221. https://doi.org/10.1016/j.ijheh.2017.08.004.

Gluckman, P.D., Buklijas, T., Hanson, M.A., 2016. The developmental origins of health and disease (DOHaD) concept. In: The Epigenome and Developmental Origins of Health and Disease. Elsevier, pp. 1–15. https://doi.org/10.1016/B978-0-12-801383-0.00001-3.

Gluckman, P.D., Hanson, M.A., Beedle, A.S., 2007. Early life events and their consequences for later disease: a life history and evolutionary perspective. Am. J. Hum. Biol. 19, 1–19. https://doi.org/10.1002/ajhb.20590.

Gray, C., Gibbons, R., Larouche, R., Sandseter, E., Bienenstock, A., Brussoni, M., Chabot, G., Herrington, S., Janssen, I., Pickett, W., Power, M., Stanger, N., Sampson, M., Tremblay, M., 2015. What is the relationship between outdoor time and physical activity, sedentary behaviour, and physical fitness in children? A systematic review. Int. J. Environ. Res. Publ. Health 12, 6455–6474. https://doi.org/10.3390/ijerph120606455.

Hartig, T., Mitchell, R., de Vries, S., Frumkin, H., 2014. Nature and health. Annu. Rev. Publ. Health 35, 207–228. https://doi.org/10.1146/annurev-publhealth-032013-182443.

Hay, S.I., 2000. An overview of remote sensing and geodesy for epidemiology and public health application. Adv. Parasitol. 47, 1–35.

Helbich, M., 2019. Spatiotemporal contextual uncertainties in green space exposure measures: exploring a time series of the normalized difference vegetation indices. Int. J. Environ. Res. Publ. Health 16, 852. https://doi.org/10.3390/ijerph16050852.

Helbich, M., 2018. Toward dynamic urban environmental exposure assessments in mental health research. Environ. Res. 161, 129–135. https://doi.org/10.1016/j.envres.2017.11.006.

Higgins, J.P., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions.

Hodson, C.B., Sander, H.A., 2017. Green urban landscapes and school-level academic performance. Landsc. Urban Plann. 160, 16–27. https://doi.org/10.1016/j.landurbplan.2016.11.011.

International Agency for Research on Cancer, I.A. for R. on C, 2019. IARC Monographs on the Identification of Carcinogenic Hazards to Humans: Preamble [WWW Document]. IARC Monographs on the Identification of Carcinogenic Hazards to Humans. URL. https://monographs.iarc.fr/wp-content/uploads/2019/01/Preamble-2019.pdf (accessed 11.11.19).

Jarvis, I., Gergel, S., Koehoorn, M., 2020. Greenspace access does not correspond to nature exposure: measures of urban natural space with implications for health research. Landsc. Urban Plann. 13.

Kim, J.-H., Lee, C., Sohn, W., 2016. Urban natural environments, obesity, and health-related quality of life among hispanic children living in inner-city neighborhoods. Int. J. Environ. Res. Publ. Health 13, 121. https://doi.org/10.3390/ijerph13010121.

Kuo, F.E., Faber Taylor, A., 2004. A potential natural treatment for attention-deficit/hyperactivity disorder: evidence from a national study. Am. J. Publ. Health 94, 1580–1586.

Kuo, M., Browning, M.H.E.M., Sachdeva, S., Lee, K., Westphal, L., 2018. Might school performance grow on trees? Examining the link between "greenness" and academic achievement in urban, high-poverty schools. Front. Psychol. 9 https://doi.org/10.3389/fpsyg.2018.01669.

Kweon, B.-S., Ellis, C.D., Lee, J., Jacobs, K., 2017. The link between school environments and student academic performance. Urban For. Urban Green. 23, 35–43. https://doi.org/10.1016/j.ufug.2017.02.002.

Labib, S.M., Huck, J.J., Lindley, S., 2020. Modelling and mapping eye-level greenness visibility exposure using multi-source data at high spatial resolutions. Sci. Total Environ. 143050. https://doi.org/10.1016/j.scitotenv.2020.143050.

Labib, S.M., Lindley, S., Huck, J.J., 2019. Spatial dimensions of the influence of urban green-blue spaces on human health: a systematic review. Environ. Res. 108869. https://doi.org/10.1016/j.envres.2019.108869.

Larkin, A., Hystad, P., 2018. Evaluating street view exposure measures of visible green space for health research. J. Expo. Sci. Environ. Epidemiol. 29, 447–456. https://doi.org/10.1038/s41370-018-0017-1.

Larson, L.R., Barger, B., Ogletree, S., Torquati, J., Rosenberg, S., Gaither, C.J., Bartz, J. M., Gardner, A., Moody, E., Schutte, A., 2018. Gray space and green space proximity associated with higher anxiety in youth with autism. Health Place 53, 94–102. https://doi.org/10.1016/j.healthplace.2018.07.006.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For. Urban Green. 14, 675–685. https://doi.org/10.1016/j.ufug.2015.06.006.

Liao, J., Yang, S., Xia, W., Peng, A., Zhao, J., Li, Y., Zhang, Y., Qian, Z., Vaughn, M.G., Schootman, M., Zhang, B., Xu, S., 2020. Associations of exposure to green space with problem behaviours in preschool-aged children. Int. J. Epidemiol. 49, 944–953. https://doi.org/10.1093/ije/dyz243.

Liao, J., Zhang, B., Xia, W., Cao, Z., Zhang, Y., Liang, S., Hu, K., Xu, S., Li, Y., 2019. Residential exposure to green space and early childhood neurodevelopment. Environ. Int. 128, 70–76. https://doi.org/10.1016/j.envint.2019.03.070.

Macintyre, S., Macdonald, L., Ellaway, A., 2008. Lack of agreement between measured and self-reported distance from public green parks in Glasgow, Scotland. Int. J. Behav. Nutr. Phys. Activ. 5, 26. https://doi.org/10.1186/1479-5868-5-26.

MacNaughton, P., Eitland, E., Kloog, I., Schwartz, J., Allen, J., 2017. Impact of particulate matter exposure and surrounding "greenness" on chronic absenteeism in Massachusetts public schools. Int. J. Environ. Res. Publ. Health 14. https://doi.org/10.3390/ijerph14020207.

Madzia, J., Ryan, P., Yolton, K., Percy, Z., Newman, N., LeMasters, G., Brokamp, C., 2019. Residential greenspace association with childhood behavioral outcomes. J. Pediatr. 207, 233–240. https://doi.org/10.1016/j.jpeds.2018.10.061.

Markevych, I., Fuertes, E., Tiesler, C.M.T., Birk, M., Bauer, C.-P., Koletzko, S., von Berg, A., Berdel, D., Heinrich, J., 2014a. Surrounding greenness and birth weight: results from the GINIplus and LISAplus birth cohorts in Munich. Health Place 1, 39–46. https://doi.org/10.1016/j.healthplace.2013.12.001.

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A.M., de Vries, S., Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M.J., Lupp, G., Richardson, E.A., Astell-Burt, T., Dimitrova, D., Feng, X., Sadeh, M., Standl, M., Heinrich, J., Fuertes, E., 2017. Exploring pathways linking greenspace to health: theoretical and methodological guidance. Environ. Res. 158, 301–317. https://doi.org/10.1016/j.envres.2017.06.028.

Markevych, I., Tesch, F., Datzmann, T., Romanos, M., Schmitt, J., Heinrich, J., 2018. Outdoor air pollution, greenspace, and incidence of ADHD: a semi-individual study. Sci. Total Environ. 642, 1362–1368. https://doi.org/10.1016/j.scitotenv.2018.06.167.

Markevych, I., Tiesler, C.M.T., Fuertes, E., Romanos, M., Dadvand, P., Nieuwenhuijsen, M.J., Berdel, D., Koletzko, S., Heinrich, J., 2014b. Access to urban

green spaces and behavioural problems in children: results from the GINIplus and LISAplus studies. Environ. Int. 71, 29–35. https://doi.org/10.1016/j.envint.2014.06.002.
Mårtensson, F., Boldemann, C., Söderström, M., Blennow, M., Englund, J.-E., Grahn, P., 2009. Outdoor environmental assessment of attention promoting settings for preschool children. Health Place 15, 1149–1157. https://doi.org/10.1016/j.healthplace.2009.07.002.
Mavridis, D., Salanti, G., 2014. Exploring and accounting for publication bias in mental health: a brief overview of methods. Evid. Base Ment. Health 17, 11–15. https://doi.org/10.1136/eb-2013-101700.
Mazumdar, S., Dunshea, A., Chong, S., Jalaludin, B., 2020. Tree canopy cover is best associated with perceptions of greenspace: a short communication. IJERPH 17, 6501. https://doi.org/10.3390/ijerph17186501.
McCormick, R., 2017. Does access to green space impact the mental well-being of children: a systematic review. J. Pediatr. Nurs. 37, 3–7. https://doi.org/10.1016/j.pedn.2017.08.027.
McCracken, D.S., Allen, D.A., Gow, A.J., 2016. Associations between urban greenspace and health-related quality of life in children. Prev Med Rep 3, 211–221. https://doi.org/10.1016/j.pmedr.2016.01.013.
McCrorie, P.R., Fenton, C., Ellaway, A., 2014. Combining GPS, GIS, and accelerometry to explore the physical activity and environment relationship in children and young people - a review. Int. J. Behav. Nutr. Phys. Activ. 11, 93. https://doi.org/10.1186/s12966-014-0093-0.
McEachan, R.R.C., Yang, T.C., Roberts, H., Pickett, K.E., Arseneau-Powell, D., Gidlow, C. J., Wright, J., Nieuwenhuijsen, M., 2018. Availability, use of, and satisfaction with green space, and children's mental wellbeing at age 4 years in a multicultural, deprived, urban area: results from the Born in Bradford cohort study. The Lancet Planetary Health 2, e244–e254. https://doi.org/10.1016/S2542-5196(18)30119-0.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6, 6.
Morgan, R.L., Whaley, P., Thayer, K.A., Schünemann, H.J., 2018. Identifying the PECO: a framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environ. Int. 121, 1027–1031. https://doi.org/10.1016/j.envint.2018.07.015.
Mygind, L., Kurtzhals, M., Nowell, C., Melby, P.S., Stevenson, M.P., Nieuwenhuijsen, M., Lum, J.A.G., Flensborg-Madsen, T., Bentsen, P., Enticott, P.G., 2021. Landscapes of becoming social: a systematic review of evidence for associations and pathways between interactions with nature and socioemotional development in children. Environ. Int. 146, 106238. https://doi.org/10.1016/j.envint.2020.106238.
Nutsford, D., Reitsma, F., Pearson, A.L., Kingham, S., 2015. Personalising the viewshed: visibility analysis from the human perspective. Appl. Geogr. 62, 1–7. https://doi.org/10.1016/j.apgeog.2015.04.004.
Pettorelli, N., Schulte to Bühne, H., Shaprio, A., Glover-Kapfer, P., 2018. Satellite Remote Sensing for Conservation, vol. 4. [WWF] World Wildlife Fund Conservation Technology Series, Switzerland.
Poore, K.R., Hanson, M.A., Faustman, E.M., Neira, M., 2017. Avoidable early life environmental exposures. The Lancet Planetary Health 1, e172–e173. https://doi.org/10.1016/S2542-5196(17)30048-7.
Punch, S., Tisdall, E.K.M., 2012. Exploring children and young people's relationships across Majority and Minority Worlds. Child Geogr. 10, 241–248. https://doi.org/10.1080/14733285.2012.693375.
Reichert, M., Braun, U., Lautenbach, S., Zipf, A., Ebner-Priemer, U., Tost, H., Meyer-Lindenberg, A., 2020. Studying the impact of built environments on human mental health in everyday life: methodological developments, state-of-the-art and technological frontiers. Current Opinion in Psychology, Socio-Ecological Psychology 32, 158–164. https://doi.org/10.1016/j.copsyc.2019.08.026.
Reuben, A., Arseneault, L., Belsky, D.W., Caspi, A., Fisher, H.L., Houts, R.M., Moffitt, T. E., Odgers, C., 2019. Residential neighborhood greenery and children's cognitive development. Soc. Sci. Med. 230, 271–279. https://doi.org/10.1016/j.socscimed.2019.04.029.
Rhew, I.C., Vander Stoep, A., Kearney, A., Smith, N.L., Dunbar, M.D., 2011. Validation of the normalized difference vegetation index as a measure of neighborhood greenness. Ann. Epidemiol. 21, 946–952. https://doi.org/10.1016/j.annepidem.2011.09.001.
Richardson, E.A., Pearce, J., Shortt, N.K., Mitchell, R., 2017. The role of public and private natural space in children's social, emotional and behavioural development in Scotland: a longitudinal study. Environ. Res. 158, 729–736. https://doi.org/10.1016/j.envres.2017.07.038.
Russell, R., Chung, M., Balk, E.M., Atkinson, S., Giovannucci, E.L., Ip, S., Mayne, S.T., Raman, G., Ross, A.C., Trikalinos, T., Keith P West, J., Lau, J., 2009. Issues and Challenges in Conducting Systematic Review to Support Development of Nutrient Reference Values: Workshop Summary: Nutrition Research Series, vol. 2. Agency for Healthcare Research and Quality (US).
Schutte, A.R., Torquati, J.C., Beattie, H.L., 2017. Impact of urban nature on executive functioning in early and middle childhood. Environ. Behav. 49, 3–30. https://doi.org/10.1177/0013916515603095.
Schwartz, A.J., Dodds, P.S., O'Neil-Dunne, J.P.M., Danforth, C.M., Ricketts, T.H., 2019. Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. People and Nature pan 3, 10045. https://doi.org/10.1002/pan3.10045.
Sivarajah, S., Smith, S.M., Thomas, S.C., 2018. Tree cover and species composition effects on academic performance of primary school students. PloS One 13, e0193254. https://doi.org/10.1371/journal.pone.0193254.
Smith, M., Cui, J., Ikeda, E., Mavoa, S., Hasanzadeh, K., Zhao, J., Rinne, T.E., Donnellan, N., Kyttä, M., 2021. Objective measurement of children's physical activity geographies: a systematic search and scoping review. Health Place 67, 102489. https://doi.org/10.1016/j.healthplace.2020.102489.
Söderström, M., Boldemann, C., Sahlin, U., Mårtensson, F., Raustorp, A., Blennow, M., 2013. The quality of the outdoor environment influences childrens health - a cross-sectional study of preschools. Acta Paediatr. 102, 83–91. https://doi.org/10.1111/apa.12047.
Tallis, H., Bratman, G.N., Samhouri, J.F., Fargione, J., 2018. Are California elementary school test scores more strongly associated with urban trees than poverty? Front. Psychol. 9, 2074. https://doi.org/10.3389/fpsyg.2018.02074.
Tamayo-Uria, I., Maitre, L., Thomsen, C., Nieuwenhuijsen, M.J., Chatzi, L., Siroux, V., Aasvang, G.M., Agier, L., Andrusaityte, S., Casas, M., de Castro, M., Dedele, A., Haug, L.S., Heude, B., Grazuleviciene, R., Gutzkow, K.B., Krog, N.H., Mason, D., McEachan, R.R.C., Meltzer, H.M., Petraviciene, I., Robinson, O., Roumeliotaki, T., Sakhi, A.K., Urquiza, J., Vafeiadi, M., Waiblinger, D., Warembourg, C., Wright, J., Slama, R., Vrijheid, M., Basagaña, X., 2019. The early-life exposome: description and patterns in six European countries. Environ. Int. 123, 189–200. https://doi.org/10.1016/j.envint.2018.11.067.
Tillmann, S., Clark, A.F., Gilliland, J.A., 2018. Children and nature: linking accessibility of natural environments and children's health-related quality of life. Int. J. Environ. Res. Publ. Health 15. https://doi.org/10.3390/ijerph15061072.
Tost, H., Reichert, M., Braun, U., Reinhard, I., Peters, R., Lautenbach, S., Hoell, A., Schwarz, E., Ebner-Priemer, U., Zipf, A., Meyer-Lindenberg, A., 2019. Neural correlates of individual differences in affective benefit of real-life urban green space exposure. Nat. Neurosci. 22, 1389–1393. https://doi.org/10.1038/s41593-019-0451-y.
van den Bosch, M., Östergren, P.-O., Grahn, P., Skärbäck, E., Währborg, P., 2015. Moving to serene nature may prevent poor mental health—results from a Swedish longitudinal cohort study. Int. J. Environ. Res. Publ. Health 12, 7974–7989. https://doi.org/10.3390/ijerph120707974.
Vanaken, G.-J., Danckaerts, M., 2018. Impact of green space exposure on children's and adolescents' mental health: a systematic review. IJERPH 15, 2668. https://doi.org/10.3390/ijerph15122668.
Ward, J.S., Duncan, J.S., Jarden, A., Stewart, T., 2016. The impact of children's exposure to greenspace on physical activity, cognitive development, emotional wellbeing, and ability to appraise risk. Health Place 40, 44–50. https://doi.org/10.1016/j.healthplace.2016.04.015.
Wells, N.M., 2000. At home with nature: effects of "greenness" on children's cognitive functioning. Environ. Behav. 32, 775–795. https://doi.org/10.1177/00139160021972793.
Wells, N.M., Evans, G.W., 2003. Nearby nature: a buffer of life stress among rural children. Environ. Behav. 35, 311–330. https://doi.org/10.1177/0013916503035003001.
White, M.P., Alcock, I., Grellier, J., Wheeler, B.W., Hartig, T., Warber, S.L., Bone, A., Depledge, M.H., Fleming, L.E., 2019. Spending at least 120 minutes a week in nature is associated with good health and wellbeing. Sci. Rep. 9, 7730. https://doi.org/10.1038/s41598-019-44097-3.
Wu, C.-D., McNeely, E., Cedeño-Laurent, J.G., Pan, W.-C., Adamkiewicz, G., Dominici, F., Lung, S.-C.C., Su, H.-J., Spengler, J.D., 2014. Linking student performance in Massachusetts elementary schools with the "greenness" of school surroundings using remote sensing. PloS One 9, e108548. https://doi.org/10.1371/journal.pone.0108548.
Wu, J., Jackson, L., 2017. Inverse relationship between urban green space and childhood autism in California elementary school districts. Environ. Int. 107, 140–146. https://doi.org/10.1016/j.envint.2017.07.010.
Yang, B.-Y., Zeng, X.-W., Markevych, I., Bloom, M.S., Heinrich, J., Knibbs, L.D., Dharmage, S.C., Lin, S., Jalava, P., Guo, Y., Jalaludin, B., Morawska, L., Zhou, Y., Hu, L.-W., Yu, H.-Y., Yu, Y., Dong, G.-H., 2019. Association between greenness surrounding schools and kindergartens and attention-deficit/hyperactivity disorder in children in China. JAMA Netw Open 2, e1917862. https://doi.org/10.1001/jamanetworkopen.2019.17862.
Yu, C.-Y., Zhu, X., 2015. Impacts of residential self-selection and built environments on children's walking-to-school behaviors. Environ. Behav. 47, 268–287. https://doi.org/10.1177/0013916513500959.