Out of the

Institute and Clinic for Occupational, Social and Environmental Medicine, University Hospital

Spatial Modelling for Health and Exposure Data- Child stunting and Blood Lead in Kabwe, Zambia.

Doctoral Thesis

for the awarding of a Doctor of Philosophy (Ph.D.)

at the Medical Faculty of

Ludwig-Maximilians-Universität, Munich

submitted by

Given Moonga

born in

Choma, Zambia

submitted in

2022.

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

First evaluator (1. TAC member): Prof. Dr. med. Dennis Nowak
Second evaluator (2. TAC member): Prof. Dr. med. Stephan Boese-O'Reilly
Third evaluator: Prof. Dr. Bernhard Liebl
Fourth evaluator: Prof. Dr. Berthold Koletzko

Dean: **Prof. Dr. med. Thomas Gudermann**

Date of the defense:

05 December 2022

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Dean's Office
Medical Faculty

MMRS
MUNICH MEDICAL RESEARCH SCHOOL

Affidavit

Moonga Given
Surname, first name

Street

Zip code, town

Zambia
Country

I hereby declare, that the submitted thesis entitled

Spatial Modelling for Health and Exposure Data- Child stunting and Blood Lead in Kabwe, Zambia.

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Munich, 16 December 2022
Place, date

Moonga Given
Signature doctoral candidate

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Dean's Office
Medical Faculty

Confirmation of congruency between printed and electronic version of the doctoral thesis

Moonga Given

Surname, first name

Street

Zip code, town

Zambia

Country

I hereby declare that the electronic version of the submitted thesis, entitled

Spatial Modelling for Health and Exposure Data- Child stunting and Blood Lead in Kabwe, Zambia.

is congruent with the printed version both in content and format.

Munich, 16 December 2022

Place, date

Moonga Given

Signature doctoral candidate

Key Words

Spatial, Georeferenced, Modelling, Stunting, Lead, Zambia

Contents

Contents

List of Abbreviations

BDR	Bayesian Distributional regression
CNS	Central Nervous System
GIS	Geographic Information Systems
HMIS	Health Management Information Systems
MAP	Malaria Atlas Project
NCD	Non-Communicable Disease
SDG	Sustainable Development Goals
SES	Social Economic Status
WHO	World Health Organization
ZDHS	Zambia demographic and Health Survey

List of Publications

Moonga, G., Böse-O'Reilly, S., Berger, U., Harttgen, K., Michelo, C., Nowak, D., Siebert, U., Yabe, J. and Seiler, J., 2021. Modelling chronic malnutrition in Zambia: A Bayesian distributional regression approach. Plos one, 16(8), p.e0255073.

Moonga, G., Chisola, M.N., Berger, U., Nowak, D., Yabe, J., Nakata, H., Nakayama, S., Ishizuka, M. and Bose-O'Reilly, S., 2022. Geospatial approach to investigate spatial clustering and hotspots of blood lead levels in children within Kabwe, Zambia. Environmental research, 207, p.112646.

Rakete, S., **Moonga, G**., Wahl, A.M., Mambrey, V., Shoko, D., Moyo, D., Muteti-Fana, S., Tobollik, M., Steckling-Muschack, N. and Bose-O'Reilly, S., 2022. Biomonitoring of arsenic, cadmium and lead in two artisanal and small-scale gold mining areas in Zimbabwe. Environmental Science and Pollution Research, 29(3), pp.4762-4768.

My Contribution to the Publications

Contribution to Paper I

The PhD candidate created the initial study concept and research proposal. The proposal was distributed to the rest of the supervisory team for review. The candidate wrote a formal request for third-party data after receiving approval and ethical clearance for the study. Following the collection of all data, the candidate was in charge of data cleaning, analysis, and interpretation. The candidate then presented the study findings to the supervisory team. The first manuscript was then drafted and reviewed by the rest of the team before being submitted for publication. This process was overseen by the candidate, who took into account feedback from the co-authors. When the manuscript was complete, the candidate was designated as the corresponding author and was in charge of the publication process.

The candidate was also in charge of overseeing the entire project plan and providing bimonthly updates on the status of the work.

Contribution to Paper II

The second manuscript's dataset came from a project in which the candidate was involved. The candidate asked for permission to use project data. Following approval, the candidate proceeded with data cleaning, analysis, and result interpretation. The findings were compiled and presented to the supervisory team for review. Following that, the candidate wrote the manuscript with the help of the supervisory team. The candidate submitted the manuscript for publication as the first author after finishing it. Until the article was published, the candidate handled all publication queries with input from the supervisory team.

Contribution to Paper III (Appendix)

For the third manuscript, the candidate was in charge of data analysis and interpretation. The candidate also took part in responding to journal reviews.

CHAPTER 1

INTRODUCTORY SUMMARY

1.1 Introduction

The global burden of malnutrition remains unacceptably high. Malnutrition is estimated to contribute to more than one-third of all child deaths, with more than half of children under the age of five stunted, and one in five underweight. These proportions represent a global burden of 150 million stunted children, 50.5 million wasted, 38.3 million overweight, and 38.3 million over- weight (World Health Organization, 2018). This is despite the "Sustainable Development Goals" (SDG) target of eliminating all forms of malnutrition by 2030 (SDG, 2018, Alao et al., 2021).

Undernutrition is most prevalent in low and middle-income countries, whereas overweight and obesity are more prevalent in higher-income countries (Micha et al., 2020). The Sub-Saharan Africa region alone accounts for more one-third of undernourished children globally (Akombi et al., 2017). Moreover, the region has the highest burden of infectious diseases such as HIV and TB, as well as rising rates of "Non-Communicable Diseases" (NCD). The double burden of poverty and disease only worsens the situation in this region. There are also significant health disparities within countries that must be addressed in order to meet "universal health coverage" goals.

Zambia is one of the countries in Sub-Saharan Africa severely affected by malnutrition. Approximately 40 % of under five children in Zambia are stunted, 5% wasted, 15% underweight, and 9% overweight (Mzumara et al., 2018, Jonah et al., 2018). The distribution of undernutrition across the country reveals striking regional disparities, with patterns closely related to observed socioec- onomic inequalities (Moonga et al., 2021).

Childhood stunting is a marker of chronic malnutrition and has severe health consequences in adult life. These include suboptimal cognitive development, degenerative disease, reduced economic output, increased morbidity and mortality. Stunting is therefore of great public health concern, with far-reaching individual and societal consequences (Menon et al., 2018).

Zambia has a long history of mining; the sector has been the mainstay to the country's economy. Copper mining in the Copperbelt Province has contributed the largest mineral export since independence (Unceta, 2021). Despite the economic boom based on a mining bedrock, mining activi- ties have serious environmental and human health consequences. Much of this has been caused by insufficient health and safety measures, as well as inadequate environmental remediation fol- lowing the closure of mining operations. Kabwe is one example of a former mining town in Zambia where significant environmental and human health risks have resulted from unregulated closure of min- ing operations. Because of a lack of comprehensive environmental control of tailing hills since the mine's closure in 1994, communities have been exposed to lead-rich dust from tailing (Yabe et al., 2020). This situation greatly threatens the community's' health, especially children (Bose- O'Reilly et al., 2018). Poverty, malnutrition, and the underlying burden of HIV only heightens the burden on these communities' health. (Yabe et al., 2015).

Susceptibility to lead toxicity is increased in children particularly because of their hand-to-mouth activities, and higher gastrointestinal absorption and bioavailability given their smaller bodies (Calabrese et al., 1997). Lead induced toxicity in children occurs even at low levels of exposure. The "central nervous system" (CNS) is more sensitive to lead toxicity during the early (neonatal) and subsequent developmental stages as compared to adults years (Naranjo et al., 2020). The neurological effects include diminished cognition and intelligence quotient (IQ), memory loss, and attention difficulty (Liu et al., 2011). Aside from its effects on cognitive functions, lead neurotoxicity has far-reaching psychosocial consequences, including unacceptable social attributes and ag- gressive behaviour (Ara and Usmani, 2015).

Exposure and the ability to mitigate the negative health effects of mine pollution and malnutrition are linked by socio-demographic factors. As a result of existing inequalities, there are disparities in vulnerability within these communities. Therefore, when providing interventions, efforts should be made to identify and prioritize the most vulnerable individuals. Spatial analytical methodologies and the use of "Geographic Information Systems" (GIS) can be useful in this regard. GIS enables researchers to investigate potentially modifiable ecological explanations for disease clusters, which may help clarify the aetiology of health-related events. Furthermore, the approach can aid in understanding the spatial differences in exposure and disease outcomes, which is critical in informing intervention implementation (Chen et al., 2008).

In this study we used "Bayesian Distributional Regression" (BDR) methodology that allows for the investigation of non-linear association and interaction effects while also allowing for the inclusion of spatial predictors (Moonga et al., 2021). These models, which are based on the Markov property, also allow for the modelling of the entire response distribution rather than just the mean. This is especially important in some medical research scenarios. Z–scores below and above 2-standard deviations, for example, are of interest in nutrition modelling because they represent stunting and obesity, respectively. Quantile distribution regression is an alternative method that could be used for this investigation. Based on a Bayesian Distributional Regression, on the other hand, goes far beyond quantiles and mean estimates to investigate the entire distribution. In order to meet one of our objectives, we also used spatial autocorrelation methods to investigate additional spatial variation.

1.2 Rationale and Objectives

Timely and accurate health data is critical for informing better health-care systems. In order to promote inclusive health service provision, a comprehensive “Health Information System” (HIS) must account for inequalities at the subnational level. Health policies must therefore aim to address the identified structural and social inequalities that impede equitable health access. Spatial analysis of health outcomes is thus critical to enable this process. This method accounts for geographically distinct environmental and social health barriers.

Several studies have modelled the determinants of malnutrition and child mortality in Zambia (Adebayo, 2003, Dwyer-Lindgren et al., 2014, Masiye et al., 2010, Kandala et al., 2009, Mzumara et al., 2018). The effects of mining deposits (heavy metals) have also been estimated (Yabe et al., 2020). However, no spatial analysis has identified the small area spatial variances in the exposure and distribution of the outcome. Therefore, the purpose of this study was to demonstrate health disparities at subnational level and to investigate predicting spatial covariates. In addition, we used spatial methodology to look into variations in blood lead levels in Kabwe. Finally, we used these concepts to demonstrate disparities in heavy metal exposure in two mining communities in Zimbabwe.

This PhD project demonstrates the opportunity that spatial data provides in understanding exposures and health outcomes at sub- national level. The general objective of this study was to investigate the small area differentials in malnutrition and magnitude of spatial explanatory covariates for lead exposure in children. The specific objectives were:

I. To estimate the spatial distribution of under-five stunting in Zambia.
II. To investigate environmental and socio-demographic factors associated with child malnutrition in Zambia.
III. To investigate spatial autocorrelated covariates that explain small area variation in blood lead level in Kabwe Zambia.

1.3 Methods

1.3.1 Description of Data sources

We used mainly two types of data sources namely: I) “Zambia Demographic Health Survey” (ZDHS). The ZDHS is a national sample survey designed to provide up to date information on fertility rates, breastfeeding practices, nutritional status of mothers and children, early childhood mortality and maternal mortality. II) “Project for Visualization of Impact of Chronic/ Latent Chemical Hazard and Geo-Ecological Remediation in Zambia” (KAMPAI Project). The project collected data on child blood lead levels 0-4 and over 4 years old, gender, age, BLL and residential coordinates. The project also collected soil lead levels in selected residential areas in Kabwe.

The study also utilised remote sensed data from several sources such as the “Malaria Atlas Project” (MAP).

1.3.2 Data analysis

Based on a BDR model, we examined the spatial distribution of stunting and the effects of socioeconomic and remote sensing covariates on anthropometric outcomes for children. To test for spatial dependencies and small area variation of blood lead levels, we used spatial autocorrelation and the hotspot analysis method. R Studio version 4.0.2 was used for all the data analysis.

1.3.3 Ethical considerations

We obtained approval for the study protocol from the ethics committee at the LMU Medical faculty, approval number 19-780. The study observed the principles of medical research, and the ethical considerations. Some of the data we used included geolocations of individual participant households. Therefore, special considerations were made in the storage and analysis of this data. These procedures are detailed in the published articles (Rakete et al., 2022, Moonga et al., 2022, Moonga et al., 2021).

1.3.4 Team

The study team consisted of the principal investigator (author), three supervisors from LMU-CIH, and one supervisor from Zambia (home country). All team members contributed in the conceptualization, proposal design, analysis and discussion of results and in the manuscript publication. Before the final submission, the thesis was also shared with the team members.

1.4 Results

This study lends support to the use of spatial data in resource-constrained settings. We met the study objectives and published three papers in peer reviewed journal as part of our efforts to disseminate our findings (Rakete et al., 2022, Moonga et al., 2022, Moonga et al., 2021).

Stunting and lead exposure were found to have distinct spatial distributions in Zambia, whereas heavy metal exposure was disproportionate in two mining districts in Zimbabwe. The study also identified cold spots and hot spots of blood lead in Kabwe. The study relied on secondary data, which may be a limitation when data quality is not guaranteed, or when there is a lot of missing information. To counteract this, robust statistical approaches were used. Furthermore, to validate the findings, one section of the study used national datasets and two compared surveys (Moonga et al., 2021).

The first and second objectives of the study are discussed in detail in the article "Modelling Chronic Malnutrition in Zambia: A Bayesian Distributional Regression Approach" (Moonga et al., 2021). Stunting rates in Zambia decreased from a mean z-score of -1.59 CI (-1.63; -1.55) to -1.47 CI (-1.49; -1.44) between 2007 and 2013, but remain unacceptably high. Disparities were observed at the provincial and district levels. The Northern region, which is primarily industrialized and a mining belt had higher levels of stunting than the agricultural Southern and Central provinces. The asset index, number of vaccinations received, mothers' BMI, and mothers' years of education all had a non-linear effect on the mean (μ) and variance in the height-for-age z-score for children in Zambia (Moonga et al., 2021).

The article "Geospatial approach to investigate spatial clustering and hotspots of blood lead levels in children within Kabwe, Zambia" (Moonga et al., 2022), addressed the third objective. Lead levels in blood were significantly autocorrelated (Moran's Index was 0.62 ($p = 0.001$), with distinct cold and hot spot areas (μ =7 *g/dL). Child blood lead levels are seen to be high in Kabwe town, indicating a serious health issue that requires immediate attention (Moonga et al., 2022).
The third published article, titled "Biomonitoring of arsenic, cadmium and lead in two artisanal and small-scale gold mining areas of Zimbabwe", shows regional variation in heavy metal exposure in two mining districts in Zimbabwe (Rakete et al., 2022).

1.5 Discussion

The majority of programmatic and routine data collected in developing countries remain underutilized. This is partly due to a lack of expertise in conducting comprehensive and informative data analysis. There is also observed scarcity of high-quality data and repositories from which to obtain linked data. Spatial data, which has the potential to be extremely useful in the development of targeted interventions, is underutilized in particular. Besides this advantage, remote sensing data can be easily collected and used to account for hard-to-reach areas. This has the potential to be a game changer in terms of reducing inequalities between and within countries, particularly in Sub-Saharan Africa, and thus contributing to the achievement of the “Sustainable Development Goals” (SDG).

Furthermore, remote sensing environmental data is extremely useful for gaining a broad understanding of exposure pathways. Rainfall patterns, temperature, and pollution, for example, are linked to malnutrition. Temperature, prevailing winds, and runoff surface water are all important factors in heavy metal exposure pathways in polluted communities. All such variables and many others can be assessed remotely and then applied in real-world scenarios. In this regard, this study is novel in that it makes use of spatial data for low- and middle-income contexts.

Three aspects of the study are unique; firstly, it shows how geographic information system (GIS) data can be used to identify underserved and high-risk communities. Secondly, the application of Bayesian distributional regression models demonstrates the feasibility of combining socio-economic and remote sensing variables. Thirdly, the method allows for the investigation of non-linear predictor effects as well as the modelling of the entire distribution rather than just the mean. This approach is more realistic because most health indicators do not necessarily follow a linear pattern. Linear models, on the other hand, would overestimate some aspects of health.

While the use of remote sensing data is becoming more common in the developed world, it is still underutilized in most developing countries. Remote sensing provides a quick method to collect widespread data, even for hard to reach areas. Owing to technical and economic constraints, many underdeveloped nations may not have any data that is consistently gathered. As a result, satellite data bridges this chasm. Because of data quality issues in the majority of these settings, a combination of socioeconomic variables and remote sensed data explains variability in health outcomes better than remote sensed data alone. This field remains unexplored, necessitating more funding and research.

Previous studies in Zambia have estimated malnutrition rates (Mzumara et al., 2018), but there has been no investigation into the spatial variation. This study fills that gap by demonstrating nutrition differences in children at the district level and between demographic health surveys conducted five years apart. Through this subnational analysis, the study identifies specific areas that are more affected by malnutrition, as well as districts with the highest standard deviation, which is an indicator of potential inequalities. Mining areas, which are frequently more urban, were found to have higher stunting levels than agricultural areas. This suggests a complicated scenario, which is related in part to household expenditures on healthy foods, and the benefits of access to agricultural products, which is more likely in farming towns in the central and southern provinces.

The study also established some non-linear effects of mothers' education on child stunting. After the eighth grade, there is significant effect on both the mean z-score and standard deviations. This finding is particularly significant for Sub-Saharan African countries, which continue to struggle with teen pregnancies and school dropouts. Dropping out of high school is associated with stunted children and a high prevalence of HIV and sexually transmitted diseases. Therefore, this finding will help to inform education policy about the importance of keeping girls in school, with a focus on finishing secondary school. In some contexts, education is viewed as a proxy for "social economic status” (SES). As a result, others may attribute the observed differences to economic status. While this may be true, the study suggests that the benefits of education, such as improved nutrition and hygiene knowledge, may be the most predictive factors. This is in light of Zambia's high unemployment rate, which would reduce economic disparities between educated and uneducated mothers.

The striking disparity in stunting levels between rural and urban areas reflects socioeconomic disparities between these areas. Rural communities are underserved and marginalized in the majority of developing countries. They have limited access to health services and essential commodities, such as clean water, and food. Geopolitical factors frequently influence a country's regional distribution of resources. This in part explains why the nutritional status of certain regions has not improved over time, as some areas remain side-lined. This demonstrates unequal distribution of national resources or a lack of deliberate efforts to identify and target such most affected areas (Fekete and Weyers, 2016). Highlighting these inequalities is essential in the identification of most affected communities, thus providing a voice for the poor and marginalized. This information is also beneficial for the effective implementation of remedial measures.

Africa is rich in mineral resources, and mining is still a major economic driver. The health and environmental consequences of mining activities, on the other hand, are rarely fully addressed, both during and after they are completed. Communities have to deal with the fallout, with children often bearing the brunt of the burden. Blood lead levels in Kabwe were found to have distinct hotspots and cold spots, as well as significant spatial autocorrelation. This is explained in part by the residential areas' proximity to the mine site. Within this community, residential locations are linked to social and economic status. This emphasizes the importance of protecting vulnerable communities and facilitating deliberate interventions to mitigate the effects of structural issues. Disparities in lead exposure in Zimbabwean mining towns also demonstrates inequities in health and safety standards. An understanding of such trends is critical, especially when it comes to directing corrective actions.

The study finding have been disseminated, and are being utilised by an intervention programme aimed at treating children intoxicated with lead in Kabwe. Furthermore, environmental remediation programs now prioritize the underserved communities that were identified in this study. The findings of this study have been presented at international conferences. For example, the "INCHES conference" on lead and children's health held on 6th January 2022.

As a follow-up to this study, another study is being conducted to investigate heavy metal exposure pathways via home-grown vegetables, free-range chickens, and various sources of drinking water. This one-health approach will close gaps in lead exposure pathways. Further, the University of Zambia developed an elective course in Spatial Epidemiology and invited the author to contribute in the curriculum development. The course will foster further use of georeferenced data and increase in the competence to utilise spatial data within the country.

In conclusion, we find the integration of remote sensing data and socioeconomic characteristics to be a novel aspect of this work. This approach has great potential for quick access to data, particularly in meeting the SDGs, particularly Goal 10 that aims at reducing inequality within and between countries. Because of the identified benefits, the study advocates for greater use of spatial data. Furthermore, we recommend research into predictive models based on spatial data to help developing countries plan ahead.

Reference

ADEBAYO, S. B. 2003. Modelling childhood malnutrition in Zambia: an adaptive Bayesian splines approach. *Statistical Methods and Applications,* 12**,** 227-241.

AKOMBI, B. J., AGHO, K. E., MEROM, D., RENZAHO, A. M. & HALL, J. J. 2017. Child malnutrition in sub-Saharan Africa: a meta-analysis of demographic and health surveys (2006-2016). *PloS one,* 12**,** e0177338.

ALAO, R., NUR, H., FIVIAN, E., SHANKAR, B., KADIYALA, S. & HARRIS-FRY, H. 2021. Economic inequality in malnutrition: a global systematic review and meta-analysis. *BMJ Global Health,* 6**,** e006906.

ARA, A. & USMANI, J. A. 2015. Lead toxicity: a review. *Interdisciplinary toxicology,* 8**,** 55-64.

BOSE-O'REILLY, S., YABE, J., MAKUMBA, J., SCHUTZMEIER, P., ERICSON, B. & CARAVANOS, J. 2018. Lead intoxicated children in Kabwe, Zambia. *Environmental Research,* 165**,** 420-424.

CALABRESE, E. J., STANEK, E., JAMES, R. C. & ROBERTS, S. M. J. E. H. P. 1997. Soil ingestion: a concern for acute toxicity in children. 105**,** 1354-1358.

CHEN, J., ROTH, R. E., NAITO, A. T., LENGERICH, E. J. & MACEACHREN, A. M. 2008. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of US cervical cancer mortality. *International journal of health geographics,* 7**,** 57.

DWYER-LINDGREN, L., KAKUNGU, F., HANGOMA, P., NG, M., WANG, H., FLAXMAN, A. D., MASIYE, F., GAKIDOU, E. J. S. & EPIDEMIOLOGY, S.-T. 2014. Estimation of district-level under-5 mortality in Zambia using birth history data, 1980–2010. 11**,** 89-107.

FEKETE, C. & WEYERS, S. 2016. Social inequalities in nutrition: evidence, causes and interventions. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz,* 59**,** 197-205.

JONAH, C. M., SAMBU, W. C. & MAY, J. D. 2018. A comparative analysis of socioeconomic inequities in stunting: a case of three middle-income African countries. *Archives of Public Health,* 76**,** 1-15.

KANDALA, N. B., FAHRMEIR, L., KLASEN, S. & PRIEBE, J. 2009. Geo-additive models of childhood undernutrition in three sub-Saharan African countries. *Population, Space and Place,* 15**,** 461-473.

LIU, J., MCCAULEY, L., COMPHER, C., YAN, C., SHEN, X., NEEDLEMAN, H. & PINTO-MARTIN, J. A. J. E. H. 2011. Regular breakfast and blood lead levels among preschool children. 10**,** 28.

MASIYE, F., CHAMA, C., CHITAH, B. & JONSSON, D. J. Z. S. S. J. 2010. Determinants of child nutritional status in Zambia: An analysis of a national survey. 1**,** 4.

MENON, P., HEADEY, D., AVULA, R. & NGUYEN, P. H. 2018. Understanding the geographical burden of stunting in India: A regression-decomposition analysis of district-level data from 2015–16. *Maternal & child nutrition,* 14**,** e12620.

MICHA, R., MANNAR, V., AFSHIN, A., ALLEMANDI, L., BAKER, P., BATTERSBY, J., BHUTTA, Z., CHEN, K., CORVALAN, C. & DI CESARE, M. 2020. 2020 global nutrition report: action on equity to end malnutrition.

MOONGA, G., BÖSE-O'REILLY, S., BERGER, U., HARTTGEN, K., MICHELO, C., NOWAK, D., SIEBERT, U., YABE, J. & SEILER, J. 2021. Modelling chronic malnutrition in Zambia: A Bayesian distributional regression approach. *Plos one,* 16**,** e0255073.

MOONGA, G., CHISOLA, M. N., BERGER, U., NOWAK, D., YABE, J., NAKATA, H., NAKAYAMA, S., ISHIZUKA, M. & BOSE-O'REILLY, S. 2022. Geospatial approach to investigate spatial clustering and hotspots of blood lead levels in children within Kabwe, Zambia. *Environmental research,* 207**,** 112646.

MZUMARA, B., BWEMBYA, P., HALWIINDI, H., MUGODE, R. & BANDA, J. 2018. Factors associated with stunting among children below five years of age in Zambia: evidence from the 2014 Zambia demographic and health survey. *BMC nutrition,* 4**,** 1-8.

NARANJO, V. I., HENDRICKS, M. & JONES, K. S. 2020. Lead toxicity in children: an unremitting public health problem. *Pediatric Neurology,* 113**,** 51-55.

RAKETE, S., MOONGA, G., WAHL, A.-M., MAMBREY, V., SHOKO, D., MOYO, D., MUTETI-FANA, S., TOBOLLIK, M., STECKLING-MUSCHACK, N. & BOSE-O'REILLY, S. 2022. Biomonitoring of arsenic, cadmium and lead in two artisanal and small-scale gold mining areas in Zimbabwe. *Environmental Science and Pollution Research,* 29**,** 4762-4768.

SDG, U. 2018. Sustainable development goals. *United Nations*.

UNCETA, R. A. 2021. The economic and social impact of mining-resources exploitation in Zambia. *Resources Policy,* 74**,** 102242.

WORLD HEALTH ORGANIZATION 2018. The state of food security and nutrition in the world 2018: building climate resilience for food security and nutrition. *WHO,*.

YABE, J., NAKAYAMA, S. M., IKENAKA, Y., YOHANNES, Y. B., BORTEY-SAM, N., OROSZLANY, B., MUZANDU, K., CHOONGO, K., KABALO, A. N. & NTAPISHA, J. J. C. 2015. Lead poisoning in children from townships in the vicinity of a lead–zinc mine in Kabwe, Zambia. 119**,** 941-947.

YABE, J., NAKAYAMA, S. M., NAKATA, H., TOYOMAKI, H., YOHANNES, Y. B., MUZANDU, K., KATABA, A., ZYAMBO, G., HIWATARI, M. & NARITA, D. 2020. Current trends of blood lead levels, distribution patterns and exposure variations among household members in Kabwe, Zambia. *Chemosphere,* 243**,** 125412.

Chapter II

Modelling Chronic Malnutrition in Zambia: A Bayesian Distributional Regression Approach

PLOS ONE

RESEARCH ARTICLE

Modelling chronic malnutrition in Zambia: A Bayesian distributional regression approach

Given Moonga[1,2,3☯]*, Stephan Böse-O'Reilly[2,4‡], Ursula Berger[5‡], Kenneth Harttgen[6‡], Charles Michelo[3‡], Dennis Nowak[4‡], Uwe Siebert[2‡], John Yabe[7‡], Johannes Seiler[2,8☯]

1 Center for International Health, Ludwig Maximilian University of Munich, Munich, Germany, **2** Department of Public Health, Health Services Research and Health Technology Assessment, UMIT—University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, **3** Department of Epidemiology and Biostatistics, University of Zambia, Lusaka, Zambia, **4** Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Clinical Centre of the Ludwig Maximilian University of Munich, Munich, Germany, **5** Institute for medical Information Processing, Biometry, and Epidemiology, Ludwig Maximilian University of Munich, Munich, Germany, **6** Department of Humanities, Social and Political Sciences, ETH Zurich, Zurich, Switzerland, **7** School of Veterinary Medicine, University of Zambia, Lusaka, Zambia, **8** Department of Statistics, University of Innsbruck, Innsbruck, Austria

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* Given.Moonga@lrz.uni-muenchen.de

Citation: Moonga G, Böse-O'Reilly S, Berger U, Harttgen K, Michelo C, Nowak D, et al. (2021) Modelling chronic malnutrition in Zambia: A Bayesian distributional regression approach. PLoS ONE 16(8): e0255073. https://doi.org/10.1371/journal.pone.0255073

Editor: Tzai-Hung Wen, National Taiwan University, TAIWAN

Received: February 23, 2020

Accepted: July 10, 2021

Published: August 4, 2021

Data Availability Statement: Data are third party data available from the following sites: https://crudata.uea.ac.uk/cru/data/drought/ https://malariaatlas.org/ https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11/data-download https://dhsprogram.com/data/.

Funding: Johannes Seiler acknowledges financial support from the University of Innsbruck through a postdoctoral scholarship. The funders had no role

Abstract

Background

The burden of child under-nutrition still remains a global challenge, with greater severity being faced by low- and middle-income countries, despite the strategies in the Sustainable Development Goals (SDGs). Globally, malnutrition is the one of the most important risk factors associated with illness and death, affecting hundreds of millions of pregnant women and young children. Sub-Saharan Africa is one of the regions in the world struggling with the burden of chronic malnutrition. The 2018 Zambia Demographic and Health Survey (ZDHS) report estimated that 35% of the children under five years of age are stunted. The objective of this study was to analyse the distribution, and associated factors of stunting in Zambia.

Methods

We analysed the relationships between socio-economic, and remote sensed characteristics and anthropometric outcomes in under five children, using Bayesian distributional regression. Georeferenced data was available for 25,852 children from two waves of the ZDHS, 31% observation were from the 2007 and 69% were from the 2013/14. We assessed the linear, non-linear and spatial effects of covariates on the height-for-age z-score.

Results

Stunting decreased between 2007 and 2013/14 from a mean z-score of 1.59 (credible interval (CI): -1.63; -1.55) to -1.47 (CI: -1.49; -1.44). We found a strong non-linear relationship for the education of the mother and the wealth of the household on the height-for-age z-score. Moreover, increasing levels of maternal education above the eighth grade were associated with a reduced variation of stunting. Our study finds that remote sensed covariates alone

in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: No authers have competing interests.

explain little of the variation of the height-for-age z-score, which highlights the importance to collect socio-economic characteristics, and to control for socio-economic characteristics of the individual and the household.

Conclusions

While stunting still remains unacceptably high in Zambia with remarkable regional inequalities, the decline is lagging behind goal two of the SDGs. This emphasises the need for policies that help to reduce the share of chronic malnourished children within Zambia.

Introduction

The burden of child malnutrition still remains a global challenge, with greater severity being faced by low-and middle-income countries [1–3]. Globally, malnutrition is amongst the most important risk factors associated with illness and death, affecting hundreds of millions of pregnant women and young children [3–6]. Stunting in early childhood is strongly associated with numerous short-term and long-term consequences, including increased childhood morbidity and mortality, delayed growth and motor development and long-term educational and economic consequences later in life [7]. Undernourishment causes children to start life at mentally suboptimal levels [8].

Assessment of childhood malnutrition commonly relies on standard anthropometric measures for insufficient height-for-age (stunting) indicating chronic undernutrition, insufficient weight-for-height (wasting), indicating acute undernutrition; and insufficient weight-for-age (underweight), an indicator commonly used to asses, both, chronic, and acute undernutrition [9, 10].

Anthropometric measurements are practical techniques for assessing children's growth patterns during the first years of life. The measurements also provide useful insights into the nutrition and health situation of entire population groups. Anthropometric indicators are less accurate than clinical and biochemical techniques in assessing individual nutritional status. However in resources limited settings, the measurements are a useful screening tool to identify individuals at risk of undernutrition, who can later be referred to subsequent possible confirmatory investigation [11].

It is estimated that globally 52 million children under-five years of age are wasted, 17 million are severely wasted and 155 million are stunted. Around 45% of deaths among children under-five years of age, most of which occur in the sub-Saharan Africa are linked to undernutrition [3, 9]. It is also estimated that four out of ten children under the age of five in Zambia are stunted [12]. This paper will therefore focus on childhood stunting in Zambia.

Global prevalence of stunting in children younger than five years declined during the past two decades, but still remain unacceptably high in South Asia and sub-Saharan Africa regions [5]. If current trends remain unchecked, projections indicate that 127 million children under five years of age will be stunted in 2025 [1]. There is therefore need to heighten various interventions in these affected region and to investigate possible area specific determinants of stunting.

There are already fairly well documented perspectives on determinants of malnutrition. The treatise on these determinants mainly relies on the United Nations Children's Fund (UNICEF) conceptual framework on malnutrition which has evolved over time as more knowledge and evidence on the causes, consequences and impacts of undernutrition is generated. The framework distinguishes between immediate, intermediate and underlying determinants of malnutrition [5, 13–15].

The immediate causes of undernutrition include inadequate dietary intake and disease, while the underlying causes could include household food insecurity, inadequate care and feeding practices for children, unhealthy household and surrounding environments, and inaccessible and often inadequate health care. Basic causes of poor nutrition encompasses the societal structures and processes that neglect human rights and perpetuate poverty, constraints faced by populations to essential resources [13].

Several studies done within sub-Saharan Africa investigated determinants such as the mother's level of education, income levels and these factors have been linked to malnutrition [9, 12, 16, 17]. The source of the drinking water, the wealth of the household, the area of residence, age of the child, the sex of the child, the breastfeeding duration, the age of the mother has also been investigated and were observed to be significant correlates of stunting [12, 18]. Within Zambia, stunting was observed to be more likely among children of less educated mothers (45%) and those from the poorest households (47%) [19]. The determinates of malnutrition are related to each other and the differences and direction between these levels of determinism as indicated in the UNICEF framework are often not discrete but in reality related. As discussed by Kandala [17] for example, the mother's level of education might be influencing child care practises- an intermediate determinant—and the resources available to the household—an underlying determinant.

Previous studies elsewhere have observed that stunting tends to show regional variation [4, 9, 16]. We see this trend in Zambia as well, where the decline of stunting has been only gradual and unacceptable, with higher prevalence in Northern province where 50% of the children being stunted, and stunting being less common in Lusaka, Copperbelt, and Western provinces where 36% of children are stunted [19]. We see this regional variation of stunting in Fig 1 which shows stunting in Zambia in the 2007 and 2013/14 waves of the Zambian Demographic and Health Surveys (ZDHS). The ZDHS is a national-wide survey which is representative at a sub-national level and contains information on trends in fertility, childhood mortality, use of family planning methods, and maternal and child health indicators including HIV and AIDS [19]. The figure shows the height-for-age z-score, with Western province better than Northern province for the 2007 wave. We see a slight difference in 2013/14 as stunting seemed to get worse in parts of the Western province.

Much of the work done on the determinants of stunting in Zambia, have considered socio-economic characteristics and have assessed the linear effects of these determinants on the conditional mean [12, 20], using models specifications such as; linear models, generalized linear models (GLMs) and generalized additive models (GAMs) [21]. These aproaches are useful and have the advantage of being easy to estimate and to interpret. However, they may risk model misspecification and draw inaccurate estimates, when heterogeneity is present, or when extreme values in the response are present and when a linear relationship is not plausible. In the analysis of certain outcomes, like stunting for example, the interest is not only in the conditional mean, but also in extreme values (height-for-age z-scores), or other parameters of the response. Quantile regression is one possibility to model beyond the conditional mean, with the interest to show variation of the outcome at a quantile level, without making any assumptions of the response distribution. This method for example has been applied in child malnutrition studies [16]. However, distributional regression offers advantage over quantile regression, as it provides the possibility to characterize the complete probabilistic distribution of the response in one joint model [21, 22]. Moreover, distributional regression is more efficient, if prior knowledge on specific aspects of the response distribution is available, or can be

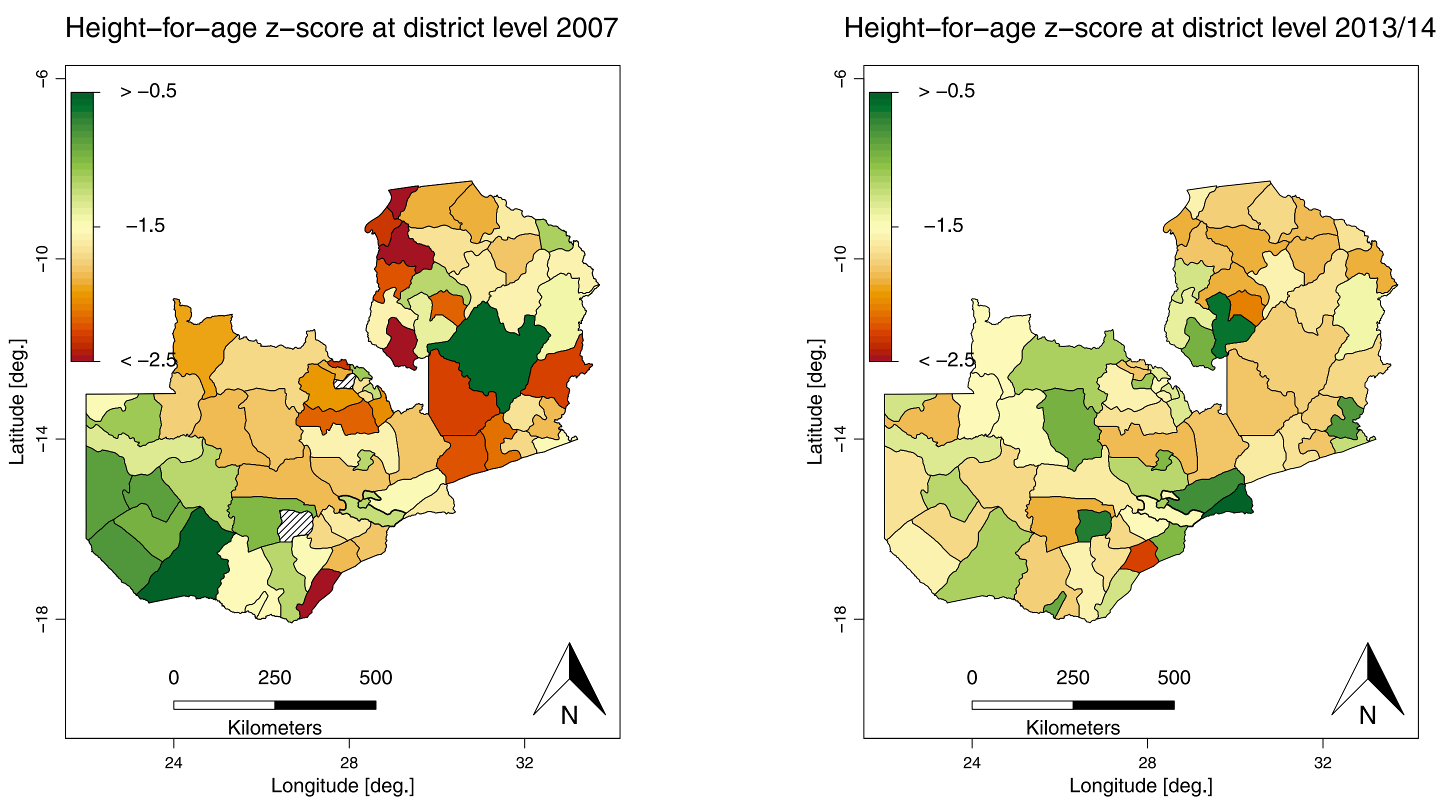

Fig 1. Levels of stunting over time in Zambia. The panel shows the average height-for-age z-score at district level for 2007 (left) and 2013/14 (right) of Zambia. *Source*: Demographic and Health Surveys (data) and Database of Global Administrative Areas (boundary information); calculation by authors. The shapefile used to create these maps is republished from [54] under CC BY license, with permission from Robert J. Hijmans, original copyright [2021].

https://doi.org/10.1371/journal.pone.0255073.g001

estimated [23]. Furthermore, the characterization of the whole distribution of the response is more informative.

The study by Kandala [24] which focused on stunting in sub-Saharan countries found that there are distinct spatial patterns of malnutrition that are not explained by the socio-economic determinants or other well-known correlates alone [16, 25]. As such, our study includes spatial covariates, since we aim at investigating spatial differences of stunting in Zambia at sub-district level while jointly analysing socio-economic and environmental characteristics.

The following three reasons make our study novel compared to previous work [18]. Firstly, we jointly analysed remote sensed data and socio-economic covariates at sub-national level. This is made possible due to availability of georeferenced data at the primary sampling unit a household pertains to in the recent two ZDHS datasets. The Demographic and Health Surveys rely in most cases on a two-stage survey, and the primary sampling units corresponds to the enumeration areas from the most recent completed census that has been selected. Georeferenced data is important as it generates more specific information which can facilitate targeted interventions. Secondly, we used Bayesian distributional regression which allows us to model all parameters of the underlying response distribution. Lastly, we used two waves of the demographic health surveys to control for spatio temporal interactions. Therefore, this study demonstrates small area variation in stunting in Zambia and analyse possible inequalities and deprivation at the sub-district level.

Data sources

Socio-economic and georeferenced covariates

We used data from the 2007 and 2013/14 ZDHS. The ZDHS is a national-wide survey which is representative at a sub-national level and contains information on trends in fertility, childhood mortality, use of family planning methods, and maternal and child health indicators including HIV and AIDS. For these population health indicators, data is collected for women aged 15–49, men aged 15–59 and children below five years of age [19].

The ZDHS provide besides information on the district a household pertains to, also information about the geolocation of the primary sampling unit a household belongs to, and from which the data was collected. The location of the primary sampling unit is the spatial information used in the empirical analysis. During data processing, GPS coordinates are displaced to ensure that respondent confidentiality is maintained. The displacement is randomly applied so that rural points contain a minimum of 0 and a maximum of 5 km of positional error. Urban points contain a minimum of 0 and a maximum of 2 km of error. A further 1% of the rural sample points are offset a minimum of 0 and a maximum of 10 km [26].

Demographic Health Surveys have documented weakness for estimation of individual anthropometric measurements. Potential threats to high data quality may occur across various research stages, from survey design to data analysis. There is also often a substantial amount of missing or implausible anthropometric data across surveys [27].

Furthermore, there is caution over the use of stunting as an individual classifier in epidemiologic research or its interpretation as a clinically meaningful health outcome. Stunting should be used as originally designed to be from its original use as a population level indicator of community well-being [28], as it reflects past health and nutrition conditions; and an indication of socio-economic development of a country [1].

Despite the above highlighted limitations of DHS and anthropometric indicators, they remain useful national wide measurements that can be used to estimate child health. Moreover, in general anthropometric measures are a good indicator for planning as they can provide a lot of information to policy makers to answer, how, where and which type of intervention would be favourable in specific settings.

Socio-economic and spatial determinants

The effects of socio-economic factors, such as the education of the mother, household size, wealth of the household on the health status of children are well documented [12, 29]. We calculated an index representing the wealth of the household based on the household's assets using Principal Components Analysis (PCA) following Filmer and Pritchett, and Sahn [30, 31]. Previous studies have shown that household wealth status was a predictor of childhood malnutrition. Children from poor households are more likely to be stunted than those from richer households [29].

In our analysis we investigated the impact of different socio-economic factors, which impact on height-for age Z-score has been discussed in literature. Table 1 gives an overview and the according references.

Remote sensed covariates

We obtained remote sensed data on drought severity, malaria incidence, and population density. The description, and source to these data sets is provided in Table 2.

For example, the malaria incidence data was obtained from the Malaria Atlas Project (MAP). The project collects malaria data on malaria cases reported by surveillance systems,

Table 1. Included covariates, their source, and effect on the height-for-age z-score found in the literature.

Covariate	Used data source	Effect on stunting found in literature	Reference
Asset Index	DHS	Household wealth inequality associated with childhood stunting	[29]
Age mother at birth	DHS	Increasing non-linearly	[16]
Age child	DHS	Decreasing non-linearly	[17]
Birth order	DHS	Being born forth or higher significantly more stunted	[16]
Breastfeeding duration	DHS	Breastfeeding interval ≤ associated with low level of stunting	[25]
Education mother	DHS	Stunting improves non-linearly with the educational level	[32]
Household size	DHS	Increases linearly	[12]
Mothers' BMI	DHS	U-shape relationship with childhood stunting	[17]
Number of vaccinations	DHS	Lower levels of stunting when fully vaccinated	[25]
Drought severity index	See Table 2	Not further specified	[33]
Malaria incidence	See Table 2	No clear pattern	[34]
Population density	See Table 2	Not further specified	[33]

https://doi.org/10.1371/journal.pone.0255073.t001

nationally representative cross-sectional surveys of parasite rate, and satellite imagery capturing global environmental conditions that influence malaria transmission [35].

Methodology

We assessed the relationships between socio-economic and remote sensed characteristics and anthropometric outcomes using the Bayesian Distributional Regression (BDR). BDR models all parameters of the response distribution based on structured additive predictors and allows to incorporate for example, non-linear effects of metric covariates, spatial effects, or varying effects. Applications of structured additive regression models to topics in Global Public Health are found in several publications [39–42]. This approach permits us to fully analyse the whole distribution [41, 43] and our analysis was not restricted to assessing the conditional mean of the height-for-age z-score. Instead suspected heterogeneity across socio-economic and georeferenced factors and the anthropometric measure can be directly captured. In the context of growth failures this is of particular importance, as previous studies highlighted high levels of heterogeneity related to growth failures [33].

Bayesian distributional regression

Relying on Bayesian distributional regression requires to specify the distribution of the response variable. Assuming the response distribution to be Gaussian permits to model besides the conditional mean also the variance or standard deviation of the response variable. Graphical analysis using amongst others randomised quantile residuals [44] strengthens that a Gaussian model is plausible. See also Fig 2, for more details. In the left- hand panel of Fig 2 the histogram of the height-for-age z-score together with the underlying density illustrates why the normal distribution seems to be an appropriate choice. This is further confirmed in the second and third panel, where the histogram of the quantile residuals including the underlying kernel density estimate, respectively, the QQ-plot of the randomised quantile residuals are shown.

Table 2. Source of remote sensed covariates.

Covariate	Description	Source	Reference
Drought index	scPDSI CRU4.03	Climate Research Unit	[36, 37]
Malaria incidence	Plasmodium falciparum incidence	Malaria Atlas Project	[35]
Population density	Number of people per km^2	Socioeconomic Data and Applications Center	[38]

https://doi.org/10.1371/journal.pone.0255073.t002

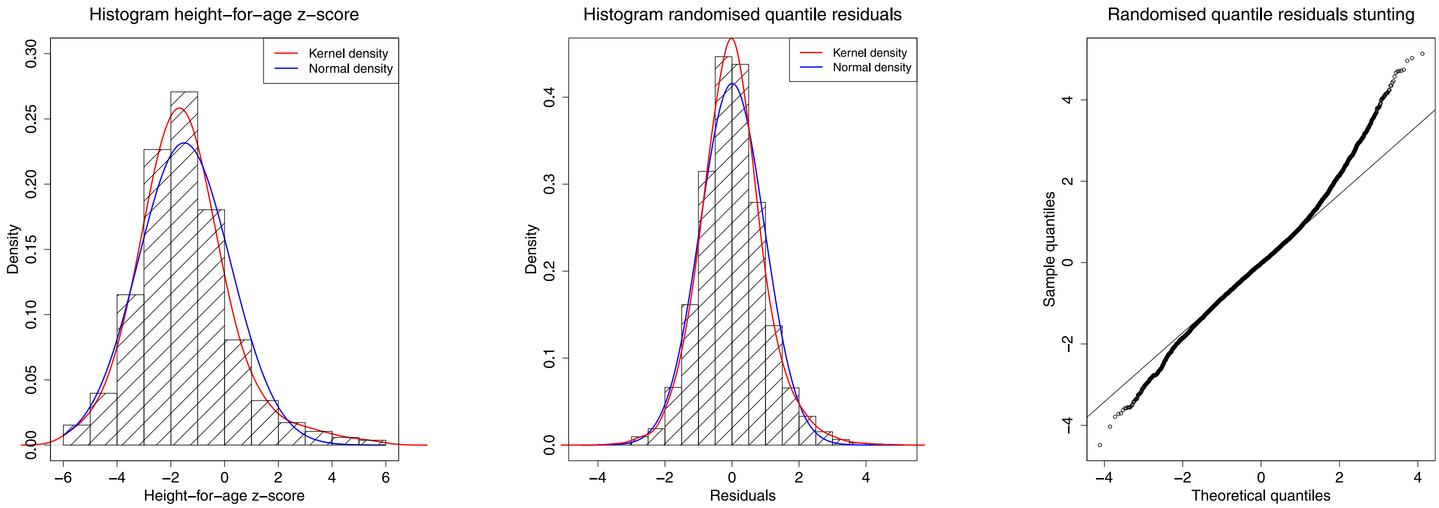

Fig 2. Histogram of the response and histogram and QQ-plot of the randomised quantile residuals. The left-hand panel shows the histogram and kernel density estimates of the height-for-age z-score, the middle panel shows the histogram of the randomised quantile residuals together with the normal density estimates, and the right-hand panels depicts the QQ-plot of the randomised quantile residuals. *Source*: Demographic and Health Surveys (data); calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.g002

Assuming the response distribution of the height-for-age z-score to be Gaussian, both the mean μ and the standard deviation σ are related to a structured additive predictor. Accordingly, the $\texttt{z} - \texttt{score}_{is} \sim \mathcal{N}(\mu_{is}, \sigma_i)$ can be specified as Gaussian response, with $i = 1, \ldots, I$ being the number of children, and $s = 1, \ldots, S$ be the location of the primary sampling unit the child pertains to. Using the same notation as in the methodology manual of BayesX the regression model can be written as follows [45]:

$$\begin{aligned} \boldsymbol{\mu} &= h_\mu(\boldsymbol{\eta}_{\boldsymbol{\mu}}) = \boldsymbol{\eta}_{\boldsymbol{\mu}}, \\ \boldsymbol{\sigma} &= h_\sigma(\boldsymbol{\eta}_{\boldsymbol{\sigma}}) = \exp(\boldsymbol{\eta}_{\boldsymbol{\sigma}}). \end{aligned} \tag{1}$$

Here both parameters of the normal distribution the mean μ and the standard deviation σ are related to the set of covariates specified further in Tables 1 and 2. Accordingly, the response functions h_μ, and h_σ link the two parameters to their structured additive predictors which is specified as follows:

$$\begin{aligned} \boldsymbol{\eta}_{\boldsymbol{\mu}} = \quad & f_1(\texttt{Asset index}) + f_2(\texttt{Birth order, Age mother at birth}) + \\ & f_3(\texttt{Age child, Breastfeeding duration}) + f_4(\texttt{Education mother}) + \\ & f_5(\texttt{Household size}) + f_6(\texttt{BMI mother}) + \\ & f_7(\texttt{Number of vaccinations}) + f_8(\texttt{Drought severity index}) + \\ & f_9(\texttt{Malaria incidence}) + f_{10}(\log(1 + \texttt{Population density})) \\ & f_{11}(\texttt{Spatial, Time}) + \boldsymbol{x}'\boldsymbol{\beta}, \\ \boldsymbol{\eta}_{\boldsymbol{\sigma}} = \quad & f_1(\texttt{Asset index}) + f_2(\texttt{Birth order, Age mother at birth}) + \\ & f_3(\texttt{Age child, Breastfeeding duration}) + f_4(\texttt{Education mother}) + \\ & f_5(\texttt{Household size}) + f_6(\texttt{BMI mother}) + \\ & f_7(\texttt{Number of vaccinations}) + f_8(\texttt{Drought severity index}) + \\ & f_9(\texttt{Malaria incidence}) + f_{10}(\log(1 + \texttt{Population density})) \\ & f_{11}(\texttt{Spatial, Time}) + \boldsymbol{x}'\boldsymbol{\beta}, \end{aligned} \tag{2}$$

where $f_1(\cdot)$ to $f_{10}(\cdot)$ are potential non-linear effects of socio-economic and remote sensed covariates approximated using Bayesian penalised Splines (P-Splines) first described by Lang and Brezger [46]. Bayesian P-Splines are based on P-Splines as introduced by Eilers and Marx [47], and use an approach based on basis functions. As smoothness priors of the unknown regression parameters $\boldsymbol{\beta}_j$ a random walk prior of the form $\boldsymbol{\beta}_j|\gamma_j^2 \propto \exp\left(-\frac{1}{2\gamma_j^2}\boldsymbol{\beta}_j'\boldsymbol{K}_j\boldsymbol{\beta}_j\right)$ is specified, with γ_j^2 being random variance parameters and $\boldsymbol{K}_j$ is an appropriate penalty matrix, see also Lang and Brezger [46] for an elaborate discussion. Relying on Bayesian P-Splines allows to incorporate different model terms such as non-linear effects for continuous variables, varying coefficients [48], or spatial effect. In addition, Bayesian P-Splines allow to decompose the predictor additively and are known for having good mixing properties. See S1 to S8 Figs in the Supporting Information for convergence diagnosctics of the sampling paths of the parameters included in the final model that is based on a Markov chain Monte Carlo (MCMC) algorithm. $f_{11}(\cdot)$ is the spatio-temporal effect included in the model to account for unexplained heterogeneity by incorporating a Markov random field prior [46]. In more detail, following Lang and Brezger [46], the basic Markov random field prior for the regression coefficients β_s of the spatially correlated effect f_s is defined as follows: $\beta_s|\beta_{s'}, s \neq s' \sim \mathcal{N}(\frac{1}{N_s}\sum_{s'\in\delta_s}\beta_{s'}, \frac{\tau_s^2}{N_s})$. Where N_s is the number of neighbouring sites of location s, s′ belongs to the set of neighbouring sites δ_s of location s, and τ_s^2 being the spatially adaptive variance parameteres [47]. Incorporating the spatial effect on a Markov random field proposal allows to account for the remaining heterogeneity that is not explained by the included covariates. See also Chapter 4 of the methodology manual of **BayesX** [45], and Seiler and colleagues [42] for a more elaborate discussion on the incorporation of the spatio- temporal effect based on a Markov random field prior. This variables are further specified in Table 1. $\boldsymbol{x}'\boldsymbol{\beta}$ subsumes the vector of included effect coded categorical covariates that are the gender, the place of living, and the survey wave. For an illustration of effect coding see for example Fahrmeier and colleagues [49]. From a Bayesian perspective the categorical variables are considered to be random variables for which a diffuse prior of the form $p(\gamma_j) \propto$ const is assigned [46].

Model selection

The fit of the models are compared by relying on the Deviance Information Criterion (DIC) [50] and Widely Applicable Information Criterion (WAIC) [51], and are summarised in Table 3. As a rule of thumb can be seen that the model with the lowest value describes the data best. We specified six distinct models, aiming to identify the importance of, for instance, socio-economic or georeferenced factors. In more detail, the differences between these models are summarised in Table 4.

Table 3. Estimation results: DIC and WAIC.

Model	DIC	WAIC
Model 1	95588.2	95550.7
Model 2	91027.1	91566.4
Model 3	91375.7	91884.0
Model 4	96263.1	96541.9
Model 5	**90992.4**	**91527.7**
Model 6	91012.7	91543.3

Values of the DIC and the WAIC for different model specifications. *Source*: DHS; calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.t003

Table 4. Specification of estimated models.

Model term	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$\boldsymbol{\eta}_\mu$	yes (y)	y	y	y	y	y
$\boldsymbol{\eta}_\sigma$	n (n)	y	y	y	y	y
$\boldsymbol{x}'\boldsymbol{\beta}$	y	y	y	y	y	y
f_1(Asset index)	y	y	y	n	y	y
f_2(Birth order, Age mother at birth)	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	n	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$
f_3(Age child, Breastfeeding duration)	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	n	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$	y $\boldsymbol{\eta}_\mu$, n $\boldsymbol{\eta}_\sigma$
f_4(Education mother)	y	y	y	n	y	y
f_5(Household size)	y	y	y	n	y	y
f_6(BMI mother)	y	y	y	n	y	y
f_7(Number of vaccinations)	y	y	y	n	y	y
f_8(Drought severity index)	y	y	y	n	y	y
f_9(Malaria incidence)	y	y	y	n	y	y
f_{10}(log(1 + Population density))	y	y	y	n	y	n
f_{11}(Spatial, Time)	y	y	y	n	y	y

Table of specified models indicating the differences between models and the included model terms and covariates. Note that after evaluating the sampling paths of the resulting Markov chains of the MCMC simulations, in $\boldsymbol{\eta}_\sigma$ $f_2(\cdot)$ and $f_3(\cdot)$ had to be omitted due bad mixing.

https://doi.org/10.1371/journal.pone.0255073.t004

Results

In the following Section we will discuss the results of Model 5, omitting insignificant terms, as Model 5 has both the lowest DIC and WAIC. Result of the included covariates are however, similar throughout all specifications.

Descriptive analysis

Table 5 shows the baseline characteristics of selected covariates in the population between the two ZDHS survey of 2007 and 2013/14, and remote sensed data aggregates for these waves.

Data was available for 25,852 children from the two waves, 31% observation were from the 2007 and 69% were from the 2013/14 ZDHS. Levels of stunting decreased between 2007 and 13/14 from a mean z-scores of -1.59 CI(-1.63; -1.55) to -1.47 CI(-1.49; -1.44). The breastfeeding duration declined from 16.22 to 15.69. There was a notable increase in the number of received vaccinations by children from 5.6 to 7.5 vaccinations. There was a slight increase in the number of years the mother spent in school from 7.2 to 7.8. Malaria incidence rates (plasmodium falciparum incidence) declined from 26% to 20%. Night-time light increased from 2.75, to 3.72 (observed values were log transformed), a possible indication of increase in urbanisation. Night-time light was highly correlated ($\rho = 0.73$) to population density as such it was omitted in subsequent analysis.

High disparities in the height-for-age z-score have been observed at the district and provincial level within Zambia. There was a drift in the spatial pattern of malnutrition in the 2013/14 wave compared to the previous survey, indicating a general improvement. See also Fig 1 for a more detailed, descriptive, analysis of the spatial patterns of the height-for-age z-score within Zambia.

We observed that in the 2007 wave, stunting was lowest in the Western and Muchinga province. In the Southern province generally, low values were also observed, except for the Sinazongwe district. For Eastern province, Nyimba, Katete, Petauke and Lundazi districts had high levels. In the Luapula province, high levels of stunting were observed in the districts of

Table 5. Descriptive statistics covariates.

	2007 ZDHS		2013/14 ZDHS	
Response	Mean (95% CI)	*n*	Mean (95% CI)	*n*
Stunting	-1.59 (-1.63; -1.55)	7,936	-1.47 CI (-1.49; -1.44)	17,916
Covariates	Mean (95% CI), %	SD	Mean (95% CI), %	SD
Proportion of male children (= 1)	49.62%		50.08%	
Age children in months	29.14 (28.76; 29.51)	17.03	29.84 (29.58; 30.09)	17.20
Breastfeeding duration in months	16.22 (16.07; 16.38)	7.04	15.69 (15.58; 15.79)	7.11
Birth order within household	2.99 (2.93; 3.04)	2.38	2.88 (2.84; 2.91)	2.37
Number of vaccinations	5.64 (5.59; 5.69)	2.40	7.45 (7.42; 7.49)	2.17
Age mother at birth in years	24.30 (24.15; 24.45)	6.82	24.08 (23.98; 24.18)	6.94
BMI mother	22.35 (22.27; 22.43)	3.42	22.57 (22.51; 22.62)	3.75
Years of education mother	7.18 (7.10; 7.26)	3.53	7.81 (7.76; 7.87)	3.64
Urban place of living (= 1)	38.73%		43.01%	
Size of the household	6.27 (6.21; 6.32)	2.49	6.60 (6.56; 6.64)	2.77
Asset index deviation regional mean	0.00 (-0.02; 0.02)	0.88	-0.01 (-0.02; 0.01)	0.88
Malaria incidence	0.26 (0.25; 0.26)	0.11	0.20 (0.20; 0.20)	0.12
Population density	260.07 (241.79; 278.36)	831.22	321.87 (305.78; 337.96)	1098.58
Drought severity index	-0.59 (-0.60; -0.58)	0.66	0.37 (0.36; 0.39)	0.91

Descriptive statistics of categorical and continuous covariates. Note that the Drought severity index corresponds to the self-calibrating Palmer Drought Severity Index (scPDSI). *Source*: DHS and other sources (see Table 2 for detailed information); calculations by authors.

https://doi.org/10.1371/journal.pone.0255073.t005

Milenge, Mwense, Kawambwa, Nchelenge and Chiengi. Stunting was severe in some parts of the Copperbelt province which is predominantly a mining region and the Northern province. Central province had moderate levels, except for Serenje district.

Linear effects

With respect to the linear effects, Table 6 shows the effect of the gender and the area of residence on the posterior mean of the response variable. Considering the posterior mean of the height-for-age z-score of -1.70, boys were found to more stunted compared to girls. Stunting was also found to be higher in children from rural households compared to urban areas. Two patterns well documented in the literature for other countries and also Zambia [12, 17, 52].

Table 6. Estimation results: Linear effects of Model 5.

	Covariate	Posterior mean	95% Credible interval
η_μ	Intercept	-1.70	-1.84; -1.57
	Boys	-0.10	-0.12; -0.08
	Urban	0.04	0.00; 0.07
	Wave ZDHS 2007	-0.07	-0.13; -0.01
η_σ	Intercept	0.23	0.17; 0.29
	Boys	-0.00	-0.01; 0.01
	Urban	0.01	-0.01; 0.04
	Wave ZDHS 2007	0.05	0.01; 0.09

Results of linear covariates included in the Model 5. *Source*: DHS; calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.t006

Socio-economic characteristics

Fig 3 shows the non-linear effect of the asset index, and the years of education of the mother on the mean μ and standard deviation σ of the response variable. Undernutrition has been associated to poverty [4], we observed that children living in poor household showed worse outcomes compare to children living in wealthier households, i.e the z-score is linearly increasing with increasing asset index. The effect of the asset index on the standard deviation does not notably vary across the range of the asset index, indicating a homogenous effect of wealth.

The bottom panel of Fig 3 shows that an increase in the level of education of the mother above eight years of education is associated with an appreciable increase in the height-for-age z-score. Notably also is that, there is not much difference in this effect for mother's education less than 8 years. This entails that primary school education does not improve the nutrition outcome of the children as much. Above eight years of schooling, we see clearly that increase in the years has positive effect on the z-score. Moreover, the variation in the height-for-age z-score is higher with less years of schooling, whereas the variation gradually decreases with increasing levels of education of the mother.

Fig 4 shows the non-linear effect of the number of vaccinations the child received and mother's BMI on the mean and the standard deviation z- score. The top graph shows that there was a positive effect on the mean z-score with increase in the number of vaccinations the child received. There is also greater variation in stunting levels among children who received less than 2 vaccinations.

Low values of the mothers BMI are negatively associated with the height-for-age z-score of the child, while for increasing values of the BMI also an increase in the posterior mean of the z/score can be observed. For values above 40 for the BMI of their mother the results are inconclusive indicated by the widening of the credible intervals. Low values of the BMI of the mother are associated with less variation compared to high values.

Fig 5 shows that increasing malaria incidence about 0.3 had a negative effect on the z-score, however we do not see any meaningful differences in the standard deviation over the spectrum the malaria incidences.

Due to the high correlation of breastfeeding and age of the child, an interaction between these two variables can be presumed for which one has to account for. Fig 6 shows that children below 12 months of age who were breastfeed, were not malnourished. Accordingly, malnutrition mostly seems to be a process that comes to effect as children grow. Stunting was high for children above 36 months of age and who were breastfeed. On the right side, the figure shows that stunting was low in children whose mothers were around 30 years and with respect to birth order, which emphasizes that especially children of very young mothers are those most vulnerable.

Georeferenced characteristics

Fig 7 emphasizes the pronounced north and south pattern after adjusting for all the other variables, in particular after adjusting for wealth and rurality which was already described in the descriptive analysis. The highest variation in the height-for-age z-score was also observed in north (Luapula province) in both waves as can be seen on the right side of the figure.

Discussion & conclusion

Using the two waves of the ZDHS, we modelled the height-for-age z-score by using socio-economic and remote sensed information. To analyse the whole distribution and not just focusing on the conditional mean, we used a Bayesian distributional regression approach accounting for heterogeneity as well.

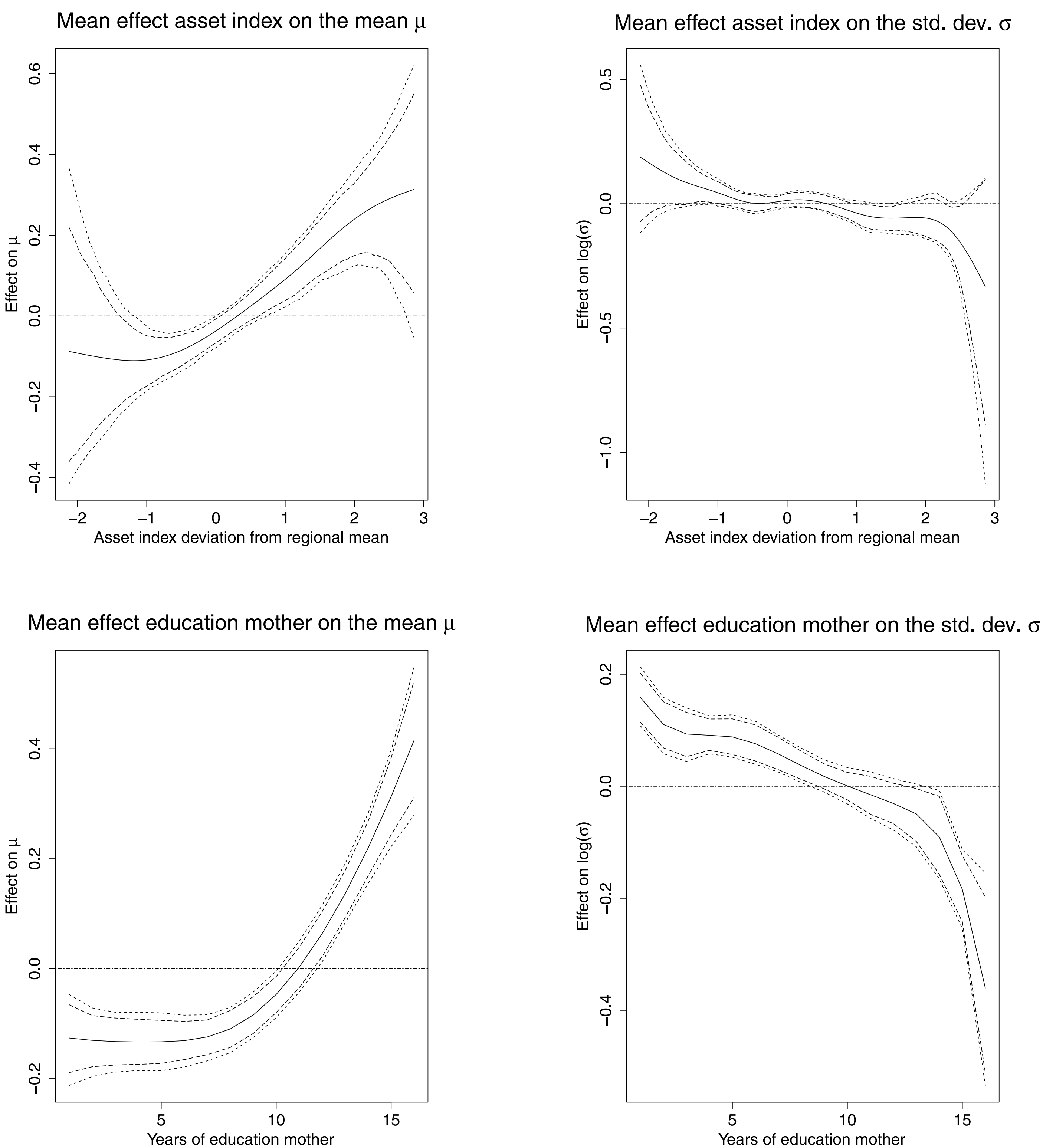

Fig 3. Non-linear effects of the asset index and the years of education. The Figure depicts the mean effects on the mean μ (left), and the standard deviation σ (right) together with 80 per cent and 95 per cent simultaneous credible intervals for the asset index (top), and the years of education of the mother (bottom). *Source*: Demographic and Health Surveys (data); calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.g003

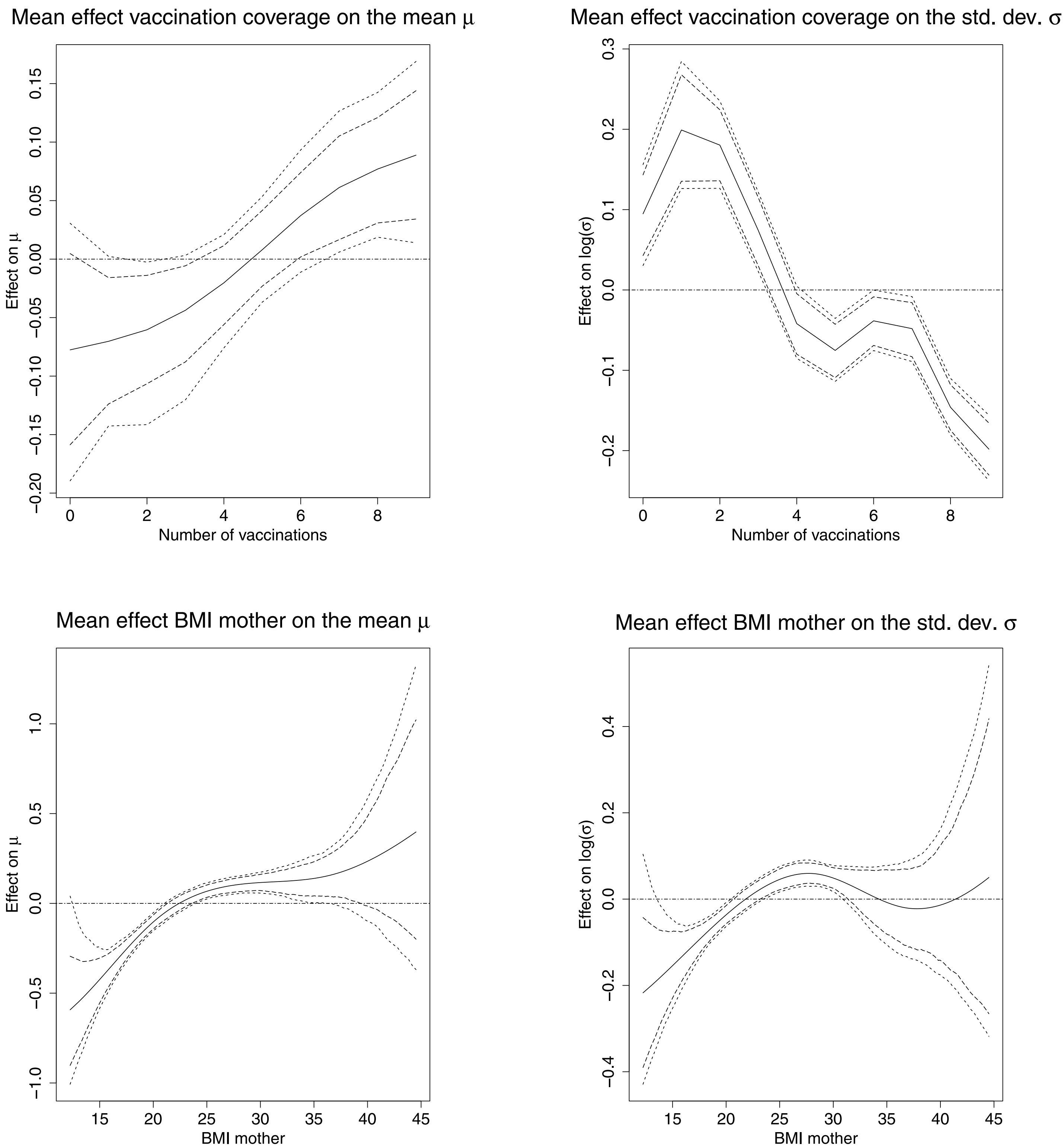

Fig 4. Non-linear effects of the vaccination coverage and the BMI of the mother. The Figure depicts the mean effects on the mean μ (left), and the standard deviation σ (right) together with 80 per cent and 95 per cent simultaneous credible intervals for the vaccination coverage (top), and the BMI of the mother (bottom). *Source*: Demographic and Health Surveys (data); calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.g004

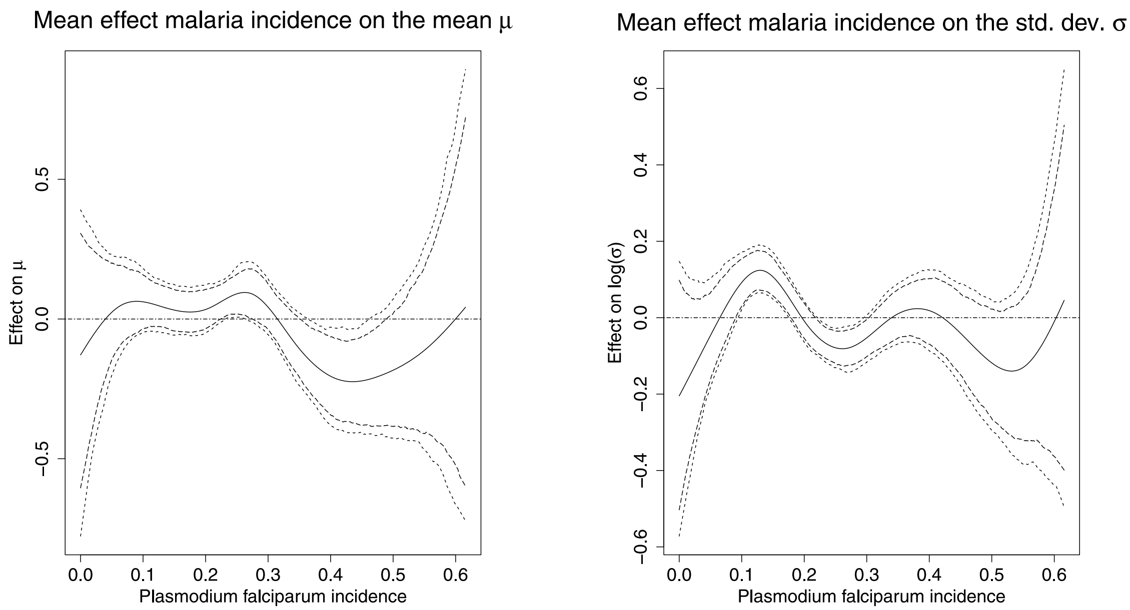

Fig 5. Non-linear effects of the malaria incidence. The Figure depicts the mean effects on the mean μ (left), and the standard deviation σ (right) together with 80 per cent and 95 per cent simultaneous credible intervals for the malaria incidence. *Source*: Demographic and Health Surveys (data); calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.g005

Using Bayesian distributional regression, we assessed the relationship of socio-economic, and remote sensed covariates and stunting. Bayesian distributional regression, presents an advantage in terms of model flexibility allowing to incorporate, amongst others, non-linear effects and spatial effects. This however comes also with the drawback of data intensity and computational complexity.

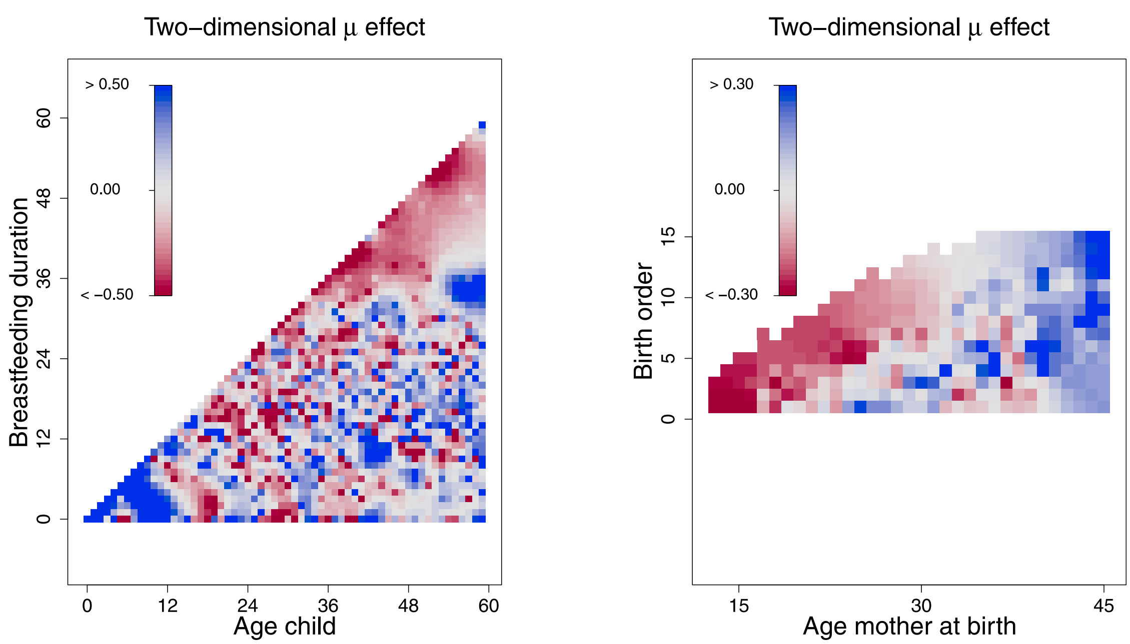

Fig 6. Smooth effects of the interaction of the age of the child and breastfeeding duration, and interaction the age of the mother and the birth order. The Figure depicts the mean effects on the mean μ for the interactions of the age of the child and breastfeeding duration, and the interaction the age of the mother and the birth order, respectively. *Source*: Demographic and Health Surveys (data); calculation by authors.

https://doi.org/10.1371/journal.pone.0255073.g006

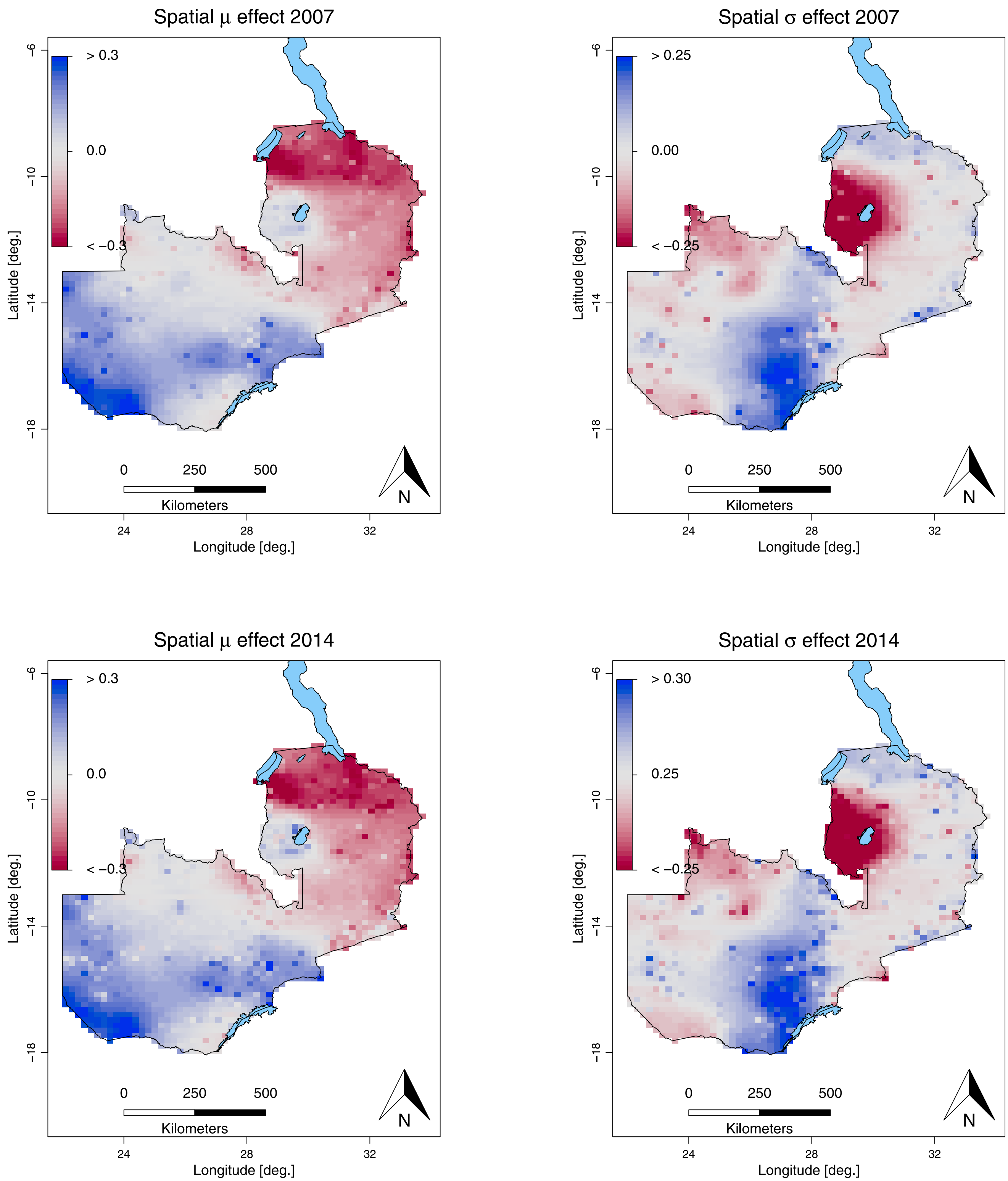

Fig 7. Smooth spatial effect. The Figure depicts the mean spatial effect of the mean μ (left), and the standard deviation σ (right) for the year 2007 (left) and the year 2013/14 (right). *Source*: Demographic and Health Surveys (data) and Database of Global Administrative Areas (boundary information); calculation by authors. The shapefile used to create these maps is republished from [54] under CC BY license, with permission from Robert J. Hijmans, original copyright [2021].

https://doi.org/10.1371/journal.pone.0255073.g007

Remote sensed techniques can be useful for future research on community health assessment as these techniques provide an advantage to take measurements quickly for remote and hard to reach areas. The data also enable to make meaningful analyses at sub-national levels which can improve targeting of interventions due to high levels of geographic specificity [26, 33], however they do not give a full picture. Therefore, it is important to account for other covariates such as socio-economic characteristics at the individual or household level.

When relying on remote sensed information to asses anthropometric measures or biophysical developments, great caution should be taken with respect to data quality. Our study finds that remote sensed covariates alone explain little of the variation of the response, this emphasizes the need to control also for socio-economic characteristics. We find that the combination of remote sensed data and socio-economic characteristics explain more of the variation of the response, compared to solely focusing on one of the two sources of explanatory variables. In addition this also highlights the strong influence of socio-economic covariates or can be seen as an indicator of poor quality of the available remote sensed information.

Clear non-linear patterns emerged with respect to the years of education of the mother, and number of vaccinations. There was a clear non-linear tendency among children whose mothers had up to eight years of schooling having a low height-for-age z-score. For children of mothers with secondary or higher education the height-for-age z-score starts to improve strongly. This trend is consistent with what has been observed in others studies were odds of stunting were higher among children from mothers who had few years of education [12, 52] and lowest among those who had advanced years in education [52]. Higher education level has been associated with increased income levels and improved knowledge among mothers who are usually the primary caregivers. As such educated mothers are more likely to take better care of their children by making informed nutritional decisions [24, 52, 53]. Increasing number of vaccinations showed improved z-score among children. Even though the effect was significant, the same size of the effect might not be relevant in practice.

Moreover, considering the full distribution like we did shows that the variation is highest for low levels of education and decreases with increasing years of education. This study did not consider the association of paternal education and the child z-score, however in another study, it was found that it was Maternal education that had a positive impact on children's nutritional status [17].

We observe differences in levels of malnutrition in various regions in Zambia. One consistent pattern is that of discrepancy between the rural areas which are worse off compared to urban areas and confirms socio-economic inequalities between rural and urban areas. This may suggest social and economic inequalities between such areas. This has already been documented in other studies [12, 14, 24]. Furthermore, in terms of a spatial distribution, when you consider a smooth spatial effect, there is a clear regional variation in addition to the effect of rurality. Even after accounting for economic activities, the farming southern regions tend to be well off compared to the more industrialised northern areas. There is need to investigate further the underlying factors that contribute to the variation in the height-for-age z-score.

The present study shows that stunting still remain high in Zambia with remarkable regional inequalities and the decline is gradual which is unacceptable. There is need therefore to address the socio-economic indicators if this status is to improve.

Supporting information

S1 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_μ of the effect of the asset index, the two-dimensional effect of the birth order and the age of the mother at birth, the two-dimensional effect of the age of the child and breastfeeding duration, and the

effect of the mothers years of education. Demographic and Health Surveys calculation by authors.
(EPS)

S2 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_μ of the household size, the linear effects, and the malaria incidence. Demographic and Health Surveys calculation by authors.
(EPS)

S3 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_μ of the maternal BMI, the population density, the aridity index, and the spatial effect. Demographic and Health Surveys calculation by authors.
(EPS)

S4 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_μ of the effect of the number of vaccination. Demographic and Health Surveys calculation by authors.
(EPS)

S5 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_σ of the effect of the asset index, and the effect of the mothers years of education. Demographic and Health Surveys calculation by authors.
(EPS)

S6 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_σ of the household size, the linear effects, and the malaria incidence. Demographic and Health Surveys calculation by authors.
(EPS)

S7 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_σ of the maternal BMI, the population density, the aridity index, and the spatial effect. Demographic and Health Surveys calculation by authors.
(EPS)

S8 Fig. Sampling paths. The Figure depicts the sampling paths of the parameters in η_σ of the effect of the number of vaccination. Demographic and Health Surveys calculation by authors.
(EPS)

Acknowledgments

The computational results presented have been achieved (in part) using the HPC infrastructure LEO of the University of Innsbruck.

Author Contributions

Conceptualization: Given Moonga, Stephan Böse-O'Reilly, Johannes Seiler.

Data curation: Kenneth Harttgen, Johannes Seiler.

Formal analysis: Given Moonga, Johannes Seiler.

Methodology: Given Moonga, Ursula Berger, Johannes Seiler.

Supervision: Stephan Böse-O'Reilly, Ursula Berger, Kenneth Harttgen, Charles Michelo, Dennis Nowak, Uwe Siebert, John Yabe.

Visualization: Given Moonga, Johannes Seiler.

Writing – original draft: Given Moonga, Johannes Seiler.

Writing – review & editing: Given Moonga, Stephan Böse-O'Reilly, Ursula Berger, Kenneth Harttgen, Dennis Nowak, Uwe Siebert, Johannes Seiler.

References

1. World Health Organization. Global nutrition targets 2025: Stunting policy brief. World Health Organization; 2014.
2. Dwyer-Lindgren L, Kakungu F, Hangoma P, Ng M, Wang H, Flaxman AD, et al. Estimation of district-level under-5 mortality in Zambia using birth history data, 1980–-2010. Spatial and Spatio-temporal Epidemiologyl. 2014; 11:89–107. https://doi.org/10.1016/j.sste.2014.09.002
3. World Health Organization. The state of food security and nutrition in the world 2018: building climate resilience for food security and nutrition. Food & Agriculture Organization; 2018.
4. Müller O, Krawinkel M. Malnutrition and health in developing countries. Cmaj. 2005; 173(3):279–286. https://doi.org/10.1503/cmaj.050342 PMID: 16076825
5. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, de Onis M, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. The Lancet. 2013; 382(9890):427–451. https://doi.org/10.1016/S0140-6736(13)60937-X PMID: 23746772
6. Tette EM, Sifah EK, Nartey ET. Factors affecting malnutrition in children and the uptake of interventions to prevent the condition. BMC pediatrics. 2015; 15(1):189. https://doi.org/10.1186/s12887-015-0496-3 PMID: 26586172
7. Menon P, Headey D, Avula R, Nguyen PH. Understanding the geographical burden of stunting in India: A regression-decomposition analysis of district-level data from 2015–16. Maternal & child nutrition. 2018; 14(4):e12620. https://doi.org/10.1111/mcn.12620 PMID: 29797455
8. Svedberg P. Undernutrition in Sub-Saharan Africa: Is there a gender bias? The Journal of Development Studies. 1990; 26(3):469–486. https://doi.org/10.1080/00220389008422165
9. Adebayo SB. Modelling childhood malnutrition in Zambia: an adaptive bayesian splines approach. Statistical Methods and Applications. 2003; 12(2):227–241. https://doi.org/10.1007/s10260-003-0057-z
10. Marx S, Phalkey R, Aranda-Jan CB, Profe J, Sauerborn R, Höfle B. Geographic information analysis and web-based geoportals to explore malnutrition in Sub-Saharan Africa: a systematic review of approaches. BMC public health. 2014; 14(1):1189. https://doi.org/10.1186/1471-2458-14-1189 PMID: 25409548
11. Gorstein J, Akre J. The Use of Anthropometry to Assess Nutritional Status. World health statistics quarterly Rapport trimestriel de statistiques sanitaires mondiales. 1988; 41:48–58. PMID: 3176514
12. Mzumara B, Bwembya P, Halwiindi H, Mugode R, Banda J. Factors associated with stunting among children below five years of age in Zambia: evidence from the 2014 Zambia demographic and health survey. BMC Nutrition. 2018; 4(1):51. https://doi.org/10.1186/s40795-018-0260-9 PMID: 32153912
13. UNICEF. Strategy for Improved Nutrition of Children and Women in Developing Countries. A UNICEF Policy Review. New York: UNICEF; 1990.
14. Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al. Maternal and child undernutrition: global and regional exposures and health consequences. The Lancet. 2008; 371(9608):243–260. https://doi.org/10.1016/S0140-6736(07)61690-0 PMID: 18207566
15. UNICEF. Improving child nutrition: the achievable imperative for global progress. New York; 2013.
16. Gayawan E, Adebayo SB, Komolafe AA, Akomolafe AA. Spatial distribution of malnutrition among children under five in Nigeria: A Bayesian quantile regression approach. Applied Spatial Analysis and Policy. 2019; 12(2):229–254. https://doi.org/10.1007/s12061-017-9240-8
17. Kandala NB, Madungu TP, Emina JB, Nzita KP, Cappuccio FP. Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter? BMC public health. 2011; 11(1):261. https://doi.org/10.1186/1471-2458-11-261 PMID: 21518428
18. Kandala NB, Fahrmeir L, Klasen S, Priebe J. Geo-additive models of childhood undernutrition in three sub-Saharan African countries. Population, Space and Place. 2009; 15(5):461–473. https://doi.org/10.1002/psp.524
19. Central Statistical Office (CSO)[Zambia] Ministry of Health (MOH)[Zambia] and ICF International. Zambia demographic and health survey 2013–14; 2014.
20. Ngoma C, Mayimbo S. The Negative Impact of Poverty on the Health of Women and Children. Annals of Medical and Health Sciences Research. 2017; 7(6).

21. Umlauf N, Kneib T. A primer on Bayesian distributional regression. Statistical Modelling. 2018; 18(3-4):219–247. https://doi.org/10.1177/1471082X18759140

22. Kneib T. Beyond mean regression. Statistical Modelling. 2013; 13(4):275–303. https://doi.org/10.1177/1471082X13494159

23. Klein N. An Introduction to Bayesian Structured Additive Distributional Regression; 2013.

24. Kandala NB, Lang S, Klasen S, Fahrmeir L. Semiparametric Analysis of the Socio-Demographic and Spatial Determinants of Undernutrition in Two African Countries; 2001.

25. Belitz C, Hübner J, Klasen S, Lang S. In: Determinants of the Socioeconomic and Spatial Pattern of Undernutrition by Sex in India: A Geoadditive Semi-parametric Regression Approach. Heidelberg: Physica-Verlag HD; 2010. p. 155–179.

26. Brown ME, Grace K, Shively G, Johnson KB, Carroll M. Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change. Population and environment. 2014; 36(1):48–72. https://doi.org/10.1007/s11111-013-0201-0 PMID: 25132700

27. Corsi DJ, Perkins JM, Subramanian S. Child anthropometry data quality from Demographic and Health Surveys, Multiple Indicator Cluster Surveys, and National Nutrition Surveys in the West Central Africa region: are we comparing apples and oranges? Global health action. 2017; 10(1):1328185. https://doi.org/10.1080/16549716.2017.1328185 PMID: 28641057

28. Perumal N, Bassani DG, Roth DE. Use and misuse of stunting as a measure of child health. The Journal of nutrition. 2018; 148(3):311–315. https://doi.org/10.1093/jn/nxx064 PMID: 29546307

29. Hong R, Banta JE, Betancourt JA. Relationship between household wealth inequality and chronic childhood under-nutrition in Bangladesh. International Journal for Equity in Health. 2006; 5(1):15. https://doi.org/10.1186/1475-9276-5-15 PMID: 17147798

30. Filmer D, Pritchett LH. Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India. Demography. 2001; 38(1):115–132. https://doi.org/10.1353/dem.2001.0003 PMID: 11227840

31. Sahn DE, Stifel D. Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. Review of Income and Wealth. 2003; 49(4):463–489. https://doi.org/10.1111/j.0034-6586.2003.00100.x

32. Frongillo Edward A Jr, de Onis M, Hanson KMP. Socioeconomic and Demographic Factors Are Associated with Worldwide Patterns of Stunting and Wasting of Children. The Journal of Nutrition. 1997; 127(12):2302–2309. https://doi.org/10.1093/jn/127.12.2302

33. Osgood-Zimmerman A, Millear AI, Stubbs RW, Shields C, Pickering BV, Earl L, et al. Mapping child growth failure in Africa between 2000 and 2015. Nature. 2018; 555:41–47. https://doi.org/10.1038/nature25760 PMID: 29493591

34. Amoah B, Giorgi E, Heyes DJ, van Burren S, Diggle PJ. Geostatistical modelling of the association between malaria and child growth in Africa. International Journal of Health Geographics. 2018; 17(1):7. https://doi.org/10.1186/s12942-018-0127-y PMID: 29482559

35. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. Nature. 2015; 526:207–211. https://doi.org/10.1038/nature15535 PMID: 26375008

36. van der Schrier G, Barichivich J, Briffa KR, Jones PD. A scPDSI-based global data set of dry and wet spells for 1901–2009. Journal of Geophysical Research: Atmospheres. 2013; 118(10):4025–4048.

37. Barichivich J, Osborn TJ, Harris I, van der Schrier G, Jones PD. Drought. In: Hartfield G, Blunden J, Arndt DS, editors. State of the Climate in 2018. vol. 100. Bulletin of the American Meteorological Society; 2018.

38. Center for International Earth Science Information Network—CIESIN—Columbia University. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 10; 2017.

39. Adebayo SB, Fahrmeir L. Analysing child mortality in Nigeria with geoadditive discrete-time survival models. Statistics in Medicine. 2005; 24(5):709–728. https://doi.org/10.1002/sim.1842 PMID: 15696506

40. Osei FB, Duker AA, Stein A. Bayesian structured additive regression modeling of epidemic data: application to cholera. BMC Medical Research Methodology. 2012; 12(1):118. https://doi.org/10.1186/1471-2288-12-118 PMID: 22866662

41. Klein N, Kneib T, Klasen S, Lang S. Bayesian structured additive distributional regression for multivariate responses. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2015; 64(4):569–591. https://doi.org/10.1111/rssc.12090

42. Seiler J, Harttgen K, Kneib T, Lang S. Modelling children's anthropometric status using Bayesian distributional regression merging socio-economic and remote sensed data from South Asia and sub-

Saharan Africa. Economics & Human Biology. 2021; 40:100950. https://doi.org/10.1016/j.ehb.2020.100950 PMID: 33321408

43. Klein N, Kneib T, Lang S, Sohn A. Bayesian structured additive distributional regression with an application to regional income inequality in Germany. Ann Appl Stat. 2015; 9:1024–1052. https://doi.org/10.1214/15-AOAS823

44. Dunn PK, Smyth GK. Randomized Quantile Residuals. Journal of Computational and Graphical Statistics. 1996; 5(3):236–244. https://doi.org/10.1080/10618600.1996.10474708

45. Belitz C, Brezger A, Klein N, Kneib T, Lang S, Umlauf N. BayesX: Software for Bayesian Inference in Structured Additive Regression Models. Version 3.0.2; 2015. Available from: http://www.BayesX.org/.

46. Lang S, Brezger A. Bayesian P-Splines. Journal of Computational and Graphical Statistics. 2004; 13 (1):183–212. https://doi.org/10.1198/1061860043010

47. Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. Statististical Science. 1996; 11 (2):89–121. https://doi.org/10.1214/ss/1038425655

48. Hastie TJ, Tibshirani RJ. Varying-Coefficient Models. Journal of the Royal Statistical Society Series B (Methodological). 1993; 55(4):757–796. https://doi.org/10.1111/j.2517-6161.1993.tb01939.x

49. Fahrmeir L, Kneib T, Lang S, Marx B. Regression: Models, Methods and Applications. 1st ed. Springer-Verlag Berlin Heidelberg; 2013.

50. Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A. Bayesian Measures of Model Complexity and Fit. Journal of the Royal Statistical Society Series B (Statistical Methodology). 2002; 64(4):583–639. https://doi.org/10.1111/1467-9868.00353

51. Watanabe S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research. 2010; 11:3571–3594.

52. Angdembe MR, Dulal BP, Bhattarai K, Karn S. Trends and predictors of inequality in childhood stunting in Nepal from 1996 to 2016. International journal for equity in health. 2019; 18(1):42. https://doi.org/10.1186/s12939-019-0944-z PMID: 30836975

53. Svefors P, Rahman A, Ekström EC, Khan AI, Lindström E, Persson LÅke, et al. Stunted at 10 years. Linear growth trajectories and stunting from birth to pre-adolescence in a rural Bangladeshi cohort. PloS one. 2016; 11(3):e0149700. https://doi.org/10.1371/journal.pone.0149700 PMID: 26934484

54. Hijmans, Robert J, University of California, Berkeley, Museum of Vertebrate Zoology. Global Administrative Areas (GADM), version 3.6; 2018. Available from: http://gadm.org.

Chapter III

Geospatial Approach to Investigate Spatial Clustering and Hotspots of Blood Lead Levels in Children within Kabwe, Zambia

Environmental Research 207 (2022) 112646

Contents lists available at ScienceDirect

Environmental Research

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

Geospatial approach to investigate spatial clustering and hotspots of blood lead levels in children within Kabwe, Zambia

Given Moonga[a,b,c,h,*], Moses N. Chisola[d], Ursula Berger[f], Dennis Nowak[a], John Yabe[e], Hokuto Nakata[g], Shouta Nakayama[g], Mayumi Ishizuka[g], Stephan Bose-O'Reilly[a,h,i]

[a] Institute and Clinic for Occupational-, Social- and Environmental Medicine, LMU University Hospital Munich, Ziemssenstr. 1, D-80336, Munich, Germany
[b] CIH[LMU] Center for International Health, LMU University Hospital, Munich, Germany
[c] Department of Epidemiology and Biostatistics, University of Zambia, Lusaka, Zambia
[d] Department of Geography and Environmental Studies, University of Zambia, Lusaka, Zambia
[e] School of Veterinary Medicine, University of Zambia, Lusaka, Zambia
[f] Institute for Medical Information Processing, Biometry and Epidemiology, Ludwig Maximilian University Munich, Munich, Germany
[g] Laboratory of Toxicology, Department of Environmental Veterinary Sciences, Faculty of Veterinary Medicine, Hokkaido University, Sapporo, Japan
[h] Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT (Private University for Health Sciences, Medical Informatics and Technology), Hall i.T, Austria
[i] University Children's Hospital Regensburg (KUNO-Clinics), University of Regensburg, Clinic St. Hedwig, Regensburg, Germany

ARTICLE INFO

Keywords:
Spatial cluster
Geospatial
Blood lead
Hot spots
Moran's I

ABSTRACT

Background: Communities around Kabwe, Zambia are exposed to lead due to deposits from an old lead (Pb) and zinc (Zn) mining site. Children are particularly more vulnerable than adults, presenting with greatest risk of health complications. They have increased oral uptake due to their hand to mouth activities. Spatial analysis of childhood lead exposure is useful in identifying specific areas with highest risk of pollution. The objective of the current study was to use a geospatial approach to investigate spatial clustering and hotspots of blood lead levels in children within Kabwe.

Methods: We analysed existing data on blood lead levels (BLL) for 362 children below the age of 15 from Kabwe town. We used spatial autocorrelation methods involving the global Moran's I and local Getis-Ord Gi*statistic in ArcMap 10.5.1, to test for spatial dependency among the blood lead levels in children using the household geolocations.

Results: BLL in children from Kabwe are spatially autocorrelated with a Moran's Index of 0.62 ($p < 0.001$). We found distinct hotspots (mean 51.9 μg/dL) in communities close to the old lead and zinc-mining site, lying on its western side. Whereas coldspots (mean 7 μg/dL) where observed in areas distant to the mine and traced on the eastern side. This pattern suggests a possible association between observed BLL and distance from the abandoned lead and zinc mine, and prevailing winds.

Conclusion: Using geocoded data for households, we found clustering of childhood blood lead and identified distinct hotspot areas with high lead levels for Kabwe town. The geospatial approach used is especially valuable in resource-constrained settings like Zambia, where the precise identification of high risk locations allows for the initiation of targeted remedial and treatment programs.

1. Introduction

Lead (Pb) is a toxic metal and a global health hazard. Over 815 million children worldwide are reported to have dangerously high concentrations of Pb in their bloodstream Burki (2020), (Zhang et al., 2008). Low-income countries of sub-Saharan Africa face the greatest impacts of childhood lead poisoning (Landrigan et al., 2018). Children are more vulnerable to Pb related negative health outcomes, compared to adults, as their still developing central nervous system is more susceptible to Pb exposure, mainly during the primary developmental

* Corresponding author. Institute and Clinic for Occupational-, Social- and Environmental Medicine, LMU University Hospital Munich, Ziemssenstr. 1, D-80336, Munich, Germany.
E-mail address: Given.Moonga@lrz.uni-muenchen.de (G. Moonga).

https://doi.org/10.1016/j.envres.2021.112646
Received 14 March 2021; Received in revised form 26 December 2021; Accepted 28 December 2021
Available online 31 December 2021

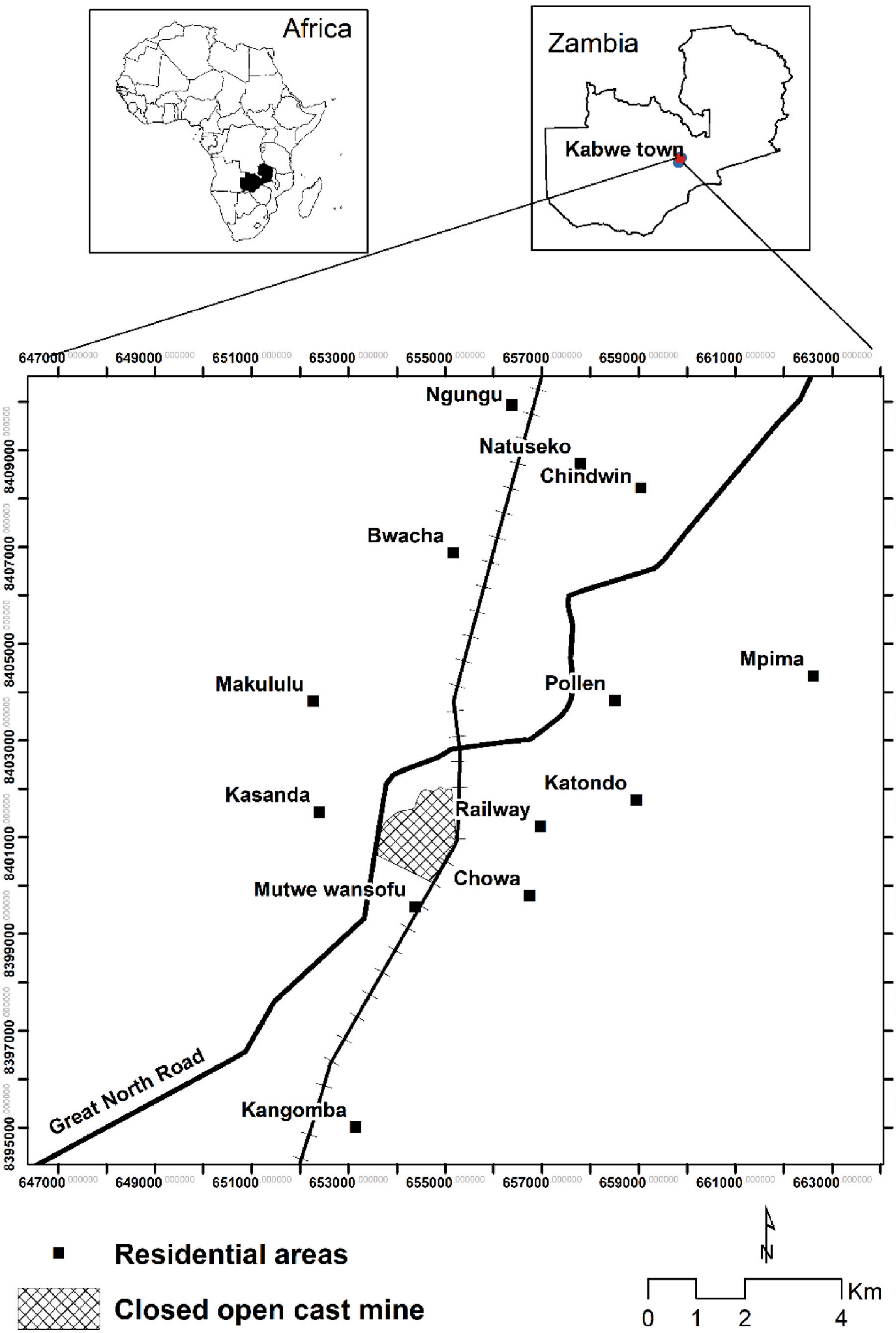

Fig. 1. Study area.

stages (Rooney et al., 2018; Scheuplein et al., 2002). Pb exposure, even at low levels, has been associated with deficits in cognitive functioning and intelligence quotient (IQ) in children (Liu et al., 2011; Lanphear et al., 2005). No level of Pb exposure appears to be safe (Bellinger et al., 1991). High Pb levels of exposure exceeding 80 μg/dL can lead to anaemia, seizures, coma, encephalopathy, and death (World Health Organization, 2010; Wang et al., 2009).

While the physiological mechanisms for Pb dose–response may be similar for all children (Mielke et al., 2019), there are varied exposure pathways. In most settings, blood lead levels (BLL) are related to Pb soil contamination (Matte et al., 1991) and social economic factors (Mielke et al., 2019). The impact of socio-economic and environmental covariates on the Pb exposure in children has been well documented (Stark et al., 1982; Mielke et al., 2019).

Globally childhood Pb poisoning remains a major environmental health concern in cities and communities with Pb-contaminated soils (Ikem et al., 2008; Lo et al., 2012). Contaminated soil, especially if the soil is dry and dusty, can be ingested and absorbed. Children are particularly vulnerable to Pb intoxication due to their hand-to-mouth activities, as such they are more likely to ingest more Pb from contaminated soils. Also, vulnerability is increased due to their much higher gastrointestinal absorption relatively to adults (World Health Organization, 2010; Plumlee et al., 2013).

Kabwe town, located in the Central Province of Zambia had a long history of lead (Pb), zinc (Zn) and cadmium (Cd) mining dating back to the 1900's. The operations came to a stop in 1994, leaving behind ineliminable impact of pollution on the environment (Nakayama et al., 2011). The open cast mining and the resulting big tailing hill still contain a lot of Pb. Neither the tailing hill nor the open pit side were ever properly rehabilitated. Most roads in Kabwe are unpaved; the backyards of the housing areas are dry and dusty. During the 9 months period of dry season, Pb is easily transported towards the windward lying areas and communities.

Previous Pb pollution remediation initiatives in Kabwe have not been sustainable. For example, from 2003 to 2011, the World Bank funded the Copperbelt Environment Project (CEP). The project aimed at cleaning up affected communities, and the treatment of children with high blood Pb levels. Environmental remedial activities included removal of top soil and planting of grass in the affected communities. These activities halted once the project ended in 2011 due to lack of resources and local capacity (Bank, 2011). Currently there is an ongoing project, "the Zambia Mining and Environmental Remediation and Improvement Project (ZMERIP)" aimed at environmental remediation and treatment of 10, 000 children with high BLL. The project has an environmental remediation component with activities including; removal of top soil and planting of grass in affected communities.

Ettler et al. (2020) analysed soil samples from Kabwe townships and main roads and reported that soil Pb levels were above recommended levels for residential areas. They also observed that the geometric mean for soil Pb in townships closer to the mining sites were higher than far off areas. Highly polluted townships were those immediately adjacent to the former Kabwe mining complex and homes downwind from the smelter and the tailings (Bose-O'Reilly, Yabe et al., 2018a,b).

Former reports and scientific publications showed high BLL for people living in Kabwe, due to their continued exposure to Pb. In most compounds, BLL in children is reported to be above 65 μg/dL (Yabe et al., 2015, 2020). The Pb contaminated dust that emanates from the mine dump is the main source of this observed pollution, affecting mostly children (Bose-O'reilly et al., 2018a).

1.1. Spatial patterns of Pb exposure in Kabwe

Analysis of the Pb spatial distribution is particularly important in the identification of areas with high risk of exposure (Akkus and Ozdenerol, 2014; Miranda et al., 2002). In resource-limited settings such as Zambia, use of Geographic Information System (GIS) tools to identify affected populations is essential in ensuring that remediation and treatment programmes target communities at greatest risk. The integration of GIS to understand Pb exposure patterns and coverage of interventions also strengthens the implementation of control programmes. Spatial analytical methods such as cluster and hot spot analysis are ideal in this regard (Zhang et al., 2008; Akkus and Ozdenerol, 2014). Moreover, hotspot analysis can help understand disparities in exposure and health outcomes at a lower administrative level. This precision is valuable as it enables quantification of inequalities and identification of successes and failures of programmes and policies at the local level (Osgood-Zimmerman et al., 2018).

Previous studies observed that BLL in Kabwe varied depending on the townships, and relative distance from the mine site (Yabe et al., 2020, Bose-O'reilly et al., 2018b). Children in residential areas (Kasanda, Makululu, Chowa) closer to the mine site have higher average BLL while those from further located areas (Hamududu) have lower BLL (Yabe et al., 2020). Another study observed this trend, in domestic dogs, where the blood Pb concentrations were higher in dogs from communities that were located near the mine than the far-flung residential areas (Toyomaki et al., 2020).

Despite the indication of regional variation in the distribution of BLL in Kabwe, until now no study has investigated the household spatial distribution of BLL to identify hotspot and coldspot areas. Hotspot areas can be clustered (spatial clusters) or exist individually (spatial outliers). Spatial clusters in this context are areas (households) with high BLL values surrounded by observations with also high values. Whereas spatial outliers are households with high BLL surrounded by samples with normal or low values (Zhang and Lin, 2006). The aim of this study therefore was to investigate clustering of BLL and identify hotspot areas using GIS techniques.

2. Methodology

2.1. Study population

We re-analysed BLL data that was collected in 2017 by the University of Zambia with collaborators from Hokkaido University under the Kabwe Mine Pollution Amelioration Initiative (KAMPAI) Project. Forty (40) Standard Enumeration Areas (SEA) falling within the catchment area of health facilities were randomly selected, from which 25 households in each were randomly selected and geo-coordinates recorded. Blood was collected from the father, mother and two children. Data was collected on 1000 households with a total of 1190 household members, of which 291 were younger children (three months to three years old), 271 older children (four to nine years old), 412 mothers, and 216 fathers. A detailed description on how the data was collected has been provided by Yabe et al. (2020).

We analysed data for 362 children below the age of 15, each sampled from a single household from the 40 SEAs. Using location parameters (geo-coordinates), we analysed data from these households. We selected the youngest child from each household that was enlisted in the study by Yabe et al. (2020). We focused our analysis on children as they are reported to be the most vulnerable to impacts of lead poisoning (Bose-O'reilly et al., 2018b). Moreover, this is the age range reported to have highest BLL. Fig. 1, shows the residential areas that were included in current study.

2.2. Laboratory methods

Pb analysis in whole blood samples was done on-site immediately after blood sample collection using a point-of-care blood Pb testing analyser, LeadCare© II (Magellan Diagnostics, USA). The LeadCare II Analyser used had limits of quantification of 3.3–65 μg/dL, as such, precise levels out of this range could not be determined. BLLs below instrument detection limit were therefore treated as 1.65 μg/dL, the mean of 0 and 3.3. For samples above 65 μg/dL, a 3 times dilution was done using 0.1% HCl. Detailed laboratory procedures are described elsewhere (Yabe et al., 2020).

2.3. Statistical methods

We used spatial autocorrelation methods involving the global Moran's I and local Getis-Ord Gi*statistic to assess the spatial patterns in the children's BLL. These methods are briefly discussed below.

2.3.1. Test for spatial dependency

The global Moran's I was implemented in ArcMap 10.5.1 (Release, 2012) to test for a general spatial dependency among the BLL in children in Kabwe, i.e. to examine whether high or low levels of BLL show spatial clusters or whether they are scattered in a random pattern.

The global Moran's I is given by the formula:

$$I = \frac{n}{S_o} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} r_i r_j}{\sum_{i=1}^{n} r_i^2} \tag{1}$$

where r_i is the deviation of the child's BLL value at area i from its mean $(x_i - \mu)$, $w_{i,j}$ is the spatial weight between area ,i and j, n are the numbers of observations and S_o is the sum of all the spatial weights:

$$S_o = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \tag{2}$$

2.3.2. Hotspot analysis

To identify the spatial location (coordinates) of cluster of high and of low BLL levels (hotspots and coldspots), we used the local Getis-Ord Gi*statistic (Getis and Ord, 2010). We examined the BLL observation with respect to neighbouring BLL observations. An observation with a high value or low value does not necessarily imply a hotspot or a coldspot, respectively, unless it is surrounded by observations with high values (hotspot) or low values (coldspot). Thus, a large positive Gi*statistic is obtained when the local sum of an observation and its neighbours is larger than the expected local sum indicating clustering of high values, while a small value of the Gi*statistic indicates clustering of low BLL values. In addition, we evaluated whether the Gi*statistic significantly differs from 0, i.e., whether a cluster has a significantly elevated or significantly low BLL level, leading to the definition of hotspot and coldspot, respectively. The Gi*statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \mu \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}{n-1}}} \tag{3}$$

where x_j is the BLL value for the child in area j, $w_{i,j}$ is the spatial weight between area i and j, n is the number of observations, mu is the mean BLL level and S is the standard deviation of x, i.e.:

$$\mu = \frac{\sum_{j=1}^{n} x_j}{n} \tag{4}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - (u)^2} \tag{5}$$

All spatial analyses were performed using ArcMap 10.5.1. All test decisions were based on a significance level of 0.05.

2.4. Ethical clearance

The University of Zambia Research Ethics Committee (UNZAREC; REF. No. 012-04-16) approved the study. The Ministry of Health through the Zambia National Health Research Ethics Board and the Kabwe District Medical Office granted further approvals (Yabe et al., 2020). In accordance to ethical guidelines and data protection, until the point of data integration and analysis the location coordinates were stored separately from attribute data, all attribute data were de-identified.

Table 1
Blood lead level (BLL) distribution by age.

	All	0–3 years	4–9 years	10–15 years
	n = 362	n = 173	n = 170	n = 19
Mean (μg/dL)	30.1	31.9	28.7	26.9
SD	25.8	28.2	23.4	23.3
Median (μg/dL)	23.8	24.6	21.9	29.7
Minimum (μg/dL)	3.3	3.3	3.3	3.3
Maximum (μg/dL)	162.3	162.3	94.8	67.2

Table 2
Blood lead level (BLL) distribution in cold and hot spots.

	Mean (μg/dL)	Median (μg/dL)	SD	Min (μg/dL)	Max (μg/dL)
Coldspot 90%	15.2	10.9	14.9	3	94.8
Coldspot 95%	15.7	10.7	11.9	3	41.2
Coldspot 99%	7	5.5	6.4	3	38.7
Hot spot 99%	51.9	49.8	20.9	12	162.3

Table 3
Blood lead level (BLL) distribution by residential areas.

Clinic	Mean (μg/dL)	Median μg/dL)	SD	Maximum (μg/dL)	Minimum (μg/dL)
Bwacha	11.1	6.7	16.8	94.8	3
Chowa	24.6	24	10	48.3	7.4
Hamududu	4.5	3	3.2	18.4	3
Kangomba	10.6	10.1	8.8	37.2	3
Kasanda	60.2	56.9	21.8	162.3	30.4
Katondo	11.6	6.7	12.2	38.7	3
Mahatma Ghandi	5.9	5	2.2	9	3.9
Makululu	43.7	40.2	21.1	118.5	9.1
Mpima prison	7.3	6.7	4.7	23.3	3
Natuseko	11.5	10.4	7.1	30.7	3.9
Ngungu	6	4.7	3.9	14.2	3
Pollen	5.3	4.7	1.7	8.3	3.9
Railway	16.4	15.4	7.2	26.2	8.8
Total	29.9	22.3	26	162.3	3

3. Results

3.1. Blood lead level distribution

Table 1 shows age categories and the distribution of BLL in these age groups. The largest age group (47.7%) was 0–3 years while the smallest (5%) was the 10–15 years group. The global mean BLL was 30.1 μg/dL and the median 23.8 μg/dL, which is comparable to the values reported for the young child by Yabe et al. (2020). The distribution of individual blood Pb ranged from a minimum of 3.3 μg/dL to maximum of 162 μg/dL. Age groups 0–3 and 10–15 years had the highest mean (31.9 μg/dL), and median (29.7 μg/dL) respectively. As shown in Table 2, the mean blood Pb in coldspots was 7 μg/dL at 99% confidence level while the hotspots had a mean of 51.9 μg/dL at a similar confidence level. Table 3 shows the distribution of blood Pb in the communities contained in our analysis. Kasanda had the highest mean (60.2 μg/dL) while the lowest mean was observed in Hamududu (4.5 μg/dL).

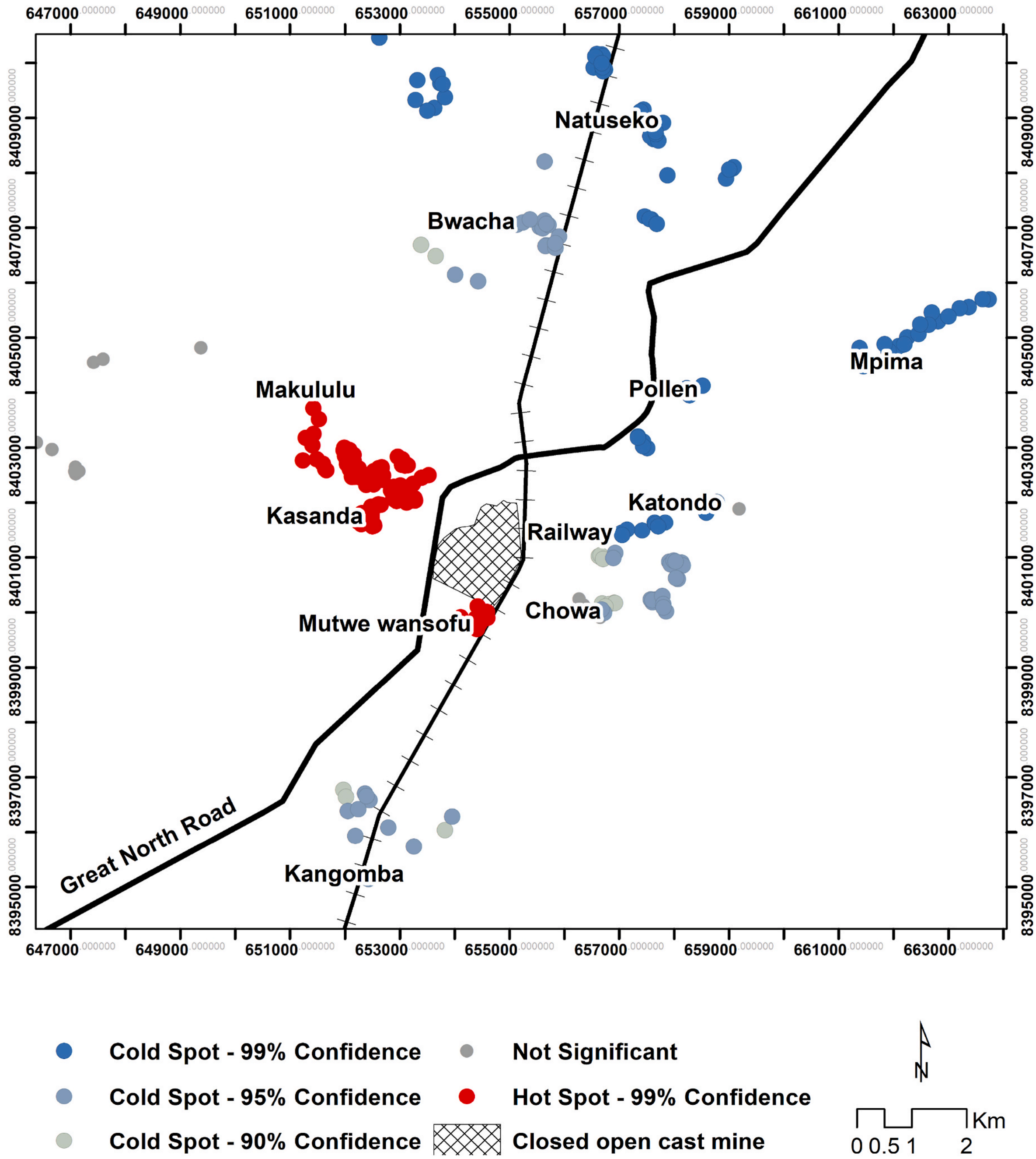

Fig. 2. Spatial distribution of blood lead levels among the children in Kabwe, Zambia.

3.2. Spatial autocorrelation analysis of blood lead levels

We observe a positive Global Moran's I (Index) of 0.63 (Z-score 26.1, p-value < 0001) indicating significant spatial clustering of children's BLL levels in Kabwe. In Fig. 2: the red dots indicate areas with high Pb concentration surrounded by other samples with a high concentration (hotspots). Whereas the blue dots indicate areas with low Pb concentration surrounded by other low Pb levels (cold spots).

We see a clear spatial pattern in the distribution of BLL. Hot spot residential areas are seen on the western side of the Pb open cast mine. These areas include Kasanda, which was a mine residential area for the lowest skilled mine workers, and Makululu an informal settlement adjacent to Kasanda. On the south side close to the mine site, Mutwe Wansofu is another hot spot area. On the other hand, the northern side is mainly characterised by cold spots. We see Natuseko on the north, Mpima (northeast) and Katondo (east) all being cold spots.

4. Discussion

Using secondary data of households with geo-coordinates, we analysed the spatial autocorrelation and identified spatial clusters of blood lead level (BLL) from children in Kabwe. This is particularly useful, as a basis for setting up targeted health and environmental interventions in affected areas (Oyana and Margai, 2010; Requia et al., 2017). The study results confirm that distance and wind direction are major factors associated with the observed hotspots. Prevailing wind direction in Kabwe is predominantly east to west. Hotspot areas lie on the western side, and in close proximity to the mine. These areas include Kasanda and Mutwe Wansomfu. Conversely, we found distinct cold spots areas further away from the mine site (Mpima, Natuseko, Kangombe) and generally more on the eastern side.

Wind direction and distance are particularly important because much of the Pb pollution in Kabwe is due to the open pit mine and the tailing hills. Wind blows loose soil particles from the open pit area and the remaining tailing hills. The soil lead levels in these communities on the windward side are high (Bose-O'reilly et al., 2018b; Ettler et al., 2020). These areas generally lack greenness and remain dusty for most of the months. Children are exposed to Pb as they play within these communities and areas close to the mining site.

Social economic differentials are other important factors in relation to the Pb exposure. Historically, Kasanda, Chowa and Luangwa townships belonged to the mine during the operational years. On the western side, Kasanda was for the less skilled workers while Makululu is an informal settlement that sprung up from immigrants, mostly less skilled who were searching for jobs in the mines. These low-income neighborhoods are the least developed, characterised by unpaved roads, and houses made of mad bricks. The Pb exposure is high in these communities. On the other hand, Chowa and Luangwa residential areas are on the eastern side of the old mine and characterised by cold spots. These residential areas were for the skilled workers, expatriates, and rank higher in terms of the social economic class. A high proportion of children from the low-income communities are malnourished and parents are unable to meet the hospital bills. As such, the Pb pollution poses a greater burden on the poorer communities.

The age 0–3 years had the highest average, and maximum BLL. This could be attributed to the increased hand to mouth activity in this age group. Kabwe Soil Pb levels are high and negatively correlated with distance from the mine site. There is still a relative high risk of daily ingestion by young children across age group, as they play around the mine (Nakayama et al., 2011; Smolders et al., 2019).

5. Conclusion

The geospatial approach used in the present study has provided insight in spatial patterns of blood lead levels in the children of Kabwe. The study has established clustering effect of BLL and identified hotspot areas. Clearly, the BLL are dependent on distance and windward direction from the old mine site. Remedial and treatment interventions should consider this, and prioritize these affected communities. The relationship of the observed soil Pb levels and distributions of BLL is yet to be established. Further research is needed to develop a model using lead soil data to predict dangerous childhood Pb exposure. Soil Pb levels are relatively easy and cheap to collect. With a similar spatial analysis on soil Pb levels, hotspot areas could be detected, and the progress of interventions could be very well followed and documented. Spatial modelling would be ideal to establish for Kabwe the still unknown attributable fraction of each exposure pathway (inhalation, ingestion) and each source of exposure (soil, air, water, food).

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 acknowledge support from grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan awarded to M. Ishizuka (No. 16H0177906, 18K1984708), S.M.M. Nakayama (No. 16K16197, 17KK0009), and H. Nakata (No. 19K2047209) and the foundation of JSPS Bilateral Open Partnership Joint Research Projects (JPJSBP120209902). We also acknowledge financial support from The Nihon Seimei Foundation, The Japan Prize Foundation, and The Akiyama Life Science Foundation. This research was also supported by JST/JICA, SATREPS (Science and Technology Research Partnership for Sustainable Development; No. JPMJSA1501) and JST aXis (Accelerating Social Implementation for SDGs Achievement).

References

Akkus, C., Ozdenerol, E., 2014. Exploring childhood lead exposure through GIS: a review of the recent literature. Int. J. Environ. Res. Publ. Health 11, 6314–6334.

Bank, W., 2011. Implementation Completion and Results Report (IDA-37410 IDA-3741A IDA-H0260).

Bellinger, D., Sloman, J., Leviton, A., Rabinowitz, M., Needleman, H.L., Waternaux, C., 1991. Low-level lead exposure and children's cognitive function in the preschool years. Pediatrics 87, 219–227.

Bose-O'reilly, S., Yabe, J., Makumba, J., Schutzmeier, P., Ericson, B., Caravanos, J., 2018a. Lead intoxicated children in Kabwe, Zambia. Environ. Res. 165, 420–424.

Bose-O'reilly, S., Yabe, J., Makumba, J., Schutzmeier, P., Ericson, B., Caravanos, J., 2018b. Lead intoxicated children in Kabwe, Zambia. Environ. Res. 165, 420–424.

Burki, T., 2020. Report says 815 million children have high blood lead levels. Lancet 396, 370.

Ettler, V., Štěpánek, D., Mihaljevič, M., Drahota, P., Jedlicka, R., Kříbek, B., Vaněk, A., Penížek, V., Sracek, O., Nyambe, I., 2020. Slag dusts from Kabwe (Zambia): contaminant mineralogy and oral bioaccessibility. Chemosphere 127642.

Getis, A., Ord, J.K., 2010. The analysis of spatial association by use of distance statistics. In: Perspectives on Spatial Data Analysis. Springer.

Ikem, A., Campbell, M., Nyirakabibi, I., Garth, J.J.E.M., Assessment, 2008. Baseline Concentrations of Trace Elements in Residential Soils from Southeastern Missouri, vol. 140, pp. 69–81.

Landrigan, P.J., Fuller, R., Acosta, N.J.R., Adeyi, O., Arnold, R., Basu, N.N., Balde, A.B., Bertollini, R., Bose-O'reilly, S., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., Mcteer, M.A., Murray, C.J.L., Ndahimananjara, J.D., Perera, F., Potocnik, J., Preker, A.S., Ramesh, J., Rockstrom, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W. A., Van Schayck, O.C.P., Yadama, G.N., Yumkella, K., Zhong, M., 2018. The Lancet Commission on pollution and health. Lancet 391, 462–512.

Lanphear, B.P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D.C., Canfield, R.L., Dietrich, K.N., Bornschein, R., Greene, T., 2005. Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. Environ. Health Perspect. 113, 894–899.

Liu, J., Mccauley, L., Compher, C., Yan, C., Shen, X., Needleman, H., Pinto-Martin, J.A.J. E.H., 2011. Regular Breakfast and Blood Lead Levels Among Preschool Children, vol. 10, p. 28.

Lo, Y.-C., Dooyema, C.A., Neri, A., Durant, J., Jefferies, T., Medina-Marino, A., De Ravello, L., Thoroughman, D., Davis, L., Dankoli, R.S., 2012. Childhood lead poisoning associated with gold ore processing: a village-level investigation—Zamfara State, Nigeria, October–November 2010. Environ. Health Perspect. 120, 1450–1455.

Matte, T.D., Figueroa, J.P., Ostrowski, S., Burr, G., jackson-Hunt, L., Baker, E.L., 1991. Relationship between soil lead levels and blood lead levels among children living near a lead smelter in Jamaica. Chem. Speciat. Bioavailab. 3, 173–177.

Mielke, H.W., Gonzales, C.R., Powell, E.T., Laidlaw, M.A., Berry, K.J., Mielke, P.W., Egendorf, S.P., 2019. The concurrent decline of soil lead and children's blood lead in New Orleans. Proc. Natl. Acad. Sci. Unit. States Am. 116, 22058–22064.

Miranda, M.L., Dolinoy, D.C., Overstreet, M.A., 2002. Mapping for prevention: GIS models for directing childhood lead poisoning prevention programs. Environ. Health Perspect. 110, 947–953.

Nakayama, S.M., Ikenaka, Y., Hamada, K., Muzandu, K., Choongo, K., Teraoka, H., Mizuno, N., Ishizuka, M., 2011. Metal and metalloid contamination in roadside soil and wild rats around a Pb–Zn mine in Kabwe, Zambia. Environ. Pollut. 159, 175–181.

Osgood-Zimmerman, A., Millear, A.I., Stubbs, R.W., Shields, C., Pickering, B.V., Earl, L., Graetz, N., Kinyoki, D.K., Ray, S.E., Bhatt, S., 2018. Mapping child growth failure in Africa between 2000 and 2015. Nature 555, 41–47.

Oyana, T.J., Margai, F.M., 2010. Spatial patterns and health disparities in pediatric lead exposure in Chicago: Characteristics and profiles of high-risk neighborhoods. Prof. Geogr. 62, 46–65.

Plumlee, G.S., Durant, J.T., Morman, S.A., Neri, A., Wolf, R.E., Dooyema, C.A., Hageman, P.L., Lowers, H.A., Fernette, G.L., Meeker, G.P., 2013. Linking geological and health sciences to assess childhood lead poisoning from artisanal gold mining in Nigeria. Environ. Health Perspect. 121, 744–750.

Release, E.A., 2012. 10.1. Environmental Systems Research Institute, Redlands, CA.

Requia, W.J., Dalumpines, R., Adams, M.D., Arain, A., Ferguson, M., Koutrakis, P., 2017. Modeling spatial patterns of link-based PM2.5 emissions and subsequent human exposure in a large canadian metropolitan area. Atmos. Environ. 158, 172–180.

Rooney, J.P., Woods, N.F., Martin, M.D., Woods, J.S., 2018. Genetic polymorphisms of GRIN2A and GRIN2B modify the neurobehavioral effects of low-level lead exposure in children. Environ. Res. 165, 1–10.

Scheuplein, R., Charnley, G., Dourson, M., 2002. Differential sensitivity of children and adults to chemical toxicity: I. Biological basis. Regul. Toxicol. Pharmacol. 35, 429–447.

Smolders, E., Roels, L., Kuhangana, T.C., Coorevits, K., Vassilieva, E., Nemery, B., Lubaba Nkulu, C.L.B., 2019. Unprecedentedly high dust ingestion estimates for the general population in a mining district of DR Congo. Environ. Sci. Technol. 53, 7851–7858.

Stark, A.D., Quah, R.F., Meigs, J.W., Delouise, E.R., 1982. The relationship of environmental lead to blood-lead levels in children. Environ. Res. 27, 372–383.

Toyomaki, H., Yabe, J., Nakayama, S.M., Yohannes, Y.B., Muzandu, K., Liazambi, A., Ikenaka, Y., Kuritani, T., Nakagawa, M., Ishizuka, M., 2020. Factors associated with lead (Pb) exposure on dogs around a Pb mining area, Kabwe, Zambia. Chemosphere 125884.

Wang, Q., Zhao, H., Chen, J., Gu, K., Zhang, Y., Zhu, Y., Zhou, Y., Ye, L., 2009. Adverse health effects of lead exposure on children and exploration to internal lead indicator. Sci. Total Environ. 407, 5986–5992.

World Health Organization, 2010. Childhood Lead Poisoning.

Yabe, J., Nakayama, S.M., Ikenaka, Y., yohannes, Y.B., Bortey-Sam, N., Oroszlany, B., Muzandu, K., Choongo, K., Kabalo, A.N., Ntapisha, J., Mweene, A., Umemura, T., Ishizuka, M., 2015. Lead poisoning in children from townships in the vicinity of a lead-zinc mine in Kabwe, Zambia. Chemosphere 119, 941–947.

Yabe, J., Nakayama, S.M., Nakata, H., Toyomaki, H., Yohannes, Y.B., Muzandu, K., Kataba, A., Zyambo, G., Hiwatari, M., Narita, D., 2020. Current trends of blood lead levels, distribution patterns and exposure variations among household members in Kabwe, Zambia. Chemosphere 243, 125412.

Zhang, C., Luo, L., Xu, W., Ledwith, V., 2008. Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. Sci. Total Environ. 398, 212–221.

Zhang, T., Lin, G., 2006. A supplemental indicator of high-value or low-value spatial clustering. Geogr. Anal. 38, 209–225.

Annex: Paper IV

Biomonitoring of Arsenic, Cadmium and Lead in two Artisanal and Small-scale Gold Mining Areas in Zimbabwe

Environmental Science and Pollution Research (2022) 29:4762–4768
https://doi.org/10.1007/s11356-021-15940-w

RESEARCH ARTICLE

Biomonitoring of arsenic, cadmium and lead in two artisanal and small-scale gold mining areas in Zimbabwe

Stefan Rakete[1] · Given Moonga[1,2,3] · Anna-Maria Wahl[1] · Viola Mambrey[1] · Dennis Shoko[4] · Dingani Moyo[5,6] · Shamiso Muteti-Fana[7] · Myriam Tobollik[8] · Nadine Steckling-Muschack[1] · Stephan Bose-O'Reilly[1,9,10]

Received: 26 March 2021 / Accepted: 9 August 2021 / Published online: 19 August 2021

Abstract

People living and working in artisanal and small-scale gold mining (ASGM) areas are frequently exposed to elemental mercury (Hg), which is used for gold extraction. However, additional exposure to other toxic metals such as arsenic (As), cadmium (Cd) and lead (Pb) may result from mining-related activities and could be ingested via dust, water or food. In these areas, only limited biomonitoring data is available for toxic metals other than Hg. In particular, data about the exposure to As, Cd and Pb is unavailable for the Zimbabwean population. Therefore, we conducted a cross-sectional study in two ASGM areas in Zimbabwe to evaluate the internal exposure to these metals. In total, urine and blood samples from 207 people that identified themselves as miners were collected and analysed for As and Cd in urine as well as Pb in blood by GF-AAS. Median levels (interquartile ranges in μg/l) of As and Pb were 9.7 μg/l (4.0, 18.5) and 19.7 μg/l (12.5, 34.5), respectively. The 25th percentile and the median for Cd were below the limit of detection (0.5 μg/l); the 75th percentile was at 0.9 μg/l. The results were compared to reference values found for the general population in the USA and Germany, and a significant number of participants exceeded these values (As, 33 %; Cd, 27 %; Pb, 32 %), indicating a relevant exposure to toxic metals. Although not representative for the Zimbabwean population, our results demonstrate that the exposure to toxic metals is relevant for the public health in Zimbabwe and requires further investigation.

Keywords Biomonitoring · Toxic metals · Arsenic · Cadmium · Lead · Artisanal and small-scale gold mining · Zimbabwe

Responsible Editor: Lotfi Aleya

✉ Stefan Rakete
stefan.rakete@med.uni-muenchen.de

1 Institute and Clinic for Occupational-, Social- and Environmental Medicine, University Hospital, LMU Munich, Ziemssenstr. 1, D-80336 Munich, Germany

2 Center for International Health, LMU University Hospital Munich, Munich, Germany

3 Department of Epidemiology and Biostatistics, University of Zambia, Lusaka, Zambia

4 Tailjet Consultancy Services, Harare, Zimbabwe

5 School of Public Health, Faculty of Health Sciences, Occupational Health Division, University of the Witwatersrand, Johannesburg, Republic of South Africa

6 Faculty of Medicine & Faculty of Social Sciences, Midlands State University, Gweru, Zimbabwe

7 Department of Community Medicine, UZ College of Health Sciences, Harare, Zimbabwe

8 Section Environmental Medicine and Health Effects Assessment German Environment Agency, Berlin, Germany

9 Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT (Private University for Health Sciences, Medical Informatics and Technology), Hall in Tirol, Austria

10 University Children's Hospital Regensburg (KUNO-Clinics), Clinic St. Hedwig, University of Regensburg, Regensburg, Germany

Introduction

Artisanal and small-scale gold mining (ASGM) is predominantly an informal, poverty-driven, poorly resourced, comparatively inefficient and transient sector. Nonetheless, it provides an income to people in many developing countries rich in gold resources, and it is estimated that about 15 million miners are involved globally (Seccatore et al. 2014). ASGM in Zimbabwe has been growing very rapidly over the last 40 years. This growth was even faster over the last 2 decades, from an estimated 300,000 miners in 2000 to over 1.5 million today (Mkodzongi and Spiegel 2020, Mudzwiti et al. 2015). This has been due to a variety of factors including a persistently shrinking economy across multiple industries along with recurrent droughts during the last 2 to 3 decades. Consequently, there has been a growth in terms of tonnages mined and processed as well as gold produced. However, this resulted in the expansion of mining sites, increased used of process chemicals such as mercury and the exposure of new rock surfaces. This has led to the worsening of adverse environmental and public health impacts due to the release of toxic metals, mainly due the use of mercury (Hg) in the mining process (Billaud et al. 2004; Mudzwiti et al. 2015).

More than 90% of gold deposits in Zimbabwe are associated with the largely mafic/basic and cratonic greenstone belts, while the rest is located within the Limpopo Mobile Belt to the south and the Lomagundi meta-sediments in the north. The genesis of most gold deposits in Zimbabwe (>75%) involved the generation of superheated (>2000 °C) aqueous hydrothermal sulphide complexes at depth, in the upper mantle to lower crust. Gold is mainly transported in solution as aqueous complexes of hydrogen sulphides/bi-sulphides and chlorides as well as metal complexes (Foster 1985). Various elements, and especially heavy metal elements, have an affinity for dissolved hydrogen sulphides and chlorides in hydrothermal ore-forming solutions. This includes, but is not limited to, As, silver (Ag), Cd, Hg, Pb, selenium (Se), antimony (Sb), tellurium (Te), thallium (Tl) and zinc (Zn). These metals and many others are commonly associated with gold ores and are routinely used as pathfinders in gold exploration geochemistry (Saunders et al. 2014). Especially As, Cd and Pb are relatively toxic compared to other metals and were categorized within the top ten chemicals of public health concern by the World Health Organization (International Programme on Chemical Safety (IPCS) 2010). Mining activities such as excavation, crushing and milling, which are used in ASGM, may result in the increased liberation of these toxic metals. While the precious gold is collected at the end of the mining process, the other metals may end up in the tailings dumps at mining locations and thus represent an exposure hazard for people living and working in these mining areas.

Although, biomonitoring of As (Adu-Poku et al. 2019; Basu et al. 2011; Nyanza et al. 2019; Obiri et al. 2016), Cd (Basu et al. 2011; Obiri et al. 2016) and Pb (Gottesfeld et al. 2019; Nyanza et al. 2019; Obiri et al. 2016) has been conducted in several ASGM areas, exposure data is still relatively limited, particularly in Zimbabwe. Therefore, the purpose of this study was the analysis of As and Cd in urine as well as Pb in blood samples collected during a cross-sectional study involving people that identified themselves as miners in two ASGM areas in Zimbabwe. Data on Hg levels and health-related quality of life from this study were published elsewhere (Butscher et al. 2020; Mambrey et al. 2020; Wahl et al. 2021).

Materials and methods

Study design

This cross-sectional study was conducted within 2 weeks in March 2019 at two hospitals in the gold mining towns of Kadoma and Shurugwi Districts (Zimbabwe), respectively. Inclusion criteria for participation were a minimum age of 18 years. All females and males that identified themselves as miners and worked for at least a month were included. There were no specific exclusion criteria other than age. Participants were recruited using snowball sampling where participants recruit further participants among their colleagues. This sampling technique was used due to a widespread and hard-to-reach target population. Each participant signed an informed consent form and material transfer agreement, prior to the data and sample collection. All documents were available in the three main languages English, Shona and Ndebele spoken in Zimbabwe. Altogether, 207 participants consented to participate (131 from Kadoma and 76 from Shurugwi). Participation in the study was voluntary. Participants were asked to fill out questionnaires concerning general information on demographics. To ensure a confidential analysis, the samples and data were pseudonymized. More information about the study was previously published (Mambrey et al. 2020).

Urine and blood collection

All sample containers were labelled with the participant's code for future allocation. Spot urine samples were collected using disposable urine collection cups. For transport and analysis, aliquots of the urine samples were transferred into a Urine Monovette (Sarstedt®). To prevent bacterial growth and degradation, the samples were acidified with nitric acid to a pH of approximately 2, which was tested with a pH strip. Trained health professionals took venous blood samples into 7 ml lithium-heparin-coated tubes for trace metal analyses (Sarstedt®) from all participants. All samples were continuously stored at 4 °C. Once located in the laboratory, samples were stored at –18 °C until analysis.

Analysis of trace metals in urine and blood

All samples were at least analysed in duplicate after thawing on a roll mixer. For quality control (QC), certified reference materials (ClinChek®, Recipe, Munich, Germany) for whole blood and urine were analysed daily, and the sample analysis was only continued if the QC results were within the given specifications. As, Cd and Pb were analysed by graphite furnace atomic absorption spectroscopy (GF AAS, AAnalyst 600, Perkin Elmer, Rodgau, Germany) with the furnace programmes according to the recommendations of the manufacturer. The individual specifications for each element can be found below. The quantitation of all elements based on the standard addition method and the limit of detection (LOD) was calculated from the blank signal.

Analysis of total As in urine Total As in urine was analysed at a detection wavelength of 193.7 nm. Urine samples were diluted sixfold with 0.01% Triton-X in 0.13% nitric acid. Twenty microlitres of this dilution were automatically pipetted into the graphite tube of the GF-AAS. Five micrograms of Pd (as Pd $(NO_3)_2$) and 3 μg $Mg(NO_3)_2$ were added as matrix modifiers. For standard addition, 10 and 20 pg As were directly added to the sample in the graphite tube, respectively. The LOD was at 0.5 μg/l.

Analysis of Cd in urine Cd in urine was analysed at a detection wavelength of 228.8 nm. Urine samples were diluted fourfold with 0.01% Triton-X in 0.13% nitric acid. Twenty microlitres of this dilution were automatically pipetted into the graphite tube of the GF-AAS. Fifty micrograms of $NH_4H_2PO_4$ and 3 μg $Mg(NO_3)_2$ were added as matrix modifiers. For standard addition, 0.5 and 1 pg Cd were directly added to the sample in the graphite tube, respectively. The LOD was at 0.5 μg/l.

Analysis of Pb in blood Pb in blood was analysed at a detection wavelength of 193.7 nm. Blood samples were diluted tenfold with 0.05% Triton-X. Twenty microlitres of this dilution were automatically pipetted into the graphite tube of the GF-AAS. Fifty micrograms of NH_4H2PO_4 and 3 μg $Mg(NO_3)_2$ were added as matrix modifiers. For standard addition, 5 and 10 pg Pb were directly added to the sample in the graphite tube, respectively. The LOD was at 1.0 μg/l.

Analysis of creatinine in urine

Creatinine in urine samples was determined for creatinine-corrected levels of toxic metals in urine. Creatinine-corrected urine values were considered in order to account for the influence of the effect of urine dilution on the exposure indicator. Urine samples were sent to the central laboratory of University Hospital of LMU and analysed with Cobas C702 using the Jaffé method. Creatinine-corrected values from urine samples with creatinine levels < 0.3 g/l and > 3.0 g/l were excluded from statistical analysis.

Statistical analysis

All data analyses were performed with SPSS (version 26, IBM). One participant was excluded from statistical analysis, as the urine sample apparently contained blood. For samples below the LOD, the result was set to ½ LOD for further statistical analysis. Descriptive analysis included the geometric mean, minimum, maximum, median, 25th percentile, 75th percentile and 95th percentile. Differences in toxic metal concentrations between groups (gender, living area and self-reported fish consumption) were tested using the Mann-Whitney *U* test. Continuous variables (toxic metals, age, area years and mining years) were correlated using the Spearman correlation. Results for Hg in blood in urine and blood from the same participants used for correlation analyses were previously published elsewhere (Mambrey et al. 2020).

Results

Demographic information on the study population can be found in Table 1. One sample had to be excluded due to obvious sample contamination. Consequently, the levels of As and Cd in urine as well as the levels of Pb in blood were analysed for 206 participants, and the results are given in Table 2. Although urine samples with creatinine levels < 0.3 or > 3.0 g/l should preferably not be used for biomonitoring, we decided to only refrain from creatinine correction, as sampling could not be repeated (Cocker et al. 2011). Therefore, thirteen urine samples with urinary creatinine levels outside of this range were excluded from creatinine correction. The participants' levels of As in urine and Pb in blood follow a log-normal distribution (Figure S1), which was expected for this population. For Cd in urine, a log-normal distribution is anticipated, too. However, the majority of samples were below the LOD. The results were further stratified by gender (82 % male), living area (63% Kadoma) and self-reported fish consumption (80 % at least once a week), and differences were tested for significance (Table S1). For gender, urinary

Table 1 Demographic information on the study population

Age	*N*	207	
	Median (min.–max.)	38 (18–77)	
		N	**(%)**
Gender	Males	169	(81.6)
	Females	38	(18.4)
Living area	Kadoma	131	(63.3)
	Shurugwi	76	(36.7)
Fish consumption	< once a week	42	(20.3)
	> once a week	165	(79.7)

Table 2 Descriptive analysis of biomonitoring results for urinary levels of As and Cd and blood levels of Pb. Creatinine correction was not applied for urine samples with creatinine levels below 0.3 or above 3.0 g/l. Results were compared to international reference and threshold values for As and Cd in urine and Pb in blood, and the percentage of exceedances in this study is given in the brackets. Reference values represent the actual internal exposure of a representative population. Threshold values were derived from toxicological data. *NHANES* National Health and Nutrition Examination Survey, USA, *UBA* German Environment Agency, *CDC* Centers for Disease Control and Prevention, USA, *NIOSH* National Institute for Occupational Safety and Health, USA, *HBM-II* human biological monitoring alert level (Centers of Disease Preventions and Control (CDC) 2019, Schulz et al. 2011, UBA - German Environment Agency 2019)

	Creatinine in urine	As in urine		Cd in urine		Pb in blood
	g/l	μg/l	μg/g crea.	μg/l	μg/g crea.	μg/l
N	206	206	193	206	193	206
LOD	0.1	0.5		0.5		1
< LOD	0	12		120		0
GM	1,3	7.2	5.6	0.6	0.4	21.9
Minimum	0.1	< LOD		< LOD		6.6
25th percentile	1.0	3.7	2.7	< LOD		12.5
Median	1.4	9.7	6.5	< LOD		19.7
75th percentile	2.0	17.1	13.2	0.9	0.7	34.1
95th percentile	2.8	47.1	33.6	3.6	1.8	76.4
Maximum	4.6	460.5	250.3	11.4	4.9	275.7
		As in urine [μg/l]		Cd in urine [μg/l]		Pb in blood [μg/l]
Reference values (exceedances in %)						
NHANES		49.9 (3.6)		1.1 (22)		28.9 (32)
UBA		15.0 (33)		0.8 (27)		30 (♀, 24 %) 40 (♂, 20 %)
Threshold values (exceedances in %)						
NIOSH (CDC)		n.a.		n.a		50 (11)
UBA (HBM-II)		n.a.		4.0 (4)		n.a.

LOD limit of detection, < *LOD* number of results below limit of detection, *GM* geometric mean, *n.a.* not available

As levels were significantly higher in women. In contrast, blood Pb levels were significantly higher in men. For the area of living, significantly higher levels of As in urine were found for Kadoma if corrected for urinary creatinine. For self-reported fish consumption, no significant differences were found for any toxic metal.

For further evaluation, the results were compared to available international reference and threshold values from Germany and the USA (Centers of Disease Preventions and Control (CDC) 2019, Schulz et al. 2011, UBA - German Environment Agency 2019). The reference and threshold values for As, Cd and Pb as well as the percentage of participants that exceeded these values are given in Table 2. The 95th percentiles of Cd and Pb found in this study were higher than the reference values of the NHANES and UBA data. For As in urine, the 95th percentile found in this study was higher than the value proposed by UBA but comparable to NHANES.

The results for the correlation between toxic metals (including Hg), age, area years and mining years are given in Table S2. In general, correlation of the biomonitoring results with age, area years and mining years was found to be very low. However, Hg levels in urine and blood showed a weak positive correlation with mining years. In contrast, Cd in urine showed a weak negative correlation with age. If the levels of toxic metals were correlated to each other, a strong positive relationship was found for non-corrected and corrected levels of As, Cd and Hg in urine, respectively. Furthermore, a strong correlation was found for Hg in urine and Hg in blood. Whereas As, Cd and Pb mainly showed no to very weak correlation among each other, Hg levels in urine and blood showed a relatively moderate positive correlation with As, Cd and Pb.

Discussion

As explained in the introduction, toxic metals may be associated with gold-containing ores in ASGM areas. Therefore, it seems plausible that elevated As, Cd and Pb levels in the participants are related to the mining activities in Kadoma and Shurugwi. Below, we discuss the results for the individual metals and compare them to previously published studies in

Table 3 Comparison of the study results (all values are given in µg/l) with other studies in mining areas. All values are given as 25th percentile (P25), median and 75th percentile (P75) unless marked otherwise (#)

Parameter [µg/l]	Values measured in this Study	Values measured in studies in current and former mining areas (country)	Reference
	P25–median–P75	P25–median–P75	
As in urine	4.0–10.0–18.5	4.9–9.4–15.1 (Tanzania)[1]	Nyanza et al. (2019)
		73.2–100.2–135.3 (Ghana)	Basu et al. (2011)
		11.1–16.5–19.4 (Mexico)[2]	Moreno et al. (2010)
		0.5–1.17–1.93 (Spain)[2]	Molina-Villalba et al. (2015)
		0.06 (Guatemala)	Basu et al. (2010)
Cd in urine	< LOD–< LOD–0.9	0.25–0.36–0.6 (Ghana)	Basu et al. (2011)
		0.13–0.29–0.46 (Spain)[2]	Molina-Villalba et al. (2015)
		0.11 (Guatemala)	Basu et al. (2010)
Pb in blood	12.5–19.9–34.5	16.9–25.4–33.7 (Tanzania)[1]	Nyanza et al. (2019)
		64–94–113 (Mexico)[2]	Moreno et al. (2010)
		26.7 (Guatemala)	Basu et al. (2010)
		13.0 (Zambia)	Yabe et al. (2020)
		21.5 (Nigeria)	Gottesfeld et al. (2019)
		28.0 (Ghana)#	Obiri et al. (2016)

[1] Pregnant women

[2] Children

[#] Mean

LOD limit of detection

mining areas (Table 3). Although some of the studies found comparable results, the exposure to toxic metals likely depends on multiple factors such as the study population, the local concentrations of metals in the ore, the diet and many others. Generally, living close to mining areas seems to have a significant effect on the body burden (Basu et al. 2010, Molina-Villalba et al. 2015, Nemery and Banza Lubaba Nkulu 2018, Obiri et al. 2016). One possible explanation is that contaminated mining tailings contribute to the exposure to toxic metals by causing elevated concentrations in water, food and airborne dust (Moreno et al. 2010; van Straaten 2000).

Arsenic In contrast to the other metals, the reference values for As in urine from Germany (UBA) and the USA (NHANES) differ considerably from one another (15.0 vs 49.9 µg/l). This is likely due to different exposures to As, e.g. by fish consumption and other sources such as drinking water. However, reference values represent the actual exposure in the general population and cannot be used for toxicological evaluation. Consequently, the number of participants in this study that were above reference values for As heavily depends on which value will be used. It seems that the exposure to As is more comparable to the US population. However, we did not analyse the As species for differentiate inorganic (As(III), As(V)) and organic (e.g. arsenobetaine) As species. A major source of organic As is fish. However, self-reported fish consumption had no effect on As levels in urine. Zimbabwe has no access to open sea. Consequently, the consumed fish is commonly freshwater fish which usually contains relatively low amounts of As. The elevated levels of As in women and participants from Kadoma may be explained by different dietary patterns or a generally elevated exposure to As due to mining activities. Our results were comparable to what has been found in an ASGM area in Tanzania, but also in non-active mining areas in Mexico (Moreno et al. 2010; Nyanza et al. 2019). However, Basu et al. found a tenfold higher median As level in a Ghanaian ASGM study (Basu et al. 2011). In contrast, studies from Spain and Guatemala found relatively low levels of As in the urine (Basu et al. 2010; Molina-Villalba et al. 2015).

Cadmium Most of the samples were below the LOD, indicating a generally low exposure. Still, a considerable number of participants were above the references and threshold values. Cd levels are generally elevated in smokers. Unfortunately, we do not have the data for smoking in our study. In Zimbabwe's mining areas, far more men smoke compared to women (Billaud et al. 2004). However, this was not reflected in our results, where women were more frequently above international reference and threshold values for Cd (data not shown). For Cd in urine, the studies from Ghana, Guatemala and Spain showed very similar results (Basu et al. 2010; Basu et al. 2011; Molina-Villalba et al. 2015). Therefore, Cd exposure in ASGM and other mining areas seems to be relatively low compared to other toxic metals.

Lead About a third of the participants were above international reference values. The exposure to lead could be via contaminated drinking water, although the uptake of Pb in the gastrointestinal tract of adults is considered relatively low and airborne dust due to mining activities. Besides this, Pb exposure could be due to emissions of leaded gasoline, which was used longer in Africa than in the US or Europe (Todd and Hazel 2010). Therefore, the recent exposure to these emissions might still be reflected by elevated blood Pb levels. The results of this study are comparable to what has been found in other studies (Basu et al. 2010; Gottesfeld et al. 2019; Obiri et al. 2016; Yabe et al. 2020). However, in a study from Mexico, blood Pb levels were significantly higher (Moreno et al. 2010).

Limitations

Unfortunately, we did not have access to information about smoking, diet other than fish and levels of toxic metals in water, food, soil and air, thus limiting a more thorough exposure assessment. Additionally, we were not able to speciate As due to technical and financial limitations, further limiting exposure assessment and comparison with other studies. For As and Cd, the sensitivity of the analytical method was too low for some samples. In fact, Cd could not be detected in more than half of the samples, clearly limiting Cd exposure assessment. Nevertheless, the LOD was below the used reference and threshold levels. However, a relatively high number of participants were above international reference values (22 % > NHANES, 27 % > UBA). This may be explained by the fact that the Cd levels of 51 samples were relatively close to the LOD. Consequently, these results may be subject to some uncertainty regarding absolute quantitation. Furthermore, the inclusion of a control group from a non-mining area in Zimbabwe may have provided information if the exposure to As, Cd and Pb is actually caused by mining activities. However, this study was designed as a cross-sectional study as many studies had demonstrated that mining activities are associated with increased exposure to toxic metals.

Conclusions

This is, to the best of our knowledge, the first study that analysed the urinary levels of As and Cd as well as the blood levels of Pb in people identifying themselves as artisanal and small-scale gold miners in Zimbabwe. A high proportion of the participants had As, Cd and Pb levels above international reference levels. Therefore, the exposure to toxic metals in the two ASGM areas in Zimbabwe is relevant to public health and should be the subject of further investigation to clarify the influence of possible confounders, e.g. the diet. Furthermore, the exposure to toxic metals should be assessed in the general population of Zimbabwe to investigate if the results found in this study are related to ASGM activities.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11356-021-15940-w.

Acknowledgements The authors would like to thank the funding institution and our team in Zimbabwe and Germany. Especially we would like to thank Stefan Gröbmair, Evans Rudzvidzo, Dr. Godknows Madziva, Brian Chizaza, Josephine Mashamba, Jana Becker, James Rooney, and Friederike Butscher.

Author contribution Conceptualization, SR, VM, DS, DN, SM, NS and SB; data curation, GM, VM and SB; formal analysis, SR, GM and VM; funding acquisition, VM, DS, DM, NS and SB; investigation, SR, VM, DS, DM, SM, MT and SB; methodology, SR, GM, VM, DM, SM, NS and SB; project administration, VM, DS, DM, NS and SB; supervision, SR, NS and SB; validation, SR; visualization, GM, AW and SB; writing (original draft), SR, GM, AW, VM, DS, DM, SM, MT, and SB; writing (review and editing), SR, GM, AW, VM, DS, DM, SM, MT, NS and SB.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU IG II 2 45083 - 4/0, SB); SB and NS received funding from the European Union's Seventh Programme under grant agreement No 603946 (Health and Environment-Wide Associations Based on Large Population Surveys, HEALS).

Data availability The datasets used and/or analysed during the current study may be made available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate Approval for conducting the study was gained from the Ministry of Health in Zimbabwe and the respective local and regional authorities. The Ethics Committee of the Medical Research Council and the Zimbabwe Research Council (MRCZ/A/2367, September 26, 2018, and February 25, 2019) and the Ludwig Maximilian University of Munich (18-421, October 15, 2018) approved the study protocols and gave their permissions. In agreement with the Helsinki Declaration of Ethics Code for experiments with human subjects, the study was performed. Each participant signed an informed consent form and material transfer agreement, prior to the data and sample collection. All documents were available in the three main languages: English, Shona and Ndebele spoken in Zimbabwe.

Consent for publication The participants signed informed consent regarding publishing their data.

Competing interests The authors declare no competing interests.

References

Adu-Poku B, Asiedu N, Akoto O, Ataki J (2019) Modelling the distribution of arsenic and mercury in urine using chemometric tools. Cogent Chem 5:1586064. https://doi.org/10.1080/23312009.2019.1586064

Basu N, Abare M, Buchanan S, Cryderman D, Nam D-H, Sirkin S, Schmitt S, Hu H (2010) A combined ecological and epidemiologic investigation of metal exposures amongst Indigenous peoples near the Marlin Mine in Western Guatemala. Sci Total Environ 409:70–77. https://doi.org/10.1016/j.scitotenv.2010.09.041

Basu N, Nam D-H, Kwansaa-Ansah E, Renne EP, Nriagu JO (2011) Multiple metals exposure in a small-scale artisanal gold mining community. Environ Res 111:463–467. https://doi.org/10.1016/j.envres.2011.02.006

Billaud P, Laperche V, Maury-Brachet R, Boudou A, Shoko D, Kahwai S, Freyssinet P (2004) Removal of barriers to the introduction of cleaner artisanal gold mining and extraction technologies in Kadoma, Zimbabwe - Final report, part A Environmental assessment. UNIDO Project EG/GLO/01/G34 No.03/089. BRGM Project Nr 53320-FR BRGM (Bureau de Recherches Géologiques et Minières), www.unites.uqam.ca/gmf/intranet/gmp, Orleans, France

Butscher F-M, Rakete S, Tobollik M, Mambrey V, Moyo D, Shoko D, Muteti-Fana S, Steckling-Muschack N, Bose-O'Reilly S (2020) Health-related quality of life (EQ-5D + C) among people living in artisanal and small-scale gold mining areas in Zimbabwe: a cross-sectional study. Health Qual Life Outcomes 18:284. https://doi.org/10.1186/s12955-020-01530-w

Centers of Disease Preventions and Control (CDC) 2019: Fourth National Report on Human Exposure to Environmental Chemicals - Updated Tables

Cocker J, Mason HJ, Warren ND, Cotton RJ (2011) Creatinine adjustment of biological monitoring results. Occup Med 61:349–353. https://doi.org/10.1093/occmed/kqr084

Foster RP (1985) Major controls of Archaean gold mineralization in Zimbabwe. Trans Geol Soc S Afr 88:109–133

Gottesfeld P, Meltzer G, Costello S, Greig J, Thurtle N, Bil K, Mwangombe BJ, Nota MM (2019) Declining blood lead levels among small-scale miners participating in a safer mining pilot programme in Nigeria. Occup Environ Med 76:849–853. https://doi.org/10.1136/oemed-2019-105830

International Programme on Chemical Safety (IPCS), (2010): Ten chemicals of major public health concern. https://www.who.int/ipcs/assessment/public_health/chemicals_phc/en/ Accessed July, 5th 2021

Mambrey V, Rakete S, Tobollik M, Shoko D, Moyo D, Schutzmeier P, Steckling-Muschack N, Muteti-Fana S, Bose-O'Reilly S (2020) Artisanal and small-scale gold mining: a cross-sectional assessment of occupational mercury exposure and exposure risk factors in Kadoma and Shurugwi, Zimbabwe. Environ Res 184:109379

Mkodzongi G, Spiegel SJ (2020) Mobility, temporary migration and changing livelihoods in Zimbabwe's artisanal mining sector. Extr Ind Soc 7:994–1001. https://doi.org/10.1016/j.exis.2020.05.001

Molina-Villalba I, Lacasaña M, Rodríguez-Barranco M, Hernández AF, Gonzalez-Alzaga B, Aguilar-Garduño C, Gil F (2015) Biomonitoring of arsenic, cadmium, lead, manganese and mercury in urine and hair of children living near mining and industrial areas. Chemosphere 124:83–91. https://doi.org/10.1016/j.chemosphere.2014.11.016

Moreno ME, Acosta-Saavedra LC, Meza-Figueroa D, Vera E, Cebrian ME, Ostrosky-Wegman P, Calderon-Aranda ES (2010) Biomonitoring of metal in children living in a mine tailings zone in Southern Mexico: a pilot study. Int J Hyg Environ Health 213:252–258. https://doi.org/10.1016/j.ijheh.2010.03.005

Mudzwiti P, Mukwakwami N, Mungoni M, I M (2015) A golden opportunity: scoping study of artisanal and small scale mining in Zimbabwe. Harare, Zimbabwe

Nemery B, Banza Lubaba Nkulu C (2018) Assessing exposure to metals using biomonitoring: achievements and challenges experienced through surveys in low- and middle-income countries. Toxicol Lett 298:13–18. https://doi.org/10.1016/j.toxlet.2018.06.004

Nyanza EC, Bernier FP, Manyama M, Hatfield J, Martin JW, Dewey D (2019) Maternal exposure to arsenic and mercury in small-scale gold mining areas of Northern Tanzania. Environ Res 173:432–442. https://doi.org/10.1016/j.envint.2019.105450

Obiri S, Yeboah PO, Osae S, Adu-Kumi S (2016) Levels of arsenic, mercury, cadmium, copper, lead, zinc and manganese in serum and whole blood of resident adults from mining and non-mining communities in Ghana. Environ Sci Pollut Res 23:16589–16597. https://doi.org/10.1007/s11356-016-6537-0

Saunders J, Hofstra A, Goldfarb R, Reed MH (2014) Geochemistry of hydrothermal gold deposits. In Heinrich D. Holland, Karl K (Eds.) Treatise on Geochemistry: Second Edition 13, 383–424. https://doi.org/10.1016/B978-0-08-095975-7.01117-7

Schulz C, Wilhelm M, Heudorf U, Kolossa-Gehring M (2011) Update of the reference and HBM values derived by the German Human Biomonitoring Commission. Int J Hyg Environ Health 215:26–35. https://doi.org/10.1016/j.ijheh.2011.06.007

Seccatore J, Veiga M, Origliasso C, Marin T, De Tomi G (2014) An estimation of the artisanal small-scale production of gold in the world. Sci Total Environ 496:662–667. https://doi.org/10.1016/j.scitotenv.2014.05.003

Todd DT, Hazel (2010) Outcome and Influence Evaluation of the UNEP Partnership for Clean Fuels and Vehicles (PCFV) United Nations Environment Programme (UNEP). https://wedocs.unep.org/bitstream/handle/20.500.11822/274/Outcome_and_Influence_Evaluation_of_the_UNEP_Based_Partnership_for_Clean_Fuels_and_Vehicles_(PCFV).pdf?sequence=1&isAllowed=y

UBA - German Environment Agency (2019) Reference values (RV95) for antimony, arsenic and metals (Pb, Cd, Ni, Hg, Pt, Tl, U) in urine or in the blood. https://www.umweltbundesamt.de/sites/default/files/medien_/4031/dokumente/tab._referenzwerte_-_metalle_30._september_2019_aktualisiert.pdf, Accessed: July, 5th 2020

van Straaten P (2000) Mercury contamination associated with small-scale gold mining in Tanzania and Zimbabwe. Sci Total Environ 259:105–113. https://doi.org/10.1016/S0048-9697(00)00553-2

Wahl A-M, Bose-O'Reilly S, Mambrey V, Rooney JPK, Shoko D, Moyo D, Muteti-Fana S, Steckling-Muschack N, Rakete S (2021) Analysis of the mercury distribution in blood as a potential tool for exposure assessment - results from two artisanal and small-scale gold mining areas in Zimbabwe. Biol Trace Elem Res. https://doi.org/10.1007/s12011-021-02714-1

Yabe J, Nakayama SMM, Nakata H, Toyomaki H, Yohannes YB, Muzandu K, Kataba A, Zyambo G, Hiwatari M, Narita D, Yamada D, Hangoma P, Munyinda NS, Mufune T, Ikenaka Y, Choongo K, Ishizuka M (2020) Current trends of blood lead levels, distribution patterns and exposure variations among household members in Kabwe, Zambia. Chemosphere 243:125412. https://doi.org/10.1016/j.chemosphere.2019.125412

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

CHAPTER V

CONCLUSION

5.1 Conclusion

The study established three objectives, each of which was thoroughly addressed. The spatial distribution of stunting among children under the age of five in Zambia was demonstrated, and the environmental and socio-demographic factors associated with child malnutrition were identified. To address the observed disproportionate levels of stunting, deliberate nutritional programs should be implemented.

Childhood lead exposure in Kabwe requires immediate attention, and interventions should prioritize the identified hotspots. A one-health approach that incorporates multiple exposure pathways should be encouraged. This will reduce the amount of lead exposure from multiple sources, including dust, water, and crops.

The spatial methodology used in this study can be applied in similar states or to answer similar questions. The approach is useful in resource-constrained settings to facilitate targeted intervention.

5.2 Acknowledgements.

I owe my heartfelt gratitude to the extraordinary people who inspired me to embark on this journey and persevere with it to a successful completion of my PhD studies.

I am grateful to Prof. Dr. med. Dennis Nowak, my habilitated supervisor, for his encouragement and mentoring. I cannot thank him enough for creating an environment conducive to expand my intellectual imagination, for his support in sourcing funding and for guiding to complete this work.

I am thankful to my direct supervisor, PD. Dr. med. Stephan Bose-O' Reilly, for encouraging me to begin my PhD journey and for being there for me at every step of the way. I would like to express my heartfelt gratitude to him for being my life coach, academic and moral guide, and for guiding me through various challenges, I encountered along the way. I am forever indebted to him for his patience, motivation and mentorship.

Thanks to PD. Dr. Ursula Berger for her invaluable advice and inspiration and for providing me with many opportunities along the way. I thank her most sincerely for her guidance in statistical analysis of this work.

Dr. John Yabe, my local supervisor, deserves special recognition for his encouragement and belief in me. More importantly, he was willing to share project data with us in order to complete this work.

I would like to express my gratitude to Dr. Johannes Seiler, who became my statistical analysis colleague. I am most grateful for his mentorship and amiable attitude during the course of this work. I gained a brother in him in the process of this journey.

Sincerely thanks to the Centre for International Health (CIH-LMU) for admitting me into the PhD program and for their support in all spheres of my school life. Dr. Gunther Froschl, Dr. Arlett Heiber, Ms. Lisa Hoffäller, and Dr. Sarah Scholze deserve special mention for their exceptional assist in ensuring a comfortable stay during the program. I cannot imagine a better team to support a PhD program.

Special thanks to my classmates for creating a friendly academic and social environment, and a memorial PhD journey. I am also grateful to my alma mater, the University of Zambia, for providing a foundation for my academic career.

Finally, I am most grateful to my family for their unwavering support and belief in my potential, which has encouraged and strengthened me throughout the process.