

# SCIENTIFIC REPORTS

OPEN

## The risk of selenium deficiency in Malawi is large and varies over multiple spatial scales

Felix P. Phiri<sup>1,2</sup>, E. Louise Ander<sup>3</sup>, Elizabeth H. Bailey<sup>1</sup>, Benson Chilima<sup>4</sup>, Allan D. C. Chilimba<sup>5</sup>, Jellita Gondwe<sup>4</sup>, Edward J. M. Joy<sup>6</sup>, Alexander A. Kalimira<sup>7</sup>, Diriba B. Kumssa<sup>1</sup>, R. Murray Lark<sup>1</sup>, John C. Phuka<sup>8</sup>, Andrew Salter<sup>1</sup>, Parminder S. Suchdev<sup>9</sup>, Michael J. Watts<sup>3</sup>, Scott D. Young<sup>1</sup> & Martin R. Broadley<sup>1</sup>

Selenium (Se) is an essential human micronutrient. Deficiency of Se decreases the activity of selenoproteins and can compromise immune and thyroid function and cognitive development, and increase risks from non-communicable diseases. The prevalence of Se deficiency is unknown in many countries, especially in sub-Saharan Africa (SSA). Here we report that the risk of Se deficiency in Malawi is large among a nationally representative population of 2,761 people. For example, 62.5% and 29.6% of women of reproductive age (WRA,  $n = 802$ ) had plasma Se concentrations below the thresholds for the optimal activity of the selenoproteins glutathione peroxidase 3 (GPx3;  $< 86.9 \text{ ng mL}^{-1}$ ) and iodothyronine deiodinase (IDI;  $< 64.8 \text{ ng mL}^{-1}$ ), respectively. This is the first nationally representative evidence of widespread Se deficiency in SSA. Geostatistical modelling shows that Se deficiency risks are influenced by soil type, and also by proximity to Lake Malawi where more fish is likely to be consumed. Selenium deficiency should be quantified more widely in existing national micronutrient surveillance programmes in SSA given the marginal additional cost this would incur.

More than 800 million people have insufficient dietary energy intake globally, with  $> 2$  billion people at risk of one or more micronutrient deficiencies (MNDs), or *Hidden Hunger*<sup>1</sup>. Global disease burdens attributed to MNDs are based on iron (Fe), zinc (Zn), iodine (I) and vitamin A, which are considered to be the MNDs of greatest public health concern by the World Health Organization (WHO)<sup>1</sup>. However, recent estimates of dietary mineral supplies, derived from Food Balance Sheets (FBS) combined with food composition tables, indicate that other MNDs, such as calcium (Ca) and selenium (Se) may also have considerable public health significance at a regional level<sup>2-6</sup>.

Selenium has many roles in human health. Deficiency of Se, which is typically due to inadequate dietary Se intake, decreases the expression and activity of several essential selenoproteins and this has been associated with reduced immune function, thyroid function and cognitive development, and increased risks from non-communicable diseases<sup>7-9</sup>. A high prevalence of inadequate dietary Se has been inferred from FBS for many countries in sub-Saharan Africa (SSA)<sup>3</sup>. In Malawi, an estimated 70% of the population consumed inadequate Se based on nationally-representative household-level food consumption data<sup>6</sup>. Small-scale cross-sectional studies of food intakes<sup>10,11</sup>, composite diets<sup>10</sup>, and Se status in blood plasma among the Malawi population are consistent with predictions that the majority of the Malawian population are Se deficient<sup>10,12</sup>.

<sup>1</sup>School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, LE12 5RD, UK. <sup>2</sup>Department of Nutrition, HIV and AIDS, Ministry of Health, Lilongwe, Malawi. <sup>3</sup>Inorganic Geochemistry, Centre for Environmental Geochemistry, British Geological Survey, Nottingham, NG12 5GG, UK. <sup>4</sup>Community Health Sciences Unit, Ministry of Health, Private Bag 65, Lilongwe, Malawi. <sup>5</sup>The Department of Agricultural Research Services, P.O. Box 30799, Lilongwe 3, Malawi. <sup>6</sup>Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, UK. <sup>7</sup>Department of Human Nutrition and Health, Faculty of Food and Human Sciences, Bunda Campus, Lilongwe University of Agriculture and Natural Resources, P.O. Box 219, Lilongwe, Malawi. <sup>8</sup>School of Public Health and Family Medicine, College of Medicine, University of Malawi, Private Bag 360, Chichiri, Blantyre 3, Malawi. <sup>9</sup>Department of Pediatrics and Hubert Department of Global Health, Emory University, Atlanta, Georgia, 30322, USA. Correspondence and requests for materials should be addressed to F.P.P. (email: [felixphiri8@gmail.com](mailto:felixphiri8@gmail.com)) or M.R.B. (email: [martin.broadley@nottingham.ac.uk](mailto:martin.broadley@nottingham.ac.uk))

Characteristic	n	Median	Mean	SD	Min	Max
		ng mL <sup>-1</sup>				
<b>PSC</b>						
Females	494	57.5	60.9	25.1	9.9	207
Males	496	57.7	63.4	27.3	12.5	198
<b>SAC</b>						
Females	385	69.2	74.6	37.4	11.0	284
Males	369	67.1	72.5	31.7	17.6	281
<b>WRA</b>	802	78.4	83.7	38.4	11.1	374
<b>Men</b>	215	81.9	83.3	27.9	17.5	198
<b>All</b>	2761	68.2	73.2	33.9	9.9	374

**Table 1.** Plasma selenium (Se) concentration among demographic groups in Malawi. n = sample number in cohort; SD = standard deviation; Min = minimum value in the cohort; Max = maximum value in the cohort; PSC = preschool children; SAC = school-aged children; WRA = women of reproductive age (aged 15–49 years); Men (aged 20–54 years).

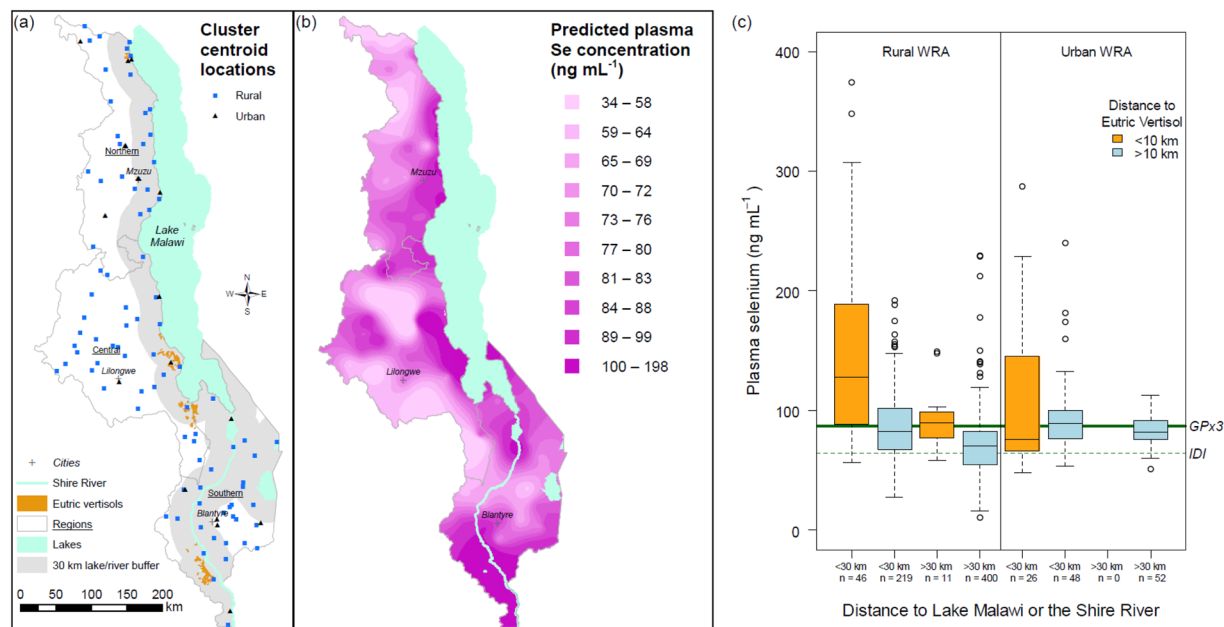
Characteristic	n	Prevalence (%) below the plasma Se concentration threshold			
		GPx3	IDI	KD	
<b>WRA</b>	Urban	126	52.3	12.7	0.0
	Rural	676	64.3	32.7	1.8
	All	802	62.5	29.6	1.5
<b>Men</b>	Urban	28	50.0	17.9	0.0
	Rural	187	58.3	23.0	1.1
	All	215	57.2	22.3	0.9

**Table 2.** Prevalence of selenium (Se) deficiency among women of reproductive age (WRA, aged 15–49 years) and men (aged 20–54 years) in Malawi, based on plasma Se concentration thresholds for optimal activity of glutathione peroxidase 3 (GPx3; <84.9 ng mL<sup>-1</sup>), iodothyronine deiodinase (IDI; <64.8 ng mL<sup>-1</sup>), and for Keshan Disease (KD; <30 ng mL<sup>-1</sup>) which is a cardiomyopathy linked to Se status reported in China. Thresholds from Thomson (ref.<sup>9</sup>).

We tested if Se deficiency is prevalent among the Malawian population, using blood plasma Se as an accepted population biomarker of Se status<sup>7–9</sup>. Plasma samples were collected during the previously planned Malawi Micronutrient Survey (MNS) and Malawi Demographic and Health Survey (MDHS) of 2015–16<sup>13</sup>. The MNS comprised a nationally representative sample of preschool children (PSC, aged 6–59 months), school-aged children (SAC, aged 5–14 years), women of reproductive age (WRA, aged 15–49 years), and men (aged 20–54 years). The scope of the MNS/MDHS, including details of the sampling design, ethics, informed consent, and individual confidentiality procedures are described in the Methods. The Se concentration of plasma was determined using inductively coupled plasma-mass spectrometry (ICP-MS). In total, data were obtained for 2,761 individuals from 1,233 households, located in 102 clusters, 84 and 18 of which were located in rural and urban areas, respectively. Data were analysed using a linear mixed model to quantify long- and short-range spatial variation in Se status<sup>14</sup>. Conditional distributions of Se status could be computed from this model at unsampled sites and, from these, both the expected status and the probabilities that plasma Se concentration of WRA was <84.9 ng mL<sup>-1</sup> and <64.8 ng mL<sup>-1</sup>. These threshold values were defined based on optimal activities of the selenoproteins glutathione peroxidase 3 (GPx3) and iodothyronine deiodinase (IDI), respectively, in adults<sup>9</sup>. These probabilities were presented as maps, using a verbal scale to visualise the probability that plasma Se falls below the thresholds of interest.

## Results and Discussion

**Prevalence of selenium deficiency.** Selenium deficiency is widespread among the Malawi population. Across all demographic groups, the mean and median plasma Se concentration was 73.2 ng mL<sup>-1</sup> and 68.2 ng mL<sup>-1</sup>, respectively (standard deviation, SD, 33.9 ng mL<sup>-1</sup>; range 9.9–374 ng mL<sup>-1</sup>; Table 1). Plasma Se concentration increased with age, ranging from a median of 57.7 ng mL<sup>-1</sup> ( $\pm 26.2$  ng mL<sup>-1</sup> SD) in PSC, to 81.9 ng mL<sup>-1</sup> ( $\pm 27.9$  ng mL<sup>-1</sup> SD) in adult men (Table 1). To estimate Se deficiency prevalence, we determined the proportion of WRA and men whose plasma Se concentration was less than thresholds previously established for the optimal activities of two essential selenoproteins, GPx3 and IDI, involved in oxidative stress response and thyroid function, respectively<sup>9</sup>. Plasma Se concentration was below the threshold for optimal GPx3 activity in 62.5% of WRA and 57.2% of men (Table 2). Plasma Se concentration was below the threshold for optimal IDI activity in 30% of WRA and 22% of men. Whilst thresholds for the optimal activity of GPx3 and IDI were used here to indicate Se deficiency, other selenoproteins, including Selenoprotein P, whose optimal activity is at blood plasma Se concentration >100 ng mL<sup>-1</sup><sup>17</sup>, could be used for this type of analysis and would indicate a much greater overall deficiency prevalence. Thus, further research is required to identify the most suitable threshold for a particular geographical and demographic context<sup>7</sup>. The subsequent spatial statistical analysis in this study focused on

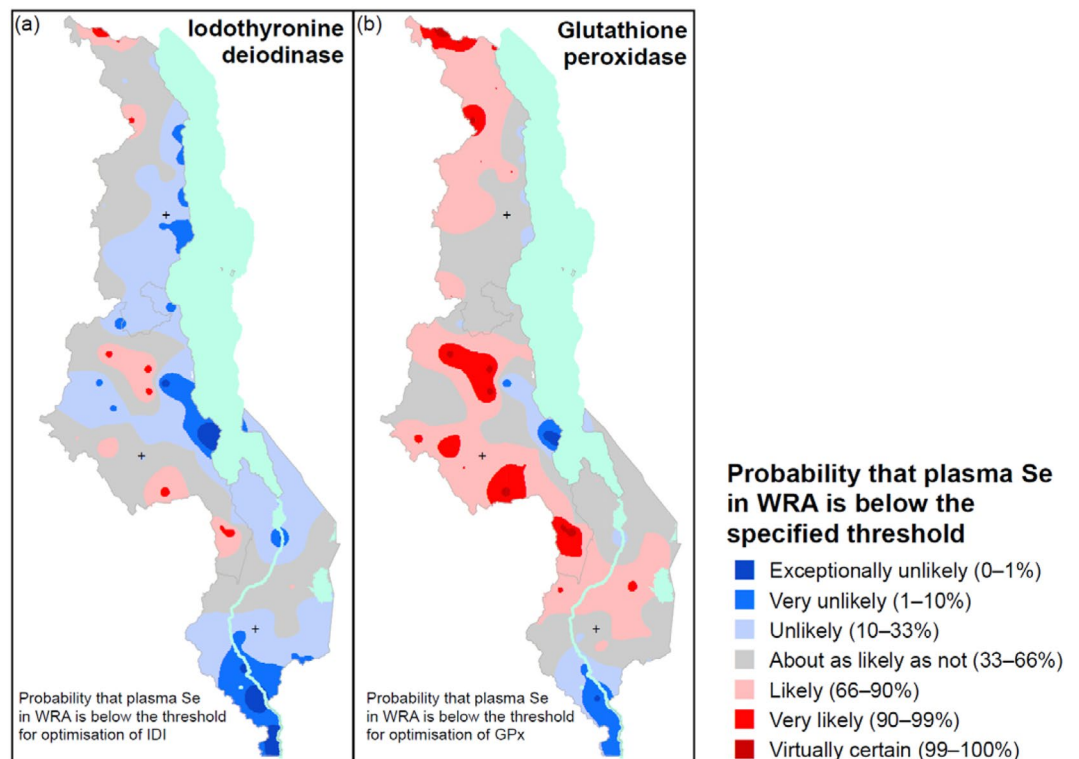


**Figure 1.** Selenium (Se) status of the Malawi population: (a) location of rural and urban Enumeration Areas from which study participants were recruited, (b) median unbiased predicted plasma Se concentration of women of reproductive age (WRA), (c) plasma Se concentration among WRA according to proximity to Eutric Vertisol soil classification, and to Lake Malawi or the Shire River; the threshold for the optimal activity of iodothyronine deiodinase (IDI) and glutathione peroxidase 3 (GPx3) are indicated by dashed and solid lines, respectively.

WRA, due to the larger sample size and accepted deficiency thresholds for adults. However, if we assume the same thresholds of deficiency among children as for adults, then 86% and 64% of PSC are at risk of suboptimal expression of GPx3 and IDI, respectively. Given the potential implications of Se deficiency for immune function, thyroid function, and cognitive development, there is an urgent need to establish thresholds for Se deficiency in children.

**Spatial determinants of Se status.** A much greater proportion of the variation in plasma Se concentration arose between clusters than within clusters, showing that Se status is under strong geospatial control (Fig. 1). The variation between cluster means shows spatial dependence up to a distance of 120 km. We can hypothesise that a spatial dependency in Se status among the Malawi population is due primarily to variation in maize grain Se concentration and subsequently dietary Se supplies. In Malawi, maize represented 48% of dietary energy supply in 2013<sup>15</sup> and contributed ~70% of energy consumption in the poorest quintile of households<sup>6</sup>. Food production and consumption is highly localised in Malawi, e.g. in the 2011 Third Integrated Household Survey (IHS3) it was reported that 57% of maize flour consumed by households came from subsistence production, while the remainder (38% from purchases and 5% from gifts) was likely to be sourced locally in rural locations<sup>6,16</sup>. The consumption of foods known to be good sources of Se (e.g. animal products including fish) is limited in Malawi although is greater in populations living closer to Lake Malawi, other lakes, and the Shire River<sup>6,8,16</sup>. Whilst strong links between the Se concentration in food crops, dietary Se intakes, and Se status have previously been reported in small cross-sectional studies in Malawi<sup>10</sup>, more research is needed to establish the relative contribution of different plant and animal sources to Se intake and status, including bioavailability studies, at a range of geographical scales.

Plasma Se concentrations were much greater in populations living in proximity to vertisol soil types in the south and east of Malawi, and in populations living closer to Lake Malawi and the Shire River (Fig. 1). For example, plasma Se concentrations among rural dwelling WRA living within 10 km of the nearest mapped vertisol, and within 30 km of the lake or river, were  $148 \pm 76 \text{ ng mL}^{-1}$  (mean  $\pm$  SD;  $n = 46$ ). Vertisol soils in Malawi are typically located close to Lake Malawi and the Shire River, where fish consumption are likely to be greater. However, the main effect driving Se intakes and status is likely to be soil factors. Notably, the greatest plasma Se concentrations among WRA were observed in Salima District, an area next to the lake where vertisol soils also occur. In Salima, the plasma Se concentration of some individual WRA exceeded  $300 \text{ ng mL}^{-1}$  (range  $104\text{--}374 \text{ ng mL}^{-1}$ ;  $n = 11$ ). These observations are consistent with a previous cross-sectional study that showed much greater plasma Se concentration of  $119 \pm 23 \text{ ng mL}^{-1}$  (mean  $\pm$  SD;  $n = 60$ ) among WRA living close to vertisol soils in Mikalango Extension Planning Area (EPA), compared to  $54 \pm 10 \text{ ng mL}^{-1}$  (mean  $\pm$  SD;  $n = 60$ ) for those living in Zombwe EPA<sup>10</sup>. Previously, the Se concentration of maize grain growing on vertisol soils in Malawi was shown to be 10–15-fold greater than in grain grown on acidic soils which dominate Malawi croplands<sup>2</sup>. In acidic tropical soils, Se is largely unavailable to plants probably due to its strong adsorption to iron and aluminium oxides as  $\text{Se}^{(\text{IV})}$ . In vertisol soils, which cover ~0.5% of the land area in Malawi, Se is more readily available to plants because  $\text{Se}^{(\text{IV})}$



**Figure 2.** Probability that measured plasma selenium (Se) concentration of women of reproductive age (WRA) would fall below the threshold for the optimal activity of (a) iodothyronine deiodinase (IDI), (b) glutathione peroxidase 3 (GPx3). The legend units are defined verbally, and by a range of probability values, expressed as percentages.

forms are adsorbed less strongly at higher soil pH. Furthermore,  $\text{Se}^{(\text{VI})}$  forms are more stable at higher soil pH under aerobic soil conditions and are taken up more efficiently than  $\text{Se}^{(\text{IV})}$  by plant roots<sup>17</sup>.

The output from this spatial analysis allows us to visualise the probability of Se deficiency at a national scale, using a verbal scale that shows the likelihood of an individual being below a threshold (Fig. 2). These maps and the underlying statistical model could be used to plan further investigations. For example, one might plan future national-scale sampling so as to reduce uncertainties in the predictions where these are large, and one might optimise the balance of sample effort between the inclusion of new clusters or the amount of replication at household or within-household level.

**Socioeconomic determinants of Se status.** The prevalence of Se deficiency was associated with household wealth. Plasma Se concentration was below the threshold for optimal GPx3 activity among 70%, 68%, 59%, 59% and 56% of WRA from households in the poorest to wealthiest quintiles, respectively. These trends are qualitatively consistent with estimates of inadequate dietary Se supplies among 92%, 86%, 81%, 72% and 51% of households from the poorest to wealthiest quintiles, respectively, based on household-level food consumption and spatially disaggregated food composition data<sup>6</sup>. The prevalence of Se deficiency among WRA in urban areas was lower than in rural areas, i.e. 34% and 62%, respectively, were below the threshold for optimal GPx3 activity. Households in urban areas typically had greater expenditure on food<sup>16</sup>, while urban markets are likely to provide more diverse foods of wider geographical origin.

**Implications for micronutrient surveillance programmes and policies.** There is evidence that Se status is under spatial control in other countries in SSA, albeit based on much smaller cross-sectional geographical studies, including Cote d'Ivoire, the Democratic Republic of the Congo, and Ethiopia<sup>18–23</sup>. Taken together, these data indicate that prevalence of Se deficiency is likely to be widespread in SSA<sup>3</sup>, and that the variability at subnational scales needs to be explored and communicated at finer resolution, using 'GeoNutrition' type spatial frameworks as described in this study. Such frameworks also need to include the collection of disaggregated data on the Se (and other mineral) supply characteristics of soils and food crops. These data are not yet available in the same form as the plasma Se concentration data reported in this study<sup>2,5,6</sup>. Conducting a MNS is expensive, time-consuming and logistically-challenging. Given that population Se deficiency can be inferred readily from plasma Se concentration data, there is a strong case for adapting MNS in SSA to quantify Se deficiency given that no additional burden is placed on participants. There is also a clear need to establish thresholds for Se deficiency in children, especially given the evidence that children are less likely than adults to have access to foods that are good sources of Se within households<sup>24</sup>.



The present study provides the strongest evidence to-date of widespread Se deficiency in a national population in SSA. The prevalence of Se deficiency is likely to remain high in Malawi without a public health intervention. There is scope to alleviate Se deficiency through increasing dietary diversity. Thus, whilst Se deficiency is prevalent in all socio-economic groups, the risk of Se deficiency is lower amongst wealthier groups and increased consumption of animal products is one dietary diversification option. However, unless animals have access to Se supplements in feeds, the concentration of Se in livestock products, as well as livestock productivity and health, will be constrained by low concentrations of Se in forages due to soil factors. Leaves of the tree crop *Moringa oleifera* and *M. stenopetala* (Moringaceae) have been shown to be a good potential dietary source of Se across a range of soil types from surveys in Malawi<sup>5</sup>, Kenya<sup>25</sup> and Ethiopia<sup>25</sup>. Agronomic biofortification (agro-fortification) of maize using Se-enriched fertilisers has been proposed as a cost-effective strategy to alleviate Se deficiency in Malawi<sup>2,6,26,27</sup>; such an approach has been adopted successfully in Finland since 1984<sup>28</sup>. The ethics, efficacy, effectiveness, and safety of interventions to promote dietary diversification and/or micronutrient fertilisers would, of course, need to be assessed carefully by national policy makers in the context of other public health spending.

## Methods

**Aim of study.** The aim of the study was to determine Se status of individuals at a nationally representative scale in Malawi, using blood plasma Se concentration as a population-level biomarker. The study was conducted using plasma samples and data collected during the Malawi Micronutrient Survey (MNS) and Demographic and Health Survey (MDHS) of 2015–16<sup>13,29</sup>. The wider purpose of the MNS was to provide a range of data for planning national nutritional policies. The study comprised a nationally representative sample of preschool children (PSC, aged 6–59 months), women of reproductive age (WRA, aged 15–49), school-aged children (SAC, aged 5–14), and men (aged 20–54). Details of the scope of the full MNS, including details of the design, informed consent, and individual confidentiality procedures are described elsewhere<sup>13,29,30</sup>. The study was conducted under informed consent. Ethical approval for blood sampling and testing for mineral micronutrient biomarkers as part of the MNS was granted by the Malawi National Health Sciences Research Committee (NHSRC), reference number NHSRC 15/5/1436. Because the MNS had not originally planned to analyse blood samples for Se, a material transfer agreement (MTA) linked to NHSRC 15/5/1436 was considered and approved by the NHSRC (11<sup>th</sup> January 2016), to include Se as an additional mineral micronutrient biomarker. There were no modifications to the MNS sample collection procedures. All methods were performed in accordance with the relevant guidelines and regulations.

**Sampling.** The MNS represented a subsample of the wider MDHS. The MDHS was designed as a cross-sectional study, with a two-stage cluster sampling design to enable indicators to be obtained nationally and for each of the 28 Districts of Malawi, for both urban and rural populations. The sampling frame was based on the Malawi Population and Housing Census of 2008. For the first stage of the MDHS, a total of 850 Enumeration Areas (EAs), referred to as clusters, were selected from a total of >9,000 EAs in Malawi using a probability proportional to population size. For the second stage of the MDHS, 30 and 33 households were selected at random from each urban and rural cluster, respectively, using an updated list of the eligible 27,531 households within the 850 MDHS clusters.

For the MNS, 105 target clusters were selected randomly from the MDHS clusters so that 8 urban and 27 rural clusters were selected from each region (North, Central and South); a total target of 24 and 81 urban and rural clusters, respectively. From each of the urban and rural MNS clusters, 10 and 11 households, respectively, were excluded from the MNS because these households had been selected for HIV testing within the MDHS. Therefore, the MNS comprised 20 and 22 households from each urban and rural cluster, respectively, giving a target sample size of 480 and 1782 urban and rural households, respectively.

Within each of the households selected for the MNS, eligible participants (defined as having spent the night prior to that survey in that household) were invited to participate as follows: PSC from all households; WRA from 9 households selected randomly from the PSC households; SAC from 6 of the 9 WRA households; men from 4 of the 6 SAC households. Based on estimates of the numbers of each demographic group per household, the likely sample size within each demographic group was estimated to be approximately 1459 PSC, 750 WRA, 762 SAC, and 252 men. These sample sizes were estimated to provide adequate statistical power to detect changes in a range of nutritional indicators for PSC from previous MNS, including vitamin A deficiency, anaemia, zinc and iron deficiency, and stunting, based on 90% household and individual participation<sup>13,29,30</sup>. In each of the target population groups, individuals deemed ineligible to participate in the MNS included those<sup>1</sup>: chronically ill<sup>2</sup>; outside the target age groups<sup>3</sup>; diagnosed with chronic disease or long term illness requiring treatment<sup>4</sup>; with acute or chronic infection which may alter serum zinc status<sup>5</sup>; on regular prescribed medicine<sup>6</sup>; using oral contraceptive pills or hormone replacement therapy<sup>7</sup>; regularly using dietary supplements including iodine, zinc or selenium supplements<sup>8</sup>; children with a physical disability that would affect either height or weight measurement. These exclusion criteria were based on the requirements of the MNS and not for the purposes of studying plasma Se status. Survey participants identified as having severe anaemia (hemoglobin <7 g dL<sup>-1</sup>), malaria (based on positive results from a rapid diagnostic kit), moderate or severe acute malnutrition (using age-appropriate mid- upper arm circumference (MUAC) cutoffs), haematuria (based on urine dipstick in all groups except WRA), or severe illness as identified by a team nurse, were referred to a local hospital for evaluation.

**Data and sample collection.** The MDHS and MNS teams visited clusters sequentially; the MDHS collected data electronically using tablet computers; the MNS used a paper questionnaire. The tablets were programmed so that MDHS teams were notified of eligible households and demographic groups for the MNS. For each eligible household, an MDHS team placed a household label with a unique barcode on the questionnaire and entered the barcode number into the tablet. This barcode number enabled data collected independently by the MDHS and the

MNS teams to be linked. Once MDHS data collection was completed within a cluster, the questionnaire booklets were given to the MNS team. The MNS team would typically proceed to a cluster within 2–3 days of the MDHS data collection.

The MNS team visited each household that had indicated to the MDHS team that they were willing to participate in the MNS. The head of each MNS household was asked for written consent so potential participants from that household could be issued with uniquely barcoded bracelets, linked to the questionnaire booklet, and accompanied by MNS staff to the field laboratory. Individual interviews and specimen collection was then conducted, subject to further individual consent from the individual or parent/guardian<sup>13,29,30</sup>. If eligible households, or individuals within households, were unavailable, up to two follow-up visits were made to each household; households were not replaced.

Whole blood venous samples (~5 mL) were collected into trace-element free purple top vacutainers through venipuncture, using a butterfly needle, by experienced phlebotomists. All individuals (PSC, SAC, WRA and men) were seated prior to blood collection, skin was wiped with 70% alcohol at the area of antecubital vein, and the arm was restricted with a tourniquet for maximum of 1 minute. During blood collection, the tourniquet was loosened before removing the needle from the arm. Once blood was collected, the vacutainer was inverted 10 times to allow blood to mix with the EDTA in the tube to prevent clotting. Blood samples in vacutainers were centrifuged at 3,500 rpm for 10 minutes in mobile centrifuges, and plasma was aliquoted into sterile labeled cryovials; ~3 mL was dispensed into transparent, flat-based 8 mL polypropylene tubes (57 × 16.5 mm; Sarstedt, Nümbrecht, Germany), capped using transparent high-density polyethylene, and racked into groups of 50. The samples were stored in portable freezers at –20 °C, transferred to District-level hospital storage, and then to the central laboratory at Community Health Sciences Unit (CHSU), Lilongwe, for storage in boxes of 100. The plasma samples were transferred on dry ice for Se analyses, to the University of Nottingham, United Kingdom.

Strict quality control measures were followed during blood collection and transportation using well-trained nurses/clinicians and laboratory technicians under the supervision of the Centers for Disease Control & Disease Prevention (CDC) technical team. At every stage of sample collection, laboratory technicians ensured that all containers for the collection and transportation of blood samples were clean and free from any contaminants and interfering substances. The health professionals involved in blood collection followed strict sterile procedures and medical ethics. Throughout the sample collection, transportation, analysis, and disposal, all blood samples were treated as potentially infectious for HIV and other infectious agents, and disposable polyethylene powder free gloves were used for blood handling.

**Analysis of blood plasma for Se.** The concentrations of Se in plasma samples were measured using inductively coupled plasma-mass spectrometry (ICP-MS; Thermo Fisher Scientific iCAPQ, Thermo Fisher Scientific, Bremen, Germany). Samples were introduced from an autosampler incorporating an ASXpress™ rapid uptake module (Cetac ASX-520, Teledyne Technologies Inc., Omaha, NE, USA) through a perfluoroalkoxy (PFA) Microflow PFA-ST nebuliser (Thermo Fisher Scientific, Bremen, Germany). External Se calibration standards at 0, 5, and 10 ng mL<sup>-1</sup> (SCP Science ‘PlasmaCal’; certified concentration 1003 ± 6 µg mL<sup>-1</sup>) were prepared in 0.5% HNO<sub>3</sub> (trace analysis grade (TAG), Fisher Scientific UK Ltd, Loughborough, UK). Samples were diluted 1 in 31 in 0.5% HNO<sub>3</sub> (TAG). The ICP-MS was operated in ‘collision-reaction cell mode’, with kinetic energy discrimination, using H<sub>2</sub> as the cell gas to maximise the sensitivity for Se determination. Correction for drift was undertaken using three internal standards (<sup>72</sup>Ge, <sup>103</sup>Rh, <sup>185</sup>Re in 0.5% HNO<sub>3</sub> and 4% methanol) introduced to the sample flow via a Y-piece; methanol enhances Se ionisation in the plasma. The limit of detection (LOD) for Se was 0.0011 ng mL<sup>-1</sup>, representing 3 × the standard deviation (SD) of 10 operational blanks. Accuracy was verified by the use of appropriate reference materials (Seronorm-1 and Seronorm-2; Nycomed Pharma AS, Billingstad, Norway). Mean recoveries were 96.5% (SD = 2.3%) and 96.9% (SD = 2.9%) for Seronorm-1 (Lot 1309438; accredited value 86.5 ng/mL) and Seronorm-2 (Lot 1309416; accredited value 137 ng mL<sup>-1</sup>), respectively (n = 47). All plasma samples with Se concentrations >200 ng mL<sup>-1</sup> (n = 27) were re-run for confirmation; the mean difference was 3%, excluding one outlier. Samples were also analysed for gadolinium, used in MRI scanning, to confirm that the doubly charged species (<sup>156</sup>Gd<sup>++</sup>; m/z = 78), presented a negligible interference to the analyte <sup>78</sup>Se.

**Data analyses.** In this paper, we report the analysis of plasma Se concentrations for PSC, SAC, WRA, and men. We focused our spatial analysis on WRA because thresholds linked to optimal activities of selenoproteins glutathione peroxidase 3 (GPx3) and iodothyronine deiodinase (IDI), respectively are only applicable to adults<sup>9</sup>, and due to the greater numbers of WRA in the sample. Summary statistics for WRA plasma Se concentration data are shown in Supplementary Information. Given the marked skewness of these data; these were transformed to natural logarithms, and the summary statistics on this scale are also shown in Supplementary Information. These log<sub>e</sub>-transformed data were used for the remaining analysis.

As described above, the sampling units are individuals from within households selected at random within clusters. This sampling structure was accounted for in the analysis with a linear mixed model (LMM) in which the only fixed effect was a constant mean. The variation in plasma Se concentration between individuals within households (which includes analytical error) was treated as a random effect of mean zero with homogeneous variance over all households. The between-household within-cluster variation was also treated as an independent random effect of mean zero and with homogeneous variance over all clusters. The between-cluster variation was modelled as a spatially correlated second-order Gaussian random variable with a Matérn correlation function<sup>31</sup>. Exploratory analysis of the mean values for the clusters suggested that the assumption of stationary variation about a fixed mean was reasonable, with no pronounced trends. It also suggested that it was reasonable to treat the spatial dependence of the cluster means as isotropic (i.e. dependent only on the distance between any two clusters and not on the direction of the vector that joins them). The variance components, and the parameters of the Matérn correlation function were estimated by residual maximum likelihood (REML). Following the

recommendation of Diggle and Ribeiro<sup>32</sup>, the smoothness parameter of the Matérn correlation was estimated by the profile likelihood.

The between-individuals (within-households) variance (log variable) was 0.028 and the between-household (within cluster) variance was 0.018. The between-cluster variance was much larger (0.15). The smoothness parameter estimated by the maximum profile likelihood (equivalent to the familiar exponential correlation function) was 0.5 and the spatial parameter was 39.4 km which means that the variation between cluster means shows spatial dependence up to a distance of about 120 km. To validate this model, each observation was predicted from the rest by cross-validation. The assumption of normally distributed cross-validation errors was plausible (Supplementary Information), and the standardised squared prediction error of the cross-validation predictions had a median value of 0.33 which falls within the 95% confidence interval for this statistic under prediction with a valid covariance model and normal errors<sup>33</sup>.

To map plasma Se concentrations for WRA, ordinary kriging estimates of individual concentrations were computed at nodes of a 500 × 500 m square grid. At each node, the probability was computed that the true value of this variable took values below thresholds for optimal plasma Se <84.9 and <64.8 ng mL<sup>-1</sup>, the lower thresholds linked to optimal activities of the selenoproteins GPx3 and IDI, respectively<sup>9</sup>. This probability was obtained from the prediction distribution of log-transformed plasma Se concentration at each location on the grid. This is a conditional distribution, dependent on the fitted model and the neighbouring observations. The ordinary kriging (OK) prediction at each grid node was treated as the mean value of the prediction distribution and the OK variance as the prediction variance. The OK predictions of plasma Se concentration were back-transformed from the log scale by exponentiation which gives a median-unbiased central value of the prediction distribution on the original scale<sup>34</sup>. The probabilities of falling below the thresholds were mapped using interval values based on a verbal scale for communicating uncertain information, supplemented with numerical values in the legend, following Lark *et al.*<sup>35,36</sup>.

## Data Availability

The datasets generated during the current study are available from the corresponding author on reasonable request.

## References

- Wang, H. *et al.* Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1459–1544 (2016).
- Chilimba, A. D. C. *et al.* Maize grain and soil surveys reveal suboptimal dietary selenium intake is widespread in Malawi. *Sci. Rep.* **1**, 72 (2011).
- Joy, E. J. M. *et al.* Dietary mineral supplies in Africa. *Physiol. Plantarum* **151**, 208–229 (2014).
- Kumssa, D. B. *et al.* Dietary calcium and zinc deficiency risks are decreasing but remain prevalent. *Sci. Rep.* **5**, 10974 (2015).
- Joy, E. J. M. *et al.* Soil type influences crop mineral composition in Malawi. *Sci. Total Environ.* **505**, 587–595 (2015).
- Joy, E. J. M. *et al.* Dietary mineral supplies in Malawi: spatial and socioeconomic assessment. *BMC Nutr.* **1**, 42 (2015).
- Fairweather-Tait, S. J. *et al.* Selenium in human health and disease. *Antioxid. Redox Signal.* **14**, 1337–1383 (2011).
- Rayman, M. P. Selenium and human health. *Lancet*. **379**, 1256–1268 (2012).
- Thomson, C. D. Assessment of requirements for selenium and adequacy of selenium status: a review. *Eur. J. Clin. Nutr.* **58**, 391–402 (2004).
- Hurst, R. *et al.* Soil-type influences human selenium status and underlies widespread selenium deficiency risks in Malawi. *Sci Rep.* **3**, 1425 (2013).
- Eick, F., Maleta, K., Govasmark, E., Duttaroy, A. K. & Bjune, A. G. Food intake of selenium and sulphur amino acids in tuberculosis patients and healthy adults in Malawi. *Int. J. Tuberc. Lung Dis.* **13**, 1313–1315 (2009).
- Van Lettow, M., West, C. E., van der Meer, J. W. M., Wieringa, F. T. & Semba, R. D. Low plasma selenium concentrations, high plasma human immunodeficiency virus load and high interleukin-6 concentrations are risk factors associated with anemia in adults presenting with pulmonary tuberculosis in Zomba district, Malawi. *Eur. J. Clin. Nutr.* **59**, 526–532 (2005).
- National Statistical Office (NSO) & The DHS Program (ICF). *Malawi Demographic and Health Survey 2015–16* (NSO, Zomba, Malawi, and ICF, Rockville, Maryland, USA, 2017).
- Verbeke, G. Linear mixed models for longitudinal data. In *Linear Mixed Models in Practice. Lecture Notes in Statistics*, Vol. 126 63–153 (Springer, New York, USA, 1997).
- Food and Agriculture Organization of the United Nations. *National-Level Food Supply Data, Food Balance Sheets*. Available, <http://www.fao.org/faostat> [accessed October 2017].
- National Statistics Office (NSO) of the Republic of Malawi. *Malawi Third Integrated Household Survey (IHS3): Malawi and World Bank Living Standards and Measurements Surveys* (NSO, Zomba, Malawi, 2012).
- White, P. J. Selenium accumulation by plants. *Ann. Bot.* **117**, 217–235 (2015).
- Arnaud, J., Malvy, D., Richard, M.-J., Faure, H. & Chaventré, A. Selenium status in an iodine deficient population of the West Ivory Coast. *J. Physiol. Anthropol. Appl. Human Sci.* **20**, 81–84 (2001).
- Tiahou, G. *et al.* Lack of oxidative stress in a selenium deficient area in Ivory Coast. *Eur. J. Nutr.* **43**, 367–374 (2004).
- Goyens, P., Golstein, J., Nsombola, B., Vis, H. & Dumont, J. E. Selenium deficiency as a possible factor in the pathogenesis of myxoedematous endemic cretinism. *Acta Endocrinol. (Copenh.)* **114**, 497–502 (1987).
- Vanderpas, J. B. *et al.* Iodine and selenium deficiency associated with cretinism in northern Zaire. *Am. J. Clin. Nutr.* **52**, 1087–1093 (1990).
- Ngo, D. B. *et al.* Selenium status in pregnant women of a rural population (Zaire) in relationship to iodine deficiency. *Trop. Med. Int. Health* **2**, 572–581 (1997).
- Gashu, D. *et al.* Association of serum selenium with thyroxin in severely iodine-deficient young children from the Amhara region of Ethiopia. *Eur. J. Clin. Nutr.* **70**, 929–934 (2016).
- Benemariya, H., Robberecht, H. & Deelstra, H. Daily dietary intake of copper, zinc and selenium by different population groups in Burundi, Africa. *Sci. Total Environ.* **136**, 49–76 (1993).
- Kumssa, D. B. *et al.* Variation in the mineral element concentration of *Moringa oleifera* Lam. and *M. stenopetala* (Bak. f.) Cuf.: Role in human nutrition. *PLoS ONE* **12**, e0175503 (2017).
- Chilimba, A. D. C. *et al.* Agronomic biofortification of maize with selenium (Se) in Malawi. *Field Crops Res.* **125**, 118–128 (2012).
- Chilimba, A. D. C. *et al.* Assessing residual availability of selenium applied to maize crops. *Field Crops Res.* **134**, 11–18 (2012).
- Alfthan, G. *et al.* Effects of nationwide addition of selenium to fertilizers on foods, and animal and human health in Finland: From deficiency to optimal selenium status of the population. *J. Trace Elem. Med. Biol.* **31**, 142–147 (2015).

29. National Statistical Office (NSO) [Malawi], Community Health Services Unit (CHSU) of the Ministry of Health, Department of Nutrition, HIV and AIDS [Malawi], Centers for Disease Control & Prevention (CDC) [USA]. *Micronutrient Survey Key Indicators Report 2015–16*, <http://dhsprogram.com/pubs/pdf/FR319/FR319m.pdf> (NSO, Zomba, Malawi; CHSU Lilongwe, Malawi; CDC Atlanta, Georgia, USA; 2017).
30. McGann, P. T. *et al.* Prevalence of inherited blood disorders and associations with malaria and anemia in Malawian children. *Blood Adv.* **2**, 3035–3044 (2018).
31. Matérn, B. *Spatial Variation. 2<sup>nd</sup> Edition* (Springer, Berlin, Germany, 1986).
32. Diggle, P. J. & Ribeiro, P. J. *Model-Based Geostatistics* (Springer, New York, USA, 2007).
33. Lark, R. M. A comparison of some robust estimators of the variogram for use in soil survey. *Eur. J. Soil Sci.* **51**, 137–157 (2000).
34. Pawlowsky-Glahn, V. & Olea, R. A. *Geostatistical Analysis of Compositional Data* (Oxford University Press, New York, USA, 2004).
35. Lark, R. M. *et al.* Mapping trace element deficiency by cokriging from regional geochemical soil data: a case study on cobalt for grazing sheep in Ireland. *Geoderma* **226–227**, 64–78 (2014).
36. Lark, R. M., Ander, E. L. & Broadley, M. R. Combining two national-scale data sets to map soil properties, the case of available magnesium in England and Wales. *Eur. J. Soil Sci.* **70**, 361–377 (2019).

## Acknowledgements

Funding for this study was provided by the University of Nottingham (UoN), the British Geological Survey (BGS) and the Royal Society-UK Department for International Development (DFID) under project AQ140000, “*Strengthening African capacity in soil geochemistry for agriculture and health*”. We thank The Government of Malawi (Ministry of Health; Ministry of Agriculture, Irrigation and Water Development; National Statistics Office), and the Centers for Disease Control & Prevention (CDC), notably Katie Tripp, Anne Williams, and Carine Mapango, for enabling this work to be conducted, together with all of the participating field teams and volunteers.

## Author Contributions

F.P.P., E.L.A., E.H.B., B.C., A.D.C.C., E.J.M.J., A.A.K., J.C.P., A.S., M.J.W., S.D.Y. and M.R.B. designed the research. F.P.P. led sample collection, supported by B.C., J.G., P.S. and M.R.B., F.P.P., E.H.B. and S.D.Y. conducted sample analyses. F.P.P., E.L.A., E.J.M.J., D.B.K., R.M.L. and M.R.B. analysed the data. E.L.A. and R.M.L. produced maps and final data tables. F.P.P. and M.R.B. led manuscript writing, with contributions from all authors.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-43013-z>.

**Competing Interests:** The authors declare no competing interests.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019