

Association of maternal prenatal selenium concentration and preterm birth: a multicountry meta-analysis

Nagendra Monangi ^{1,2,3}, Huan Xu,^{2,4} Rasheda Khanam,⁵ Waqasuddin Khan ⁶, Saikat Deb,^{7,8} Jesmin Pervin,⁹ Joan T Price,¹⁰ INTERBIO-21st Study Consortium, Stephen H Kennedy,¹¹ Abdullah Al Mahmud,¹² Yuemei Fan,¹³ Thanh Q Le,¹⁴ Angharad Care,¹⁵ Julio A Landero,¹⁶ Gerald F Combs,¹⁷ Elizabeth Belling,⁴ Joanne Chappell,^{2,4} Fansheng Kong,⁴ Criag Lacher,¹⁸ Salahuddin Ahmed,¹⁹ Nabidul Haque Chowdhury,¹⁹ Sayedur Rahman,¹⁹ Furqan Kabir,⁶ Imran Nisar,⁶ Aneeta Hotwani,⁶ Usha Mehmood,⁶ Ambreen Nizar,⁶ Javairia Khalid,⁶ Usha Dhingra,⁷ Arup Dutta,⁷ Said Ali,⁸ Fahad Aftab,⁸ Mohammed Hamad Juma,⁸ Monjur Rahman,²⁰ Bellington Vwalika,²¹ Patrick Musonda,²² Tahmeed Ahmed,²³ Md Munirul Islam ²⁴, Ulla Ashorn,²⁵ Kenneth Maleta,²⁶ Mikko Hallman,²⁷ Laura Goodfellow,²⁸ Juhi K Gupta ²⁸, Ana Alfirevic,²⁸ Susan Murphy,²⁹ Larry Rand,³⁰ Kelli K Ryckman,³¹ Jeffrey C Murray,³² Rajiv Bahl,³³ James A Litch ³⁴, Courtney Baruch-Gravett,³⁴ Zarko Alfirevic,³⁵ Per Ashorn ^{36,37}, Abdullah Baqui,³⁸ Jane Hirst,³⁹ Cathrine Hoyo,⁴⁰ Fyezah Jehan,⁴¹ Laura L Jelliffe-Pawłowski,⁴² Anisur Rahman ⁹, Daniel E Roth,⁴³ Sunil Sazawal,^{7,8} Jeffrey Stringer ¹⁰, Ge Zhang,^{2,3,4} Louis Muglia^{2,3,4,44}

To cite: Monangi N, Xu H, Khanam R, *et al.* Association of maternal prenatal selenium concentration and preterm birth: a multicountry meta-analysis. *BMJ Global Health* 2021;**6**:e005856. doi:10.1136/bmjgh-2021-005856

Handling editor Sanni Yaya

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2021-005856>).

NM, HX, RK, WK, SD, JP, JTP, AAM, YF, TQL and AC contributed equally. ZA, PA, AB, JH, CH, FJ, LLJ-P, AR, DER, SS, JS, GZ and LM contributed equally.

NM, HX, RK, WK, SD, JP, JTP, AAM, YF, TQL and AC are joint first authors. ZA, PA, AB, JH, CH, FJ, LLJ-P, AR, DER, SS, JS, GZ and LM are joint last authors.

Received 29 March 2021
Accepted 4 August 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to Nagendra Monangi; nagendra.monangi@cchmc.org and Ge Zhang; Ge.Zhang@cchmc.org

ABSTRACT

Background Selenium (Se), an essential trace mineral, has been implicated in preterm birth (PTB). We aimed to determine the association of maternal Se concentrations during pregnancy with PTB risk and gestational duration in a large number of samples collected from diverse populations.

Methods Gestational duration data and maternal plasma or serum samples of 9946 singleton live births were obtained from 17 geographically diverse study cohorts. Maternal Se concentrations were determined by inductively coupled plasma mass spectrometry analysis. The associations between maternal Se with PTB and gestational duration were analysed using logistic and linear regressions. The results were then combined using fixed-effect and random-effect meta-analysis.

Findings In all study samples, the Se concentrations followed a normal distribution with a mean of 93.8 ng/mL (SD: 28.5 ng/mL) but varied substantially across different sites. The fixed-effect meta-analysis across the 17 cohorts showed that Se was significantly associated with PTB and gestational duration with effect size estimates of an OR=0.95 (95% CI: 0.9 to 1.00) for PTB and 0.66 days (95% CI: 0.38 to 0.94) longer gestation per 15 ng/mL increase in Se concentration. However, there was a substantial heterogeneity among study cohorts and the random-effect meta-analysis did not achieve statistical significance. The largest effect sizes were observed in UK (Liverpool) cohort, and most significant associations were observed in samples from Malawi.

Key questions

What is already known?

- Conflicting results have been reported on the association between maternal selenium (Se) levels and preterm birth (PTB) risk.
- Most previous studies were typically small or focused on a single geographic region with limited data in populations at high risk for PTB.

What are the new findings?

- Maternal prenatal plasma/serum Se concentrations varied substantially between different geographic regions. Clinically meaningful associations were observed between maternal Se concentration and PTB within specific cohorts; however, this finding was not generalisable across all the cohorts.
- The significant associations observed in specific study cohorts might be mediated or confounded by individual site-specific factors.

What do the new findings imply?

- Our results do not support a uniform association between maternal prenatal Se concentration and PTB risk.
- The significant associations observed in specific study cohorts might have potential implications for targeted Se supplementation in high-risk settings.

Interpretation While our study observed statistically significant associations between maternal Se concentration and PTB at some sites, this did not

generalise across the entire cohort. Whether population-specific factors explain the heterogeneity of our findings warrants further investigation. Further evidence is needed to understand the biologic pathways, clinical efficacy and safety, before changes to antenatal nutritional recommendations for Se supplementation are considered.

INTRODUCTION

Preterm birth (PTB), defined as delivery prior to 37 completed weeks of gestation, is the leading global cause of infant and under-5-year old childhood mortality.¹ Each year, an estimated 15 million babies are born preterm, of whom approximately 1 million die with complications of prematurity.² Most countries with reliable trend data show an increase in PTB rates over the past 20 years, with more than 60% of cases occurring in Africa and South Asia.²⁻⁴ Due to the immaturity of multiple organ systems, preterm infants are at increased risk of short-term and long-term health sequelae including cognitive disabilities, impaired motor skills, hearing loss, chronic immunologic/infectious morbidities^{5 6} and elevated risks of adulthood obesity, diabetes and hypertension.⁷⁻⁹

Despite the profound global health significance and recognition that the prevention of PTB would provide major improvements in child health, there has only been limited progress in preventing PTB. Recently, a two-stage genome-wide association study of over 50 000 women of European ancestry identified and replicated EEFSEC gene, encoding the selenocysteine tRNA (tRNA^{Seleno})-specific eukaryotic elongation factor, that was robustly associated with gestational duration.¹⁰ EEFSEC plays a critical role in incorporating selenium (Se) in the form of selenocysteine into selenoproteins such as glutathione

peroxidases, the iodothyronine 5'-deiodonases, selenoprotein P and thioredoxin reductases.¹¹ The implication of selenocysteine pathway suggests a potential benefit for further evaluation of the role of maternal Se status on PTB risk. The possible involvement of maternal Se concentration in PTB has also been suggested by previous epidemiological studies;¹²⁻¹⁶ however, the sample sizes of these studies were usually small, or focus on a single geographic area and the results are not always consistent between studies.^{17 18}

In this study, we aimed to examine the association of maternal Se concentrations during pregnancy with PTB risk and gestational duration. As dietary Se intake is highly related to its regional soil content,¹⁹ we leveraged the availability of archived biological samples from geographically diverse cohorts and tested the association between maternal Se concentrations and gestational duration in a large number of samples collected from these study cohorts with different social and ancestral background and varying degrees of Se exposures.

METHODS

Study design and participants

The International Consortium on Selenium, Genetics, and Preterm Birth is a Bill & Melinda Gates Foundation (BMGF) funded project to study the potential association between maternal Se concentration and PTB risk using existing samples and data from multiple birth studies. The consortium comprises 17 international pregnancy cohorts across a wide geographic distribution (figure 1) with Cincinnati Children's Hospital Medical Center (CCHMC) serving as the coordinating hub. Among the



Figure 1 Geographic location of study sites.

participating sites, Malawi (iLiNS-DYAD)²⁰ and Bangladesh (MDIG)²¹ cohorts were intervention trials and USA, CA (CPPOP) was a case-control study. All the other cohorts were designed to enrol women randomly at hospitals. Description and study characteristics of these participating study cohorts are provided in online supplemental text 1 and online supplemental table 1).

Samples and sampling data

Demographic, prenatal, delivery and fetal/newborn data (online supplemental table 2) as collected by the individual sites according to their local protocols were shared with the coordinating hub (CCHMC). The data collected from Bangladesh (GAPPS), Bangladesh (MDIG),²¹ Vietnam (PBB), USA (NEST; CPOP),^{22 23} and all AMANHI cohorts²⁴ were case/control (preterm/term) samples. The data collected from other sites including Malawi (iLiNS-DYAD),²⁰ Zambia (GAPPS) and the six INTERBIO sites²⁵ were random samples with a PTB rate ranging from 4.5% (INTERBIO, Kenya) to 20% (INTERBIO, Pakistan). Gestational age dating was assigned at the site level by ultrasound, last menstrual period (LMP) or both (online supplemental table 1). Preterm cases were defined as birth prior to 37 weeks of gestation and term controls as birth at 37 weeks or later. We excluded stillbirths and multigestational pregnancies.

Selenium measurement

Se status was assessed on the basis of the concentrations of Se in plasma or serum.²⁶ Plasma or serum samples obtained from participating cohorts were stored at -70°C or -80°C refrigerators before and after use at the CCHMC Biobank (online supplemental table 1). To mitigate potential batch effect, samples from each site were randomised prior to analysis in batches. Inductively coupled plasma mass spectrometry (ICP-MS) measurements of Se concentrations in serum or plasma were performed using Agilent 7700 ICP-MS (Agilent Technologies) at the laboratory of Clinical Chemistry and Biochemistry, University of Cincinnati as described in detail in the protocol (online supplemental text 2) except the samples from Bangladesh (MDIG) which were analysed at the Centers for Disease Control and Prevention (Atlanta, GA).

Statistical analysis

Phenotypic data from participating study sites were harmonised by applying a uniform data structure and consistent coding rules for phenotype variables (eg, gestational duration, maternal age, height and fetal sex). Maternal Se data generated by the laboratory were combined and merged with phenotype data. The distributions of gestational duration and Se measures for each site were visually inspected using histograms and violin plots. Outliers for gestational duration and Se measurements were detected based on fitting with appropriate probability distributions and removed from further association analysis.

To determine the covariates to be included in the association analysis, we first examined the correlation of PTB (and gestational duration) with other covariates as well as the correlation between Se concentration and other covariates in each site using Pearson correlation. Variables significantly correlated ($p < 0.05$) with either PTB or gestational duration or Se concentration were included as covariates. The DerSimonian-Laird (DSL) random-effect meta-analysis was used to combine the correlation coefficients obtained from each cohort. For each site, we estimated the association between maternal Se concentration and PTB (and gestational duration as a continuous variable) using logistic (for PTB) or linear (for gestational duration) regression analysis. Fixed-effect meta-analysis and random-effect meta-analysis were used to combine the results from different cohorts. Between-study heterogeneity was checked using Cochran's Q test. Some of the cohorts used case/control samples (online supplemental table 1) and regression analysis of gestational duration as a continuous variable without accounting for the non-random sampling could potentially introduce bias in effect size estimation. To address this problem, we conducted regression analysis weighted by inverse of sampling probability (IPW). Detailed description of this analysis can be found in online supplemental text. All analyses were done with Microsoft R Open 3.5.1.

Patient and public involvement

Patients or public were not involved in setting the research question or the outcome measures, nor were they involved in the design or conduct of the study. No participants were asked to advise on interpretation or writing up of the manuscript. For the study, individual study cohorts shared the archived biological samples from established biobanks and there is no direct patient or public involvement.

RESULTS

Gestational duration, PTB and their correlations with other covariates

Pregnancy phenotype and birth outcomes of 10 640 pregnancies were obtained from 17 study sites (online supplemental table 1). Among these, 9946 singleton livebirths had gestational duration measured in days (gday) and maternal plasma or serum samples (figure 2). The demographic characteristics of these mothers (eg, age and height) and the major birth outcomes (eg, gestational duration and birth weight) are summarised by the site (table 1). After removing three outliers, the gestational duration followed a Weibull distribution with a mean of 268 days and ranging from 147 to 312 days (distribution parameters: shape: 21.2, scale: 275.8) (online supplemental figure 1). The distributions of gestational days in term (gday ≥ 259 days) and preterm (gday < 259 days) deliveries from each site are shown in online supplemental figure 2.

We examined the correlation of PTB and gestational duration with other covariates (maternal age, height, fetal sex and gestational age at sampling) in each participant

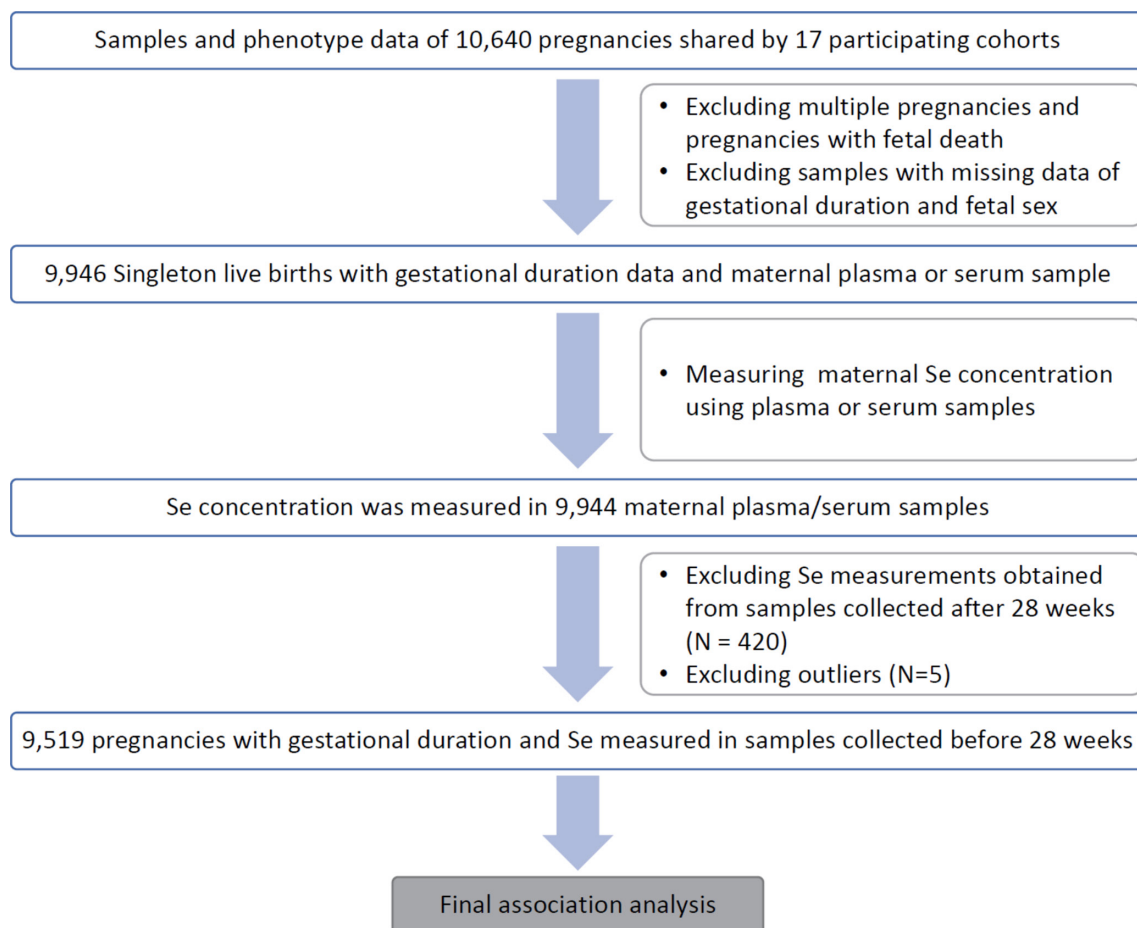


Figure 2 Flow chart of the study illustrating the total number of subjects, inclusion/exclusion criteria.

site (online supplemental figure 3). Meta-analysis using DSL method showed that PTB risk was significantly associated with maternal height and fetal sex. Similarly, gestational duration was also significantly associated with maternal height (shorter mothers had shorter gestational duration) and fetal sex (males had shorter gestational duration).

Maternal prenatal Se concentration and its correlations with other covariates

Se concentrations were successfully measured in 9944 mothers. After removing two outliers, the Se concentrations followed a normal distribution with some positive skewness (online supplemental figure 4) with a mean of 93.8 ng/mL and SD of 28.5 ng/mL. Se levels varied substantially across different sites (figure 3, online supplemental figure 5), and also across different experimental batches for each site (online supplemental figure 6). The highest average Se was observed in the Tanzania (AMANHI) cohort with a mean level of 131.4 ng/mL, and the lowest Se was observed in Zambia (GAPPS) with a mean concentration of 55.9 ng/mL. The largest variation was observed in Malawi (iLiNS-DYAD) (range: 26.1 to 228.7 ng/mL, SD=29.5) (online supplemental table 3).

We examined the correlation of maternal Se concentration with other covariates in each site. When combined across sites, the Se concentration across sites

was significantly positively correlated with maternal age ($p=0.08$, $p=2.5e-5$) and negatively correlated with gestational age at the time of sample collection ($\rho = -0.13$, $p=3.0e-5$) (online supplemental figure 7). The gestational age at sample collection varied substantially from site to site and in some sites, there were some samples collected after second trimester (≥ 28 weeks). In order to minimise the bias introduced by these samples (eg, exclusion of extremely PTB and reduction of maternal Se concentration), we excluded 416 samples which were collected at 28 weeks of gestational age or later and four samples without known date of sample collection) from the final association analysis (online supplemental figure 8).

Association of maternal selenium concentration with PTB and gestational duration

We examined the association of maternal prenatal Se (before third trimester with gestational duration at sample collection <28 weeks) with PTB and gestational duration in each individual site and then combined the results using meta-analysis (figure 4). In total, the associations were tested in 9519 pregnancies (figure 2). The following factors found to be significantly associated ($p<0.05$) with either gestational duration or Se concentration were incorporated as covariates. These include maternal age (mage), maternal height (ht), fetal sex (fsex) and

Table 1 Demographic characteristics of study subjects

Site	Sample size	Term	Preterm	Male	Female	Gday at delivery	Gday at sampling	Maternal age (year)	Maternal height (cm)	Birth weight (g)
Bangladesh (AMANHI)	506	253 (50%)	253 (50%)	239 (47%)	267 (53%)	260.7 (20)	95.8 (23.1)	23.6 (4.5)	149.2 (5.6)	2516.1 (495.2)
Bangladesh (GAPPS)	258	172 (67%)	86 (33%)	132 (51%)	126 (49%)	267.8 (18.7)	158.8 (4.7)	23.9 (5.9)	151.8 (5.5)	2722.7 (581.1)
Bangladesh (MDIG)	208	138 (66%)	70 (34%)	106 (51%)	102 (49%)	265.9 (14.2)	143.9 (13.5)	23.4 (4.5)	151.4 (5.6)	2664.2 (375.7)
Brazil (INTERBIO)	389	344 (88%)	45 (12%)	212 (54%)	177 (46%)	270.4 (10.8)	132.2 (53.9)	28.5 (5.4)	162.5 (6.4)	3147.1 (464.3)
Kenya (INTERBIO)	553	528 (96%)	25 (4%)	293 (53%)	260 (47%)	278.3 (11.1)	112.5 (40.6)	30.4 (4.1)	161.8 (5.8)	3267 (463.8)
Malawi (ILINS-DYAD)	1212	1126 (93%)	86 (7%)	587 (48%)	625 (52%)	276 (14.3)	117.7 (14.9)	25.2 (6.2)	156.1 (5.7)	2976.6 (449.5)
Pakistan (AMANHI)	348	233 (67%)	115 (33%)	189 (54%)	159 (46%)	265.5 (16.5)	95.1 (24.6)	26.3 (5.1)	154.8 (6.1)	2684.4 (500.1)
Pakistan (INTERBIO)	516	413 (80%)	103 (20%)	251 (49%)	265 (51%)	264.9 (13.5)	103.5 (35.6)	30.1 (4.6)	158 (5.9)	2876.5 (480.7)
South Africa (INTERBIO)	352	299 (85%)	53 (15%)	181 (51%)	171 (49%)	269.4 (17.5)	88.2 (19.6)	30.2 (5.8)	159 (6.9)	2940.3 (588.8)
Tanzania (AMANHI)	351	234 (67%)	117 (33%)	174 (50%)	177 (50%)	267.5 (19.6)	99.1 (23.1)	27.9 (6.6)	155.2 (5.9)	3111.9 (592.5)
Thailand (INTERBIO)	514	485 (94%)	29 (6%)	266 (52%)	248 (48%)	275.6 (11.5)	114.4 (37.1)	26.2 (6.1)	151.8 (5.1)	2965.8 (457.3)
UK (INTERBIO)	648	594 (92%)	54 (8%)	342 (53%)	306 (47%)	275.9 (14.6)	89.9 (20.3)	31.1 (4.8)	165.3 (6.5)	3301.1 (586)
UK (Liverpool)	525	424 (81%)	101 (19%)	271 (52%)	254 (48%)	267 (21.7)	140.8 (9.5)	30.6 (4.9)	164.8 (6.3)	3141.2 (730.3)
USA, California (CPPOP)	966	484 (50%)	482 (50%)	505 (52%)	461 (48%)	249.8 (29.9)	115.7 (7.7)	30 (6.1)	161.6 (7.3)	2763.1 (923)
USA, North Carolina (NEST)	657	438 (67%)	219 (33%)	363 (55%)	294 (45%)	263.1 (22.5)	161.1 (90.5)	28.3 (6.1)	162.8 (7.7)	2999.6 (750.4)
Vietnam (PBB)	970	651 (67%)	319 (33%)	495 (51%)	475 (49%)	264.4 (18.9)	149 (6.2)	29.1 (4.6)	156 (4.8)	2959.9 (613.5)
Zambia (GAPPS)	973	853 (88%)	120 (12%)	478 (49%)	495 (51%)	271.6 (18.2)	137.3 (27.2)	27.7 (5.8)	160.4 (6.5)	3008.7 (591.7)
All	9946	7669 (77.1%)	2277 (22.9%)	5084 (51.1%)	4862 (48.9%)	267.8 (20.2)	122.1 (39.9)	28 (5.9)	158.5 (7.6)	2967.4 (637.4)

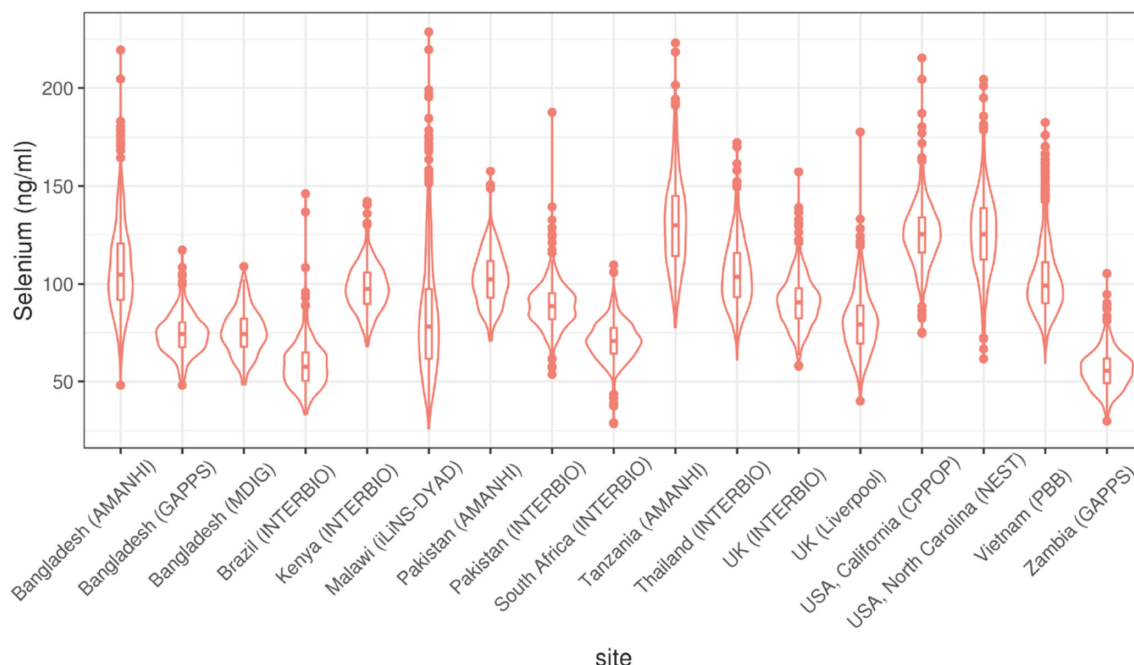


Figure 3 Selenium concentration by participating sites.

gestational days at sample collection (gday(sample)) and experimental batch (batch).

The fixed-effect meta-analysis across the 17 cohorts showed that Se concentration was significantly associated with PTB and gestational duration. The associated effect size estimates were an OR=0.95 (95% CI: 0.9 to 1.00) for PTB or 0.66 days (95% CI: 0.38 to 0.94) longer gestation per 15 ng/mL increase in Se concentration. However, there was substantial between cohort heterogeneity as shown by the forest plots (figure 4) and the significant p values for Cochran's Q statistic (p=0.0037 for PTB and p=6.03e-5 for gestational duration). Given the

site

enrichment of preterm cases in the case-control studies that could potentially introduce bias, we conducted the IPW analysis in the eight case/control data sets and the results were similar to the meta-analysis of gestational age without adjustments (online supplemental figure 9).

The largest effect sizes were observed in UK (Liverpool) cohort, and highly significant associations were observed in Malawi (iLINS-DYAD) samples (figure 4). Other than these two, only the Tanzania (AMANHI) cohort showed associations with marginally smaller p-values (PTB: p=0.026 and gestational duration: p=0.049). After excluding the Malawi (iLINS-DYAD) and

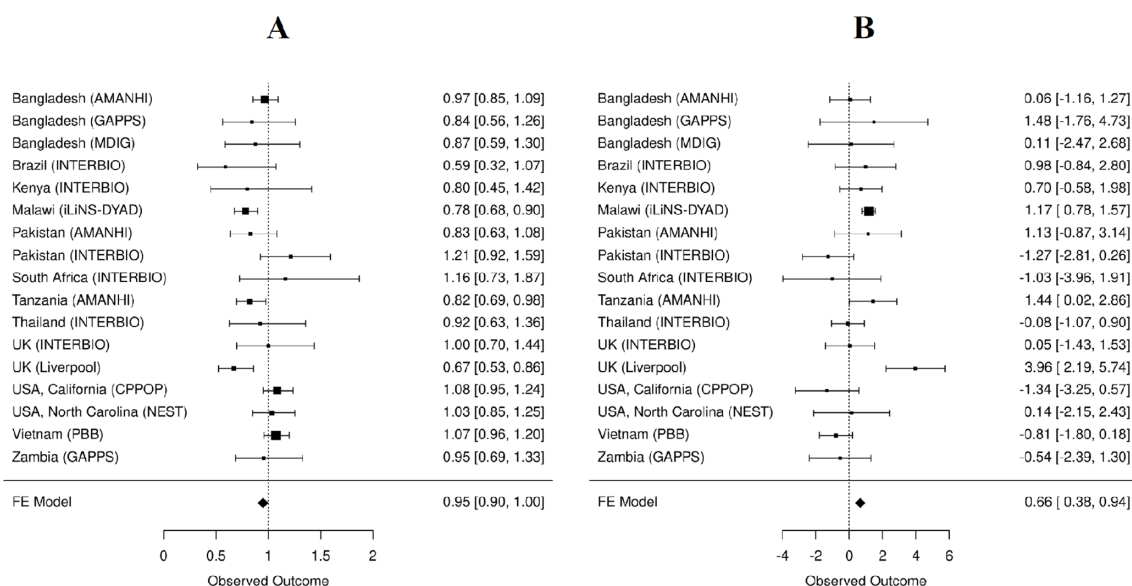


Figure 4 Meta-analysis of the association of maternal Se concentration with PTB (A) and gestational duration (B). (A) The estimated association between Se concentration and PTB is shown as OR per 15 ng/mL increase in Se concentration. (B) The estimated association between Se concentration and gestational duration is shown as change in gestational days per 15 ng/mL increase in Se concentration. PTB, preterm birth; Se, selenium.

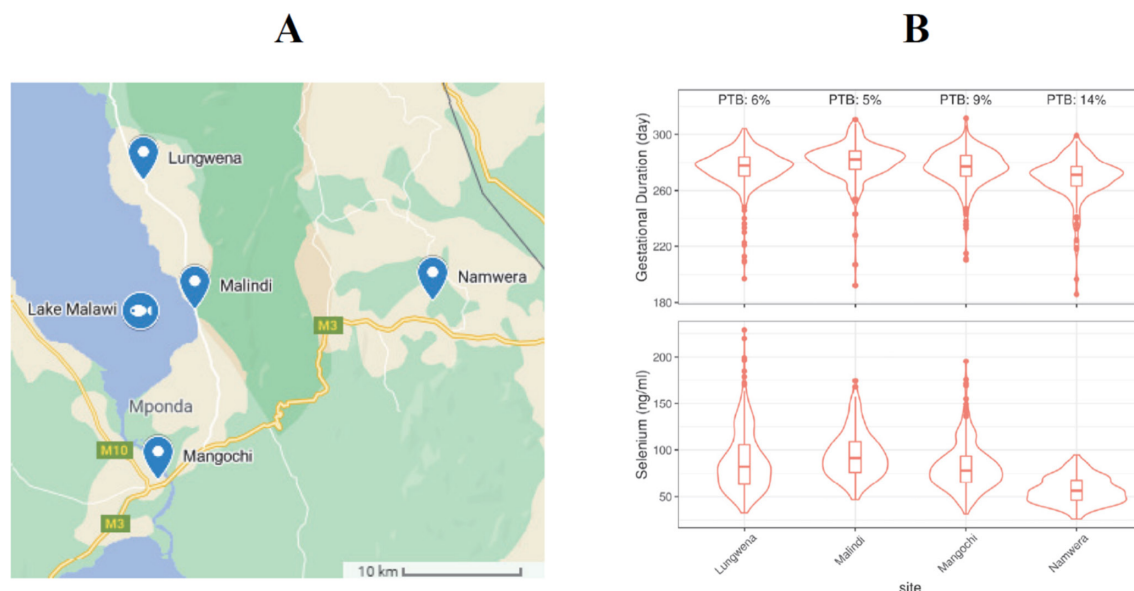


Figure 5 Geographic distribution of the four Malawi subsites (A) and distributions of gestational duration and maternal Se concentration at these four sample recruitment sites (B). PTB, preterm birth; Se, selenium.

UK (Liverpool) cohorts, the fixed-effect meta-analysis was no longer significant (PTB: $p=0.82$ and gestational duration: $p=0.92$). Given the large between study heterogeneity, we also conducted a random-effect meta-analysis of all 17 cohorts. The associations of maternal Se concentrations with PTB and gestational duration in this model did not achieve statistical significance (PTB: $p=0.081$ and gestational duration: $p=0.24$).

Stratified analysis of Malawi cohort

As noted above, the Malawi (iLiNS-DYAD) cohort showed the most significant associations between Se concentrations and PTB ($p=0.00062$) and gestational duration ($p=7.7e-9$) (figure 4). Given these findings, we attempted to investigate possible factors that drove these associations. Participants of the Malawi cohort were enrolled from four health facilities that covered mostly one continuous area near Lake Malawi (figure 5A). Lungwena, Malindi and Mangochi subsites are along the banks of Lake Malawi and close to Namizimu forest reserve. The Namwera subsite is in the mountains and relatively distant from the other three sites. The demographic characteristics of the mothers (eg, age and height) and the major birth outcomes (eg, gestational duration and birth weight) separated by geographic distribution were summarised by the subsites (online supplemental table 4). Compared with the other sites, participants from Namwera had a higher PTB rate (14%) and lower mean gestational duration and birth weight (figure 5B and online supplemental table 4). The lowest mean Se (mean=56.8 ng/mL and SD=14.4 ng/mL) was also observed in Namwera samples (figure 5B and online supplemental table 5). When subsite of sample collection was included as a covariate the effect size estimations of maternal Se concentration were OR=0.85 (95% CI: 0.72 to 1.00) for PTB and 0.49 longer days of gestation (95%

CI: 0.07 to 0.92) per 15 ng/mL increase in Se concentration (online supplemental figure 10). Although still significant, these estimates were smaller than the estimates obtained without adjustment for subsites (OR=0.78 (CI: 0.68 to 0.90) or 1.17 days (95% CI: 0.78 to 1.57)) (figure 4).

Stratified analysis of UK (Liverpool) cohort

The largest effect sizes for Se concentration on gestational duration (3.96 days longer gestation per 15 ng/mL increase in Se) or PTB risk (OR=0.67 per 15 ng/mL increase in Se) were observed in the Liverpool cohort (figure 4). This cohort included 272 high-risk mothers who had a previous PTB and 253 low-risk mothers who did not have previous history of PTB. Among the 272 high-risk mothers, 97 (36%) of them had a subsequent PTB (online supplemental table 6). The average Se concentration was lower in the high-risk mothers than the low-risk mothers (mean=77.6 vs 82.7, t -test $p=0.0002$) (figure 6A and online supplemental table 7). In both low-risk and high-risk groups, Se concentration was positively associated with gestational duration (1.93 days (95% CI: 0.63 to 3.23) per 15 ng/mL increase in Se concentration (figure 6B)) and Se concentration was also associated with PTB risk in the high-risk group (OR=0.74 per 15 ng/mL increase in Se concentration, CI: 0.56 to 0.98).

DISCUSSION

In this multicountry study conducted in low-resource Asian and African countries with high rates of PTB risk and high-resource European and US contexts, we studied the maternal prenatal Se concentrations and their association with the risk of PTB and gestational duration. We found that the Se concentrations varied substantially across different study cohorts. The highest levels

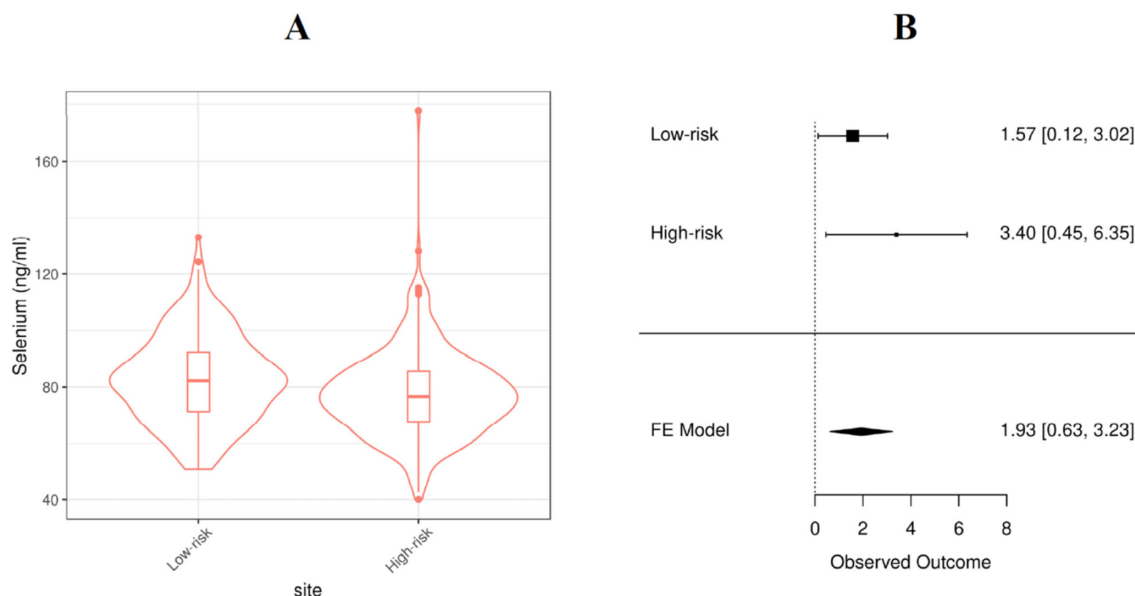


Figure 6 Selenum concentrations of Liverpool cohort (A) and the meta-analysis of maternal Se concentration associated with gestational duration (B). Se, selenum.

of maternal Se were observed in samples collected from Pemba Island (Tanzania) and in the USA (California (CPPOP) and North Carolina (NEST) cohorts. The lowest levels of maternal Se were observed in Zambia (GAPPS) and Brazil (INTERBIO). The average levels of maternal Se in the two cohorts with the lowest levels of maternal Se (<60 ng/mL) were less than half of the average levels observed in the cohorts with the highest levels of maternal Se (>125 ng/mL) (figure 3).

Possible explanation of our findings

Our result demonstrates the substantial variation in Se between different geographic regions. However, even among sites in geographic and cultural proximity (eg, the three Bangladesh or the two Pakistan sites), the maternal Se levels were still different, indicating factors other than geographic location likely influenced maternal Se levels. We also found that the Se concentration positively correlated with maternal age and negatively correlated with gestational age at the time of sample collection. It is unknown whether the decrease of maternal Se reflects the overall increased consumption of nutrients during pregnancy or if Se is utilised by specific biological processes in pregnancy. Also, it is unclear if the decrease in maternal Se with gestational age is due to relative haemodilution because of progressively greater expansion of plasma volume than the increase in red blood cell mass.

In the association analysis between maternal Se concentrations (before third trimester), PTB and gestational duration, we obtained heterogeneous results from the 17 cohorts (figure 4). We observed significant negative associations between maternal Se and PTB risk and positive associations with gestational duration in Malawi (iLiNS-DYAD), UK (Liverpool) and Tanzania (AMANHI) cohorts. However, the associations were not significant in other cohorts and the random-effect meta-analysis of

the 17 cohorts altogether also did not show significant associations.

Further stratified analysis of the Malawi (iLiNS-DYAD) cohort based on the geographic locations of the four sample collection sites (within 30 km apart, figure 5A) suggested that the observed association between Se concentration with PTB and gestational duration was mainly driven by the Namwera site (figure 5B). This site had participants with the highest PTB rate and shortest gestational duration, and lowest Se concentration compared with other three sites near the coast of Lake Malawi. After adjustment for the site of sample collection, the estimated effect size and significance of the associations substantially attenuated. This result suggests that there may be some site-specific confounding factors. However, it is also possible that the low Se concentration is a driving factor that causes the high PTB rate and shorter gestational duration in the Namwera samples because in Namwera samples alone, Se concentration is significantly associated with PTB risk and gestational duration.

The association with the largest effect size between Se and gestational duration was observed in the UK (Liverpool) cohort which included a high-risk and a low-risk group of mothers based on their previous history of PTB. The Se concentration was significantly different between these two groups (figure 6A), which suggests the Se concentration might be an indicator of some long-term risk for PTB which may have also had an effect on previous pregnancies in this cohort. Within each of the UK (Liverpool) groups, maternal Se concentration was associated with the gestational duration of the current pregnancies (figure 6B). In another UK cohort (UK INTERBIO collected at Oxford), the mean Se concentration was approximately 10 ng/mL higher but was not associated with PTB risk or gestational duration.

These disparate findings even between two UK cohorts suggests that unmeasured site-specific factors are either confounding or modifying the associations between maternal Se concentration and gestational duration.

Comparison with other studies

Our findings contribute to an emerging literature focused on the association of Se status and pregnancy outcomes, especially the risk for PTB or gestational duration. Evidence supporting the potential involvement of Se in PTB risk includes a study of Dutch women in which the lowest quartile of serum Se had twice the risk of PTB as women in the upper three quartiles.¹² Another study of pregnant women with HIV in Lagos, Nigeria showed significant associations observed between maternal Se deficiency and PTB.¹³ The Norwegian Mother, Father and Child Cohort study showed that higher Se intake from food was associated with increase in gestational length and decreased PTB risk.¹⁵ Furthermore, the Maternal Health and Birth outcomes study in South East Queensland, Australia, suggested that dietary Se concentrations were significantly higher in women birthing beyond 41 completed weeks of gestation in that cohort.¹⁴ However, there are also reports on Se metabolism with regard to gestational length that find contradicting results. The Japan Environment and Children's Study (JECS) did not find an association between serum Se concentration and PTB risk.¹⁷ The Screening for Pregnancy Endpoints (SCOPE) study in Adelaide suggested that lower circulating levels of Se may be associated with a reduced risk of pregnancy complications including PTB risk.¹⁶ Of note is the fact that these previous studies mostly focus on a single geographic region, have limited or no Se deficiency and are generally based on small sample sizes.

Our study is the most extensive investigation of the association between mid-pregnancy Se concentration and the gestational duration and PTB in global populations, including several lower-income Asian and African countries with a very high baseline PTB risk. The diversity of our study participants and the wide distribution of study sites across different geographic regions enable us to draw some general conclusions. Overall, our results do not support a ubiquitous and strong association between maternal Se concentration and PTB risk. The lack of significant associations in the cohorts with low average Se concentration also suggests Se deficiency is not the primary factor influencing PTB risk. The significant associations observed in some study cohorts might be confounded or mediated by site-specific factors. For example, Se concentration might be associated with certain dietary patterns or socioeconomic status that drive the PTB risk or certain local factors that might interact with Se and jointly influence PTB risk in some high-risk pregnancies.

Clinical importance

Several biologic mechanisms have been hypothesised to link Se status and PTB risk. Selenoproteins serve critical

cellular homeostatic functions in maintaining redox status and antioxidant defenses, and modulate inflammatory responses, which have been linked to PTB.²⁷ In some instances, preterm parturition is thought to be prompted by a cascade of inflammatory events, leading to cytokine upregulation and subsequent induction of uterine activity by promoting the expression and release of uterotonic factors.²⁸ The essential micronutrient Se, which exerts its antioxidant and anti-inflammatory properties in the form of selenoproteins such as glutathione peroxidase 3, selenoprotein P1 and thioredoxin reductase, has been shown to be protective in various inflammatory-based disease models.^{29–31} A recent study in mice showed that Se in the form of selenoproteins played an indispensable role in uterine smooth muscle contractions, and the absence of any of these proteins affected the uterine contractility.³² In vitro study of Se supplementation demonstrated that selenite suppresses key mediators involved in inflammation-induced activation of mediators involved in active labour in human fetal membranes and the myometrium.³³ Further investigations may benefit from looking in more detail at whether pregnancies exhibiting higher levels of inflammation or increased cytokine dysregulation benefit from higher levels of Se in terms of increased gestational duration or decreased PTB risk.

The hierarchy of biological activities of Se calls for biomarkers informative at different levels of Se exposure assessing Se intake, tissue Se, Se excretion and Se function.²⁶ Plasma or serum Se level provides valuable information about the Se status over a wide range of Se intake; however, there is need for additional Se speciation information particularly for assessing Se status in non-deficient individuals for whom there is high risk for PTB. Epidemiological reports and research examining the effects of different Se species and their bioavailability and bioactivity especially during pregnancy are lacking. There are recent reports suggesting that the non-linear associations between whole blood Se and plasma Se may be primarily due to accumulation of large proportion of selenoneine in red blood cells especially in coastal populations consuming marine foods.³⁴

Strengths and limitations

There are some significant limitations of the current study. Of note is the fact that the samples and phenotypic data were retrieved from existing biorepositories collected several years ago in different studies. Although we harmonised and analysed a set of key variables known to be associated with PTB and gestational duration, we were not able to include some important environmental or socioeconomic factors in the analysis due to missing or incomplete data. We excluded stillbirth due to missing data on cause-of-death, under-reporting and lack of comparability in reporting of stillbirths, especially in low-income and middle-income countries regarding the birth weight and gestational age criteria. It will be key to include these variables across cohorts in future studies.

There were differences in how gestational age was determined and distributed across cohorts. Some cohorts determined the duration by ultrasound whereas others used LMP (or both). This different dating methodology between studies may have introduced some noise into the analysis. PTB rates reported in some low-income and middle-income cohort studies appear to be low, and this might be due to under-reporting and geographic location of the recruitment site. Also, some cohorts were enriched for PTB samples, and the distribution of gestational duration did not follow a normal distribution. Although regression analysis is generally robust regardless of meeting the normality assumption, this difference may have introduced some bias in these analyses. Also, we tested PTB as primary outcome using logistic regression, which is valid to both case-control and random samples. These issues certainly point to the importance of standardising dating and sampling methods as investigations move forward.

Also, of note with respect to the study limitations is that there was large variation in gestational age at when the plasma/serum samples were collected. Also, the samples were stored for different periods of time. Given the gestational age at sample collection significantly correlated with the Se concentration, we accounted for this variance by including only the pregnant mothers with samples collected before or during the second trimester and included gestational age at sample collection as a covariate in the final association analysis. Despite these methodological adaptations, it is possible that we may not have completely accounted for the influence of gestational age at sample collection if the effect is not completely linear. More standardisation with respect to timing of collection and storage times may simplify these types of analyses in future studies.

In addition, given that the major source of Se is food, and the large proportion of it comes from the staple food items such as rice, wheat and seafood,¹⁹ it is clear that studies looking at Se, PTB risk and gestational duration would benefit from more data on diet and nutrition. While information regarding the dietary intake of Se or other supplements during the pregnancy was not readily available for the majority of the enrolled subjects from the study sites, collection of such data in future studies would be hugely beneficial.

Finally, the generalisability of our results from the Se measurements to all study populations was likely limited due to the considerable variation in the local factors that influence the Se levels or modify the effect of Se during pregnancy. This highlights the need for larger coordinated studies examining extraneous factors that may be associated with Se levels, risk of PTB and gestational duration in pregnant women across different geographic settings.

CONCLUSIONS

We studied maternal prenatal Se concentration and tested whether it is associated with the risk of PTB and gestational duration using data and samples collected from

17 international birth cohorts with diverse ethnic background and geographic distribution. Our study observed statistically significant associations between maternal Se concentration and PTB at some sites; however, this did not generalise across the entire cohort, which might lower the enthusiasm for wide use of Se supplements as a general strategy to prevent PTB or increase gestational duration. The significant associations observed in some cohorts and not others suggest local confounding factors or other risk modifiers. Effects of Se supplementation on PTB in high-risk populations with low Se in food (like Namwera region in Malawi) or in high-risk mothers with previous history of PTB need to be confirmed, ideally through a double-blind, placebo controlled clinical trial. Future studies that expand and refine sampling in populations that are found to have the greatest variations in Se intake and Se deficiency, along with Se speciation analysis will shed further light on the whether there is a potential relationship between Se, PTB risk and gestational duration.

Author affiliations

¹Division of Neonatology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

²Center for Prevention of Preterm Birth, Perinatal Institute, Cincinnati Children's Hospital Medical Center and March of Dimes Prematurity Research Center Ohio Collaborative, Cincinnati, Ohio, USA

³Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

⁴Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

⁵International Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA

⁶Biorepository and Omics Research Group, Department of Pediatrics and Child Health, Faculty of Health Sciences, Medical College, The Aga Khan University, Karachi, Sindh, Pakistan

⁷Center for Public Health Kinetics, New Delhi, India

⁸Research Division, Public Health Laboratory, Center for Public Health Kinetics, Chake Chake, Tanzania

⁹Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Dhaka District, Bangladesh

¹⁰Obstetrics and Gynecology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

¹¹INTERBIO-21st Study Consortium, Nuffield Department of Women's & Reproductive Health, University of Oxford, Oxford, UK

¹²Nutrition and Clinical Services Division, International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Dhaka District, Bangladesh

¹³Center for Child Health Research, Faculty of Medicine and Health Technology, Tampere University, Tampere, Pirkanmaa, Finland

¹⁴Benh Vien Tu Du, Ho Chi Minh City, Viet Nam

¹⁵Department of Women's and Children's Health, University of Liverpool, Liverpool, UK

¹⁶Department of Chemistry, University of Cincinnati, Cincinnati, Ohio, USA

¹⁷Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Medford, Massachusetts, USA

¹⁸Grand Forks Human Nutrition Research Center, USDA ARS, Grand Forks, North Dakota, USA

¹⁹Prohahnmo Research Foundation, Dhaka, Bangladesh

²⁰Nutritional and Clinical Services Division, International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Dhaka District, Bangladesh

²¹School of Medicine, University of Zambia, Lusaka, Zambia

²²School of Public Health, University of Zambia, Lusaka, Zambia

²³Cardiology, ICDDR,B, Dhaka, Bangladesh

²⁴Nutrition and Clinical Services Division, International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Bangladesh

- ²⁵University of Tampere, Tampere, Pirkanmaa, Finland
- ²⁶School of Public Health, University of Malawi College of Medicine, Blantyre, Malawi
- ²⁷Medical Research Centre Oulu, PEDEGO Research Unit, University of Oulu, Oulu, Pohjois-Pohjanmaa, Finland
- ²⁸Department of Women's and Children's Health, University of Liverpool, Liverpool, UK
- ²⁹Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina, USA
- ³⁰Department of Obstetrics, Gynecology and Reproductive Sciences, University of California San Francisco, San Francisco, California, USA
- ³¹Department of Epidemiology, University of Iowa, Iowa City, Iowa, USA
- ³²Department of Pediatrics, University of Iowa, Iowa City, Iowa, USA
- ³³Department of Medicine, World Health Organization, Geneva, Switzerland
- ³⁴Global Alliance to Prevent Prematurity and Stillbirth, Lynnwood, Washington, USA
- ³⁵Division of Perinatal Medicine, University of Liverpool, Liverpool, UK
- ³⁶Center for Child Health Research, Faculty of Medicine and Health Technology, University of Tampere, Tampere, Pirkanmaa, Finland
- ³⁷Department of Pediatrics, Tampere University Hospital, Tampere, Finland
- ³⁸International Center for Maternal and Newborn Health, Department of International Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
- ³⁹Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK
- ⁴⁰Department of Biological Sciences and Center for Human Health and the Environment, North Carolina State University, Raleigh, North Carolina, USA
- ⁴¹Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan
- ⁴²Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA
- ⁴³Department of Paediatrics, The Hospital for Sick Children, Toronto, Ontario, Canada
- ⁴⁴Burroughs Wellcome Fund, Research Triangle Park, North Carolina, USA

Twitter Imran Nisar @imran_nisar

Acknowledgements We are grateful to all participating families and study personnel from 17 international pregnancy cohorts, who took part in this study. Payment for access to data and article-processing charges for this publication was covered by The Bill & Melinda Gates Foundation (Grant no: OPP1175128, OPP1152451). The authors would like to acknowledge the March of Dimes Prematurity Research Center Ohio Collaborative for their support to the original GWAS study that identified the EEFSEC gene and led to this study. The Aga Khan University would like to acknowledge Dr Yoshida Sachiyo and Dr Alexander Manu from the WHO, community health workers, pregnant women and their families. This study used biospecimen processing and storage services from the Discover Together Biobank at Cincinnati Children's Hospital Research Foundation. We thank the Discover Together Biobank for support of this work. The NEST study acknowledges the support from National Institute of Environmental Health Sciences, the US Environmental Protection Agency, the National Institute of Diabetes and Digestive and Kidney Diseases and the Duke Cancer Institute. The CPPOP study acknowledge support from the UCSF California Preterm Birth Initiative. The iLINS-DYAD-M trial acknowledge the support by a grant to the University of California, Davis from The Bill & Melinda Gates Foundation (OPP49817) and a grant to the University of California, Davis from the Office of Health, Infectious Diseases and Nutrition, Bureau for Global Health, US Agency for International Development (USAID) through the Food and Nutrition Technical Assistance III Project (FANTA). MDIG, AMANHI, GAPPs and INTERBIO cohorts acknowledge the support by The Bill & Melinda Gates Foundation.

Collaborators INTERBIO-21st Study Consortium.

Contributors NM, HX and GZ prepared the first draft of the manuscript. LM conceptualised and acquired the funding support for the study. The study was designed by NM, EB, JAL, GFC, JC, GZ and LM. EB, JAL and NM conducted the ICP-MS analysis of the samples. HX, EB, NM and GZ compiled the data sets. HX and GZ developed the analytical pipeline and did the statistical analysis. JC coordinated all study-related operations. All coauthors have contributed essential intellectual input, revisions of the manuscript, discussed the results, contributed to the revisions of the final manuscript and approved the final version of the manuscript. The corresponding authors, NM and GZ had full access to the data and had final responsibility for the decision to submit for publication.

Funding This work was supported by the Bill & Melinda Gates Foundation. The funders did not have any role in study design, data analysis, data interpretation, writing of the report or submission for publication. The findings and conclusions of this article are solely the responsibility of the authors. Bill & Melinda Gates Foundation (OPP1175128, OPP1152451). Funders: GZ is supported by the March of Dimes Prematurity Research Center Ohio Collaborative and by the Burroughs Wellcome Fund (Grant 10172896). Cohort Investigations. California Prediction of Poor Outcomes of Pregnancy (CPPOP): UCSF California Preterm Birth Initiative. NEST: National Institute of Environmental Health Sciences (R21ES014947, R01ES016772, P30ES025128 and P01ES022831), the US Environmental Protection Agency (RD-83543701), the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK085173) and the Duke Cancer Institute. iLINS-DYAD-M trial is funded by a grant to the University of California, Davis from the Bill & Melinda Gates Foundation (OPP49817). A grant to the University of California, Davis from the Office of Health, Infectious Diseases and Nutrition, Bureau for Global Health, US Agency for International Development (USAID) under terms of Cooperative Agreement No. AID-OAA-A-12-00005 through the Food and Nutrition Technical Assistance III Project (FANTA), managed by FHI 360. The Global Alliance to Prevention of Prematurity and Stillbirth (GAPPs) Biorepository Programme funded by the Preventing Preterm Birth Initiative grant from the Bill & Melinda Gates Foundation (OPP1033514). MDIG trial funded by The Bill & Melinda Gates Foundation (OPP1066764).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was approved by the Institute Review Board (IRB) of the Cincinnati Children's Hospital Medical Centre (URB 2017-3573) and by the corresponding Ethics Committees of each participating institution.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Nagendra Monangi <http://orcid.org/0000-0003-2498-6452>
 Waqasuddin Khan <http://orcid.org/0000-0001-7995-8163>
 Md Munirul Islam <http://orcid.org/0000-0002-8780-8760>
 Juhi K Gupta <http://orcid.org/0000-0002-8292-9846>
 James A Litch <http://orcid.org/0000-0003-0570-0308>
 Per Ashorn <http://orcid.org/0000-0003-2311-2593>
 Anisur Rahman <http://orcid.org/0000-0003-1033-5034>
 Jeffrey Stringer <http://orcid.org/0000-0002-9590-7216>

REFERENCES

- 1 Lawn JE, Kinney MV, Belizan JM, *et al*. Born too soon: accelerating actions for prevention and care of 15 million newborns born too soon. *Reprod Health* 2013;10 Suppl 1:S6.
- 2 Liu L, Oza S, Hogan D, *et al*. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the sustainable development goals. *Lancet* 2016;388:3027-35.
- 3 Howson CP, Kinney MV, McDougall L, *et al*. Born too soon: preterm birth matters. *Reprod Health* 2013;10 Suppl 1:S1.
- 4 Blencowe H, Cousens S, Oestergaard MZ, *et al*. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012;379:2162-72.

- 5 Eichenwald EC, Stark AR. Management and outcomes of very low birth weight. *N Engl J Med* 2008;358:1700–11.
- 6 Hack M, Taylor HG, Drotar D, *et al.* Chronic conditions, functional limitations, and special health care needs of school-aged children born with extremely low-birth-weight in the 1990s. *JAMA* 2005;294:318–25.
- 7 Barker DJP. The origins of the developmental origins theory. *J Intern Med* 2007;261:412–7.
- 8 Järvelin M-R, Sovio U, King V, *et al.* Early life factors and blood pressure at age 31 years in the 1966 northern Finland birth cohort. *Hypertension* 2004;44:838–46.
- 9 Sipilä-Leppänen M, Väärasmäki M, Tikanmäki M, *et al.* Cardiometabolic risk factors in young adults who were born preterm. *Am J Epidemiol* 2015;181:861–73.
- 10 Zhang G, Feenstra B, Bacelis J, *et al.* Genetic associations with gestational duration and spontaneous preterm birth. *N Engl J Med* 2017;377:1156–67.
- 11 Labunskyy VM, Hatfield DL, Gladyshev VN. Selenoproteins: molecular pathways and physiological roles. *Physiol Rev* 2014;94:739–77.
- 12 Goldenberg RL, Culhane JF, Iams JD, *et al.* Epidemiology and causes of preterm birth. *Lancet* 2008;371:75–84.
- 13 Rayman MP, Wijnen H, Vader H, *et al.* Maternal selenium status during early gestation and risk for preterm birth. *CMAJ* 2011;183:549–55.
- 14 Okunade KS, Olowoselu OF, Osanyin GE, *et al.* Selenium deficiency and pregnancy outcome in pregnant women with HIV in Lagos, Nigeria. *Int J Gynaecol Obstet* 2018;142:207–13.
- 15 McAlpine JM, McKeating DR, Vincze L, *et al.* Essential mineral intake during pregnancy and its association with maternal health and birth outcomes in South East Queensland, Australia. *Nutr Metab Insights* 2019;12:1178638819879444.
- 16 Barman M, Brantsæter AL, Nilsson S, *et al.* Maternal dietary selenium intake is associated with increased gestational length and decreased risk of preterm delivery. *Br J Nutr* 2020;123:209–19.
- 17 Wilson RL, Bianco-Miotto T, Leemaqz SY, *et al.* Early pregnancy maternal trace mineral status and the association with adverse pregnancy outcome in a cohort of Australian women. *J Trace Elem Med Biol* 2018;46:103–9.
- 18 Tsuji M, Shibata E, Morokuma S, *et al.* The association between whole blood concentrations of heavy metals in pregnant women and premature births: the Japan environment and children's study (JECS). *Environ Res* 2018;166:562–9.
- 19 Carmichael SL, Yang W, Shaw GM, *et al.* Maternal dietary nutrient intake and risk of preterm delivery. *Am J Perinatol* 2013;30:579–88.
- 20 Combs GF. Selenium in global food systems. *Br J Nutr* 2001;85:517–47.
- 21 Roth DE, Morris SK, Zlotkin S, *et al.* Vitamin D supplementation in pregnancy and lactation and infant growth. *N Engl J Med* 2018;379:535–46.
- 22 Jelliffe-Pawlowski LL, Rand L, Bedell B, *et al.* Prediction of preterm birth with and without preeclampsia using mid-pregnancy immune and growth-related molecular factors and maternal characteristics. *J Perinatol* 2018;38:963–72.
- 23 Martin CL, Jima D, Sharp GC, *et al.* Maternal pre-pregnancy obesity, offspring cord blood DNA methylation, and offspring cardiometabolic health in early childhood: an epigenome-wide association study. *Epigenetics* 2019;14:325–40.
- 24 Alliance for Maternal and Newborn Health Improvement (AMANHI) mortality study group. Population-Based rates, timing, and causes of maternal deaths, stillbirths, and neonatal deaths in South Asia and sub-Saharan Africa: a multi-country prospective cohort study. *Lancet Glob Health* 2018;6:e1297–308.
- 25 Kamng'ona AW, Young R, Arnold CD, *et al.* Provision of lipid-based nutrient supplements to mothers during pregnancy and 6 months postpartum and to their infants from 6 to 18 months promotes infant gut microbiota diversity at 18 months of age but not microbiota maturation in a rural Malawian setting: secondary outcomes of a randomized trial. *J Nutr* 2020;150:918–28.
- 26 Kennedy SH, Victora CG, Craik R, *et al.* Deep clinical and biological phenotyping of the preterm birth and small for gestational age syndromes: The INTERBIO-21st Newborn Case-Control Study protocol. *Gates Open Res* 2018;2:49.
- 27 Combs GF. Biomarkers of selenium status. *Nutrients* 2015;7:2209–36.
- 28 Romero R, Espinoza J, Gotsch F, *et al.* The use of high-dimensional biology (genomics, transcriptomics, proteomics, and metabolomics) to understand the preterm parturition syndrome. *BJOG* 2006;113 Suppl 3:118–35.
- 29 Zhang J-li, Li J-L, Huang X-dan, *et al.* Dietary selenium regulation of transcript abundance of selenoprotein N and selenoprotein W in chicken muscle tissues. *Biometals* 2012;25:297–307.
- 30 Cheng Q, Sandalova T, Lindqvist Y, *et al.* Crystal structure and catalysis of the selenoprotein thioredoxin reductase 1. *J Biol Chem* 2009;284:3998–4008.
- 31 Speckmann B, Gerloff K, Simms L, *et al.* Selenoprotein S is a marker but not a regulator of endoplasmic reticulum stress in intestinal epithelial cells. *Free Radic Biol Med* 2014;67:265–77.
- 32 Wang F-H, Peng X, Chen Y, *et al.* Se regulates the contractile ability of uterine smooth Musclevia selenoprotein N, selenoprotein T, and selenoprotein WIN mice. *Biol Trace Elem Res* 2019;192:196–205.
- 33 Kalansuriya DM, Lim R, Lappas M. In vitro selenium supplementation suppresses key mediators involved in myometrial activation and rupture of fetal membranes. *Metallomics* 2020;12:935–51.
- 34 Little M, Achouba A, Dumas P, *et al.* Determinants of selenoneine concentration in red blood cells of Inuit from Nunavik (Northern Québec, Canada). *Environ Int* 2019;127:243–52.

SUPPLEMENTARY MATERIAL

Table of Contents

Supplementary Text 1: Summary description of the parent study, site setting and sample collection.....	2
Supplementary Text 2: Protocol for Selenium analysis in plasma and serum samples.....	5
Supplementary Text 3: Regression analysis weighted by inverse of sampling probabilities.....	7
Supplementary Tables	8
Supplementary Table 1. Study characteristics of participant cohorts	8
Supplementary Table 2. Covariates (major phenotypes) requested from sites.....	9
Supplementary Table 3. Summary statistics of maternal Se concentration in different study sites	10
Supplementary Table 4. Demographic characteristics of Malawi study subjects	11
Supplementary Table 5. Summary statistics of Se concentration in Malawi study subjects	12
Supplementary Table 6. Demographic characteristics of UK (Liverpool) study subjects	13
Supplementary Table 7. Summary statistics of Se concentration in UK (Liverpool) study subjects	14
Supplementary Figures	15
Supplementary Figure 1: Distribution of gestational duration of singleton live births with spontaneous onset of labor from all sites	15
Supplementary Figure 2. Gestational duration in term and preterm deliveries by participating sites.....	16
Supplementary Figure 3. Correlation of gestational duration and preterm birth with other covariates	17
Supplementary Figure 4. Distribution of maternal Se concentration of samples from all sites	18
Supplementary Figure 5. Maternal Se concentration in term and preterm deliveries by participating sites	19
Supplementary Figure 6. Maternal Se concentration measured at different batches (colored by site)	20
Supplementary Figure 7. Correlation of maternal Se concentration with other covariates	21
Supplementary Figure 8. Gestational age (weeks) at sample collection by sites.....	22
Supplementary Figure 9. Meta-analysis of the association of maternal Se concentration with gestational duration with adjustment of case/control sampling with IPW analysis.....	23
Supplementary Figure 10. Meta-analysis of the association of Selenium concentration (unit: 15 ng/ml) with PTB (A) and gestational duration (B) among the 4 Malawi sites	24
Supplementary References	25

Supplementary Text

Supplementary Text 1: Summary description of the parent study, site setting and sample collection

Bangladesh (AMANHI) is a biobank in a population-based cohort of 3,000 pregnant women enrolled before 19 weeks of gestation and followed up to 42 days post-partum. The overarching goal is to facilitate discoveries of biomarkers of adverse pregnancy outcomes (maternal, fetal and neonatal health outcomes) as new and more feasible methods become available. An additional goal is to identify biological mechanisms underlying the causes of the adverse outcomes including preeclampsia, spontaneous preterm birth (sPTB), stillbirth and intrauterine growth restrictions (IUGR) to create a platform to generate new approaches to treatment and prevention(1). All pregnant women were identified through pregnancy surveillance conducted by making home visits every 2 months by trained community health workers (CHWs). Pregnancies were confirmed via strip-based pregnancy tests and dated through ultrasound scans carried out by trained ultrasonologists before 19 weeks of gestation. The biobank contains maternal blood and urine specimens collected two times during pregnancy (8-19 weeks and 24-28 weeks or 32-36 weeks of gestation) and once during postpartum period (Day 42 postpartum) as well as delivery samples. Trained phlebotomists obtained maternal and umbilical cord blood samples and generated aliquots of serum, plasma, and buffy coats for storage. We also have collected and stored maternal urine, placental samples, umbilical cord blood, tissue and membrane. When cord blood collection was not possible, saliva was collected from newborns. In addition, we collected infant blood at 12 months of age and paternal saliva samples. All samples were processed and stored in -80°C freezers. CHWs collected detailed phenotypic and epidemiological data from the pregnant women four times during pregnancy (at 8-19 weeks, 24-28 weeks, 32-36 weeks, and 38-40 weeks of gestation), at delivery and twice during postpartum period (<7 days and at 42 days). The study started in July, 2014 and ended in April, 2018.

Bangladesh (GAPPS) is a population based prospective cohort study of pregnant women with an objective to establish infrastructure for researchers to conduct preterm birth and pregnancy related research. The study enrolled 4220 pregnant women from Matlab district over three years (2015-2017). The research site enrolled women early in pregnancy and collected information and biological specimens during their pregnancy and delivery. Gestational age was estimated by standardized ultrasound method. All biological samples were stored immediately in the local biorepository using the standardized protocol developed by GAPPS domestic biorepository. The study was approved by the research ethics committee of the Matlab Health Research Center at the International Center for Diarrheal Disease Research, Bangladesh.

Bangladesh (MDIG) study is a randomized intervention trial of vitamin D supplementation during pregnancy and lactation (2). Participants were generally healthy pregnant women between 17 and 24 weeks of gestation that were enrolled between March 2014 and September 2015 at the Maternal and Child Health Training Institute in Dhaka, Bangladesh. Gestational age was estimated by ultrasound or LMP or both. The study was approved by research ethics committees at the Hospital for Sick Children at Toronto and the International Center for Diarrheal Disease Research, Bangladesh (icddr,b).

Brazil, Kenya, Pakistan, South Africa, Thailand and UK (Interbio) is a multicenter, population-based research initiative coordinated by the University of Oxford to assess human growth, neurodevelopment and associated behaviors from early pregnancy to 2 years of age (3). The Interbio study was conducted between February 2012 and June 2018 at sites: Pelotas (Brazil), Nairobi (Kenya), Karachi (Pakistan), Soweto (South Africa), Mae Sot (Thailand) and Oxford (UK). Gestational age was determined by standardized ultrasound method. The studies were approved by regional ethics or institutional research boards and by the Institutional review board at the University of Oxford.

Malawi (iLiNS-DYAD) is a randomized, controlled, partially blinded, parallel-group intervention trial known as the International Lipid-based Nutrient Supplements DYAD trial which was designed to study the health impacts of lipid-based nutrient supplements during pregnancy and lactation (4). The study was conducted in 2 hospitals and 2 health centers in a rural area in Mangochi district between 2011-2012. Participants of the Malawi cohort were enrolled from four health facilities that covered mostly one continuous area near Lake Malawi. Lungwena, Malindi, and Mangochi subsites are along the banks of Lake Malawi and close to Namizimu forest reserve. Gestational age was determined using the ultrasound method. The ethical clearance for the study was granted by the University of

Malawi College of Medicine Research and Ethics Committee (COMREC) and the ethics committee at Tampere University Hospital District, Finland.

Pakistan (AMANHI) is a population-based biorepository with the aim of collecting cause-specific biosamples on maternal and neonatal mortality, and stillbirths from a well-characterized cohort of pregnant women(1). The biobank enrolled 2,500 pregnant women from 2014 to 2018 with the last pregnancy outcome occurring in January 2019, after an ultrasound in early pregnancy (<20 weeks) to confirm the gestational age. Women were also visited thrice at 24-28, 32-36 and 38-42 weeks gestation during pregnancy, and 0-6 days and 42-59 days after birth to measure blood pressure, test urine for proteinuria, and ascertain reported morbidity since the previous visit. The outcome of each identified pregnancy, whether abortion, stillbirth or live birth, was carefully documented and verbal autopsies were conducted with appropriate respondents in case of maternal deaths, stillbirths or neonatal deaths. Subsequently, all live born babies were assessed for neuromuscular, physical and feeding maturity as well as neonatal anthropometry (including baby's weight and foot length) within 72 hours of birth up until 4-5 years of age, using harmonized procedures. Blood and urine samples were collected from each woman at three time points: enrolment (<20 weeks), at either 24-28 weeks or 32-36 weeks gestation, and at 42 days postpartum. These samples were processed and stored at -80°C in multiple aliquots. At delivery of a still- or live birth, maternal stool, umbilical cord blood and placental were collected, processed and stored within 30 minutes of birth. At the 42 days postpartum visit, maternal blood and urine, infant stool and paternal saliva samples were also collected for processing and storage. Infant saliva was collected if cord blood could not be obtained. Study protocols for enrolment, visits, ultrasound scans, sample collection, processing and storage were implemented by highly trained and motivated staff. Community members were found to be very cooperative and supportive of the study. The study was approved by the regional ethics review committee.

Tanzania (AMANHI) Tanzania (Pemba) is one of the AMANHI sites with bio-banked biological samples from a cohort of 4501 pregnant women and their children. The overall objective of all participating sites in the AMANHI study was similar, and all the SOP implemented for sample collection and processing were harmonized across sites(1). The study was conducted in Pemba island of the Zanzibar archipelago with an overall population of around 432000 and approximately 82,000 households. Two districts of the island were selected for the study where pregnancy surveillance was conducted every 2 months to identify pregnant women. Consent was obtained for confirmation of pregnancy by urine strip tests, thereafter gestational age was confirmed with routine ultrasound methods. All women between 8-19 weeks of gestation were consented and then enrolled in the study. Blood and urine samples were collected from the enrolled mothers at the time of enrollment and in either during 24 – 28 weeks or 32 – 36 weeks of gestation and 42 days postpartum. At delivery in addition to cord blood, tissue samples from placenta, membrane cord were collected within 30-60 minutes of delivery. All samples were processed as per harmonized SOP and stored in the biobank at -80° C. Saliva samples from the father and fecal samples from the mother and infant were also collected after delivery. For collecting the epidemiological data study team visited all the mothers during pregnancy and after delivery. Additional information on delivery was obtained from the hospital delivery records filled in by the physician in-charge. A in house designed AMANHI biobank LMIS software was used by all AMANHI sites for recording the collection, processing and storage information of the biospecimens. Stringent quality control checks were implemented for data consistency and sample quality during the study. The study started in June 2014 and the last postnatal follow-up sample from the mother was collected in December 2018.

UK (Liverpool) was started in April 2012 and ended in December 2017. The study entitled “The development of novel biomarkers for prediction of preterm labor in a high-risk population”. This study enrolled a total of 541 pregnant women with singleton pregnancies at the Liverpool Women's Hospital, UK. Gestational age was determined using the ultrasound method. The study was approved by the Institutional Review board at the University of Liverpool.

USA, CA(CPPPOP) is a nested case-control sampling study drawn from a population-based cohort of 757,853 singleton live births in the state of California (5). Women with nonfasted serum samples banked by the California biobank program from July 2009 through December 2010 were enrolled in the study. Gestational age was determined by ultrasound or LMP or both. Methods and protocols for the study were approved by the Committee for the Protection of Human Subjects within the Health and Human Services Agency of the State of California, the

Institutional Review Board of Stanford University and the Institutional Review Board of the University of California San Francisco.

USA, NC (NEST) is a perinatal cohort study of more than 2000 pregnant and their offspring mounted to investigate the role of environmental exposures and nutrition in utero on the shifts in the epigenome of newborns (6), from which we nested a case control study. The study participants were a birth cohort from women who received prenatal care in the Duke/Durham region health care system in Durham, NC between 2005 and 2009. Ultrasound is used to determine the gestational age. The study was approved by the Institutional Review Board of Duke University, North Carolina.

Vietnam (PBB) is an observational study in Ho Chi Minh city to evaluate biomarkers for spontaneous preterm birth in partnership with Sera Prognostics, OUCRU and the Gates Foundation. This is a prospective hospital-based convenience sampling of 4800 women between September 2016 through August 2018 who have antenatal care and plan to deliver at Tu Du hospital and were between 19⁺⁰ to 22⁺⁶ days of gestation when samples were taken. Gestational age was determined using ultrasound. The study was approved by the Institutional review board at the Tu Du Hospital and the University of Oxford (OXTREC 28-16).

Zambia (GAPPS) is a population based prospective cohort study of pregnant women with an objective to establish infrastructure for researchers to conduct preterm birth and pregnancy related research. The study enrolled 2000 pregnant women from Lusaka district over three years (2015-2017). The research site enrolled women early in pregnancy and collect information and biological specimens during their pregnancy and delivery. Gestational age was estimated by ultrasound at gestational weeks 16-22. All biological samples were stored immediately in the local biorepository using the standardized protocol developed by GAPPS domestic biorepository. The study was approved by the regional ethics review board.

Supplementary Text 2: Protocol for Selenium analysis in plasma and serum samples

Sample Preparation

Before analysis the serum and plasma samples were thawed on ice for 30 minutes followed by sonication in an ultrasonic bath (Fisher Scientific, CPX1800) for 5 minutes in order to mix the samples. During these processes, the sample vials remained sealed in order to prevent dilution of the sample caused by the premature opening of the sample vials before they reach room temperature. An acid digestion was performed on the samples with the following protocol: 50 µL of plasma or serum from each sample as well as the quality control serum was transferred to 15 mL metal free vials (VWR) using clear pipette tips and a calibrated electronic micropipette (Eppendorf). Both the samples and the quality control serum were mixed briefly using a mini vortex mixer (VWR) immediately prior to transferring the aliquot in order to promote better sampling of a nonhomogeneous sample matrix. The quality control serum used was obtained from UTAK Laboratories, Inc. and included normal range trace elements which was reconstituted according to manufacturer instructions. In the event that 50 uL of sample could not be obtained, 40 uL or as low as 30 uL aliquots of sample and quality control serum were used according to sample availability.

50 µL of the internal standard mixture containing 500 ppb of Sc, In, Y and Te in 0.5 M nitric acid (High Purity Standards) was then added with a repetition pipette (Eppendorf) to the 15 mL metal free vials, followed by 200 µL of concentrated trace metal grade nitric acid (Sigma-Aldrich; Fisher Scientific). The samples were placed in a dry bath with no more than 3 cm of the tubes immersed in the heating block holes in order to allow reflux of the sample. For this, aluminum foil was inserted into the block holes.

The samples were heated at 85 °C for one hour, then at 95 °C for 1.5 hours. In order to complete the digestion, 100 µl of trace metal grade hydrogen peroxide (Sigma-Aldrich) was added with a repetition pipette after the samples cooled for 5 minutes outside the heating block. Then the samples were returned to the heating block another 30 minutes at 95 °C. Once the digestion was completed, the final volume was brought up to 2.5 ml with doubly deionized water and mixed using a vortex mixer.

Final Concentration of Standard (ppb)	0	0.5	1	2	5	10	25
Volume of 100 ppb Working Standard (µL)	0	25	50	100	250	500	1250
Volume of 500 ppb Internal Standard (µL)	100	100	100	100	100	100	100
Volume of 0.5M HNO3 (µL)	4900	4875	4850	4800	4650	4400	3650

Protocol table 1. Calibration standard preparation

Calibration Preparation

The external calibration method with internal standard in-samples was used. For this the following points were used: 0 ppb, 0.5 ppb, 1 ppb, 2 ppb, 5 ppb, 10 ppb and 25 ppb. The individual volumes used are shown in Protocol supplementary table 1. The standards were prepared in 0.5 M trace metal grade nitric acid. A stock solution containing a mixture of elements of interest at 10 ppm (SPEX CertiPrep™) was used to prepare a daily working standard of 100 ppb. For this a 100x dilution was made by adding 50 µL of the stock standard to a metal free vial and adding 0.5 M trace metal grade nitric acid to 5 mL. The final volume for the calibration standards was 5 mL. 100 µL of internal standard mixture was added to each calibration standard.

Instrumentation

The instrument used for selenium analysis was an Agilent 7700 ICP-MS. The use of a collision/reaction cell with hydrogen is essential for the proper detection power and long-term signal stability of Se in the samples required for the analysis of long batches. The ⁷⁷Se and ⁷⁸Se isotopes along with the Internal Stand Isotope ¹²⁵Te were monitored with one second of integration time and three replicates in hydrogen mode. A set of typical calibration results is shown in **Protocol figure 1**.

In the sequence, a set of heated 0.5 M nitric acid blanks with internal standard were added every 60 samples, a repetition of the 1 ppb calibration standard was also included every 60 samples and the last sample to be run was a repeat of the digested quality control serum. Two non-treated certified reference materials were used to validate the accuracy of the calibration curve with a 20x dilution, Trace Metals in Drinking Water (CRM TMDW-B, High Purity Standards) and River Sediment Solution (CRM-RS-A, High Purity Standards). This is important in order to identify the source of error in the case that the digested serum CRM values were not obtained. If the non-digested samples also show error, then the calibration or instrument performance was assumed to be the problem; while if the non-digested CRMs showed good results, then the problem was assumed to be with the digestion of the samples.

Supplementary Text 3: Regression analysis weighted by inverse of sampling probabilities.

Eight cohorts included in this study used case/control sampling (Supplementary Table 1). It is known that non-random sampling could introduce bias in the estimation of the effect size on the secondary quantitative outcomes (i.e. gestational duration). To examine whether this problem can influence our analysis, we conducted regression analysis adjusted by inverse probability weights (IPW) and compared the result with the naïve analysis (without adjustment of case/control sampling). The IPW analysis corrects for selection bias by weighting the observations by the inverse of the sampling probabilities based on their case/control status. Specifically, for a case/control data set with case: control ratio (r), the relative sampling probability of controls (sampling probability of cases was set to 1) was calculated as $k/(1 - k)/r$, where k is the disease prevalence of the target population or the frequency of cases in the parental random cohort. The inverse of the sampling probabilities was then used as weights in the regression analysis to correct for the sampling bias of case/control data.

SUPPLEMENTARY TABLES

Supplementary Table 1. Study characteristics of participant cohorts

Site	Location	Study design, sample collection	Year	Data sharing format with CCHMC	GA estimation method		Type of Sample
					Ultrasound	LMP	
Bangladesh (AMANHI)	Sylhet	Population based, random	2012-2016	Case:Control (1:1)	X		Plasma
Bangladesh (GAPPS)	Matlab	Population based, random	2015-2017	Case:Control (1:2)	X		Serum
Bangladesh (MDIG)	Dhaka	Hospital based, intervention trial	2014-2015	Case:Control (1:2)	X	X	Serum
Brazil (INTERBIO)	Pelotas	Hospital based, random	2009-2014	Random	X		Plasma
Kenya (INTERBIO)	Nairobi	Hospital based, random	2009-2014	Random	X		Plasma
Malawi (iLiNS-DYAD)	Mangochi	Hospital based, intervention trial	2011-2015	Random	X		Plasma
Pakistan (AMANHI)	Karachi	Population based, random	2012-2016	Case:Control (1:2)	X		Serum
Pakistan (INTERBIO)	Karachi	Hospital based, random	2009-2014	Random	X		Plasma
South Africa (INTERBIO)	Johannesburg	Hospital based, random	2009-2014	Random	X		Plasma
Tanzania (AMANHI)	Pemba	Population based, random	2009-2016	Case:Control (1:2)	X		Plasma
Thailand (INTERBIO)	Mae Sot	Hospital based, random	2009-2014	Random	X		Plasma
UK (INTERBIO)	Oxford	Hospital based, random	2009-2014	Random	X		Plasma
UK (LIVERPOOL)	Liverpool	Hospital based, targeted recruitment	2016-2017	Random	X		Plasma
USA, CA (CPPOP)	San Francisco	Hospital based, nested case-control	2009-2010	Case:Control (1:1)	X	X	Serum
USA, NC (NEST)	Durham	Hospital based, random	2005-2009	Case:Control (1:2)	X		Plasma
Vietnam (PBB)	Ho Chi Minh City	Hospital based, random	2016-2018	Case:Control (1:2)	X	X	Serum
Zambia (GAPPS)	Lusaka	Hospital based, random	2015-2017	Random	X		Serum

Supplementary Table 2. Covariates (major phenotypes) requested from sites

Variable	Required	Desired
Baseline characteristics		
Gestational Age (at time of sample collection)	x	
Maternal Age		x
Maternal race		x
Maternal ethnicity		x
Birth history		
Gravidity (# of pregnancies)		x
Parity (# of births)		x
# prior PTB		x
# prior stillbirth		x
Pre-pregnancy BMI		
Height/weight at visit	x	
Pre-pregnancy weight (if available)	x	
Exposures		
Smoking during pregnancy		x
Alcohol during pregnancy		x
Substance use during pregnancy		x
Delivery outcomes		
Delivery date		x
Gender	x	
Gestational age at delivery	x	
Birth weight	x	
Spontaneous versus indicated delivery (if available)	x	
Infant/fetus vital status (live birth)	x	
Conditions		
Chorioamnionitis		x
Hypertensive disorder		x
Preeclampsia		x
Gestational diabetes		x
Other specified conditions		x

Supplementary Table 3. Summary statistics of maternal Se concentration in different study sites

	Se concentration (ng/ml)							
site	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	N
Bangladesh (AMANHI)	48.1	91.9	104.8	108	120.5	219.5	23.4	506
Bangladesh (GAPPS)	48.1	67.6	74.5	74.8	80.3	117.2	10.6	258
Bangladesh (MDIG)	48.5	67.8	74.4	75.3	82.3	108.8	11.4	206
Brazil (INTERBIO)	33.2	50.3	57.6	58.7	64.9	146	12.6	389
Kenya (INTERBIO)	68.1	89.9	97.5	97.9	105.8	142.1	11.9	553
Malawi (iLiNS-DYAD)	26.1	61.8	78.2	83	97.5	228.7	29.5	1210
Pakistan (AMANHI)	71	93	102.4	102.8	111.6	157.6	13.8	348
Pakistan (INTERBIO)	53.7	82	88.7	89.1	95.3	187.6	12.2	516
South Africa (INTERBIO)	28.6	64.3	70.9	70.6	77.5	109.6	10.3	352
Tanzania (AMANHI)	77.7	114.1	129.8	131.4	144.9	223.1	23.8	351
Thailand (INTERBIO)	61.4	93.3	103.7	105.4	115.8	172.2	17.3	514
UK (INTERBIO)	58	82.5	90.6	90.9	97.8	157.2	12.8	648
UK (Liverpool)	40	69.5	79.3	80	89	177.7	15.8	525
USA, California (CPOP)	74.7	115.9	125.2	125.3	134	215.5	15.3	966
USA, North Carolina (NEST)	61.7	112.3	125.2	125.6	138.7	204.4	20.6	657
Vietnam (PBB)	59.5	90.2	99.3	102.4	111	182.5	18.3	970
Zambia (GAPPS)	29.8	49.1	55.6	55.9	61.9	105.3	9.8	973
Total	26.1	72.6	92.3	93.8	112.8	228.7	28.5	9942

Supplementary Table 4. Demographic characteristics of Malawi study subjects

Site	Sample Size	Term	Preterm	Male	Female	Gday at delivery	Gday at sampling	Maternal Age (year)	Maternal Height (cm)	Birth Weight (g)
Lungwena	473	448 (95%)	25 (5%)	242 (51%)	231 (49%)	276.3 (13.2)	117.3 (13.9)	25.5 (6.4)	155.7 (5.6)	2948.8 (427.6)
Malindi	240	229 (95%)	11 (5%)	110 (46%)	130 (54%)	280.7 (14)	117.5 (15.6)	24.5 (5.9)	156.9 (5.9)	3129.3 (454.5)
Namwera	187	161 (86%)	26 (14%)	80 (43%)	107 (57%)	268.8 (15.8)	115.1 (16.2)	24.5 (6)	155.7 (5.9)	2918.3 (408.3)
Mangochi	312	288 (92%)	24 (8%)	155 (50%)	157 (50%)	276.2 (13.6)	120.3 (14.7)	25.5 (6)	156.5 (5.5)	2926.3 (474.3)
Total	1212	1126 (92.9%)	86 (7.1%)	587 (48.4%)	625 (51.6%)	276 (14.3)	117.7 (14.9)	25.2 (6.2)	156.1 (5.7)	2976.6 (449.5)

* Categorical variables are shown as count (percentage) and continuous variables are shown as mean (sd).

* Gday: gestational days. Term \geq 259 days and Preterm: gday < 259 days.

Supplementary Table 5. Summary statistics of Se concentration in Malawi study subjects

	Se concentration (ng/ml)							
Site	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	N
Lungwena	32.7	63.3	82.1	88.1	105.3	228.7	32.4	472
Malindi	46.6	76	91.3	94.8	109.3	174.4	24.6	239
Namwera	26.1	46.1	56.2	56.8	67.6	94.7	14.4	187
Mangochi	31.5	65.2	78	82	93.3	195.3	25.4	312
Total	26.1	61.8	78.2	83	97.5	228.7	29.5	1210

Supplementary Table 6. Demographic characteristics of UK (Liverpool) study subjects

Group	Sample Size	Term	Preterm	Male	Female	GA at delivery	GA at sampling	Maternal Age	Maternal Height	Birth Weight
Low-risk	253	249 (98%)	4 (2%)	125 (49%)	128 (51%)	276.4 (11.5)	141.2 (10.1)	30.9 (4.6)	165.5 (6.1)	3472.6 (495.2)
High-risk	272	175 (64%)	97 (36%)	146 (54%)	126 (46%)	258.3 (25)	140.5 (8.8)	30.3 (5.1)	164.2 (6.5)	2830.8 (778)
Total	525	424 (80.8%)	101 (19.2%)	271 (51.6%)	254 (48.4%)	267 (21.7)	140.8 (9.5)	30.6 (4.9)	164.8 (6.3)	3141.2 (730.3)

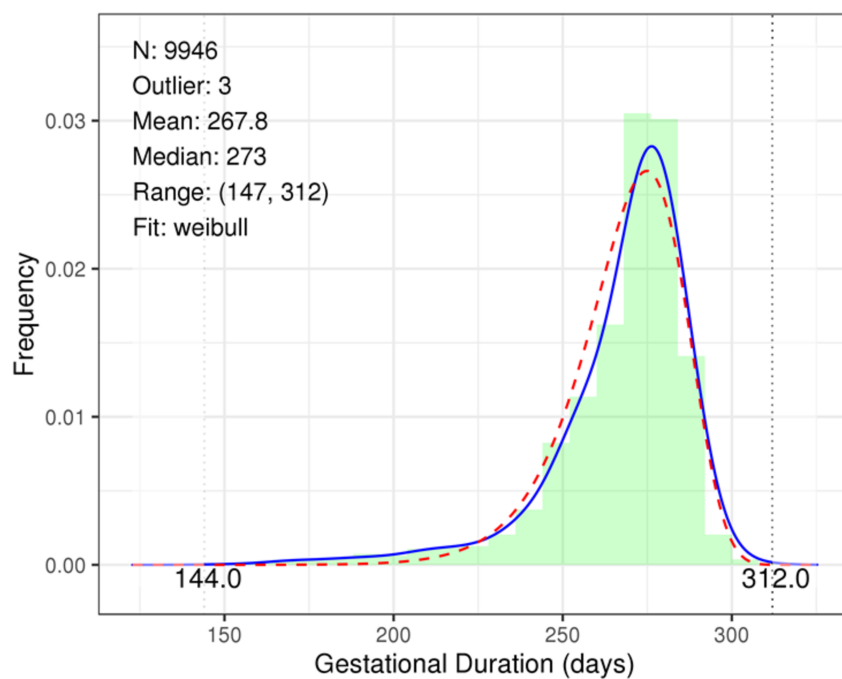
* Categorical variables are shown as count (percentage) and continuous variables are shown as mean (sd).

* Gday: gestational days. Term \geq 259 days and Preterm: gday < 259 days.

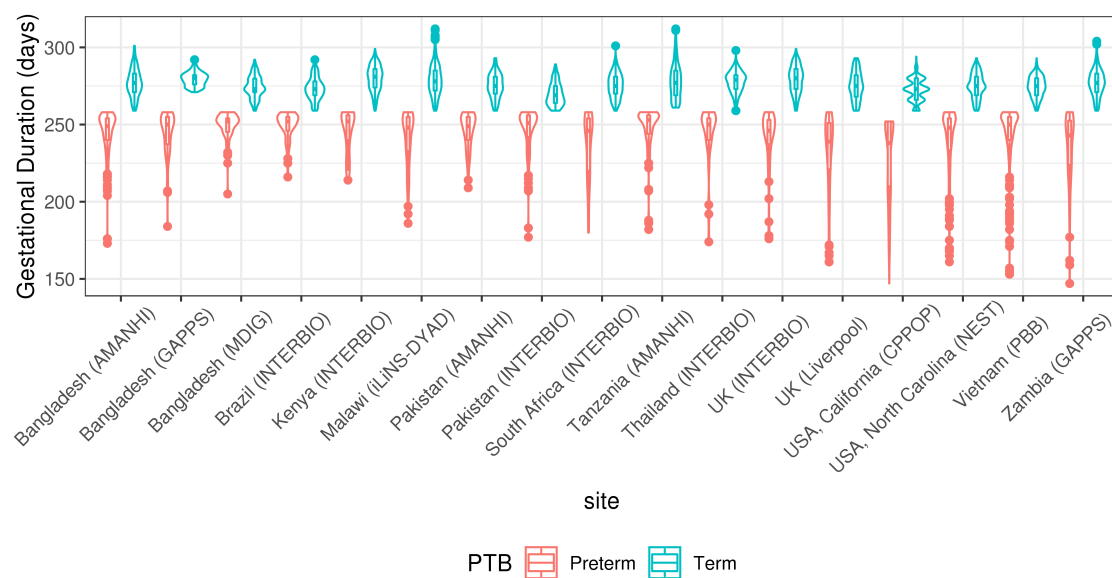
Supplementary Table 7. Summary statistics of Se concentration in UK (Liverpool) study subjects

	Se concentration (ng/ml)							
Site	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	N
Low-risk	51.0	71.3	82.2	82.7	92.3	133.1	15.6	253
High-risk	40.0	67.7	76.6	77.6	85.6	177.7	15.6	272
Total	40.0	69.5	79.3	80.0	89.0	177.7	15.8	525

SUPPLEMENTARY FIGURES

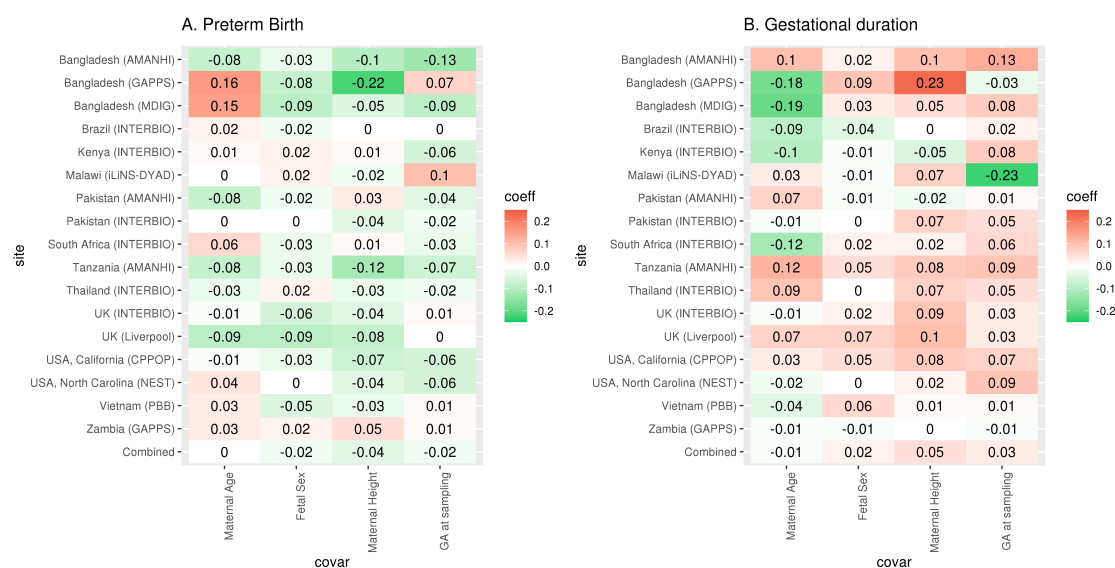


Supplementary Figure 1: Distribution of gestational duration of singleton live births with spontaneous onset of labor from all sites



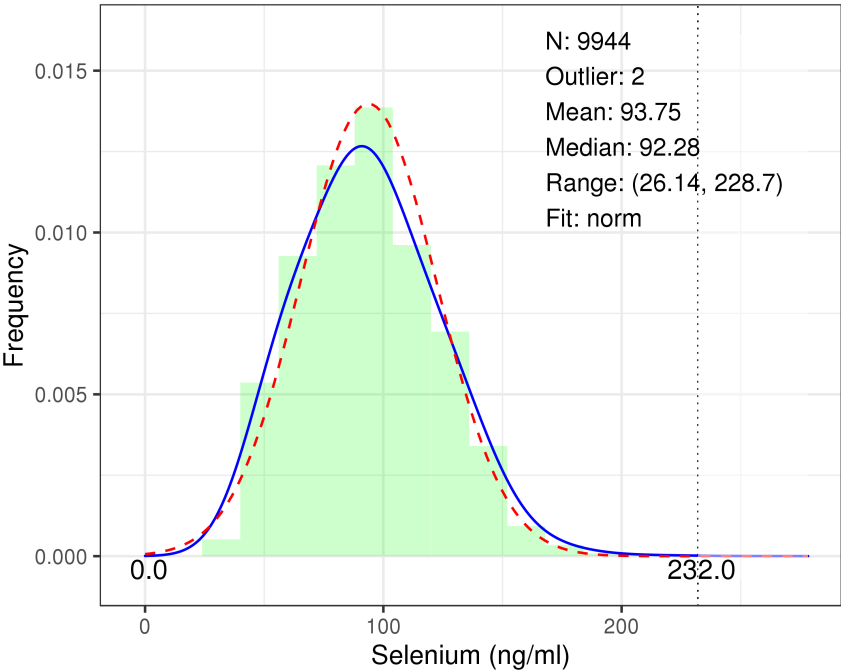
Supplementary Figure 2. Gestational duration in term and preterm deliveries by participating sites

Violin plot illustrating the distributions of gestational days in term (gday ≥ 259 days) and preterm (gday < 259 days) deliveries from each site.

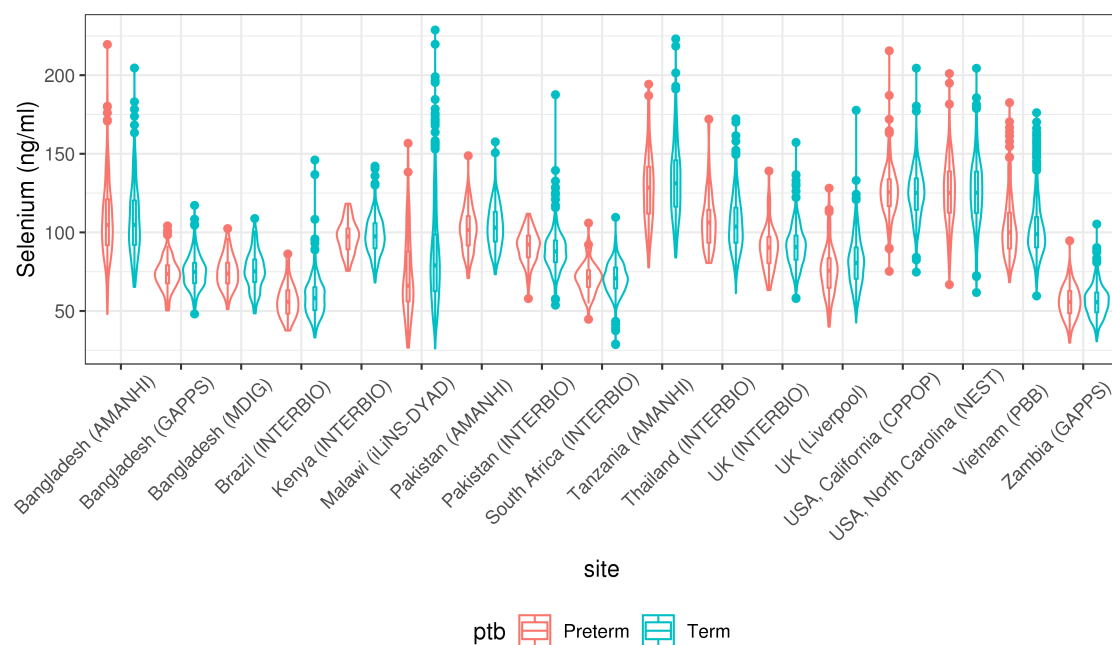


Supplementary Figure 3. Correlation of gestational duration and preterm birth with other covariates

Heat maps illustrating the correlation of preterm birth (A) and gestational duration (B) with pregnancy covariates. Red shading indicates a positive correlation and green a negative correlation, with intensity reflecting the magnitude of the correlation.

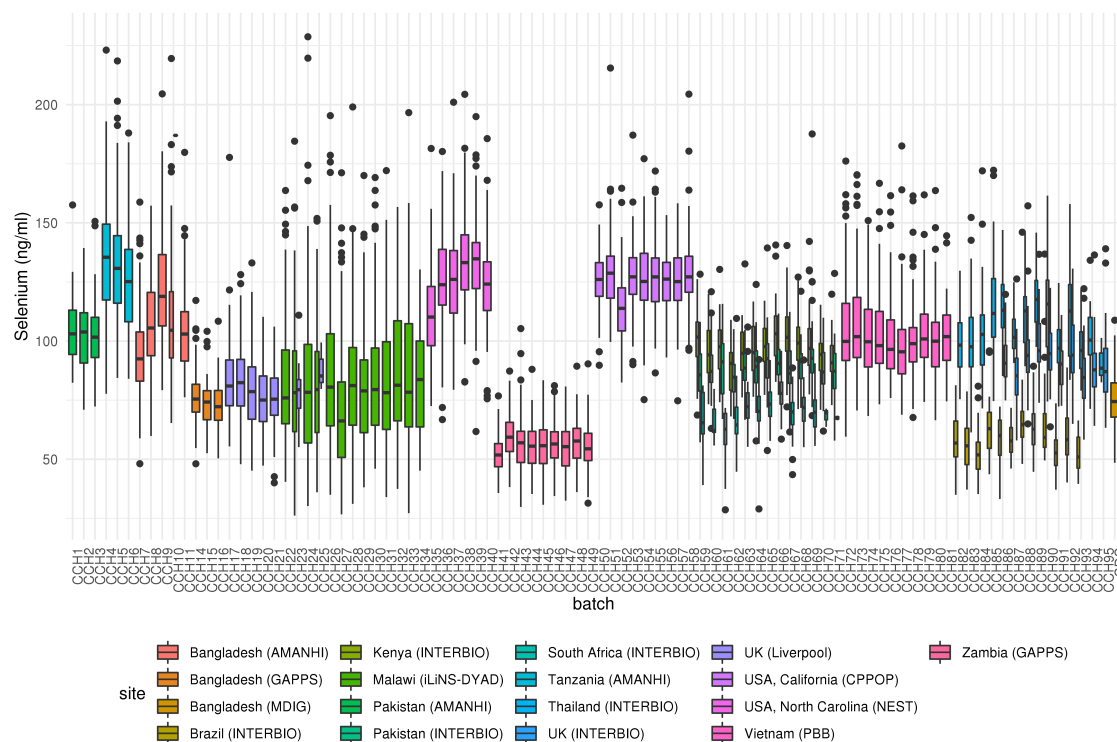


Supplementary Figure 4. Distribution of maternal Se concentration of samples from all sites

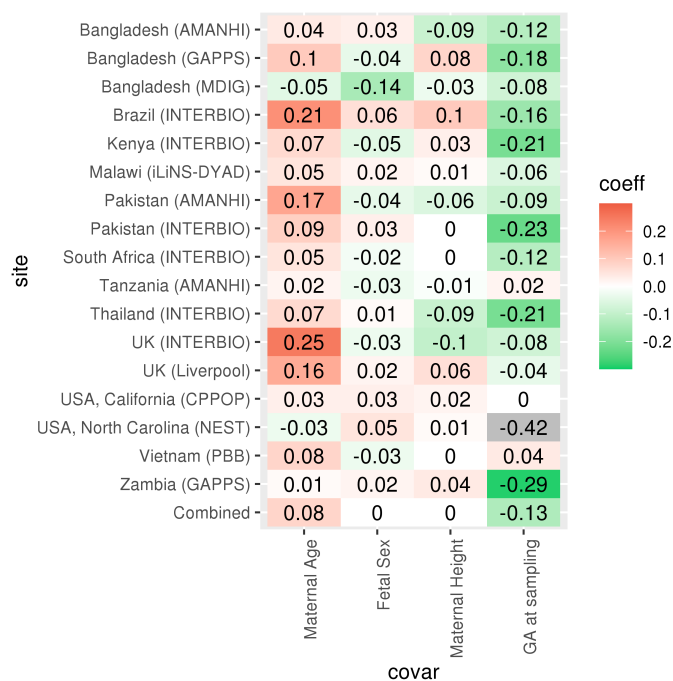


Supplementary Figure 5. Maternal Se concentration in term and preterm deliveries by participating sites

Violin plot illustrating the distributions of Se concentration in term (gday \geq 259 days) and preterm (gday < 259 days) deliveries from each site.

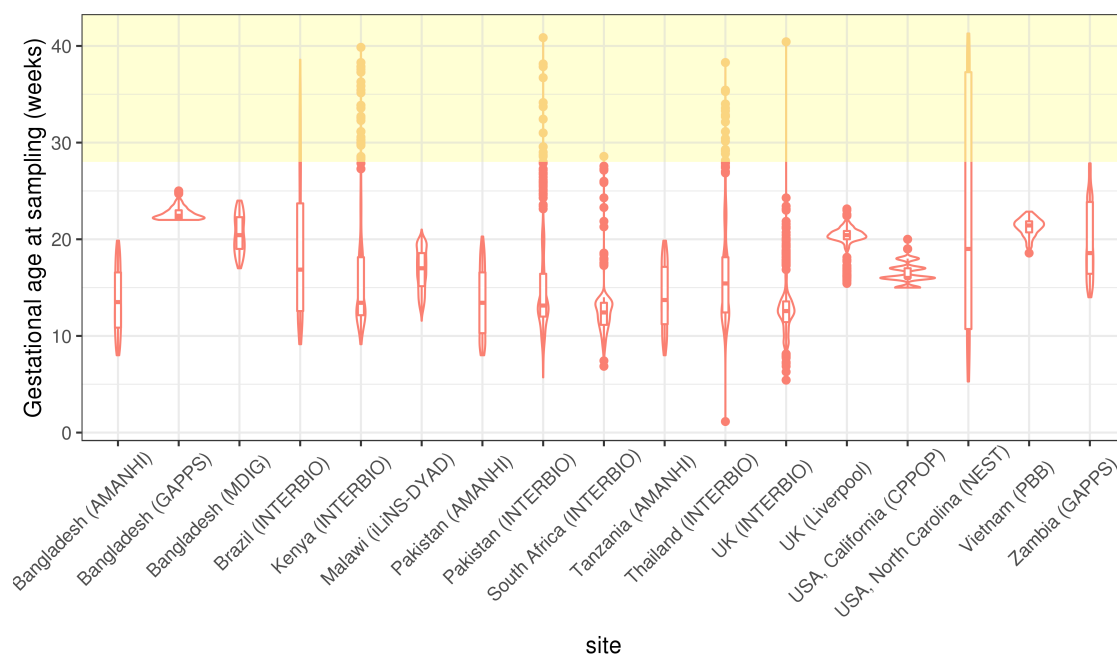


Supplementary Figure 6. Maternal Se concentration measured at different batches (colored by site)



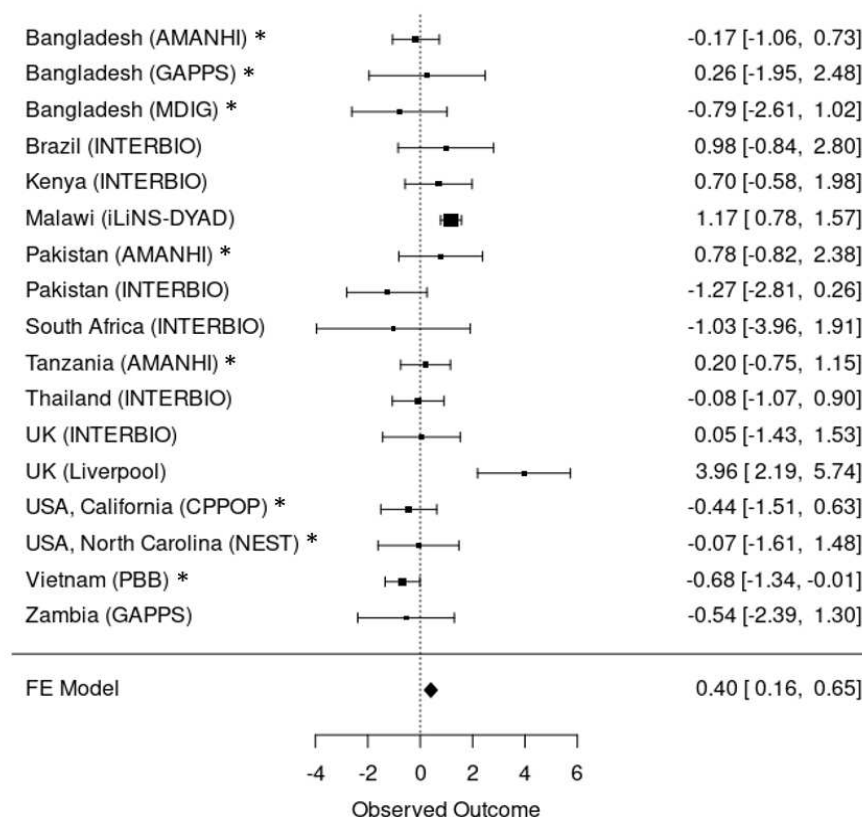
Supplementary Figure 7. Correlation of maternal Se concentration with other covariates

Heat maps illustrating the correlation of maternal Se concentration with other covariates. Red shading indicates a positive correlation and green a negative correlation, with intensity reflecting the magnitude of the correlation.



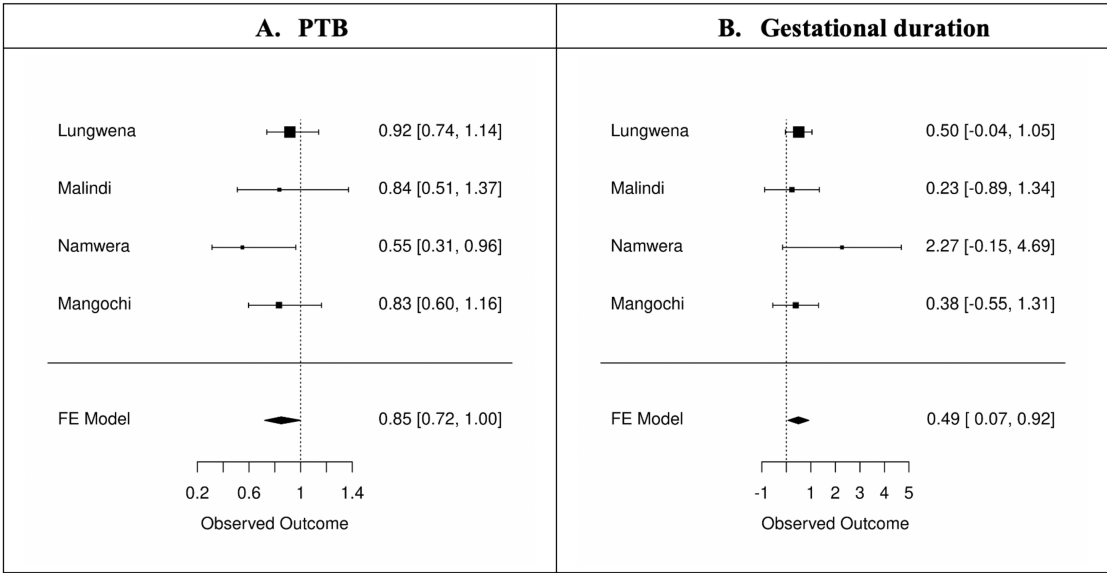
Supplementary Figure 8. Gestational age (weeks) at sample collection by sites

Violin plot illustrating the distribution of gestational age (weeks) at sample collection by sites. Yellow shaded region represents the samples collect after 2nd trimester (≥ 28 wks), which were excluded from the final association analysis.



Supplementary Figure 9. Meta-analysis of the association of maternal Se concentration with gestational duration with adjustment of case/control sampling with IPW analysis.

The sites labeled with * were case/control data sets and the sampling bias were corrected using regression analysis weighted by inverse of sampling probabilities (Supplementary Text 3).



Supplementary Figure 10. Meta-analysis of the association of Selenium concentration (unit: 15 ng/ml) with PTB (A) and gestational duration (B) among the 4 Malawi sites

(A) The estimated effect on PTB is shown as odds ratio per 15 ng/ml increase in Se concentration. (B) The estimated effect on gestational duration is shown as change in gestational days per 15 ng/ml increase in Se concentration.

SUPPLEMENTARY REFERENCES

1. Alliance for M, Newborn Health Improvement mortality study g. Population-based rates, timing, and causes of maternal deaths, stillbirths, and neonatal deaths in south Asia and sub-Saharan Africa: a multi-country prospective cohort study. *Lancet Glob Health*. 2018;6(12):e1297-e308.
2. Roth DE, Morris SK, Zlotkin S, Gernand AD, Ahmed T, Shanta SS, et al. Vitamin D Supplementation in Pregnancy and Lactation and Infant Growth. *N Engl J Med*. 2018;379(6):535-46.
3. Kennedy SH, Victora CG, Craik R, Ash S, Barros FC, Barsosio HC, et al. Deep clinical and biological phenotyping of the preterm birth and small for gestational age syndromes: The INTERBIO-21 (st) Newborn Case-Control Study protocol. *Gates Open Res*. 2018;2:49.
4. Kamng'ona AW, Young R, Arnold CD, Patson N, Jorgensen JM, Kortekangas E, et al. Provision of Lipid-Based Nutrient Supplements to Mothers During Pregnancy and 6 Months Postpartum and to Their Infants from 6 to 18 Months Promotes Infant Gut Microbiota Diversity at 18 Months of Age but Not Microbiota Maturation in a Rural Malawian Setting: Secondary Outcomes of a Randomized Trial. *J Nutr*. 2020;150(4):918-28.
5. Jelliffe-Pawlowski LL, Rand L, Bedell B, Baer RJ, Oltman SP, Norton ME, et al. Prediction of preterm birth with and without preeclampsia using mid-pregnancy immune and growth-related molecular factors and maternal characteristics. *J Perinatol*. 2018;38(8):963-72.
6. Martin CL, Jima D, Sharp GC, McCullough LE, Park SS, Gowdy KM, et al. Maternal pre-pregnancy obesity, offspring cord blood DNA methylation, and offspring cardiometabolic health in early childhood: an epigenome-wide association study. *Epigenetics*. 2019;14(4):325-40.