BMJ Open

Cohort profile: Molecular Signature in Pregnancy (MSP) – longitudinal high-frequency sampling to characterise cross-omic trajectories in pregnancy in a resource-constrained setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-041631
Article Type:	Cohort profile
Date Submitted by the Author:	13-Jun-2020
Complete List of Authors:	Brummaier, Tobias; Shoklo Malaria Research Unit, MCH Department; Schweizerisches Tropen- und Public Health-Institut, Department of Medicine Syed Ahamed Kabeer, Basirudeen; Sidra Medical and Research Center, Systems Biology and Immunology Wilaisrisak, Pornpimon ; Shoklo Malaria Research Unit Pimanpanarak, Mupawjay; Shoklo Malaria Research Unit, MCH Department Win, Aye Kyi; Shoklo Malaria Research Unit, MCH Department Pukrittayakamee, Sasithon; Mahidol University Faculty of Tropical Medicine Marr, Alexandra; Sidra Medical and Research Center Kino, Tomoshige; Sidra Medical and Research Center Al Khodor, Souhaila ; Sidra Medical and Research Center Terranegra, Annalisa; Sidra Medical and Research Center, Carrara, Verena; Shoklo Malaria Research Unit, MCH; University of Oxford Centre for Tropical Medicine and Global Health Nosten, Francois; Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Utzinger, Juerg; Schweizerisches Tropen- und Public Health-Institut, Chaussabel, Damien; Sidra Medical and Research Center Paris, Daniel; Schweizerisches Tropen- und Public Health-Institut, Department of Medicine; Universitat Basel Medizinische Fakultat McGready, Rose; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford
Keywords:	CLINICAL PHYSIOLOGY, EPIDEMIOLOGY, Reproductive medicine < GYNAECOLOGY, Fetal medicine < OBSTETRICS, Maternal medicine < OBSTETRICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Cohort profile: Molecular Signature in Pregnancy (MSP) – longitudinal high-frequency sampling to characterise cross-omic trajectories in pregnancy in a resource-constrained setting

AUTHORS

Tobias Brummaier[1,2,3,4*]
Basirudeen Syed Ahamed Kabeer[5]
Pornpimon Wilaisrisak[1]
Mupawjay Pimanpanarak[1]
Aye Kyi Win[1]
Sasithon Pukrittayakamee[6]
Alexandra K. Marr[5]
Tomoshige Kino[5]
Souhaila Al Khodor[5]
Annalisa Terranegra[5]
Verena I. Carrara[1,2]
François Nosten[1,2]
Jürg Utzinger[3,4]
Damien Chaussabel[5]
Daniel H. Paris[3,4]
Rose McGready[1,2]

1. Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand
2. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
3. Swiss Tropical and Public Health Institute, Basel, Switzerland
4. University of Basel, Basel, Switzerland
5. Sidra Medicine, Doha, Qatar
6. Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

* Corresponding author
Shoklo Malaria Research Unit
P.O. Box 46
68/30 Bann Tung Road
Mae Sot 63110
Tak Province
Thailand
Email: tobias.brummaier@gmx.at

Word count: 3,558

ABSTRACT

Purpose A successful pregnancy relies on the interplay of various biological systems. Deviations from the norm within a system or *inter-systemic* interactions may result in pregnancy associated complications and adverse pregnancy outcomes. Systems biology approaches provide an avenue of unbiased, in-depth phenotyping in health and disease. The Molecular Signature in Pregnancy (MSP) cohort was established to characterise longitudinal, cross-omic trajectories in pregnant women from a resource constrained setting. Downstream analysis will focus on characterising physiological perturbations in uneventful pregnancies, pregnancy associated complications and adverse outcomes.

Participants First trimester pregnant women of Karen or Burman ethnicity were followed prospectively throughout pregnancy, at delivery and until 3 months postpartum. Serial high frequency sampling to assess whole blood transcriptomics and microbiome composition of the gut, vagina and oral cavity, in conjunction with assessment of gene expression and microbial colonisation of gestational tissue, was done for all cohort participants.

Findings to date 381 women with live born singletons averaged 16 (IQR 15-18) antenatal visits (13,094 biological samples were collected). At 5% (19/381) the preterm birth rate was low. Other adverse events such as maternal febrile illness 7.1% (27/381), gestational diabetes 13.1% (50/381), maternal anaemia 16.3% (62/381), maternal underweight 19.2% (73/381) and a neonate born small for gestational age 20.2% (77/381) were more often observed than preterm birth.

Future plans Results from the MSP cohort will enable in-depth characterisation of cross-omic molecular trajectories in pregnancies from a population in a resource-constrained setting. Moreover, pregnancy associated complications and unfavourable pregnancy outcomes will be investigated at the same granular level, with a particular focus on population relevant needs such as effect of tropical infections on pregnancy. More detailed knowledge on multi-omic perturbations will ideally result in development of diagnostic tools and ultimately lead to targeted interventions that may disproportionally benefit pregnant women from this resource-limited population.

Registration This trial is registered under identifier NCT02797327.

KEYWORDS

Gene expression, microbiome, molecular profiling, pregnancy, resource-constrained setting.

STRENGTHS AND LIMITATIONS

- The major strength is the prospective nature of the study and frequent follow-up, coupled with high-frequency sampling and thus, availability of detailed clinical information and a considerable number of biological samples.
- High-throughput analysis, in combination with clinical data, will enable investigation of a number of pregnancy-related physiological and pathological changes.
- Populations from low-resource settings are frequently disproportionally burdened by adverse birth outcomes that may be based on exposure to different communicable diseases; hence, including them in high-end clinical research, addresses a significant research gap and may result in improvements of limited relevance to high-income countries.
- Low numbers for some phenotypes (e.g. preterm birth) may prove to be detrimental for the validity of observed differences; albeit the power will depend on the magnitude of observed differences in molecular signatures.
- In low-resource settings complete biological sample sets are often difficult to obtain, which may downsize the richness of the data.

INTRODUCTION

A successful pregnancy relies on well-timed adaptations and the interplay of multiple maternal biological systems. These interactions and temporal changes affect various organ systems, such as the cardiovascular, respiratory, endocrine system or metabolic systems, and, more recently, pivotal immunologic adaptions and changes in the human microbiome became evident.[1–5] Physiologic adaptions of the immune system during pregnancy play a central role in implantation and placentation, promotion of fetal growth and initiation of labour and delivery.[2] Deviations from the norm of this fine-tuned immune clock may lead to dysregulation in biological networks and cause various pregnancy associated complications with their immediate consequences for the mother and fetus.[6]

With a growing body of evidence of the human microbiome's significance in health and disease, investigating its role in reproductive medicine has opened up another avenue to a

deeper understanding of adverse pregnancy outcomes.[7] The placenta represents the fetomaternal link. It's crucial role during pregnancy is underlined by the fact that failure in its development and function is associated with a variety of pregnancy associated complications (e.g. preeclampsia, intrauterine growth restriction and preterm birth).[8]

In recent years investigation of these biological systems has greatly improved our understanding of their role in healthy pregnancies and in pregnancies resulting in unfavourable outcomes. However, there is a need for longitudinal, multi-omics profiling studies that may further contribute to the understanding of physiological adaptions, the interconnections of various biological systems and their significance in pregnancy associated complications.[3]

Research, in particular high-end research, is often biased towards high-income countries.[9] This imbalance is aggravated by the fact that, due to population based differences, results are often not generalisable [10], populations living in resource-constrained settings are more affected by the consequences of pregnancy associated complications (e.g. inadequate access to safe iatrogenic birth or lack of advanced neonatal care) and have different epidemiological patterns of communicable diseases. Consequently, pregnant women in low-income settings would benefit disproportionately from early identification of pregnancy associated complications and targeted interventions.

Thus, the aim of establishing this prospective Molecular Signature in Pregnancy (MSP) cohort was to characterise cross-omic trajectories in pregnant women from a resource-limited setting and describe pregnancy associated complications (e.g. preterm birth (PTB), gestational diabetes (GDM), anaemia, underweight and born too small for gestational age (SGA)). The term “molecular signature” was chosen, as it refers to molecular markers that can be used for in-depth description of a particular phenotype.[11]

Perturbations related to the immune response will be investigated by measuring the abundance of RNAs in circulating nucleated cells (i.e. leucocytes) via capillary blood sampling at multiple timepoints.[12] Microbiome profiling of the intestinal and vaginal niche in pregnancy, at delivery and postpartum will be complemented by assessment of the oral

microbiome. Lastly, placental tissue will be investigated for the presence of bacterial colonisation as well as for the relationship between the genome structure and changes in global patterns of gene expression. The generated molecular data will then be analysed in the context of clinical and patient related data. As described in the study protocol [13], an immediate focus was the investigation of PTB.

The MSP cohort profile represents the link between development of the study protocol and the results. Accordingly, cohort characteristics, including the recruitment and follow up process, demographics, pregnancy outcomes, biological samples available for analysis and preliminary results are presented.

COHORT DESCRIPTION

Setting and participants

Women with an unremarkable medical and obstetric history attending the antenatal care (ANC) facilities of Shoklo Malaria Research Unit (SMRU), a field station of the Faculty of Tropical Medicine at Mahidol University (Bangkok, Thailand) that is part of the Mahidol-Oxford Research Unit, were invited to participate. SMRU operates free-of-charge walk-in services for marginalized migrant populations in clinics at the frontier of Thailand with Myanmar (Figure 1).

Between September 2016 and July 2018, trained counsellors from the community provided general information about the trial to all pregnant women who registered at SMRU ANC. Detailed information, including visual aids, were utilized to illustrate the purpose, procedures and timing of follow-up. Potential risks and benefits of the study were explained to women meeting the following inclusion criteria: (i) pregnant woman willing and able to give written informed consent; (ii) Karen or Burman ethnicity; (iii) age 18-49 years; (iv) healthy, with viable singleton first trimester (8+0 to <14+0 weeks) pregnancy; (v) planning to deliver at the SMRU birth unit; and (vi) able and willing to comply with study requirements. If these criteria were met and appropriate consent was provided, women were followed from the first trimester, throughout pregnancy, at delivery and until 3 months postpartum.

It was estimated that approximately 400 women could be enrolled to build the cohort from this low resource environment when considering constraints of clinical capacity and budget availability. With this sample size, events such as PTB (estimated at 8%, i.e. 30 cases expected) would be feasible but also provide scope to engage with other pregnancy associated complications, tropical infections (e.g. small liver fluke infections estimated at 5%, i.e. 20 cases expected) and febrile illness (estimated at 10%, i.e. an expectation of 40 cases) with a sufficient number of uneventful pregnancies remaining to act as controls.

Figure 2 provides an overview of the number of participants recruited and followed. During the active recruitment period, 4,292 pregnant women registered to ANC. Seventy percent (2,544/3,609) were ineligible due to late presentation to ANC. There were two local conditions that limited recruitment. The first issue pertained to access; pregnant women following SMRU ANC come from villages scattered along the Thailand-Myanmar border (Figure 1) with transportation constraints affecting their ability to attend, further aggravated in the monsoon season as roads become impassable. The second issue related to staffing. As only one ultrasound machine was available at each site, there was an upper limit to the number of women who could have study related fetal growth scans done on a single day. Challenges in access and staffing issues led to the non-consideration of 178 and 475 women, respectively.

Overall, 683 women received full information about the purpose and procedures of the study. Of these, 430 agreed to participate (response rate: 63%) and consented by providing a signature or thumbprint if the event of an illiterate participant. Thirty participants were recruited in the cohort to replace women who were lost before delivery, while the study was still open to recruitment. Two participants were excluded as the fetal heartbeat (FHB) was absent after they had already agreed to the study (Figure 2).

Routine ANC procedures at SMRU are described in the study protocol.[13] Demographic characteristics of 381 enrolled women with a live birth outcome are summarised (Table 1). To demonstrate the scope of the cohort, selected subgroups are summarised alongside the characteristics for all 381 women.

Table 1. Basic characteristics of women with a live birth outcome (n=381) and subgroups of selected pregnancy associated complications in the Molecular Signature in Pregnancy (MSP) cohort.

	Overall (n=381)	Preterm (n=19)	GDM (n=50)	Anaemia (n=62)	Underweight (n=73)	SGA (n=77)
Inclusion EGA (days); median (range)	82 (56 - 97)	83 (60 - 94)	82.5 (56 - 97)	81 (60 - 94)	81 (57 - 95)	81 (62 - 97)
Maternal age groups (years)						
• 18-24	185 (48.6)	14 (73.7)	25 (50.0)	28 (45.2)	43 (58.9)	38 (49.4)
• 25-29	95 (24.9)	2 (10.5)	12 (24.0)	14 (22.6)	18 (24.7)	18 (23.4)
• 30-34	64 (16.8)	3 (15.8)	8 (16.0)	12 (19.4)	10 (13.7)	13 (16.9)
• 35-39	25 (6.6)	0	4 (8.0)	5 (8.1)	2 (2.7)	6 (7.8)
• 40-49	12 (3.1)	0	1 (2.0)	3 (4.8)	0	2 (2.6)
Residence						
• Myanmar	263 (69.0)	10 (52.6)	33 (66.0)	42 (67.7)	43 (58.9)	52 (67.5)
• Thailand	118 (31.0)	9 (47.4)	17 (34.0)	20 (32.3)	18 (24.7)	25 (32.5)
Gravidity						
• Primigravida	103 (27.0)	8 (42.1)	12 (24.0)	17 (27.4)	32 (43.8)	26 (33.8)
• 2	111 (29.1)	8 (42.1)	16 (32.0)	21 (33.9)	21 (28.8)	18 (23.4)
• 3	79 (20.7)	2 (10.5)	12 (24.0)	14 (22.6)	9 (12.3)	13 (16.9)
• 4	47 (12.3)	1 (5.3)	7 (14.0)	4 (6.5)	8 (11.0)	11 (14.3)
• ≥5	41 (10.8)	0	3 (6.0)	6 (9.7)	3 (4.1)	9 (11.7)
Parity						
• 0 (nullipara)	127 (33.3)	9 (47.4)	16 (32.0)	23 (37.1)	37 (50.7)	32 (41.6)
• 1	117 (30.7)	8 (42.1)	15 (30.0)	18 (29.0)	18 (24.7)	16 (20.8)
• 2	70 (18.4)	1 (5.3)	11 (22.0)	14 (22.6)	11 (15.1)	13 (16.9)
• 3	39 (10.2)	1 (5.3)	7 (14.0)	2 (3.2)	4 (5.5)	11 (14.3)
• ≥4	28 (7.3)	0	1 (2.0)	5 (8.1)	3 (4.1)	5 (6.5)
Literate	244 (64.0)	13 (68.4)	29 (58.0)	43 (69.4)	53 (72.6)	51 (66.2)

Smoking during in pregnancy	29 (7.6)	0	6 (12.0)	5 (8.1)	4 (5.5)	10 (13.0)
Obstetric history*						
• Miscarriage	94 (33.8)	2 (18.2)	11 (28.9)	15 (33.3)	13 (31.7)	19 (37.3)
• Stillbirth	4 (1.4)	0	0	1 (2.2)	0	0
• PIH	2 (0.7)	0	0	1 (2.2)	0	0
• Preterm rupture of membranes	8 (2.9)	2 (18.2)	0	3 (6.7)	1 (2.4)	1 (2.0)
• Preterm labour	23 (8.3)	6 (54.5)	0	9 (20.0)	3 (7.3)	3 (5.9)
• Vacuum delivery	3 (1.1)	0	0	1 (2.2)	1 (2.4)	0
• Postpartum haemorrhage	9 (3.2)	0	0	3 (6.7)	2 (4.9)	2 (3.9)
• Retained placenta	1 (0.4)	0	0	0	0	0
• Previous neonatal death	9 (3.2)	3 (27.3)	0	3 (6.7)	0	1 (2.0)
• History of GDM	2 (0.7)	0	2 (5.3)	0	0	0
• Macrosomia	2 (0.7)	0	1 (2.6)	0	0	0

Basic characteristics of pregnant women who completed per protocol follow up with a live birth outcome, compared between various pregnancy associated complications and adverse pregnancy outcomes. One study participant may be represented in multiple subgroups.
Data presented as proportion n (%) or median (range).
* Proportions calculated based on multigravida (n=278) as denominator.
Abbreviations: EGA, estimated gestational age; GDM, gestational diabetes; PIH, pregnancy induced hypertension; SGA, born too small for gestational age.

Participant follow-up

Women were followed every two weeks during pregnancy, at delivery and monthly in the first 3 months after delivery. Trial participants were encouraged to report to the study clinic if they experienced an adverse event, with special consideration for febrile episodes to maximize information on the nature of the experienced morbidity and its effect on the pregnancy.[13] Overall, 6,443 study visits during pregnancy were recorded, of which 6,140 were from women with a live birth outcome. The median number of ANC visits for women with a live birth was 16 (interquartile range [IQR] 15-18). In the uncommon event of cohort participants attending other health facilities, diagnosis and treatment was recorded at the next visit. All participants were encouraged to deliver at SMRU under the care of a skilled birth attendant. Women from the cohort achieved a low rate of home birth 6.6% (25/381) in contrast to the estimated 7 in 10 births that take place at home in rural Myanmar.[14]

As an estimator for adherence to study protocol, the number of expected and actual samples provided by women with a known pregnancy outcome were calculated. No significant difference was noted when comparing the proportion of expected and actual samples between capillary blood, stool and vaginal swab samples (p=0.070, p=0.767 and p=0.919, respectively). Missed follow-ups were the main reason for less than expected samples being provided. Additional samples were provided from participants who experienced a febrile illness, and on a few occasions, samples were erroneously taken as they were not required by the study protocol.

Early termination from the study occurred in 11.0% (47/428) of pregnancies, most of whom left ANC before birth (70.2% [33/47]) due to migration. Adverse pregnancy outcomes were recorded in 25.5% (12/47) cases and two (4.3%) participants withdrew consent, as they no longer wanted to provide vaginal swab samples (Figure 2).

Data collection

Data covering three major aspects were collected, namely (i) clinical data; (ii) genome-wide transcript abundance; and (iii) analysis of microbial composition in various anatomical sites. Table 2 summarises all study relevant procedures.

Table 2. Study procedures and timing of the Molecular Signature in Pregnancy (MSP) cohort [13]

	Screening	Baseline	Follow-up	Unwell episode	Birth	Post-partum
Viable singletons pregnancy	X					
Obstetric ultrasound*	X	X	X			
Eligibility assessment	X					
Written informed consent		X				
Demographics		X				
Medical and obstetric history		X				
Concomitant medications		X	X	X	X	X
Physical examination		X	X	X	X	X
Universal pregnancy screening, for example, thick and thin blood film for malaria diagnosis, CBC and OGTT†		X				(X)†
Sample maternal 100 µL capillary blood#		X	X‡	X§	X¶	X**
Sample vaginal swab, stool specimen and 24-hour food recall		X††	X††	X	X	X‡‡
Acceptability survey		X				X
Sample saliva			X§§		X	
Sample placenta, cord blood and maternal venous blood					X	

* Fetal growth scans on a 6-weekly basis.
# 50 µL for whole blood transcriptome analysis and 50 µL for haematocrit.
† OGTT at 24-26 weeks of gestation; repeated at 12 weeks postpartum if positive during pregnancy.
‡ 2-weekly; if the woman attended all expected 15 visits total blood is 1.5 mL.
§ If the woman attended for an unwell visit, an additional 100 µL of blood were drawn.
¶ If delivery at SMRU clinic, then an additional 100 µL of blood were drawn.
** At 1, 2 and 3 months postpartum, including maternal haematocrit.
†† In each trimester of pregnancy: 8-14, 20-22 and 34-35 weeks.
‡‡ Vaginal swab samples at 4-6 weeks and at 3 months.
§§ At 24-26 weeks of gestation.
CBC, complete blood count; OGTT, oral glucose tolerance test.

Routine ANC procedures included physical and obstetric examination, recording of concomitant medications and 6-weekly ultrasound scans to monitor fetal growth. Clinical data were collected, so that trajectories of maternal and fetal physical change would be available by the study endpoint. Two-weekly gestational weight, fetal growth and haematocrit were done as malnutrition and anaemia (and formerly malaria) are prevalent pregnancy-associated morbidities.[15–17]

Ultrasound scans to monitor fetal growth were performed 6-weekly. Throughout the study period, 2,850 scans were done; 2,693 of these were scans in women with a known, viable birth outcome (157 scans accounted for other pregnancy outcomes).

Additional details describing the study conduct and routine ANC procedures are available elsewhere.[13] Briefly, capillary blood, faecal and vaginal swab samples were taken during pregnancy, at delivery and in the post-partum period. In the event of a febrile episode, an additional set of these samples were taken together with a standardised fever screening battery. Oral microbiome was assessed in the second trimester and at delivery. Placental tissue samples were taken at delivery along with umbilical cord blood serum and maternal serum.

As SMRU birthing units do not have capacity for caesarean sections (CS), participants were referred to the closest public hospitals when indicated, and hence, delivery samples were not available, in common with births at home or in other health facilities. Even though gestational tissue samples were not available for these participants, outcome measures and neonatal anthropometry were mostly available.

Lastly, in an effort to understand the acceptance of the high frequency sampling in combination with the dense follow-up schedule and the perception of invasiveness of sampling procedures, an acceptability survey at study enrolment and completion was conducted.

Biological samples

Standard operating procedures (SOPs) focused on detailed description of the sample collection procedures were drafted before study commencement and are available on request. To accommodate the high number of samples, a unique barcoded sample sticker was placed on each tube, containing the participants study code that was assigned at enrolment, as well as date and time when samples were collected. All samples were transferred from the study sites to SMRU's central laboratory daily and stored according to SOPs. Freezerworks™ (Dataworks Development, Inc., Mountlake Terrace, WA, USA), a biorepository software was used for sample management and tracking.

Including women lost to follow-up and pregnancies that resulted in a miscarriage or stillbirth, 13,536 biological samples were taken between study enrolment and 3-month

postpartum. Considering multiple aliquots for vaginal swab samples (4 aliquots at each sampling time-point), saliva samples (n=4) and placenta samples (n=12), the total number of samples available for testing is 25,816.

Table 3 provides an overview of biological samples available for analysis and Figure 3 depicts the distribution of capillary blood, faecal and vaginal swab samples in relation to the estimated gestation of pregnancy.

All biological samples will be analysed and interpreted in the context of clinical data. To investigate physiological trajectories and deviations from these in the event of pregnancy associated complications, a nested case-control approach will be applied, and potential confounding factors will be adjusted for. Results will be published in original research articles alongside more detailed background information pertaining to the respective content knowledge, precise methods of sampling procedures and analytical steps of respective laboratory-based methods to enable the scientific community to draw their conclusions on appropriateness. The MSP cohort profile and the previously published study protocol [13] will serve as a cross-reference for all results.

Table 3. Overview of available biological samples for women with a live birth outcome (n=381) and selected subgroups of the Molecular Signature in Pregnancy (MSP) cohort.

	Overall (n= 381)	Preterm (n=19)	GDM (n=50)	Anaemia (n=62)	Underweight (n=73)	SGA (n=77)
Capillary blood						
• Pregnancy	5,263	226	672	871	1,000	1,070
• Delivery	346	19	44	57	68	70
• Postpartum	1,037	48	132	157	196	210
Stool						
• Pregnancy	1,159	52	149	189	220	233
• Delivery	299	18	43	47	58	59
• Postpartum	877	34	115	129	147	166
Vaginal swab‡						
• Pregnancy	1,175	55	151	191	223	236
• Delivery	329	19	45	55	67	65
• Postpartum	686	30	84	105	128	136
Saliva						
• Pregnancy	371	19	49	62	71	74
• Delivery	336	19	44	57	66	67
Placenta	301	16	41	50	63	63
Maternal serum	322	17	43	55	66	66
Cord blood serum	299	15	41	48	63	64
Cord blood EDTA*	294	14	41	49	59	60
Malaria tests in pregnancy	5,662	226	819	988	1,112	1,240
HCT tests in pregnancy	5,674	227	817	989	1,115	1,242
Stool microscopy in pregnancy	620	29	84	107	99	122

‡ Number refers to sampling time-points; four swabs were taken at each time-point.
* Less samples as amended to the protocol at a later stage.
Abbreviations: EDTA, ethylenediaminetetraacetic acid; GDM, gestational diabetes mellitus; HCT, haematocrit; SGA, born too small for gestational age.

The molecular signatures pertaining to biological samples will be analysed by collaborators at Sidra Medicine (Doha, Qatar). To assure appropriate sample handling, shipment is delegated to World Courier (AmerisourceBergen Corporation; Chesterbrook, PA, USA), a leading, global specialty and logistics provider. All samples are kept on dry ice and shipped in freezer boxes with thick styrofoam walls that are equipped with temperature monitors. At the time of writing, 16,916 biological samples have been successfully transferred.

Participant involvement

Study participants were not directly involved in the study design, the elaboration of research questions and outcome measures or the recruitment and conduct of the study. However, the protocol was reviewed and approved by a local community advisory board whose members are representatives of the local community and act as a bridge between researchers and the local population.[18] Their role was to advise on ethical and operational aspects of the study and they will provide a channel to inform the community of the results of this study.

FINDINGS TO DATE

Demographic data, findings during pregnancy and relevant outcome data for the overall cohort population with a live birth are presented in Tables 1 and 4, respectively. According to study objectives, data of commonly observed pregnancy associated complications (GDM, anaemia and underweight) and adverse pregnancy outcomes (PTB and SGA) are presented alongside the overall cohort.

At a rate of 5.0% (19/381), PTB (defined as delivery before 37 weeks of gestation), was considerably less than the worldwide estimate of approximately 10.6% (uncertainty interval 9.0-12.0%) of live-born infants.[19] When considering all pregnancies followed at SMRU during the study period, the observed proportion of PTB was roughly 6.8%, a number still below the worldwide average. The estimated gestational age (EGA) at delivery of PTB cases ranged from 30^{+2} to 36^{+6} (weeks^{+days}) with the majority of cases (89.5%) being moderate/late PTB (i.e. ≥32 and <37 weeks of gestation). Of the 19 PTB cases, one was iatrogenic with induction for severe preeclampsia at an EGA of 36^{+3}, while the remainder were spontaneous, without any discernible causative factor. Prioritizing healthy women at

the outset and frequent visits with health care providers might have contributed to earlier identification of health issues and timely intervention. Population based factors that might be associated with the low PTB rate are the low rate of sexually transmitted diseases, and preliminary data suggest that prevalence of *Gardnerella vaginalis*, a facultative anaerobic coccobacillus that is associated with PTB, is low in this population. This low PTB rate is unique in itself and will be thoroughly assessed by a holistic investigation of available demographic, clinical and cross-omic data.

There were six spontaneous miscarriages (median EGA 16^{+3} [IQR 16^{+1} – 17^{+2}]) and one pregnancy was terminated due to a congenital abnormality (severe hydrops fetalis) at an EGA of 20^{+1}. There were four stillbirths; all were identified as intrauterine fetal deaths, antepartum events before completion of 37 weeks of gestation. One was unexplained with an absent FHB at 28^{+1} weeks, one was associated with pre-eclampsia (absent FHB at 31^{+1}), one had a tight cord around the neck at birth (absent FHB at 35^{+3}), and in one case the mother ingested organophosphate with suicidal intent (absent FHB at 36^{+3}). One unexpected neonatal death was recorded in a term neonate born to a mother with an unremarkable pregnancy and delivery.

Seven of 381 (1.8%) women experienced *Plasmodium vivax* malaria during pregnancy and one woman was diagnosed with *P. vivax* malaria at delivery. Soil-transmitted helminth infections are common in this population and were detected in 21.3% (81/381) of pregnant women with a live birth outcome. Hookworm was detected in 15.0% (57/381) cases, *Trichuris trichiura* in 6.3% (24/381) and *Ascaris lumbricoides* in 4.2% (16/381). Fourteen women had more than one soil-transmitted helminth infection concurrently. Food-borne trematode infections (e.g. small liver flukes) were detected in 6.8% (26/381) of the participants. The immunologic fingerprint of helminthic infections and their effect on pregnancy is poorly understood. Hence, effects of helminth infections on the immune system or the intestinal microbiome in pregnancy will be studied with data from this cohort.

Generally, communicable diseases are often associated with adverse pregnancy outcomes. Fever, which is often the cardinal symptom of an infectious disease, was recorded in 7.1% (27/381) of pregnant women. Whether longitudinal, high-frequency sampling has potential

for early detection, or detection prior to the onset of symptoms, of communicable diseases in pregnancy, will be determined. Availability of data pertaining to communicable disease will also allow appropriate adjustment for investigation of other objectives (e.g. PTB).

Table 4. Pregnancy outcome and neonatal data of women with a life birth outcome (n=381) and selected subgroups in the Molecular Signature in Pregnancy (MSP) cohort.

	Overall (n= 381)	Preterm (n=19)	GDM (n=50)	Anaemia (n=62)	Underweight (n=73)	SGA (n=77)
Outcome EGA (days); median (IQR)	277 (271-282)	253 (244-255)	275 (268-279)	275 (268-283)	277 (271-281)	279 (274-283)
Preterm categories						
• Term (≥37 weeks)	362 (95.0)	0	49 (98.0)	57 (91.9)	68 (93.2)	73 (94.8)
• Moderate/late PTB (≥32)	17 (4.5)	17 (89.5)	1 (2.0)	3 (4.8)	5 (6.8)	4 (5.2)
• Very PTB (28 – <32)	2 (0.5)	2 (10.5)	0	2 (3.2)	0	0
• Extremely PTB (<28)	0	0	0	0	0	0
Infant sex (male)	185 (48.6)	11 (57.9)	26 (52.0)	18 (29.0)	33 (45.2)	25 (32.5)
Apgar after 1 min; median (IQR)	9 (9-9)	9 (9-9)	9 (9-9)	9 (9-9)	9 (9-9)	9 (9-9)
Apgar after 5 min; median (IQR)	10 (10-10)	10 (10-10)	10 (10-10)	10 (10-10)	10 (10-10)	10 (10-10)
Resuscitation						
• Yes	8 (2.1)	1 (5.3)	0	2 (3.2)	1 (1.4)	5 (6.5)
• No	359 (94.2)	17 (89.5)	49 (98.0)	59 (95.2)	71 (97.3)	70 (90.9)
• Unknown	14 (3.7)	1 (5.3)	1 (2.0)	1 (1.6)	1 (1.4)	2 (2.6)
Abnormal newborn exam[$]	5 (1.3)	1 (5.3)	0	2 (3.2)	1 (1.4)	0
Birthweight (g); median (IQR)*	2,955 (2,740-3,200)	2,273 (2,025-2,440)	3,010 (2,840-3,375)	2,870 (2,620-3,220)	2,823 (2,611-2,965)	2,600 (2,400-2,740)
Birthweight categories*						
• Small for GA	77 (21.8)	4 (22.2)	7 (14.3)	14 (23.0)	23 (33.8)	77 (100.0)
• Appropriate for GA	269 (76.2)	14 (77.8)	39 (79.6)	47 (77.0)	45 (66.2)	0
• Large for GA	7 (2.0)	0	3 (6.1)	0	0	0
Infant length (cm); median (IQR)[‡]	48.2 (47.0-49.4)	45.7 (44.1-46.6)	48.4 (47.4-49.4)	48.1 (46.5-49.0)	47.6 (46.5-48.5)	46.7 (45.6-48.0)
Length categories[‡]						
• Short for GA	70 (19.9)	2 (11.1)	7 (14.3)	14 (23.0)	23 (33.8)	44 (57.1)
• Appropriate for GA	266 (75.6)	16 (88.9)	37 (75.5)	44 (72.1)	43 (63.2)	32 (41.6)

• Tall for GA	16 (4.5)	0	5 (10.2)	3 (4.9)	2 (2.9)	0
Head circumference at birth; median (IQR)‡	32.9 (32.0-33.6)	31.0 (30.1-31.6)	33.1 (32.5-34.2)	32.4 (31.6-33.6)	32.4 (31.5-33.3)	32.0 (31.4-32.8)
Head circumference categories‡						
• Small HC for GA	106 (30.1)	4 (22.2)	8 (16.3)	21 (34.4)	30 (44.1)	48 (63.2)
• Appropriate HC for GA	240 (68.2)	14 (77.8)	37 (75.5)	39 (63.9)	37 (54.4)	28 (36.8)
• Large HC for GA	6 (1.7)	0	4 (8.2)	1 (1.6)	1 (1.5)	0
Caesarean section	22 (5.8)	0	1 (2.0)	5 (8.1)	2 (2.7)	5 (6.5)
Breech delivery	5 (1.3)	1 (5.3)	0	0	1 (1.4)	3 (3.9)
Vacuum delivery	11 (2.9)	0	1 (2.0)	4 (6.5)	3 (4.1)	3 (3.9)
Delivered by						
• Midwife	328 (86.1)	17 (89.5)	47 (94.0)	52 (83.9)	66 (90.4)	70 (90.9)
• TBA	25 (6.6)	2 (10.5)	2 (4.0)	3 (4.8)	5 (6.8)	1 (1.3)
• Doctor	24 (6.3)	0	1 (2.0)	6 (9.7)	2 (2.7)	5 (6.5)
• Other§	4 (1.0)	0	0	1 (1.6)	0	1 (1.3)
Place of delivery						
• SMRU clinic	304 (79.8)	16 (84.2)	43 (86.0)	51 (82.3)	63 (86.3)	63 (81.8)
• Home	29 (7.6)	2 (10.5)	2 (4.0)	4 (6.5)	5 (6.8)	2 (2.6)
• Thai Hospital	35 (9.2)	0	4 (8.0)	7 (11.3)	3 (4.1)	10 (13.0)
• Myanmar Hospital	4 (1.0)	0	0	0	1 (1.4)	0
• Other†	9 (2.4)	1 (5.3)	1 (2.0)	0	1 (1.4)	2 (2.6)
Induction of labour	23 (6.0)	1 (5.3)	3 (6.0)	4 (6.5)	4 (5.5)	4 (5.2)
Augmentation of labour	36 (9.4)	0	5 (10.0)	5 (8.1)	9 (12.3)	6 (7.8)
Postpartum haemorrhage	19 (5.0)	0	1 (2.0)	2 (3.2)	4 (5.5)	2 (2.6)
Estimated blood loss (mL)	150 (100-240)	100 (100-150)	115 (100-200)	150 (100-200)	140 (100-200)	138 (100-200)
Neonatal death	1 (0.3)	0	0	0	0	0

One study participant may be represented in multiple subgroups.
Data presented as proportion n (%) or median (interquartile range).

$ Spina bifida (n=2), cleft lip (n=1), Down's syndrome (n=1), congenital heart disease (n=2).
* 28 cases with missing information (e.g. home delivery).
‡ 29 cases with missing information.
§ Delivered by *other*: spontaneous self-delivery (n=2), husband (n=1), village health worker (n=1).
† Place of delivery *other*: delivery on the way to the SMRU clinics (n=6), in another clinic (n=3).
Abbreviations: EGA, estimated gestational age; GA, gestational age; GDM, gestational diabetes mellitus; HC, head circumference; IQR, interquartile range; min, minutes; PTB, preterm birth; SGA, born too small for gestational age; SMRU, Shoklo Malaria Unit; TBA, traditional birth attendant

367

Overall, only four cases (1.0%) of preeclampsia and one case (0.3%) of eclampsia were recorded which is also at the lower end of the epidemiological range (0.2-9.2).[20] A common pregnancy associated complication was GDM, which affected 13.1% (50/381) pregnancies with a live birth. GDM is the most common metabolic disorder in pregnancy and is of increasing concern in high-income settings around the globe but also in this resource-limited population.[21,22] While the proportion of pregnant women diagnosed with underweight in this population is decreasing, 19.2% (73/381) of pregnancies were still complicated by a diagnosis of underweight.[17] Maternal anaemia defined as haematocrit of less than 30% complicated 16.3% [62/381] of pregnancies and SGA (20.2% [77/381]) according to standards as published by the Intergrowth-21st consortium [23] was also commonly observed and remains a solicitude in this pregnant population. As shown in Figure 3, cohort enrolment in early pregnancy, coupled with the high-frequency sampling provides a plethora of cross-omic data. This will enable to identify signatures preceding the development of these complications, their impact on pregnancy outcome and to characterise how they plateau in the postpartum period.

Ultra-low capillary blood samples are safe and a reliable method to monitor longitudinal gene expression trajectories. Samples acquired from this MSP cohort were used to develop a modified SOP for extraction of RNA from small volume whole blood samples preserved in Tempus^{TM} solution [24] and the performance of ultra-low blood samples was benchmarked previously.[25] Figure 4 shows a comparison between different stages in pregnancy (early pregnancy and late pregnancy), and the non-pregnant control (3-month postpartum) by exploratory analysis of gene expression patterns (RNASeq) in 19 MSP women with uneventful, term pregnancies. The preliminary finding of clustering of different timepoints in pregnancy and postpartum respectively, follows patterns reported by Ngo *et al.* who were able to predict EGA by analysing cell free RNA transcripts in maternal blood.[26]

STRENGTHS AND LIMITATIONS

The major strength of this cohort is the prospective and the frequent follow-up along the trajectory from first trimester to 3 months postpartum. Coupled with the availability of a large number of well characterised biological samples, the high-throughput analysis, in

combination with clinical data, we anticipate that a number of pregnancy-related physiological and pathological changes can be investigated and described in detail. Systematic approaches enable the researchers to assess the entirety of a biological system and are thought to minimize the bias introduced through selective parameters.

Populations originating from low-resource settings do not often have the same level of access to advanced research when compared to high-income settings. While some data on intestinal microbiome are available [27], data generated from the MSP cohort will characterise composition and perturbations of the vaginal microbiome for the first time in this population.

The small number of PTB was not anticipated and is lower than rates reported internationally. While a larger figure would be desirable from an analytical point of view, comparison of molecular markers and associated population-based factors to populations with a higher risk of PTB will provide novel insight for this low observed proportion. Generally, a selection and exclusion bias at participant enrolment might have contributed to the low number of PTB, as women with an unremarkable medical and obstetric history were given priority to minimize the risk of potential loss of follow-up and subsequent sample loss. Hence, inference on the power of the result cannot be made at this point, as the power will depend on the magnitude of observed differences in molecular signatures between term and preterm pregnancies.

More faecal samples were missed compared to capillary blood or vaginal swab samples. In an effort to reduce sample loss for potential future research, a nested project was conceived to assess whether collection and transfer of fresh faecal samples in the long-term storage tube at home is feasible, sample quality is affected, and acceptability of self-collection is favourable.

Larger cohorts would be preferable to address critical research questions such as PTB. However, with the high-frequency sampling and the overall cost of the study conduct and sample analysis, the researchers enrolled not more participants than deemed to be necessary to address the objectives of this project.

COLLABORATION

Proposals for collaborations and other ideas to enhance scientific output from data generated in this cohort are welcome. SMRU is part of the Mahidol Oxford Tropical Medicine Research Unit and follows data sharing policies, as published by the Bioethics and Engagement Department. Data from the MSP cohort will be made available through the Data Access Committee at Mahidol Oxford Tropical Medicine Research Unit. Data sharing policies are available at: http://www.tropmedres.ac/data-sharing. An application form can be downloaded under this link. Gene expression data will be deposited on the NCBI Gene Expression Omnibus (GEO) and data from 16S rRNA sequencing will be stored as a bio-project in NCBI for public Sequence Read archives.

FURTHER DETAILS

Data sharing

Deidentified data from the MSP cohort will be accessible through the Data Access Committee at Mahidol Oxford Tropical Medicine Research Unit. Gene expression data and data from 16S rRNA sequencing will be deposited on public platforms.

Funding

This work was supported by Sidra Medicine, a member of the Qatar Foundation for Education, Science and Community Development, which is the funder of this cohort study. The Shoklo Malaria Research Unit is part of the Wellcome Trust Mahidol University Oxford Tropical Medicine Research Programme supported by the Wellcome Trust of Great Britain (Major Overseas Programme).

Acknowledgements

We would like to express our gratitude to the cohort participants: without the contribution of the women following ANC visits at SMRU this work would not be possible. Moreover, we

do acknowledge the continuous effort of staff in the recruitment centres for the excellent participant management and their pivotal role in data collection.

Collaborators

The MSP cohort is a collaboration between Shoklo Malaria Research Unit, a field station of the Faculty of Tropical Medicine, Mahidol University (Bangkok, Thailand) that is part of the Mahidol-Oxford Research Unit; Sidra Medicine (Doha, Qatar), which is part of Qatar Foundation for Education, Science and Community Development; and the Swiss Tropical and Public Health Institute, an associated institute of the University of Basel (Basel, Switzerland).

Author Contributions

Conceptualization: FN, DC and RM.

Ideas: TB, VIC, FN, RM and DC

Data curation: TB, BAK, PW, MP, AKW, AKM, TK, SAK, AT, VIC, DC and RM.

Formal analysis: TB, BSAK, DC and RM.

Funding acquisition: TK, SAK, AT, FN, DC and RM.

Investigation: TB, BAK, PW, MP, AKW, AKM, TK, SAK, AT, VIC, DC and RM.

Methodology: TB, BAK, AKM, TK, SAK, AT, VIC, DC, DHP and RM.

Project administration: TB, BSAK, PW, DC and RM.

Resources: TB, BSAK, AKM, TK, SAK, AT, DC and DHP.

Software: Not applicable.

Supervision: SP, FN, JU, DHP and RM.

Validation: FN, JU, DC, DHP and RM.

Visualization: TB

Writing – original draft: TB.

Writing – review & editing: TB, BSAK, TK, SAK, AT, VIC, FN, JU, DC, DHP and RM.

The contributor's roles listed above follow the Contributor Roles Taxonomy (CRediT) managed by The Consortia Advancing Standards in Research Administration Information (CASRAI) (https://casrai.org/credit/). All authors read and approved the final manuscript.

Competing interests

All authors declare no conflict of interests.

REFERENCES

1 Soma-Pillay P, Catherine N-P, Tolppanen H, *et al.* Physiological changes in pregnancy. *Cardiovasc J Afr* 2016;27:89–94. doi:10.5830/CVJA-2016-021

2 Mor G, Aldo P, Alvero AB. The unique immunological and microbial aspects of pregnancy. *Nat Rev Immunol* 2017;17:469–82. doi:10.1038/nri.2017.64

3 Peterson LS, Stelzer IA, Tsai AS, *et al.* Multiomic immune clockworks of pregnancy. *Semin Immunopathol* Published Online First: 4 February 2020. doi:10.1007/s00281-019-00772-1

4 Koren O, Goodrich JK, Cullender TC, *et al.* Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 2012;150:470–80. doi:10.1016/j.cell.2012.07.008

5 MacIntyre DA, Chandiramani M, Lee YS, *et al.* The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep* 2015;5:8988. doi:10.1038/srep08988

6 Aghaeepour N, Ganio EA, Mcilwain D, *et al.* An immune clock of human pregnancy. *Sci Immunol* 2017;2. doi:10.1126/sciimmunol.aan2946

7 Ursell LK, Metcalf JL, Parfrey LW, *et al.* Defining the human microbiome. *Nutr Rev* 2012;70 Suppl 1:S38-44. doi:10.1111/j.1753-4887.2012.00493.x

8 Morgan TK. Role of the Placenta in Preterm Birth: A Review. *Am J Perinatol* 2016;33:258–66. doi:10.1055/s-0035-1570379

9 Acharya KP, Pathak S. Applied Research in Low-Income Countries: Why and How? *Front Res Metr Anal* 2019;4. doi:10.3389/frma.2019.00003

10 Spielman RS, Bastone LA, Burdick JT, *et al.* Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet* 2007;39:226–31. doi:10.1038/ng1955

11 Nilsson R, Bjorkegren J, Tegner J. On reliable discovery of molecular signatures. *BMC Bioinformatics* 2009;10:38. doi:10.1186/1471-2105-10-38

12 Chaussabel D, Pascual V, Banchereau J. Assessing the human immune system through blood transcriptomics. *BMC Biol* 2010;8:84. doi:10.1186/1741-7007-8-84

13 Brummaier T, Syed Ahamed Kabeer B, Lindow S, *et al.* A prospective cohort for the investigation of alteration in temporal transcriptional and microbiome trajectories preceding preterm birth: a study protocol. *BMJ Open* 2019;9:e023417. doi:10.1136/bmjopen-2018-023417

14 Han SM, Rahman MM, Rahman MS, *et al.* Progress towards universal health coverage in Myanmar: a national and subnational assessment. *Lancet Glob Health* 2018;6:e989–97. doi:10.1016/S2214-109X(18)30318-8

15 McLean E, Cogswell M, Egli I, *et al.* Worldwide prevalence of anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993-2005. *Public Health Nutr* 2009;12:444–54. doi:10.1017/S1368980008002401

16 Gilder ME, Simpson JA, Bancone G, *et al.* Evaluation of a treatment protocol for anaemia in pregnancy nested in routine antenatal care in a limited-resource setting. *Glob Health Action* 2019;12:1621589. doi:10.1080/16549716.2019.1621589

17 Hashmi AH, Solomon N, Lee SJ, *et al.* Nutrition in transition: historical cohort analysis summarising trends in under- and over-nutrition among pregnant women in a marginalised population along the Thailand-Myanmar border from 1986 to 2016. *Br J Nutr* 2019;121:1413–23. doi:10.1017/S0007114519000758

18 Cheah PY, Lwin KM, Phaiphun L, *et al.* Community engagement on the Thai-Burmese border: rationale, experience and lessons learnt. *Int Health* 2010;2:123–9. doi:10.1016/j.inhe.2010.02.001

19 Chawanpaiboon S, Vogel JP, Moller A-B, *et al.* Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob Health* 2019;7:e37–46. doi:10.1016/S2214-109X(18)30451-0

20 Umesawa M, Kobashi G. Epidemiology of hypertensive disorders in pregnancy: prevalence, risk factors, predictors and prognosis. *Hypertens Res* 2017;40:213–20. doi:10.1038/hr.2016.126

21 Gilder ME, Zin TW, Wai NS, *et al.* Gestational diabetes mellitus prevalence in Maela refugee camp on the Thai-Myanmar border: a clinical report. *Glob Health Action* 2014;7:23887. doi:10.3402/gha.v7.23887

22 Chiefari E, Arcidiacono B, Foti D, *et al.* Gestational diabetes mellitus: an updated overview. *J Endocrinol Invest* 2017;40:899–909. doi:10.1007/s40618-016-0607-5

23 Villar J, Cheikh Ismail L, Victora CG, *et al.* International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet* 2014;384:857–68. doi:10.1016/s0140-6736(14)60932-6

24 Syed Ahamed Kabeer B, Tomei S, Mattei V, *et al.* A protocol for extraction of total RNA from finger stick whole blood samples preserved with TempusTM solution. *F1000Research* 2018;7:1739. doi:10.12688/f1000research.16622.1

25 Speake C, Whalen E, Gersuk VH, *et al.* Longitudinal monitoring of gene expression in ultra-low-volume blood samples self-collected at home. *Clin Exp Immunol* 2017;188:226–33. doi:10.1111/cei.12916

26 Ngo TTM, Moufarrej MN, Rasmussen M-LH, *et al.* Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* 2018;360:1133–6. doi:10.1126/science.aar3819

27 Vangay P, Johnson AJ, Ward TL, *et al.* US Immigration Westernizes the Human Gut Microbiome. *Cell* 2018;175:962-972.e10. doi:10.1016/j.cell.2018.10.029

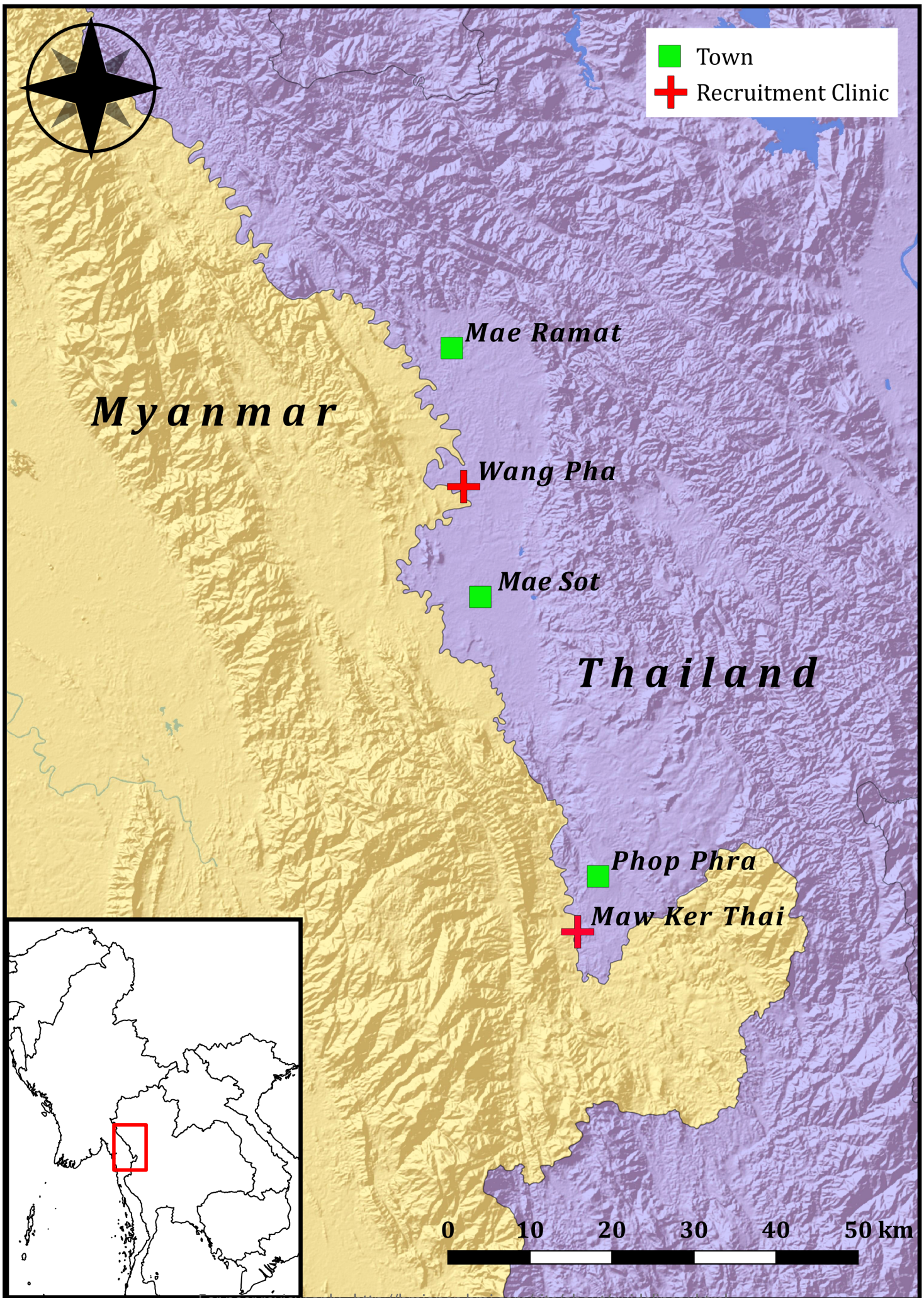
FIGURE LEGENDS

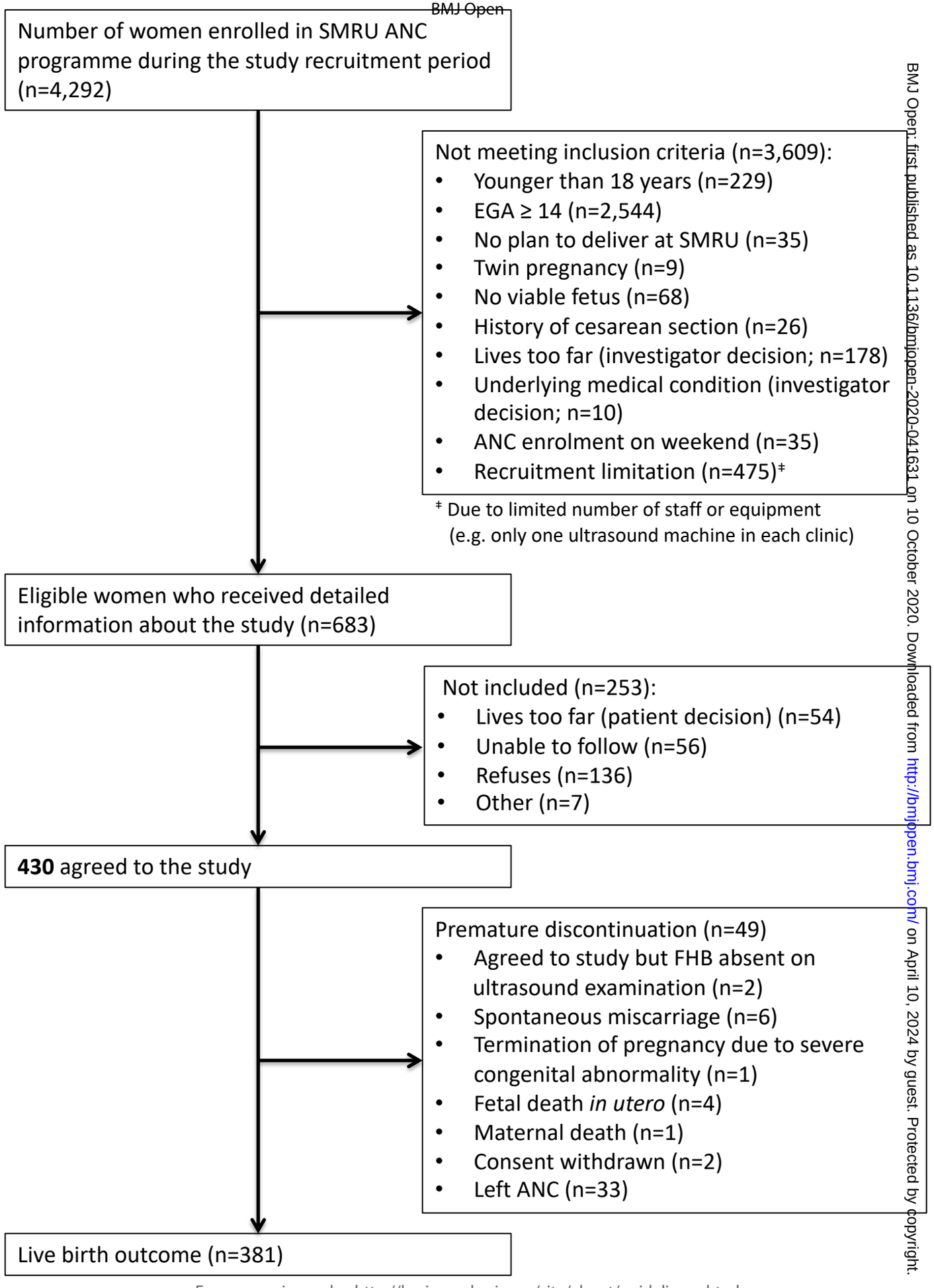
Figure 1. Setting and location of recruitment clinics.

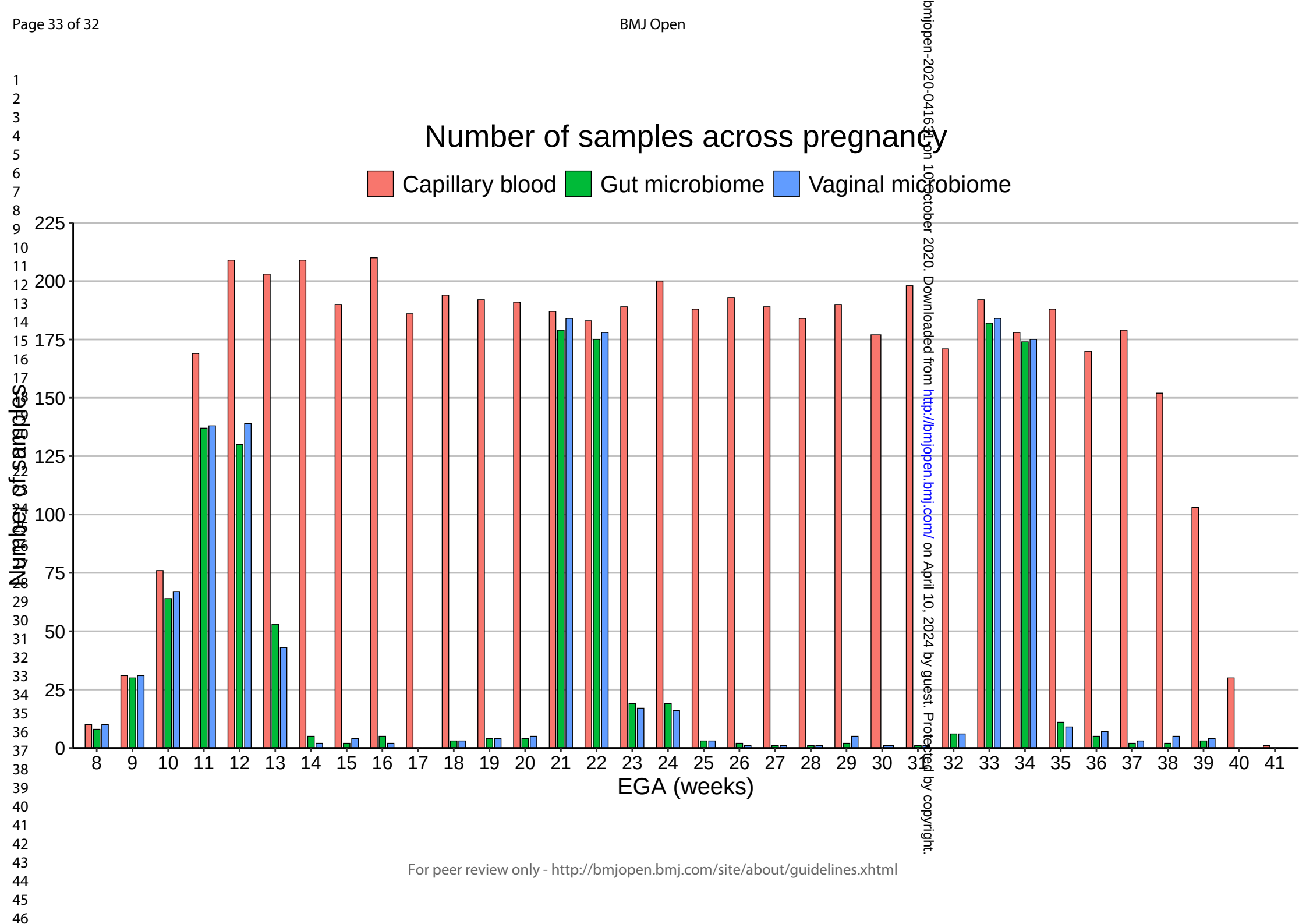
Figure 2. Screening, enrolment and outcome flowchart.
ANC – Antenatal Care; EGA – estimated gestational age; FHB – fetal heartbeat

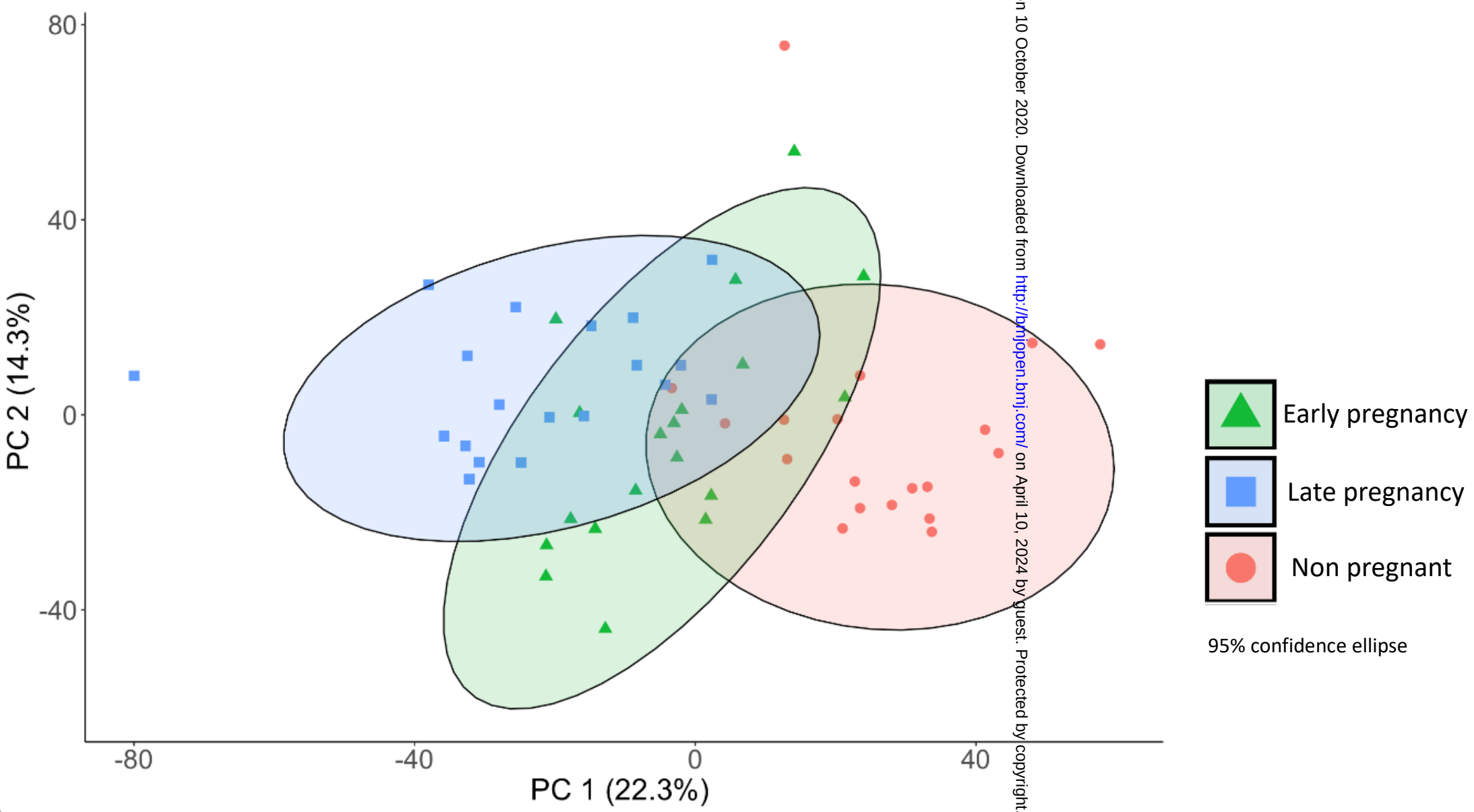
Figure 3. Number of samples over the course of pregnancy by week of gestation.

Figure 4. Principal component analysis of whole blood gene expression data (RNASeq) of 19 uneventful, term pregnancies of the MSP cohort compared between first trimester (early pregnancy), third trimester (late pregnancy) and 3-month postpartum (non-pregnant).


Town
Recruitment Clinic
Mae Ramat
Myanmar
Wang Pha
Mae Sot
Thailand
Phop Phra
Maw Ker Thai
0
10
20
30
40
50 km







BMJ Open

Cohort profile: Molecular Signature in Pregnancy (MSP) – longitudinal high-frequency sampling to characterise cross-omic trajectories in pregnancy in a resource-constrained setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-041631.R1
Article Type:	Cohort profile
Date Submitted by the Author:	31-Aug-2020
Complete List of Authors:	Brummaier, Tobias; Shoklo Malaria Research Unit, MCH Department; Schweizerisches Tropen- und Public Health-Institut, Department of Medicine Syed Ahamed Kabeer, Basirudeen; Sidra Medical and Research Center, Systems Biology and Immunology Wilaisrisak, Pornpimon ; Shoklo Malaria Research Unit Pimanpanarak, Mupawjay; Shoklo Malaria Research Unit, MCH Department Win, Aye Kyi; Shoklo Malaria Research Unit, MCH Department Pukrittayakamee, Sasithon; Mahidol University Faculty of Tropical Medicine Marr, Alexandra; Sidra Medical and Research Center Kino, Tomoshige; Sidra Medical and Research Center Al Khodor, Souhaila ; Sidra Medical and Research Center Terranegra, Annalisa; Sidra Medical and Research Center, Carrara, Verena; Shoklo Malaria Research Unit, MCH; University of Oxford Centre for Tropical Medicine and Global Health Nosten, Francois; Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Utzinger, Juerg; Schweizerisches Tropen- und Public Health-Institut, Chaussabel, Damien; Sidra Medical and Research Center Paris, Daniel; Schweizerisches Tropen- und Public Health-Institut, Department of Medicine; Universitat Basel Medizinische Fakultat McGready, Rose; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford
Primary Subject Heading:	Obstetrics and gynaecology
Secondary Subject Heading:	Epidemiology, Reproductive medicine
Keywords:	CLINICAL PHYSIOLOGY, EPIDEMIOLOGY, Reproductive medicine < GYNAECOLOGY, Fetal medicine < OBSTETRICS, Maternal medicine < OBSTETRICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Cohort profile: Molecular Signature in Pregnancy (MSP) – longitudinal high-frequency sampling to characterise cross-omic trajectories in pregnancy in a resource-constrained setting

AUTHORS

Tobias Brummaier[1,2,3,4*]
Basirudeen Syed Ahamed Kabeer[5]
Pornpimon Wilaisrisak[1]
Mupawjay Pimanpanarak[1]
Aye Kyi Win[1]
Sasithon Pukrittayakamee[6]
Alexandra K. Marr[5]
Tomoshige Kino[5]
Souhaila Al Khodor[5]
Annalisa Terranegra[5]
Verena I. Carrara[1,2]
François Nosten[1,2]
Jürg Utzinger[3,4]
Damien Chaussabel[5]
Daniel H. Paris[3,4]
Rose McGready[1,2]

1. Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand
2. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
3. Swiss Tropical and Public Health Institute, Basel, Switzerland
4. University of Basel, Basel, Switzerland
5. Sidra Medicine, Doha, Qatar
6. Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

* Corresponding author
Shoklo Malaria Research Unit
P.O. Box 46
68/30 Bann Tung Road
Mae Sot 63110
Tak Province
Thailand
Email: tobias.brummaier@gmx.at

Word count: 3,676

ABSTRACT

Purpose A successful pregnancy relies on the interplay of various biological systems. Deviations from the norm within a system or *inter-systemic* interactions may result in pregnancy associated complications and adverse pregnancy outcomes. Systems biology approaches provide an avenue of unbiased, in-depth phenotyping in health and disease. The Molecular Signature in Pregnancy (MSP) cohort was established to characterise longitudinal, cross-omic trajectories in pregnant women from a resource constrained setting. Downstream analysis will focus on characterising physiological perturbations in uneventful pregnancies, pregnancy associated complications and adverse outcomes.

Participants First trimester pregnant women of Karen or Burman ethnicity were followed prospectively throughout pregnancy, at delivery and until 3 months postpartum. Serial high frequency sampling to assess whole blood transcriptomics and microbiome composition of the gut, vagina and oral cavity, in conjunction with assessment of gene expression and microbial colonisation of gestational tissue, was done for all cohort participants.

Findings to date 381 women with live born singletons averaged 16 (IQR 15-18) antenatal visits (13,094 biological samples were collected). At 5% (19/381) the preterm birth rate was low. Other adverse events such as maternal febrile illness 7.1% (27/381), gestational diabetes 13.1% (50/381), maternal anaemia 16.3% (62/381), maternal underweight 19.2% (73/381) and a neonate born small for gestational age 20.2% (77/381) were more often observed than preterm birth.

Future plans Results from the MSP cohort will enable in-depth characterisation of cross-omic molecular trajectories in pregnancies from a population in a resource-constrained setting. Moreover, pregnancy associated complications and unfavourable pregnancy outcomes will be investigated at the same granular level, with a particular focus on population relevant needs such as effect of tropical infections on pregnancy. More detailed knowledge on multi-omic perturbations will ideally result in development of diagnostic tools and ultimately lead to targeted interventions that may disproportionally benefit pregnant women from this resource-limited population.

Registration This trial is registered under identifier NCT02797327.

KEYWORDS

Gene expression, microbiome, molecular profiling, pregnancy, resource-constrained setting.

STRENGTHS AND LIMITATIONS

- The major strength is the prospective nature of the study and frequent follow-up, coupled with high-frequency sampling and thus, availability of detailed clinical information and a considerable number of biological samples.
- High-throughput analysis, in combination with clinical data, will enable investigation of a number of pregnancy-related physiological and pathological changes.
- Populations from low-resource settings are frequently disproportionally burdened by adverse birth outcomes that may be based on exposure to different communicable diseases; hence, including them in high-end clinical research, addresses a significant research gap and may result in improvements of limited relevance to high-income countries.
- Low numbers for some phenotypes (e.g. preterm birth) may prove to be detrimental for the validity of observed differences; albeit the power will depend on the magnitude of observed differences in molecular signatures.
- In low-resource settings complete biological sample sets are often difficult to obtain, which may downsize the richness of the data.

INTRODUCTION

A successful pregnancy relies on well-timed adaptations and the interplay of multiple maternal biological systems. These interactions and temporal changes affect various organ systems, such as the cardiovascular, respiratory, endocrine system or metabolic systems, and, more recently, pivotal immunologic adaptions and changes in the human microbiome became evident.[1–5] Physiologic adaptions of the immune system during pregnancy play a central role in implantation and placentation, promotion of fetal growth and initiation of labour and delivery.[2] Deviations from the norm of this fine-tuned immune clock may lead to dysregulation in biological networks and cause various pregnancy associated complications with their immediate consequences for the mother and fetus.[6]

With a growing body of evidence of the human microbiome's significance in health and disease, investigating its role in reproductive medicine has opened up another avenue to a

deeper understanding of adverse pregnancy outcomes.[7] The placenta represents the fetomaternal link. It's crucial role during pregnancy is underlined by the fact that failure in its development and function is associated with a variety of pregnancy associated complications (e.g. preeclampsia, intrauterine growth restriction and preterm birth).[8]

In recent years investigation of these biological systems has greatly improved our understanding of their role in healthy pregnancies and in pregnancies resulting in unfavourable outcomes. However, there is a need for longitudinal, multi-omics profiling studies that may further contribute to the understanding of physiological adaptions, the interconnections of various biological systems and their significance in pregnancy associated complications.[3]

Research, in particular high-end research, is often biased towards high-income countries.[9] This imbalance is aggravated by the fact that, due to population based differences, results are often not generalisable [10], populations living in resource-constrained settings are more affected by the consequences of pregnancy associated complications (e.g. inadequate access to safe iatrogenic birth or lack of advanced neonatal care) and have different epidemiological patterns of communicable diseases. Consequently, pregnant women in low-income settings would benefit disproportionately from early identification of pregnancy associated complications and targeted interventions.

Thus, the aim of establishing this prospective Molecular Signature in Pregnancy (MSP) cohort was to characterise cross-omic trajectories in pregnant women from a resource-limited setting and describe pregnancy associated complications (e.g. preterm birth (PTB), gestational diabetes (GDM), anaemia, underweight and born too small for gestational age (SGA)). The term “molecular signature” was chosen, as it refers to molecular markers that can be used for in-depth description of a particular phenotype.[11]

Perturbations related to the immune response will be investigated by measuring the abundance of RNAs in circulating nucleated cells (i.e. leucocytes) via capillary blood sampling at multiple timepoints.[12] Microbiome profiling of the intestinal and vaginal niche in pregnancy, at delivery and postpartum will be complemented by assessment of the oral

microbiome. Lastly, placental tissue will be investigated for the presence of bacterial colonisation as well as for the relationship between the genome structure and changes in global patterns of gene expression. The generated molecular data will then be analysed in the context of clinical and patient related data. As described in the study protocol [13], an immediate focus was the investigation of PTB.

The MSP cohort profile represents the link between development of the study protocol and the results. Accordingly, cohort characteristics, including the recruitment and follow up process, demographics, pregnancy outcomes, biological samples available for analysis and preliminary results are presented.

COHORT DESCRIPTION

Setting and participants

Women with an unremarkable medical and obstetric history attending the antenatal care (ANC) facilities of Shoklo Malaria Research Unit (SMRU), a field station of the Faculty of Tropical Medicine at Mahidol University (Bangkok, Thailand) that is part of the Mahidol-Oxford Research Unit, were invited to participate. SMRU operates free-of-charge walk-in services for marginalized migrant populations in clinics at the frontier of Thailand with Myanmar (Figure 1). This mobile population mostly resides in small, rural villages. The major source of income in this community comes from work in the agricultural sector or as daily labourers, while about one in four women stay at home to look after the household.[14] As most of these daily employments provide only minimum wages coupled with the often undocumented legal status, the living conditions of most pregnant women in this community are very basic.[15] The official minimal wages in Myanmar are around 4’800 kyat (3 USD) per day, and in Thailand around 300 THB (9 USD) per day.

Between September 2016 and July 2018, trained counsellors from the community provided general information about the trial to all pregnant women who registered at SMRU ANC. Detailed information, including visual aids, were utilized to illustrate the purpose, procedures and timing of follow-up. Potential risks and benefits of the study were explained to women meeting the following inclusion criteria: (i) pregnant woman willing and able to give written informed consent; (ii) Karen or Burman ethnicity; (iii) age 18-49 years; (iv)

healthy, with viable singleton first trimester (8+0 to <14+0 weeks) pregnancy; (v) planning to deliver at the SMRU birth unit; and (vi) able and willing to comply with study requirements. If these criteria were met and appropriate consent was provided, women were followed from the first trimester, throughout pregnancy, at delivery and until 3 months postpartum.

It was estimated that approximately 400 women could be enrolled to build the cohort from this low resource environment when considering constraints of clinical capacity and budget availability. With this sample size, events such as PTB (estimated at 8%, i.e. 30 cases expected) would be feasible but also provide scope to engage with other pregnancy associated complications, tropical infections (e.g. small liver fluke infections estimated at 5%, i.e. 20 cases expected) and febrile illness (estimated at 10%, i.e. an expectation of 40 cases) with a sufficient number of uneventful pregnancies remaining to act as controls.

Figure 2 provides an overview of the number of participants recruited and followed. During the active recruitment period, 4,292 pregnant women registered to ANC. Seventy percent (2,544/3,609) were ineligible due to late presentation to ANC. There were two local conditions that limited recruitment. The first issue pertained to access; pregnant women following SMRU ANC come from villages scattered along the Thailand-Myanmar border (Figure 1) with transportation constraints affecting their ability to attend, further aggravated in the monsoon season as roads become impassable. The second issue related to staffing. As only one ultrasound machine was available at each site, there was an upper limit to the number of women who could have study related fetal growth scans done on a single day. Challenges in access and staffing issues led to the non-consideration of 178 and 475 women, respectively.

Overall, 683 women received full information about the purpose and procedures of the study. Of these, 430 agreed to participate (response rate: 63%) and consented by providing a signature or thumbprint if the event of an illiterate participant. Thirty participants were recruited in the cohort to replace women who were lost before delivery, while the study was still open to recruitment. Two participants were excluded as the fetal heartbeat (FHB) was absent after they had already agreed to the study (Figure 2).

Routine ANC procedures at SMRU are described in the study protocol.[13] Demographic characteristics of 381 enrolled women with a live birth outcome are summarised (Table 1). To demonstrate the scope of the cohort, selected subgroups are summarised alongside the characteristics for all 381 women.

Table 1. Basic characteristics of women with a live birth outcome (n=381) and subgroups of selected pregnancy associated complications in the Molecular Signature in Pregnancy (MSP) cohort.

	Overall (n=381)	Preterm (n=19)	GDM (n=50)	Anaemia (n=62)	Underweight (n=73)	SGA (n=77)
Inclusion EGA (days); median (range)	82 (56 - 97)	83 (60 - 94)	82.5 (56 - 97)	81 (60 - 94)	81 (57 - 95)	81 (62 - 97)
Maternal age groups (years)						
• 18-24	185 (48.6)	14 (73.7)	25 (50.0)	28 (45.2)	43 (58.9)	38 (49.4)
• 25-29	95 (24.9)	2 (10.5)	12 (24.0)	14 (22.6)	18 (24.7)	18 (23.4)
• 30-34	64 (16.8)	3 (15.8)	8 (16.0)	12 (19.4)	10 (13.7)	13 (16.9)
• 35-39	25 (6.6)	0	4 (8.0)	5 (8.1)	2 (2.7)	6 (7.8)
• 40-49	12 (3.1)	0	1 (2.0)	3 (4.8)	0	2 (2.6)
Residence						
• Myanmar	263 (69.0)	10 (52.6)	33 (66.0)	42 (67.7)	43 (58.9)	52 (67.5)
• Thailand	118 (31.0)	9 (47.4)	17 (34.0)	20 (32.3)	18 (24.7)	25 (32.5)
Gravidity						
• Primigravida	103 (27.0)	8 (42.1)	12 (24.0)	17 (27.4)	32 (43.8)	26 (33.8)
• 2	111 (29.1)	8 (42.1)	16 (32.0)	21 (33.9)	21 (28.8)	18 (23.4)
• 3	79 (20.7)	2 (10.5)	12 (24.0)	14 (22.6)	9 (12.3)	13 (16.9)
• 4	47 (12.3)	1 (5.3)	7 (14.0)	4 (6.5)	8 (11.0)	11 (14.3)
• ≥5	41 (10.8)	0	3 (6.0)	6 (9.7)	3 (4.1)	9 (11.7)
Parity						
• 0 (nullipara)	127 (33.3)	9 (47.4)	16 (32.0)	23 (37.1)	37 (50.7)	32 (41.6)
• 1	117 (30.7)	8 (42.1)	15 (30.0)	18 (29.0)	18 (24.7)	16 (20.8)
• 2	70 (18.4)	1 (5.3)	11 (22.0)	14 (22.6)	11 (15.1)	13 (16.9)
• 3	39 (10.2)	1 (5.3)	7 (14.0)	2 (3.2)	4 (5.5)	11 (14.3)
• ≥4	28 (7.3)	0	1 (2.0)	5 (8.1)	3 (4.1)	5 (6.5)
Literate	244 (64.0)	13 (68.4)	29 (58.0)	43 (69.4)	53 (72.6)	51 (66.2)

Smoking during in pregnancy	29 (7.6)	0	6 (12.0)	5 (8.1)	4 (5.5)	10 (13.0)
Obstetric history*						
• Miscarriage	94 (33.8)	2 (18.2)	11 (28.9)	15 (33.3)	13 (31.7)	19 (37.3)
• Stillbirth	4 (1.4)	0	0	1 (2.2)	0	0
• PIH	2 (0.7)	0	0	1 (2.2)	0	0
• Preterm rupture of membranes	8 (2.9)	2 (18.2)	0	3 (6.7)	1 (2.4)	1 (2.0)
• Preterm labour	23 (8.3)	6 (54.5)	0	9 (20.0)	3 (7.3)	3 (5.9)
• Vacuum delivery	3 (1.1)	0	0	1 (2.2)	1 (2.4)	0
• Postpartum haemorrhage	9 (3.2)	0	0	3 (6.7)	2 (4.9)	2 (3.9)
• Retained placenta	1 (0.4)	0	0	0	0	0
• Previous neonatal death	9 (3.2)	3 (27.3)	0	3 (6.7)	0	1 (2.0)
• History of GDM	2 (0.7)	0	2 (5.3)	0	0	0
• Macrosomia	2 (0.7)	0	1 (2.6)	0	0	0

Basic characteristics of pregnant women who completed per protocol follow up with a live birth outcome, compared between various pregnancy associated complications and adverse pregnancy outcomes. One study participant may be represented in multiple subgroups.
Data presented as proportion n (%) or median (range).
* Proportions calculated based on multigravida (n=278) as denominator.
Abbreviations: EGA, estimated gestational age; GDM, gestational diabetes; PIH, pregnancy induced hypertension; SGA, born too small for gestational age.

Participant follow-up

Women were followed every two weeks during pregnancy, at delivery and monthly in the first 3 months after delivery. Trial participants were encouraged to report to the study clinic if they experienced an adverse event, with special consideration for febrile episodes to maximize information on the nature of the experienced morbidity and its effect on the pregnancy.[13] Overall, 6,443 study visits during pregnancy were recorded, of which 6,140 were from women with a live birth outcome. The median number of ANC visits for women with a live birth was 16 (interquartile range [IQR] 15-18). In the uncommon event of cohort participants attending other health facilities, diagnosis and treatment was recorded at the next visit. All participants were encouraged to deliver at SMRU under the care of a skilled birth attendant. Women from the cohort achieved a low rate of home birth 6.6% (25/381) in contrast to the estimated 7 in 10 births that take place at home in rural Myanmar.[16]

As an estimator for adherence to study protocol, the number of expected and actual samples provided by women with a known pregnancy outcome were calculated. No significant difference was noted when comparing the proportion of expected and actual samples between capillary blood, stool and vaginal swab samples (p=0.070, p=0.767 and p=0.919, respectively). Missed follow-ups were the main reason for less than expected samples being provided. Additional samples were provided from participants who experienced a febrile illness, and on a few occasions, samples were erroneously taken as they were not required by the study protocol.

Early termination from the study occurred in 11.0% (47/428) of pregnancies, most of whom left ANC before birth (70.2% [33/47]) due to migration. Adverse pregnancy outcomes were recorded in 25.5% (12/47) cases and two (4.3%) participants withdrew consent, as they no longer wanted to provide vaginal swab samples (Figure 2).

Data collection

Data covering three major aspects were collected, namely (i) clinical data; (ii) genome-wide transcript abundance; and (iii) analysis of microbial composition in various anatomical sites. Table 2 summarises all study relevant procedures.

Table 2. Study procedures and timing of the Molecular Signature in Pregnancy (MSP) cohort [13]

	Screening	Baseline	Follow-up	Unwell episode	Birth	Post-partum
Viable singletons pregnancy	X					
Obstetric ultrasound*	X	X	X			
Eligibility assessment	X					
Written informed consent		X				
Demographics		X				
Medical and obstetric history		X				
Concomitant medications		X	X	X	X	X
Physical examination		X	X	X	X	X
Universal pregnancy screening, for example, thick and thin blood film for malaria diagnosis, CBC and OGTT†		X				(X)†
Sample maternal 100 µL capillary blood#		X	X‡	X§	X¶	X**
Sample vaginal swab, stool specimen and 24-hour food recall		X††	X††	X	X	X‡‡
Acceptability survey		X				X
Sample saliva			X§§		X	
Sample placenta, cord blood and maternal venous blood					X	

* Fetal growth scans on a 6-weekly basis.
# 50 µL for whole blood transcriptome analysis and 50 µL for haematocrit.
† OGTT at 24-26 weeks of gestation; repeated at 12 weeks postpartum if positive during pregnancy.
‡ 2-weekly; if the woman attended all expected 15 visits total blood is 1.5 mL.
§ If the woman attended for an unwell visit, an additional 100 µL of blood were drawn.
¶ If delivery at SMRU clinic, then an additional 100 µL of blood were drawn.
** At 1, 2 and 3 months postpartum, including maternal haematocrit.
†† In each trimester of pregnancy: 8-14, 20-22 and 34-35 weeks.
‡‡ Vaginal swab samples at 4-6 weeks and at 3 months.
§§ At 24-26 weeks of gestation.
CBC, complete blood count; OGTT, oral glucose tolerance test.

Routine ANC procedures included physical and obstetric examination, recording of concomitant medications and 6-weekly ultrasound scans to monitor fetal growth. Clinical data were collected, so that trajectories of maternal and fetal physical change would be available by the study endpoint. Two-weekly gestational weight, fetal growth and haematocrit were done as malnutrition and anaemia (and formerly malaria) are prevalent pregnancy-associated morbidities.[17–19]

Ultrasound scans to monitor fetal growth were performed 6-weekly. Throughout the study period, 2,850 scans were done; 2,693 of these were scans in women with a known, viable birth outcome (157 scans accounted for other pregnancy outcomes).

Additional details describing the study conduct and routine ANC procedures are available elsewhere.[13] Briefly, capillary blood, faecal and vaginal swab samples were taken during pregnancy, at delivery and in the post-partum period. To assess the influence of nutrition on the faecal microbiome, each faecal sample was coupled with a 24-hour food recall (Table 2). In the event of a febrile episode, an additional set of these samples were taken together with a standardised fever screening battery. Oral microbiome was assessed in the second trimester and at delivery. Placental tissue samples were taken at delivery along with umbilical cord blood serum and maternal serum.

As SMRU birthing units do not have capacity for caesarean sections (CS), participants were referred to the closest public hospitals when indicated, and hence, delivery samples were not available, in common with births at home or in other health facilities. Even though gestational tissue samples were not available for these participants, outcome measures and neonatal anthropometry were mostly available.

Lastly, in an effort to understand the acceptance of the high frequency sampling in combination with the dense follow-up schedule and the perception of invasiveness of sampling procedures, an acceptability survey at study enrolment and completion was conducted.

Biological samples

Standard operating procedures (SOPs) focused on detailed description of the sample collection procedures were drafted before study commencement and are available on request. To accommodate the high number of samples, a unique barcoded sample sticker was placed on each tube, containing the participants study code that was assigned at enrolment, as well as date and time when samples were collected. All samples were transferred from the study sites to SMRU's central laboratory daily and stored according to SOPs. Freezerworks™ (Dataworks Development, Inc., Mountlake Terrace, WA, USA), a biorepository software was used for sample management and tracking.

Including women lost to follow-up and pregnancies that resulted in a miscarriage or stillbirth, 13,536 biological samples were taken between study enrolment and 3-month postpartum. Considering multiple aliquots for vaginal swab samples (4 aliquots at each sampling time-point), saliva samples (n=4) and placenta samples (n=12), the total number of samples available for testing is 25,816.

Table 3 provides an overview of biological samples available for analysis and Figure 3 depicts the distribution of capillary blood, faecal and vaginal swab samples in relation to the estimated gestation of pregnancy.

All biological samples will be analysed and interpreted in the context of clinical data. To investigate physiological trajectories and deviations from these in the event of pregnancy associated complications, a nested case-control approach will be applied, and potential confounding factors will be adjusted for. Results will be published in original research articles alongside more detailed background information pertaining to the respective content knowledge, precise methods of sampling procedures and analytical steps of respective laboratory-based methods to enable the scientific community to draw their conclusions on appropriateness. The MSP cohort profile and the previously published study protocol [13] will serve as a cross-reference for all results.

Table 3. Overview of available biological samples for women with a live birth outcome (n=381) and selected subgroups of the Molecular Signature in Pregnancy (MSP) cohort.

	Overall (n= 381)	Preterm (n=19)	GDM (n=50)	Anaemia (n=62)	Underweight (n=73)	SGA (n=77)
Capillary blood						
• Pregnancy	5,263	226	672	871	1,000	1,070
• Delivery	346	19	44	57	68	70
• Postpartum	1,037	48	132	157	196	210
Stool						
• Pregnancy	1,159	52	149	189	220	233
• Delivery	299	18	43	47	58	59
• Postpartum	877	34	115	129	147	166
Vaginal swab‡						
• Pregnancy	1,175	55	151	191	223	236
• Delivery	329	19	45	55	67	65
• Postpartum	686	30	84	105	128	136
Saliva						
• Pregnancy	371	19	49	62	71	74
• Delivery	336	19	44	57	66	67
Placenta	301	16	41	50	63	63
Maternal serum	322	17	43	55	66	66
Cord blood serum	299	15	41	48	63	64
Cord blood EDTA*	294	14	41	49	59	60
Malaria tests in pregnancy	5,662	226	819	988	1,112	1,240
HCT tests in pregnancy	5,674	227	817	989	1,115	1,242
Stool microscopy in pregnancy	620	29	84	107	99	122

‡ Number refers to sampling time-points; four swabs were taken at each time-point.
* Less samples as amended to the protocol at a later stage.
Abbreviations: EDTA, ethylenediaminetetraacetic acid; GDM, gestational diabetes mellitus; HCT, haematocrit; SGA, born too small for gestational age.

The molecular signatures pertaining to biological samples will be analysed by collaborators at Sidra Medicine (Doha, Qatar). To assure appropriate sample handling, shipment is delegated to World Courier (AmerisourceBergen Corporation; Chesterbrook, PA, USA), a leading, global specialty and logistics provider. All samples are kept on dry ice and shipped in freezer boxes with thick styrofoam walls that are equipped with temperature monitors. At the time of writing, 16,916 biological samples have been successfully transferred.

Participant involvement

Study participants were not directly involved in the study design, the elaboration of research questions and outcome measures or the recruitment and conduct of the study. However, the protocol was reviewed and approved by a local community advisory board whose members are representatives of the local community and act as a bridge between researchers and the local population.[20] Their role was to advise on ethical and operational aspects of the study and they will provide a channel to inform the community of the results of this study.

FINDINGS TO DATE

Demographic data, findings during pregnancy and relevant outcome data for the overall cohort population with a live birth are presented in Tables 1 and 4, respectively. According to study objectives, data of commonly observed pregnancy associated complications (GDM, anaemia and underweight) and adverse pregnancy outcomes (PTB and SGA) are presented alongside the overall cohort.

At a rate of 5.0% (19/381), PTB (defined as delivery before 37 weeks of gestation), was considerably less than the worldwide estimate of approximately 10.6% (uncertainty interval 9.0-12.0%) of live-born infants.[21] When considering all pregnancies followed at SMRU during the study period, the observed proportion of PTB was roughly 6.8%, a number still below the worldwide average. The estimated gestational age (EGA) at delivery of PTB cases ranged from 30^{+2} to 36^{+6} (weeks^{+days}) with the majority of cases (89.5%) being moderate/late PTB (i.e. ≥32 and <37 weeks of gestation). Of the 19 PTB cases, one was iatrogenic with induction for severe preeclampsia at an EGA of 36^{+3}, while the remainder were spontaneous, without any discernible causative factor. Prioritizing healthy women at

the outset and frequent visits with health care providers might have contributed to earlier identification of health issues and timely intervention. Population based factors that might be associated with the low PTB rate are the low rate of sexually transmitted diseases, and preliminary data suggest that prevalence of *Gardnerella vaginalis*, a facultative anaerobic coccobacillus that is associated with PTB, is low in this population. This low PTB rate is unique in itself and will be thoroughly assessed by a holistic investigation of available demographic, clinical and cross-omic data.

There were six spontaneous miscarriages (median EGA 16^{+3} [IQR 16^{+1} – 17^{+2}]) and one pregnancy was terminated due to a congenital abnormality (severe hydrops fetalis) at an EGA of 20^{+1}. There were four stillbirths; all were identified as intrauterine fetal deaths, antepartum events before completion of 37 weeks of gestation. One was unexplained with an absent FHB at 28^{+1} weeks, one was associated with pre-eclampsia (absent FHB at 31^{+1}), one had a tight cord around the neck at birth (absent FHB at 35^{+3}), and in one case the mother ingested organophosphate with suicidal intent (absent FHB at 36^{+3}). One unexpected neonatal death was recorded in a term neonate born to a mother with an unremarkable pregnancy and delivery.

Seven of 381 (1.8%) women experienced *Plasmodium vivax* malaria during pregnancy and one woman was diagnosed with *P. vivax* malaria at delivery. Soil-transmitted helminth infections are common in this population and were detected in 21.3% (81/381) of pregnant women with a live birth outcome. Hookworm was detected in 15.0% (57/381) cases, *Trichuris trichiura* in 6.3% (24/381) and *Ascaris lumbricoides* in 4.2% (16/381). Fourteen women had more than one soil-transmitted helminth infection concurrently. Food-borne trematode infections (e.g. small liver flukes) were detected in 6.8% (26/381) of the participants. The immunologic fingerprint of helminthic infections and their effect on pregnancy is poorly understood. Hence, effects of helminth infections on the immune system or the intestinal microbiome in pregnancy will be studied with data from this cohort.

Generally, communicable diseases are often associated with adverse pregnancy outcomes. Fever, which is often the cardinal symptom of an infectious disease, was recorded in 7.1% (27/381) of pregnant women. Whether longitudinal, high-frequency sampling has potential

for early detection, or detection prior to the onset of symptoms, of communicable diseases in pregnancy, will be determined. Availability of data pertaining to communicable disease will also allow appropriate adjustment for investigation of other objectives (e.g. PTB).

Table 4. Pregnancy outcome and neonatal data of women with a life birth outcome (n=381) and selected subgroups in the Molecular Signature in Pregnancy (MSP) cohort.

	Overall (n= 381)	Preterm (n=19)	GDM (n=50)	Anaemia (n=62)	Underweight (n=73)	SGA (n=77)
Outcome EGA (days); median (IQR)	277 (271-282)	253 (244-255)	275 (268-279)	275 (268-283)	277 (271-281)	279 (274-283)
Preterm categories						
• Term (≥37 weeks)	362 (95.0)	0	49 (98.0)	57 (91.9)	68 (93.2)	73 (94.8)
• Moderate/late PTB (≥32)	17 (4.5)	17 (89.5)	1 (2.0)	3 (4.8)	5 (6.8)	4 (5.2)
• Very PTB (28 – <32)	2 (0.5)	2 (10.5)	0	2 (3.2)	0	0
• Extremely PTB (<28)	0	0	0	0	0	0
Infant sex (male)	185 (48.6)	11 (57.9)	26 (52.0)	18 (29.0)	33 (45.2)	25 (32.5)
Apgar after 1 min; median (IQR)	9 (9-9)	9 (9-9)	9 (9-9)	9 (9-9)	9 (9-9)	9 (9-9)
Apgar after 5 min; median (IQR)	10 (10-10)	10 (10-10)	10 (10-10)	10 (10-10)	10 (10-10)	10 (10-10)
Resuscitation						
• Yes	8 (2.1)	1 (5.3)	0	2 (3.2)	1 (1.4)	5 (6.5)
• No	359 (94.2)	17 (89.5)	49 (98.0)	59 (95.2)	71 (97.3)	70 (90.9)
• Unknown	14 (3.7)	1 (5.3)	1 (2.0)	1 (1.6)	1 (1.4)	2 (2.6)
Abnormal newborn exam[$]	5 (1.3)	1 (5.3)	0	2 (3.2)	1 (1.4)	0
Birthweight (g); median (IQR)*	2,955 (2,740-3,200)	2,273 (2,025-2,440)	3,010 (2,840-3,375)	2,870 (2,620-3,220)	2,823 (2,611-2,965)	2,600 (2,400-2,740)
Birthweight categories*						
• Small for GA	77 (21.8)	4 (22.2)	7 (14.3)	14 (23.0)	23 (33.8)	77 (100.0)
• Appropriate for GA	269 (76.2)	14 (77.8)	39 (79.6)	47 (77.0)	45 (66.2)	0
• Large for GA	7 (2.0)	0	3 (6.1)	0	0	0
Infant length (cm); median (IQR)[‡]	48.2 (47.0-49.4)	45.7 (44.1-46.6)	48.4 (47.4-49.4)	48.1 (46.5-49.0)	47.6 (46.5-48.5)	46.7 (45.6-48.0)
Length categories[‡]						
• Short for GA	70 (19.9)	2 (11.1)	7 (14.3)	14 (23.0)	23 (33.8)	44 (57.1)
• Appropriate for GA	266 (75.6)	16 (88.9)	37 (75.5)	44 (72.1)	43 (63.2)	32 (41.6)

• Tall for GA	16 (4.5)	0	5 (10.2)	3 (4.9)	2 (2.9)	0
Head circumference at birth; median (IQR)‡	32.9 (32.0-33.6)	31.0 (30.1-31.6)	33.1 (32.5-34.2)	32.4 (31.6-33.6)	32.4 (31.5-33.3)	32.0 (31.4-32.8)
Head circumference categories‡						
• Small HC for GA	106 (30.1)	4 (22.2)	8 (16.3)	21 (34.4)	30 (44.1)	48 (63.2)
• Appropriate HC for GA	240 (68.2)	14 (77.8)	37 (75.5)	39 (63.9)	37 (54.4)	28 (36.8)
• Large HC for GA	6 (1.7)	0	4 (8.2)	1 (1.6)	1 (1.5)	0
Caesarean section	22 (5.8)	0	1 (2.0)	5 (8.1)	2 (2.7)	5 (6.5)
Breech delivery	5 (1.3)	1 (5.3)	0	0	1 (1.4)	3 (3.9)
Vacuum delivery	11 (2.9)	0	1 (2.0)	4 (6.5)	3 (4.1)	3 (3.9)
Delivered by						
• Midwife	328 (86.1)	17 (89.5)	47 (94.0)	52 (83.9)	66 (90.4)	70 (90.9)
• TBA	25 (6.6)	2 (10.5)	2 (4.0)	3 (4.8)	5 (6.8)	1 (1.3)
• Doctor	24 (6.3)	0	1 (2.0)	6 (9.7)	2 (2.7)	5 (6.5)
• Other§	4 (1.0)	0	0	1 (1.6)	0	1 (1.3)
Place of delivery						
• SMRU clinic	304 (79.8)	16 (84.2)	43 (86.0)	51 (82.3)	63 (86.3)	63 (81.8)
• Home	29 (7.6)	2 (10.5)	2 (4.0)	4 (6.5)	5 (6.8)	2 (2.6)
• Thai Hospital	35 (9.2)	0	4 (8.0)	7 (11.3)	3 (4.1)	10 (13.0)
• Myanmar Hospital	4 (1.0)	0	0	0	1 (1.4)	0
• Other†	9 (2.4)	1 (5.3)	1 (2.0)	0	1 (1.4)	2 (2.6)
Induction of labour	23 (6.0)	1 (5.3)	3 (6.0)	4 (6.5)	4 (5.5)	4 (5.2)
Augmentation of labour	36 (9.4)	0	5 (10.0)	5 (8.1)	9 (12.3)	6 (7.8)
Postpartum haemorrhage	19 (5.0)	0	1 (2.0)	2 (3.2)	4 (5.5)	2 (2.6)
Estimated blood loss (mL)	150 (100-240)	100 (100-150)	115 (100-200)	150 (100-200)	140 (100-200)	138 (100-200)
Neonatal death	1 (0.3)	0	0	0	0	0

One study participant may be represented in multiple subgroups.
Data presented as proportion n (%) or median (interquartile range).

$ Spina bifida (n=2), cleft lip (n=1), Down's syndrome (n=1), congenital heart disease (n=2).
* 28 cases with missing information (e.g. home delivery).
‡ 29 cases with missing information.
§ Delivered by *other*: spontaneous self-delivery (n=2), husband (n=1), village health worker (n=1).
† Place of delivery *other*: delivery on the way to the SMRU clinics (n=6), in another clinic (n=3).
Abbreviations: EGA, estimated gestational age; GA, gestational age; GDM, gestational diabetes mellitus; HC, head circumference; IQR, interquartile range; min, minutes; PTB, preterm birth; SGA, born too small for gestational age; SMRU, Shoklo Malaria Unit; TBA, traditional birth attendant

375

Overall, only four cases (1.0%) of preeclampsia and one case (0.3%) of eclampsia were recorded which is also at the lower end of the epidemiological range (0.2-9.2).[22] A common pregnancy associated complication was GDM, which affected 13.1% (50/381) pregnancies with a live birth. GDM is the most common metabolic disorder in pregnancy and is of increasing concern in high-income settings around the globe but also in this resource-limited population.[23,24] While the proportion of pregnant women diagnosed with underweight in this population is decreasing, 19.2% (73/381) of pregnancies were still complicated by a diagnosis of underweight.[19] Maternal anaemia defined as haematocrit of less than 30% complicated 16.3% [62/381] of pregnancies and SGA (20.2% [77/381]) according to standards as published by the Intergrowth-21st consortium [25] was also commonly observed and remains a solicitude in this pregnant population. As shown in Figure 3, cohort enrolment in early pregnancy, coupled with the high-frequency sampling provides a plethora of cross-omic data. This will enable to identify signatures preceding the development of these complications, their impact on pregnancy outcome and to characterise how they plateau in the postpartum period.

Ultra-low capillary blood samples are safe and a reliable method to monitor longitudinal gene expression trajectories. Samples acquired from this MSP cohort were used to develop a modified SOP for extraction of RNA from small volume whole blood samples preserved in Tempus^{TM} solution [26] and the performance of ultra-low blood samples was benchmarked previously.[27] Figure 4 shows a comparison between different stages in pregnancy (early pregnancy and late pregnancy), and the non-pregnant control (3-month postpartum) by exploratory analysis of gene expression patterns (RNASeq) in 19 MSP women with uneventful, term pregnancies. The preliminary finding of clustering of different timepoints in pregnancy and postpartum respectively, follows patterns reported by Ngo *et al.* who were able to predict EGA by analysing cell free RNA transcripts in maternal blood.[28]

STRENGTHS AND LIMITATIONS

The major strength of this cohort is the prospective and the frequent follow-up along the trajectory from first trimester to 3 months postpartum. Coupled with the availability of a large number of well characterised biological samples, the high-throughput analysis, in

combination with clinical data, we anticipate that a number of pregnancy-related physiological and pathological changes can be investigated and described in detail. Systematic approaches enable the researchers to assess the entirety of a biological system and are thought to minimize the bias introduced through selective parameters.

Populations originating from low-resource settings do not often have the same level of access to advanced research when compared to high-income settings. While some data on intestinal microbiome are available [29], data generated from the MSP cohort will characterise composition and perturbations of the vaginal microbiome for the first time in this population.

The small number of PTB was not anticipated and is lower than rates reported internationally. While a larger figure would be desirable from an analytical point of view, comparison of molecular markers and associated population-based factors to populations with a higher risk of PTB will provide novel insight for this low observed proportion. Generally, a selection and exclusion bias at participant enrolment might have contributed to the low number of PTB, as women with an unremarkable medical and obstetric history were given priority to minimize the risk of potential loss of follow-up and subsequent sample loss. Hence, inference on the power of the result cannot be made at this point, as the power will depend on the magnitude of observed differences in molecular signatures between term and preterm pregnancies.

More faecal samples were missed compared to capillary blood or vaginal swab samples. In an effort to reduce sample loss for potential future research, a nested project was conceived to assess whether collection and transfer of fresh faecal samples in the long-term storage tube at home is feasible, sample quality is affected, and acceptability of self-collection is favourable.

Larger cohorts would be preferable to address critical research questions such as PTB. However, with the high-frequency sampling and the overall cost of the study conduct and sample analysis, the researchers enrolled not more participants than deemed to be necessary to address the objectives of this project.

COLLABORATION

Proposals for collaborations and other ideas to enhance scientific output from data generated in this cohort are welcome. SMRU is part of the Mahidol Oxford Tropical Medicine Research Unit and follows data sharing policies, as published by the Bioethics and Engagement Department. Data from the MSP cohort will be made available through the Data Access Committee at Mahidol Oxford Tropical Medicine Research Unit. Data sharing policies are available at: http://www.tropmedres.ac/data-sharing. An application form can be downloaded under this link. Gene expression data will be deposited on the NCBI Gene Expression Omnibus (GEO) and data from 16S rRNA sequencing will be stored as a bio-project in NCBI for public Sequence Read archives.

FURTHER DETAILS

Data sharing

Deidentified data from the MSP cohort will be accessible through the Data Access Committee at Mahidol Oxford Tropical Medicine Research Unit. Gene expression data and data from 16S rRNA sequencing will be deposited on public platforms.

Funding

This work was supported by Sidra Medicine, a member of the Qatar Foundation for Education, Science and Community Development. The Shoklo Malaria Research Unit is part of the Wellcome Trust Mahidol University Oxford Tropical Medicine Research Programme supported by the Wellcome Trust of Great Britain (Major Overseas Programme).

Acknowledgements

We would like to express our gratitude to the cohort participants: without the contribution of the women following ANC visits at SMRU this work would not be possible. Moreover, we do acknowledge the continuous effort of staff in the recruitment centres for the excellent participant management and their pivotal role in data collection.

Collaborators

The MSP cohort is a collaboration between Shoklo Malaria Research Unit, a field station of the Faculty of Tropical Medicine, Mahidol University (Bangkok, Thailand) that is part of the Mahidol-Oxford Research Unit; Sidra Medicine (Doha, Qatar), which is part of Qatar Foundation for Education, Science and Community Development; and the Swiss Tropical and Public Health Institute, an associated institute of the University of Basel (Basel, Switzerland).

Author Contributions

Conceptualization: FN, DC and RM.

Ideas: TB, VIC, FN, RM and DC

Data curation: TB, BSAK, PW, MP, AKW, AKM, TK, SAK, AT, VIC, DC and RM.

Formal analysis: TB, BSAK, DC and RM.

Funding acquisition: TK, SAK, AT, FN, DC and RM.

Investigation: TB, BSAK, PW, MP, AKW, AKM, TK, SAK, AT, VIC, DC and RM.

Methodology: TB, BSAK, AKM, TK, SAK, AT, VIC, DC, DHP and RM.

Project administration: TB, BSAK, PW, DC and RM.

Resources: TB, BSAK, AKM, TK, SAK, AT, DC and DHP.

Software: Not applicable.

Supervision: SP, FN, JU, DHP and RM.

Validation: FN, JU, DC, DHP and RM.

Visualization: TB

Writing – original draft: TB.

Writing – review & editing: TB, BSAK, TK, SAK, AT, VIC, FN, JU, DC, DHP and RM.

The contributor's roles listed above follow the Contributor Roles Taxonomy (CRediT) managed by The Consortia Advancing Standards in Research Administration Information (CASRAI) (https://casrai.org/credit/). All authors read and approved the final manuscript.

Competing interests

All authors declare no conflict of interests.

REFERENCES

1 Soma-Pillay P, Catherine N-P, Tolppanen H, *et al.* Physiological changes in pregnancy. *Cardiovasc J Afr* 2016;27:89–94. doi:10.5830/CVJA-2016-021

2 Mor G, Aldo P, Alvero AB. The unique immunological and microbial aspects of pregnancy. *Nat Rev Immunol* 2017;17:469–82. doi:10.1038/nri.2017.64

3 Peterson LS, Stelzer IA, Tsai AS, *et al.* Multiomic immune clockworks of pregnancy. *Semin Immunopathol* Published Online First: 4 February 2020. doi:10.1007/s00281-019-00772-1

4 Koren O, Goodrich JK, Cullender TC, *et al.* Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 2012;150:470–80. doi:10.1016/j.cell.2012.07.008

5 MacIntyre DA, Chandiramani M, Lee YS, *et al.* The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep* 2015;5:8988. doi:10.1038/srep08988

6 Aghaeepour N, Ganio EA, Mcilwain D, *et al.* An immune clock of human pregnancy. *Sci Immunol* 2017;2. doi:10.1126/sciimmunol.aan2946

7 Ursell LK, Metcalf JL, Parfrey LW, *et al.* Defining the human microbiome. *Nutr Rev* 2012;70 Suppl 1:S38-44. doi:10.1111/j.1753-4887.2012.00493.x

8 Morgan TK. Role of the Placenta in Preterm Birth: A Review. *Am J Perinatol* 2016;33:258–66. doi:10.1055/s-0035-1570379

9 Acharya KP, Pathak S. Applied Research in Low-Income Countries: Why and How? *Front Res Metr Anal* 2019;4. doi:10.3389/frma.2019.00003

10 Spielman RS, Bastone LA, Burdick JT, *et al.* Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet* 2007;39:226–31. doi:10.1038/ng1955

11 Nilsson R, Bjorkegren J, Tegner J. On reliable discovery of molecular signatures. *BMC Bioinformatics* 2009;10:38. doi:10.1186/1471-2105-10-38

12 Chaussabel D, Pascual V, Banchereau J. Assessing the human immune system through blood transcriptomics. *BMC Biol* 2010;8:84. doi:10.1186/1741-7007-8-84

13 Brummaier T, Syed Ahamed Kabeer B, Lindow S, *et al.* A prospective cohort for the investigation of alteration in temporal transcriptional and microbiome trajectories preceding preterm birth: a study protocol. *BMJ Open* 2019;9:e023417. doi:10.1136/bmjopen-2018-023417

14 Fellmeth G, Plugge EH, Carrara V, *et al.* Migrant perinatal depression study: a prospective cohort study of perinatal depression on the Thai-Myanmar border. *BMJ Open* 2018;8:e017129. doi:10.1136/bmjopen-2017-017129

15 Pudpong N, Durier N, Julchoo S, *et al.* Assessment of a Voluntary Non-Profit Health Insurance Scheme for Migrants along the Thai–Myanmar Border: A Case Study of the Migrant Fund in Thailand. *Int J Environ Res Public Health* 2019;16:2581. doi:10.3390/ijerph16142581

16 Han SM, Rahman MM, Rahman MS, *et al.* Progress towards universal health coverage in Myanmar: a national and subnational assessment. *Lancet Glob Health* 2018;6:e989–97. doi:10.1016/S2214-109X(18)30318-8

17 McLean E, Cogswell M, Egli I, *et al.* Worldwide prevalence of anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993-2005. *Public Health Nutr* 2009;12:444–54. doi:10.1017/S1368980008002401

18 Gilder ME, Simpson JA, Bancone G, *et al.* Evaluation of a treatment protocol for anaemia in pregnancy nested in routine antenatal care in a limited-resource setting. *Glob Health Action* 2019;12:1621589. doi:10.1080/16549716.2019.1621589

19 Hashmi AH, Solomon N, Lee SJ, *et al.* Nutrition in transition: historical cohort analysis summarising trends in under- and over-nutrition among pregnant women in a marginalised population along the Thailand-Myanmar border from 1986 to 2016. *Br J Nutr* 2019;121:1413–23. doi:10.1017/S0007114519000758

20 Cheah PY, Lwin KM, Phaiphun L, *et al.* Community engagement on the Thai-Burmese border: rationale, experience and lessons learnt. *Int Health* 2010;2:123–9. doi:10.1016/j.inhe.2010.02.001

21 Chawanpaiboon S, Vogel JP, Moller A-B, *et al.* Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob Health* 2019;7:e37–46. doi:10.1016/S2214-109X(18)30451-0

22 Umesawa M, Kobashi G. Epidemiology of hypertensive disorders in pregnancy: prevalence, risk factors, predictors and prognosis. *Hypertens Res* 2017;40:213–20. doi:10.1038/hr.2016.126

23 Gilder ME, Zin TW, Wai NS, *et al.* Gestational diabetes mellitus prevalence in Maela refugee camp on the Thai-Myanmar border: a clinical report. *Glob Health Action* 2014;7:23887. doi:10.3402/gha.v7.23887

24 Chiefari E, Arcidiacono B, Foti D, *et al.* Gestational diabetes mellitus: an updated overview. *J Endocrinol Invest* 2017;40:899–909. doi:10.1007/s40618-016-0607-5

25 Villar J, Cheikh Ismail L, Victora CG, *et al.* International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet* 2014;384:857–68. doi:10.1016/s0140-6736(14)60932-6

26 Syed Ahamed Kabeer B, Tomei S, Mattei V, *et al.* A protocol for extraction of total RNA from finger stick whole blood samples preserved with TempusTM solution. *F1000Research* 2018;7:1739. doi:10.12688/f1000research.16622.1

27 Speake C, Whalen E, Gersuk VH, *et al.* Longitudinal monitoring of gene expression in ultra-low-volume blood samples self-collected at home. *Clin Exp Immunol* 2017;188:226–33. doi:10.1111/cei.12916

28 Ngo TTM, Moufarrej MN, Rasmussen M-LH, *et al.* Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* 2018;360:1133–6. doi:10.1126/science.aar3819

29 Vangay P, Johnson AJ, Ward TL, *et al.* US Immigration Westernizes the Human Gut Microbiome. *Cell* 2018;175:962-972.e10. doi:10.1016/j.cell.2018.10.029

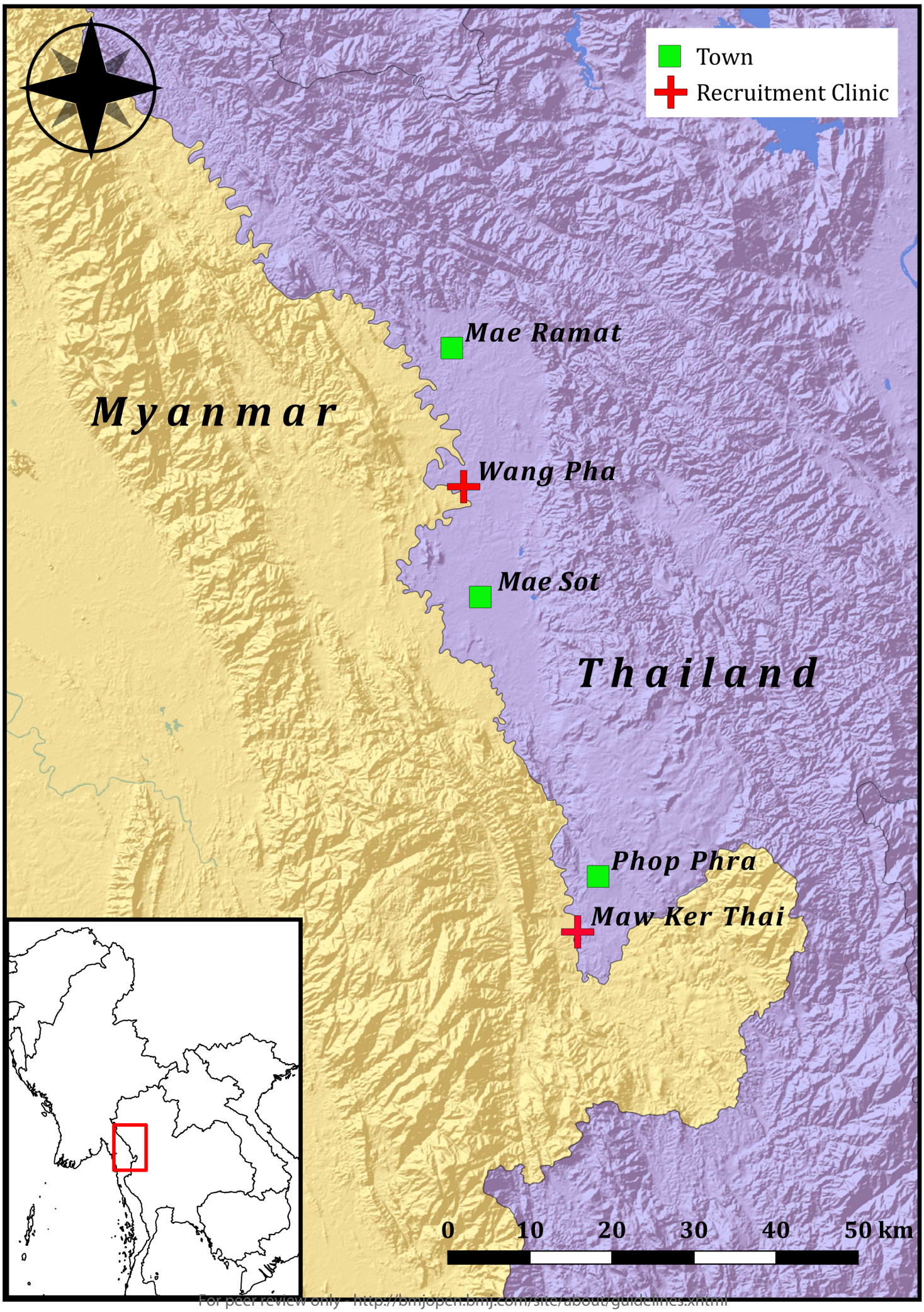
FIGURE LEGENDS

Figure 1. Setting and location of recruitment clinics.

Figure 2. Screening, enrolment and outcome flowchart.
ANC – Antenatal Care; EGA – estimated gestational age; FHB – fetal heartbeat

Figure 3. Number of samples over the course of pregnancy by week of gestation.

Figure 4. Principal component analysis of whole blood gene expression data (RNASeq) of 19 uneventful, term pregnancies of the MSP cohort compared between first trimester (early pregnancy), third trimester (late pregnancy) and 3-month postpartum (non-pregnant).


Town
Recruitment Clinic
Mae Ramat
Myanmar
Wang Pha
Mae Sot
Thailand
Phop Phra
Maw Ker Thai
0
10
20
30
40
50 km