




BMJ Open Influence of maternal body mass index on human milk composition and associations to infant metabolism and gut colonisation: MAINHEALTH – a study protocol for an observational birth cohort

Katrine Overgaard Poulsen ,^{1,2} Julie Astono,¹ Rasmus Riemer Jakobsen ,³ Niels Ulbjerg,⁴ Jens Fuglsang,⁴ Dennis Sandris Nielsen,³ Ulrik Kræmer Sundekilde ¹

To cite: Overgaard Poulsen K, Astono J, Jakobsen RR, *et al.* Influence of maternal body mass index on human milk composition and associations to infant metabolism and gut colonisation: MAINHEALTH – a study protocol for an observational birth cohort. *BMJ Open* 2022;**12**:e059552. doi:10.1136/bmjopen-2021-059552

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-059552>).

Received 24 November 2021
Accepted 09 October 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Ulrik Kræmer Sundekilde; uksundekilde@food.au.dk

ABSTRACT

Introduction Human milk provides all macronutrients for growth, bioactive compounds, micro-organisms and immunological components, which potentially interacts with and primes infant growth and, development, immune responses and the gut microbiota of the new-born. Infants with an overweight mother are more likely to become overweight later in life and overweight has been related to the gut microbiome. Therefore, it is important to investigate the mother-milk-infant triad as a biological system and if the maternal weight status influences the human milk composition, infant metabolism and gut microbiome.

Methods and analysis This study aims to include 200 mother–infant dyads stratified into one of three body mass index (BMI) categories based on mother's prepregnancy BMI. Multiomics analyses include metabolomics, proteomics, glycomics and microbiomics methods, aiming to characterise human milk from the mothers and further relate the composition to infant gut microbiota and its metabolic impact in the infant. Infant gut microbiota is analysed using 16S sequencing of faeces samples. Nuclear magnetic resonance and mass spectrometry are used for the remaining omics analysis. We investigate whether maternal pre-pregnancy BMI results in a distinct human milk composition that potentially affects the initial priming of the infant's gut environment and metabolism early in life.

Ethics and dissemination The Central Denmark Region Committees on Health Research Ethics has approved the protocol (J-nr. 1-10-72-296-18). All participants have before inclusion signed informed consent and deputy informed consent in accordance with the Declaration of Helsinki II. Results will be disseminated to health professionals including paediatricians, research community, nutritional policymakers, industry and finally the public. The scientific community will be informed via peer-reviewed publications and presentations at scientific conferences, the industry will be invited for meetings, and the public will be informed via reports in

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Large longitudinal cohort including 200 mother–infant dyads, totalling 4000 biological samples.
- ⇒ Integration of different omics techniques with results leading to a comprehensive understanding of human milk composition and its effect on establishment of the infant gut microbiota.
- ⇒ High drop-out rate due to duration of the study and requirement of breast feeding.
- ⇒ High burdens on the participants including collection of multiple samples and answering questionnaires.

science magazines and the general press. Data cleared for personal data, will be deposited at public data repositories.

Trial registration number Danish regional committee of the Central Jutland Region, journal number: 1-10-72-296-18, version 6.
Danish Data Protection Agency, journal number: 2016-051-000001, 1304.
ClinicalTrials.gov, identifier: NCT05111990.

INTRODUCTION

Human milk (HM) is the gold standard of nutrition for infants and the result of evolution through thousands of years. The WHO recommends that infants should be exclusively breastfed for the first 6 months, with partial breast feeding up to 2 years and beyond.¹ A well-nourished mother offers the infant all essential macronutrients and micronutrients through HM, which is individualised and changes dynamically over the lactation period in synchronisation with infant development.^{2–4}

Table 1 Major constituents in mature human milk and alterations associated with overweight mothers

Component	Reference content (g/100 mL) ^{*53}	Content in overweight versus normal weight mothers	References
Protein	0.7–3	Increased	54–56
Fat	2.2–3.8	Increased	54 55
Lactose	5.6–6.9	No association	55

*Range refers to differences related to increases with lactational age.⁵³

Milk components

HM contains carbohydrates, lipids, proteins and bioactive components, which are modulated in concentration during a feed,⁵ the course of lactation^{2–4} and diurnally.⁶ The composition and amount of the major classes within carbohydrates, proteins and lipids in HM is well known. Associations between maternal overweight and macronutrient content in HM have been investigated, and a reference standard of major constituents in mature milk and variations according to maternal weight are summarised in table 1. Yet, considerable weaknesses in level of analytical detail exists in the compound classes due to the complex nature of HM. Furthermore, there is little knowledge about intrinsic and extrinsic factors influencing HM composition, biological influence on the presence and concentrations of nutrients, and of the microcomponents with biological significance. A major factor known to influence milk composition is the maternal diet. In particular milk levels of fatty acids and fat-soluble vitamins.^{7 8}

Proteins and peptides

Major proteins in HM are well described.^{9–11} Less is known about peptides present in HM¹² and specifically peptides produced from proteolysis during infant digestion.¹³ Mass spectrometry (MS) analysis of the HM peptidome has revealed almost 700 endogenous peptides from 30 different proteins.¹² However, the bioactivity of these proteins and peptides remains elusive, and investigating this is a natural next step.

Most fat in HM is concentrated in the milk fat globule. It consists of a core with triacylglycerols and a tri-layered membrane, the milk fat globule membrane (MFGM).⁴ One interesting area is the proteins and peptides associated with the MFGM as little is known of their detailed functionality.¹⁴ It is known that changes in the MFGM proteins reflect changes in the functions within and metabolism of the mammary secretory cells and can be hypothesised to be affected by the energy and metabolic status of the mother. A preliminary study of six mothers with gestational diabetes mellitus showed that the human colostrum proteome was significantly different.¹⁵ We speculate if the peptidome and MFGM proteins and peptides are different in obese mothers compared with normal weight mothers.

Metabolites

The HM further contains low-molecular weight compounds/metabolites that possibly contain

bioactivity. The milk metabolome is among others influenced by maternal genotype, infant gestational age (GA), lactational age, foremilk and hindmilk and diurnal factors.^{3 16} Up to 710 different metabolites have been reported in HM.³ Some studies have associated different metabolites with increased maternal weight.^{17 18} It is therefore important to investigate if the altered milk metabolome affects the infant metabolism and gut microbiota.

Human milk oligosaccharides

The carbohydrates in HM consist of lactose and up to 200 different structures of human milk oligosaccharides (HMO) and many structures remain to be elucidated.¹⁹ From a maternal perspective, HMOs are metabolically expensive biological molecules to produce as the infants lack the enzymes to hydrolyse HMOs and can therefore not digest and take up HMOs. Hence, the function of HMOs is heavily investigated. The HMO content and composition changes dynamically over time and is dependent on the secretor-status of the mother. The secretor-status of the mother refers to the activity of the fucosyltransferases FUT2 and FUT3 and determines if she can produce HMOs containing $\alpha_{1,2}$ -linked and/or $\alpha_{1,4}$ -linked fucose, respectively.²⁰ The bioactivity of HMOs are not fully elucidated, but it is reported that they serve as substrates for certain beneficial gut microbiota species (section: infant gut microbiota),^{21 22} and some structures may serve as decoys for pathogens.²³ A few studies suggest that maternal obesity affect the HMO content in HM and the infant gut microbiota, however, the exact relationship between these factors is not established.^{24 25} It is important to account for secretor status when investigating correlations related to HMO content, as this can serve as confounding factor.²⁶

Human milk microbiome

HM also contains living microorganisms, also known as probiotics. The average infant consumes ~800 mL of HM daily containing between 10^5 and 10^6 bacteria.^{27 28} These bacteria and the HMOs may facilitate an optimal bacterial maturation of the gastrointestinal tract. The early infant microbiome is indicated to correspond closely to the milk and areolar skin microbiome of the mother. The degree of similarity increases with milk intake, however this finding needs to be validated in a larger study.²⁹

Infant gut microbiota

The initial seeding of the infant gut microbiota is affected by various factors including delivery mode, use of infant formula and maternal gut microbiota, among others.³⁰ The gut microbiota of a vaginally delivered infant consists of different bacteria, the most abundant ones being *Bifidobacterium*, *Lactobacillus*, *Bacteroides* and *Enterobacteriaceae*. The microbiota in the gut feeds of the polysaccharides not digestible to humans and in early life. This is especially the HMOs. HMOs like 2-fucosyllactose and lacto-n-neotetraose have shown to be important for growth of *Bifidobacterium*.³¹ Some microbiota can produce short-chain fatty acids, which are digestible to humans, hence these metabolites are also of importance. The gut microbiome has gained increasing interest over the past years as several diseases and obesity have been associated with altered gut microbiota. Further, it is known that maternal obesity is associated with higher risks of childhood obesity.³² Whether this association could have its offspring in infant gut microbiome colonisation is yet under investigation and the elucidation is one of our main aims.

Multiomics approach

The multiomics approach encapsulates different omics techniques on cross-sectional biospecimens. In this case, it includes metabolomics, proteomics, glycomics and microbiomics. Collectively, they thrive to elucidate

metabolite, protein, glycan and microbiota content in a biological sample. There are different analysis methods related to each omics approach. The most widely used are MS and nuclear magnetic resonance (NMR) for metabolomics, proteomics and glycomics and 16S sequencing for microbiomics.^{33 34} The advantage of the multi-omics approach lies in the data analysis and integration which requires skilled data scientists and powerful computer software. By using multiomics it is possible to investigate the mother-breastmilk-infant triad as a complete system³⁵

Beneficial effects of breast milk feeding

The link between early-stage infant nutrition and long-term health has been established, but mechanisms still lack. Long-term outcomes indicate risk reducing effects of breast-feeding in developing cardiovascular diseases,^{36 37} while positive effects on cardiovascular health and cognitive development has also been shown.^{38 39} Associations with breast feeding and lower rates of overweight and obesity in children and later in life, lower systolic blood pressure as well as a protection effect against type 2 diabetes mellitus in adolescents and adults have been found.⁴⁰ Therefore, breast feeding is the superior choice of infant nutrition compared with infant formula. The compositions of infant formula and follow-on formula are under the regulatory control of the European Food Safety Authority, which specifies the content of various components for the first year of life. Despite different formula

Table 2 Time schedule for participants starting from GA 18–20 weeks to the respective infant is 5 years old

GA18	▶ Approached by project worker receives information of the project at the AUH midwife centre
GA18-GA20	▶ Gives informed consent
GA20	▶ Have blood sample taken at AUH
Enrollment before GA 37	▶ Answers online survey for recruitment screening provided by email
GA37	▶ Answers 2×24 hours dietary recall questionnaire
Birth	▶ Vaginal sample and rectum sample collected by midwives
Within a week pp	▶ HM sample from mother
+30 days	▶ Faeces sample, urine sample, saliva sample from infant ▶ HM sample, sample of breast skin from mother ▶ Mother answers 2×24 hours dietary recall questionnaire and online survey addressing breast-feeding frequency, infant development and life-style
+60 days	▶ Faeces sample and urine sample from infant ▶ HM sample from mother ▶ Mother answers 2×24 hours dietary recall questionnaire and online survey addressing breast-feeding frequency, infant development and life-style
+90 days	▶ Faeces sample and urine sample from infant ▶ HM sample from mother ▶ Mother answers 2×24 hours dietary recall questionnaire and online survey addressing breastfeeding frequency, infant development and lifestyle
+6 months	▶ Mother answers online survey addressing breast-feeding frequency and introduction of solid foods
+1–5 years	▶ Faeces sample from infant, follow-up on infant anthropometric development once a year ▶ Mother answers online survey once a year

+ denotes time lapse from birth.
AUH, Aarhus University Hospital; GA, gestational age; HM, human milk; pp, post partum.

milk compositions being available for preterm infants, it is still far from perfect as only one formula milk is offered to other infants at relevant ages, despite the knowledge of HM variability over time. Further, using formula milk is related with higher risk of developing allergies and asthma later in life as breast feeding is known to lower the risk of developing atopy.⁴¹

The important challenge in infant nutrition remains to elucidate why HM is so superior to other sources and use this knowledge for production of the perfect formula milk.

AIM

With this longitudinal observational cohort, we aim to investigate how the mother's body mass index (BMI) affect the composition of the milk and possibly the infant metabolism and gut microbiota. This is done by integrative multiomics studies of the human milk metabolome and proteome, infant urine and faecal metabolome and infant gut microbiome. We investigate the long-term consequences of mother's health, natural human milk variation and initial gut colonisation on the child's infant gut microbiota and growth. This knowledge lies the basis for understanding the mother-milk-infant triad as a system and maybe aid in the production of improved and individualised formula milk.

Working hypotheses

The MAINHEALTH study investigates three main hypotheses:

1. Maternal obesity influence which HM nutrients are available to the infant.
2. Maternal obesity influences the infant microbiome.
3. Long-term infant gut microbiota is influenced by the early gut colonisation.

METHODS AND ANALYSIS

Study design

The MAINHEALTH cohort is a longitudinal observational study which includes 200 mother–infant dyads and allocates them into three maternal prepregnancy BMI groups; normal weight (BMI 18.5–24.99), overweight (BMI 25–30) and obese (BMI >30). Project partners are affiliated with Aarhus University—Department of Food Science (AU-FOOD), Aarhus University Hospital (AUH)—Department of Clinical Medicine, and Copenhagen University—Department of Food Science. The cohort has been approved by the Central Denmark Regional Committees on Health Research Ethics (ethical approval reference: 1-10-72-296-18) and registered on ClinicalTrials.gov (identification number: NCT05111990). The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) reporting guidelines was used in preparation of the study protocol.⁴²

The participants are recruited from week 18 to 20 in the mother's pregnancy and followed until the respective

Box 1 Criteria of inclusion for the mother and infant.

Criteria related to mother

- ⇒ Residing in Aarhus area.
- ⇒ Above 18 years of age.
- ⇒ BMI above 18.5 kg/m².
- ⇒ Intention to breastfeed the first 4–6 months following birth.
- ⇒ Be able to communicate in Danish.
- ⇒ Non-smoker.
- ⇒ Non-multiple gestation.
- ⇒ Not suffering from the following chronic diseases that demands medical treatment: diabetes mellitus, celiac disease, inflammatory bowel disease (Crohn's disease or ulcerosa colitis).
- ⇒ Do not take medicaments for irritable bowel syndrome.
- ⇒ Do not take medicine for metabolic disorders.
- ⇒ Do not take medicine for psychological disorders.
- ⇒ Not had a gastric bypass surgery.
- ⇒ Not planned caesarean section.
- ⇒ Not received antibiotics after week 12 in their pregnancy.
- ⇒ Not utilising infant formula more than 4 times a week following birth.

Criteria related to the infant post partum

- ⇒ No inborn errors of metabolism that interfere with infant eating or growth following birth.
- ⇒ Infants born after gestational age 37 weeks.
- ⇒ Infants with a birth weight between 2500 g and 5000 g. BMI, body mass index.

infant is 5 years old. Throughout this period, several types of biological material and metadata are collected (table 2).

Maternal health outcomes including BMI and blood sample parameters will be used as baseline characteristics. Correlations between measured characteristics and between metabolite concentrations in HM, infant urine and faeces, protein concentration and composition in HM, microbiome in infant faeces, infant saliva, maternal rectal and vaginal swabs from birth will be analysed.

Recruitment of participants

By May 2019, recruitment was initiated and finished December 2021 reaching 168 mother–infant dyads. Recruitment was temporarily suspended due to COVID-19 lockdown in Denmark from 12 March 2020 to 8 June 2020. Pregnant women were approached at the Aarhus Midwife Centre (Olof Palmes Alle 38, 8200 Aarhus N, Denmark). Healthy pregnant women with a BMI above 18.5 kg/m² who fulfil the criteria in box 1 are considered eligible for participation. Interested women and their partners are invited for an oral information session covering study design, associated risk and criteria for participation, that consent can be withdrawn and consideration period. Written information is handed out to address implications of participation for the mother and the infant. The mother signs informed consent for her own participation and the custody/custodies sign informed consent for the respective infant's participation. Participants are included after having signed informed consent, a blood

sample is collected around gestational week 20 and an online survey is completed.

Study management

Access to data, databases and samples are restricted to investigators. Data is secured by encryption and daily backups. Twice a year, a report of who have accessed data is compiled.

Once a year throughout the cohort period, the investigator submits a list of all serious expected and unexpected adverse reactions and all serious adverse events that occurred during the period. The report includes an assessment of the safety of the research participants. There are no expected adverse reactions or events anticipated in the current cohort as only the blood sample collected at GA20 is invasive. The cohort is monitored by the investigator, UKS.

Cohort results are communicated via publication in international journals and subsequent to participants, healthcare professionals and the public. Authorship eligibility are based on International Committee of Medical Journal Editors (ICMJE) recommendations.

Patient and public involvement

No patients involved.

Information collected from study participants

Information not related to the biological material of study participants is collected continuously throughout the study period (table 3). All information is stored online using REDCap electronic data capture tools hosted at Faculty of Health, Aarhus University,^{43 44} in which each participant is given a unique survey identification number.

Sampling of biological material

Several samples are collected from week 20 in the mother's pregnancy to the respective infant is 5 years old (summarised in table 4). All samples collected by the participants at their home are placed in their own freezer at minimum -18°C .

Blood sample collection

A blood sample is collected from the mothers at GA 20–22 and specific parameters (box 2) are measured by medical laboratory technicians employed by AUH. One

Table 3 Information gathered from participants from recruitment until 5 years post partum

Time	Description	Collection instrument
Before birth	Personal information: name, age, weight, height, address, email, marital status, level of education, ethnicity, primary income level Pregnancy related: intention to breast feed, elective caesarean section, multiparous, former pregnancies Lifestyle related: diabetes diagnosis, medication (inflammatory bowel diseases, psychological disorders, coeliac disease, irritable bowel syndrome and/or metabolic disease), previous gastric bypass operation, received antibiotics the past 12 weeks, smoking, participation in scientific projects comprising intake of prebiotics or probiotics 2×24 hours dietary recall questionnaire	REDCap Myfood24
At birth	Date of birth, gestational age at labour, parity, infant weight and length at birth, whether mother had received antibiotics, use of labour induction, suckling at hospital, caesarean section, gestational diabetes mellitus	EPJ
After birth (1–3 months)	Related to breast feeding: breastfeeding frequency, use of infant formula, any complications of breast feeding, initiation of breast feeding, symptoms of mastitis Related to infant: infant weight and length, date of the respective anthropometric measures, antibiotics use Lifestyle related: smoking habits, overall smoking in household, alcohol use, maternal medication and antibiotic use, mother's last measured weight, symptoms of COVID-19 experienced for mother or other individual in household, confirmed cases of COVID-19 for mother or other individual in household (PCR or antigen test) 2×24 hours dietary recall questionnaire	REDCap Myfood24
After birth (6 months to 5 years)	Related to infant: infant weight and height, date of the respective anthropometric measures, whether infant receives solid foods, age of infant when initiation introduction of solid foods, whether infant receives breastmilk or infant formula in addition to solid foods, frequencies of breast feeding, whether infant has received antibiotics during the past 3 months	REDCap

EPJ, electronic e-journal; REDCap, Research Electronic Data Capture.

Table 4 Biological material collected throughout the MAINHEALTH cohort including time of collection and the participant from which the material derives

Biological material	Time of collection	Samples	Participant/Person responsible for collection
Blood sample	GA 20–22	1	Mother/midwife at AUH
Vaginal and rectum swab	Labour	one each	Mother/midwife
Breast milk	Within 1 week pp, +1, +2, +3 months pp	4	Mother/mother
Urine	+1, +2, +3 month pp	3	Infant/parent
Faeces	+1, +2, +3 month pp, +1–+5 years pp	8	Infant/parent
Saliva	+1 month pp	1	Infant/parent
Swab of breast skin area	+1 month pp	1	Mother

+, a given time point PP.

AUH, Aarhus University Hospital; GA, gestational age; pp, post partum.

additional EDTA-plasma vial is frozen at -80°C for future follow-up analysis.

Vaginal and rectal swab around time of labour

At time of admission for childbirth at AUH, a midwife collects swabs from mother's vagina and rectum using Copan ESwab collection and transport system (Copan diagnostics). Amniorrhexis status is recorded. In case amniorrhexis occurs after first swab is collected, another vaginal swab is collected. ESwabs are stored at -80°C

freezers at the obstetric ward. Rectal and vaginal swabs were collected to assess mother's microbiota during birth and to correlate with the infant's microbiota.

Human milk collection

HM is collected by the participants within the first week following birth, at 1, 2 and 3 months post partum (table 4). Prior, the participants have received written instructions in collection at home. Participants are asked to avoid sampling the first drops, to manually express the foremilk (about 5 mL at first sample and about 20 mL the next samplings) into a 40 mL sterile container (Corning, Fischer Scientific) around midday and minimum 2 hours after last breast feeding.

Infant urine collection

Infant urine collection is conducted by the participants at 1, 2 and 3 months post partum (table 4). Written instructions, cotton pads (Anglamark, Coop Trading A/S, Denmark), and a 40 mL sterile container (Corning, Fischer Scientific) are handed to the participants. They place a cotton pad in the infant's diaper, collect it after urination and place it in the sterile container.

Infant faeces collection

Infant faeces is collected in the first 3 months post partum and once a year after birth until the infant is 5 years of age (table 4). Participants receive written instructions and a faeces collection tube containing a small spatula for sampling (Sarstedt, Nümbrecht, Germany). They should collect about 2 g faeces and should be taken as the first faeces after collection of the respective HM sample.

Saliva and skin swab collection

At 1-month post partum, the mothers are asked to collect a skin swab from the breast used for collecting the milk sample and a saliva sample from the infant's oral cavity. This was done to estimate microbiota transferred from the mother to the infant, and whether this originates from the milk or the breast. Participants are equipped with a Copan ESwab collection and transport system

Box 2 Blood parameters measured in samples collected at gestational age 20

Parameter measured (unit).

Hemoglobin A1c (mmol/mol).
 Alanine transaminase (IU/L).
 Albumin (g/L).
 Basic phosphatase (IU/L).
 Bilirubin ($\mu\text{mol/L}$).
 Calcium (mmol/L).
 Carbamide (mmol/L).
 high-density lipoprotein Cholesterol (mmol/L).
 low-density lipoprotein Cholesterol (mmol/L).
 C reactive protein (mg/L).
 Estimated glomerular filtration rate .
 Creatine phosphokinase (IU/L).
 Folate (nmol/L).
 Gamma-glutamyltransferase (GGT) (IU/L).
 Glucose (mmol/L).
 Iron ($\mu\text{mol/L}$).
 Lactate dehydrogenase (IU/L).
 Parathyrin (pmol/L).
 Phosphate (mmol/L).
 Potassium (mmol/L).
 Thyroid-stimulating hormone (10–3).
 Transferrin ($\mu\text{mol/L}$).
 Triglycerides (mmol/L).
 Urate (mmol/L).
 Vitamin B₁₂ (pmol/L).
 One EDTA vial.

(Copan diagnostics) for sample and written instructions for sampling.

All samples collected at the participants' homes (HM and skin swabs, infant urine, faeces, and saliva) are frozen in participants' domestic freezer and are gathered within 14 days and transported on dry ice to a -80°C freezer at AU-FOOD.

Questionnaires

Recruitment screening

Participants receive a link to an online survey after GA 20, managed by REDCap that addresses the information described in [table 3](#) (before birth).

Mother's dietary recalls

Maternal dietary intake is recorded over four periods; during pregnancy around gestational week 37, and timed to the biological sample collections at 1, 2 and 3 months. The dietary records are used to assess dietary variability in the cohort hence the timing of recalls timed with biological sampling time points. Detailed information about diet and specific nutrients is preferred over dietary records collected over a larger period. Within each period, each participant is asked to fill out the recall twice; on a weekday and on a day in the weekend. Dietary intake is recorded online using MyFood24 as a self-administered 24 hours dietary recall tool.⁴⁵

Sample collection

Questions related to the sampling of the given material are provided which address date and time of collection, whether the sample is placed in a freezer within 30 min following collection, and if the participants have other comments related to the sampling. Furthermore, in relation to the collection of infant faeces, participants should indicate whether the sample has been collected from the infant's skin and/or diaper. For both infant urine and faeces, participants should estimate the interval since urination or defecation to sample collection. For collection of HM and skin swab, mothers are asked to denote whether the sample derives from the left or right breast.

Breastfeeding pattern, infant development and maternal lifestyle

During time of sampling the first 3 months, participants are asked to fill out an online survey using REDCap addressing estimations of breast feeding frequency and use of infant formula reported as 'always breastfed/used infant formula', 'to greater extent breastfed/used infant formula', 'to lesser extent breastfed/used infant formula' and 'not breastfed/used infant formula', after which the weekly or daily frequencies of breast feeding and/or infant formula use are estimated. Included in the survey are questions related to infant and mother's lifestyle elaborated in [table 3](#). Online surveys are additionally sent to the participants when their infant is 6 months and once every year from year 1 to 5 after birth ([table 3](#)).

Chemical analyses

Metabolomics

Human milk

The HM metabolome including HMOs will be analysed by NMR as described elsewhere.⁴⁶ Briefly, samples are skimmed by centrifugation at 4000g for 15 min. To extract metabolites, the samples will be filtered using Amicon Ultra 0.5 mL 10 kDa (Millipore, Billerica, Massachusetts, USA) spin filters at 10 000g for 60 min at 4°C . ^1H NMR spectroscopy is performed on a Bruker Avance Neo 600 spectrometer, at a ^1H frequency of 600.03 MHz, equipped with a 5 mm ^1H BBI probe and SampleJet (Bruker BioSpin, Rheinstetten, Germany).

In addition, the Biocrates MxP Quant 500 Targeted Metabolomics Kit (Biocrates, Innsbruck, Austria) will be used for HM metabolomics as described.⁴⁷ The kit can identify and quantify up to 13 different small molecule classes, hexoses and 12 lipid classes using MS.

Infant urine

The cotton pad containing urine is thawed and placed on a closed, sterile Eppendorf-tube in a 50 mL tube. The tubes are centrifuged at 2000g, 4°C for 5 min to extract the urine. The extracted urine is mixed with a phosphate buffer (50 mM Na_2HPO_4 , pH 7.00) and D_2O containing 0.05% 3-(trimethylsilyl)propanoic-2,2,3,3 acid, sodium salt (TSP; Sigma-Aldrich, St. Louis, Missouri, USA) mixture (1/1, v/v). ^1H NMR spectroscopy is performed on a Bruker Avance Neo 600 spectrometer (Bruker BioSpin). Bruker IVDr methods enable automated analysis of more than 40 metabolites in urine.

Infant faeces

Tubes containing infant faeces are thawed on ice and then centrifuged for 2 min at $5000\times g$, 4°C . Approximately, 250 mg faecal material of each sample is transferred to a Eppendorf-tube and 1400 mL phosphate-buffered saline buffer is added. Tubes are vortexed until the faecal material is dispersed. Supernatant and pellet are separated by centrifugation at $10.000\times g$, 10 min, 4°C . 1 mL supernatant is transferred to a new Eppendorf-tube. Samples are stored at -80°C . The pellet fraction is intended for microbial sequencing, and supernatant fraction for ^1H NMR analyses. ^1H NMR spectroscopy is performed on a Bruker Avance Neo 600 spectrometer (Bruker BioSpin).

Microbial sequencing

All samples are analysed in three batches consisting of ~60 complete sets of mother–infant samples. Samples of the same type will be randomised before DNA extraction and sequencing to minimise the impact of batch effect.

Preprocessing

The HM stored at -80°C , is thawed on ice and then spun down at 14.000 rpm, at 4°C for 20 min to separate the sample into three phases, a cream layer (top), a watery phase (middle) and the spun down pellet. The cream layer is discarded and the supernatant is stored for future research, and the pellet is isolated.

For all other samples (infant faeces and saliva, mother's rectal swab and skin), bacterial DNA is extracted from the pellet fraction of samples after centrifugation for 2 min at 5000× g, 4°C. Supernatant fractions are stored for possible later analysis.

DNA extraction

DNA is extracted from bacterial pellets using the Beat-Bead Micro AX gravity extraction kit from A&A biotechnology (Cat. # 106-20, 106-100 Versatile, increased efficiency kit for genomic DNA purification from various sources, Mechanical lysis, V.0820). The procedure follows the kit protocol and includes heat activated enzymatic lysis, mechanical lysis in bead beating tubes and filtering through silicafilters. Each extraction includes a positive control using a mock community live bacteria and a negative water control. The DNA concentration is determined using Qubit dsDNA high-sensitivity assay measured on a Varioskan Flash (Thermo Fisher Scientific, USA) spectrophotometer.

Sequencing

A 16S rRNA gene amplicon library is constructed by amplifying the 16S rRNA gene with unique molecular identifier containing multiple forward and reverse primers targeting the 16S V1-V9 regions. PCR conditions for the amplification as follows: 95°C for 5 min, 2 cycles of 95°C for 20s, 48°C for 30s, 65°C for 10s, 72°C for 45s and a final extension at 72°C for 4 min. A second PCR step is then performed to barcode PCR amplicons with the following conditions: 95°C for 2 min, followed by 33 cycles of 95°C for 20s, 55°C for 20s, 72°C for 40s and a final extension at 72°C for 4 min. After each PCR reaction, PCR amplicons are cleaned up using Speed-Beads magnetic carboxylate (Sigma Aldrich). The size of barcoded PCR products (approximately 1500 bp) is checked by 1.5% agarose gel electrophoresis.

Sequencing libraries consisting pooled barcoded PCR products from up to 196 samples is prepared by following the ligation sequencing protocol SQK-LSK110 (Oxford Nanopore Technologies, Oxford, UK) and loaded on R V.9.1.4 flow cell for 72 hours using Grid-IONX5 (Oxford Nanopore Technologies).

Proteomics and peptidomics

Human milk

To achieve deep proteome coverage of the HM samples, they will be analysed using the Bruker timsTOF pro 2 (Bruker Daltonics, Bremen, Germany). Proteins are solubilized, reduced, alkylated and digested following a modified protocol before they are subjected to liquid chromatography followed by MS.

Another focus is on post-translation modifications (PTMs) of HM proteins. For phosphoproteomics analysis, thawed HM is digested and enriched following a modified TiO₂ protocol.⁴⁸ To investigate the effect of PTMs on the generation of bioactive peptides from digestion, samples will be digested using a simulated gastrointestinal

digestion.⁴⁹ Digested milk proteins and simulated digests will be characterised by in-house qualitative and quantitative proteomics using 1D and 2D SDS-PAGE and LC-ESI MS/MS using the Bruker timsTOF pro 2 (Bruker Daltonics).

Statistical analyses

There are currently no standard methods for sample size estimation in combined omics studies. Moreover, strong multicollinearity between metabolites and multiple testing makes sample size estimates more complicated.^{50,51} A repeated measures design can be used to increase statistical power because it allows the detection of within-person change over time compared with cross-sectional designs. A recent study identified a relationship between cognitive function at 24 months and HMOs in a cohort of 50 participants in a cross-sectional design, indicating sufficient sample sizes in a similar study.⁵²

Data analysis will be performed on different levels as results from 16S sequencing, human milk metabolomics and proteomics, infant urine and faeces metabolomics will be analysed individually to find associations to mother's prepregnancy BMI and time of sampling. Multivariate and univariate analysis will be used for this. To integrate different omics data, correlation coefficients as Spearman and Pearson will be calculated for different concentration matrices.

ETHICS AND DISSEMINATION

The Central Denmark Region Committees on Health Research Ethics has approved the protocol (J-nr. 1-10-72-296-18). All participants have before inclusion signed informed consent and deputy informed consent in accordance with the Declaration of Helsinki II. The study is registered at ClinicalTrials.gov, identification number: NCT05111990.

The proposed research is anticipated to generate a very significant body of knowledge of interest to a wide range of stakeholders, from the clinical (paediatricians and other health professionals) to the research community in the area and industrial producers of both infant milk formula and ingredients for such products but also to the public. Hence, a multilayered dissemination strategy will be followed. The international scientific community will be informed via peer-reviewed publications and presentations at scientific conferences, whereas the industry will be invited to meetings to discuss our findings.

Author affiliations

¹Department of Food Science, Aarhus University, Aarhus N, Denmark

²Sino-Danish Centre for Education and Research (SDC), University of Chinese Academy of Sciences, Beijing, People's Republic of China

³Department of Food Science, University of Copenhagen, København, Denmark

⁴Department of Clinical Medicine, Aarhus University Hospital, Aarhus N, Denmark

Acknowledgements We would like to extend our gratitude towards families including the mothers and their infants included in the study, the sample collection team and recruitment team at Aarhus University and AUH, Rikke Alkærsig Larsen, Chatrine Bodholdt Sand, Ebbe Meldgaard, and Camille Reese. Furthermore, we

would like to thank Diana Høftoft for fruitful discussions in study conceptualisation, Filiz Kesgin and Hai Qing at AUH for their assistance with blood sample collection, and midwives for assisting in sample collection during births.

Contributors Conceptualisation: KOP, NU, DSN and UKS; Methodology: KOP, RRJ, JA, JF, DSN, UKS; writing—original draft: KOP; writing—review and editing: KOP, RRJ, JA, NU, JF, DSN and UKS; project administration, UKS; funding acquisition: NU, DSN and UKS.

Funding The work was partly funded by Arla Food for Health (Viby J, Denmark; grant number: N/A) The following other funding sources have supported the project: Sino-Danish Centre for Education and Research (SDC; KOP PhD funding; grant number: N/A), Graduate School of Technical Sciences, Aarhus University (Grant number: N/A) and University of Copenhagen (Grant number: N/A).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Katrine Overgaard Poulsen <http://orcid.org/0000-0002-7478-3779>

Rasmus Riemer Jakobsen <http://orcid.org/0000-0001-7416-1358>

Ulrik Kræmer Sundekilde <http://orcid.org/0000-0003-4849-0996>

REFERENCES

- World Health Organization. Infant and young child nutrition: global strategy on infant and young child feeding report by the Secretariat, 2002. Available: http://apps.who.int/gb/archive/pdf_files/WHA55/ea5515.pdf [Accessed 30 Jun 2022].
- Asbury MR, Butcher J, Copeland JK, *et al*. Mothers of preterm infants have individualized breast milk microbiota that changes temporally based on maternal characteristics. *Cell Host Microbe* 2020;28:669–82.
- Andreas NJ, Hyde MJ, Gomez-Romero M, *et al*. Multiplatform characterization of dynamic changes in breast milk during lactation. *Electrophoresis* 2015;36:2269–85.
- Thum C, Wall C, Day L, *et al*. Changes in human milk fat globule composition throughout lactation: a review. *Front Nutr* 2022;9:835856.
- Karatas Z, Durmus Aydogdu S, Dinleyici EC, *et al*. Breastmilk ghrelin, leptin, and fat levels changing foremilk to hindmilk: is that important for self-control of feeding? *Eur J Pediatr* 2011;170:1273–80.
- Villaseñor A, Garcia-Perez I, Garcia A, *et al*. Breast milk metabolome characterization in a single-phase extraction, multiplatform analytical approach. *Anal Chem* 2014;86:8245–52.
- Innis SM. Impact of maternal diet on human milk composition and neurological development of infants. *Am J Clin Nutr* 2014;99:734S–41.
- Bravi F, Wiens F, Decarli A, *et al*. Impact of maternal nutrition on breast-milk composition: a systematic review. *Am J Clin Nutr* 2016;104:646–62.
- Lönnerdal B, Iyer S. Lactoferrin: molecular structure and biological function. *Annu Rev Nutr* 1995;15:93–110.
- Lönnerdal B. Nutritional and physiologic significance of human milk proteins. *Am J Clin Nutr* 2003;77:1537S–43.
- Andreas NJ, Kampmann B, Mehring Le-Doare K. Human breast milk: a review on its composition and bioactivity. *Early Hum Dev* 2015;91:629–35.
- Guerrero A, Dallas DC, Contreras S, *et al*. Mechanistic peptidomics: factors that dictate specificity in the formation of endogenous peptides in human milk. *Mol Cell Proteomics* 2014;13:3343–51.
- Dallas DC, Guerrero A, Khaldi N, *et al*. A peptidomic analysis of human milk digestion in the infant stomach reveals protein-specific degradation patterns. *J Nutr* 2014;144:815–20.
- Lu J, Antunes Fernandes E, Páez Cano AE, *et al*. Changes in milk proteome and metabolome associated with dry period length, energy balance, and lactation stage in postparturient dairy cows. *J Proteome Res* 2013;12:3288–96.
- Grapov D, Lemay DG, Weber D, *et al*. The human colostrum whey proteome is altered in gestational diabetes mellitus. *J Proteome Res* 2015;14:512–20.
- Poulsen KO, Sundekilde UK. The metabolomic analysis of human milk offers unique insights into potential child health benefits. *Curr Nutr Rep* 2021;10:12–29.
- Saben JL, Sims CR, Piccolo BD, *et al*. Maternal adiposity alters the human milk metabolome: associations between nonglucose monosaccharides and infant adiposity. *Am J Clin Nutr* 2020;112:1228–39.
- Isganaitis E, Venditti S, Matthews TJ, *et al*. Maternal obesity and the human milk metabolome: associations with infant body composition and postnatal weight gain. *Am J Clin Nutr* 2019;110:111–20.
- Ninonuevo MR, Park Y, Yin H, *et al*. A strategy for annotating the human milk glycome. *J Agric Food Chem* 2006;54:7471–80.
- Bode L. The functional biology of human milk oligosaccharides. *Early Hum Dev* 2015;91:619–22.
- LoCasio RG, Niñonuevo MR, Freeman SL, *et al*. Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. *J Agric Food Chem* 2007;55:8914–9.
- Sela DA. Bifidobacterial utilization of human milk oligosaccharides. *Int J Food Microbiol* 2011;149:58–64.
- Bode L. Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* 2012;22:1147–62.
- Saben JL, Sims CR, Abraham A, *et al*. Human milk oligosaccharide concentrations and infant intakes are associated with maternal overweight and obesity and predict infant growth. *Nutrients* 2021;13:446–16.
- Gilley SP, Ruebel ML, Sims C, *et al*. Associations between maternal obesity and offspring gut microbiome in the first year of life. *Pediatr Obes* 2022;17:e12921.
- van Leeuwen SS. Challenges and pitfalls in human milk oligosaccharide analysis. *Nutrients* 2019;11:2684.
- Newburg DS, Morelli L. Human milk and infant intestinal mucosal glycans guide succession of the neonatal intestinal microbiota. *Pediatr Res* 2015;77:115–20.
- Boix-Amorós A, Collado MC, Mira A. Relationship between milk microbiota, bacterial load, macronutrients, and human cells during lactation. *Front Microbiol* 2016;7:492.
- Kordy K, Gaufin T, Mwangi M, *et al*. Contributions to human breast milk microbiome and enteromammary transfer of Bifidobacterium breve. *PLoS One* 2020;15:e0219633–10.
- Moore RE, Townsend SD. Temporal development of the infant gut microbiome. *Open Biol* 2019;9:190128.
- Sprenger N, Binia A, Austin S. Human milk oligosaccharides: factors affecting their composition and their physiological significance. *Nestle Nutr Inst Workshop Ser* 2019;90:43–56.
- Gohir W, Ratcliffe EM, Sloboda DM. Of the bugs that shape us: maternal obesity, the gut microbiome, and long-term disease risk. *Pediatr Res* 2015;77:196–204.
- Casado B, Affolter M, Kussmann M. OMICS-rooted studies of milk proteins, oligosaccharides and lipids. *J Proteomics* 2009;73:196–208.
- Rebollar EA, Antwis RE, Becker MH, *et al*. Using "Omics" and Integrated Multi-Omics Approaches to Guide Probiotic Selection to Mitigate Chytridiomycosis and Other Emerging Infectious Diseases. *Front Microbiol* 2016;7:68.
- Bode L, Raman AS, Murch SH, *et al*. Understanding the mother-breastmilk-infant "triad". *Science* 2020;367:1070–2.
- Lönnerdal B. Infant formula and infant nutrition: bioactive proteins of human milk and implications for composition of infant formulas. *Am J Clin Nutr* 2014;99:712S–7.
- Lönnerdal B. Bioactive Proteins in Human Milk: Health, Nutrition, and Implications for Infant Formulas. *J Pediatr* 2016;173 Suppl:S4–9.
- Kelishadi R, Farajian S. The protective effects of breastfeeding on chronic non-communicable diseases in adulthood: a review of evidence. *Adv Biomed Res* 2014;3:3:3.
- Isaacs EB, Morley R, Lucas A. Early diet and general cognitive outcome at adolescence in children born at or below 30 weeks gestation. *J Pediatr* 2009;155:229–34.
- Victoria CG, Bahl R, Barros AJD, *et al*. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet* 2016;387:475–90.
- Esch BCAMvan, Porbahaie M, Abbring S, *et al*. The impact of milk and its components on epigenetic programming of immune function in early life and beyond: implications for allergy and asthma. *Front Immunol* 2020;11:2141.



- 42 Chan A-W, Tetzlaff JM, Göttsche PC, *et al.* Spirit 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
- 43 Harris PA, Taylor R, Minor BL, *et al.* The REDCap Consortium: building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.
- 44 Harris PA, Taylor R, Thielke R, *et al.* Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
- 45 Carter M, Albar S, Morris M, *et al.* Development of a UK online 24-h dietary assessment tool: myfood24. *Nutrients* 2015;7:4016–32.
- 46 Sundekilde UK, Downey E, O'Mahony JA, *et al.* The effect of gestational and lactational age on the human milk metabolome. *Nutrients* 2016;8:304.
- 47 Hampel D, Shahab-Ferdows S, Hossain M, *et al.* Validation and application of Biocrates AbsoluteIDQ® p180 targeted metabolomics kit using human milk. *Nutrients* 2019;11:1733.
- 48 Engholm-Keller K, Birck P, Størling J, *et al.* TiSH — a robust and sensitive global phosphoproteomics strategy employing a combination of TiO₂, SIMAC, and HILIC. *J Proteomics* 2012;75:5749–61.
- 49 Brodkorb A, Egger L, Alminger M, *et al.* INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nat Protoc* 2019;14:991–1014.
- 50 Nyamundanda G, Gormley IC, Fan Y, *et al.* MetSizeR: selecting the optimal sample size for metabolomic studies using an analysis based approach. *BMC Bioinformatics* 2013;14:338.
- 51 Blaise BJ, Correia G, Tin A, *et al.* Power analysis and sample size determination in metabolic phenotyping. *Anal Chem* 2016;88:5179–88.
- 52 Berger PK, Plows JF, Jones RB, *et al.* Human milk oligosaccharide 2'-fucosyllactose links feedings at 1 month to cognitive development at 24 months in infants of normal and overweight mothers. *PLoS One* 2020;15:e0228323.
- 53 Samuel TM, Zhou Q, Giuffrida F, *et al.* Nutritional and Non-nutritional composition of human milk is modulated by maternal, infant, and methodological factors. *Front Nutr* 2020;7:172.
- 54 Bzikowska-Jura A, Czerwonogrodzka-Senczyna A, Olędzka G, *et al.* Maternal nutrition and body composition during breastfeeding: association with human milk composition. *Nutrients* 2018;10:1379.
- 55 Nommsen LA, Lovelady CA, Heinig MJ, *et al.* Determinants of energy, protein, lipid, and lactose concentrations in human milk during the first 12 Mo of lactation: the Darling study. *Am J Clin Nutr* 1991;53:457–65.
- 56 Kugananthan S, Gridneva Z, Lai C, *et al.* Associations between maternal body composition and appetite hormones and macronutrients in human milk. *Nutrients* 2017;9:252.