

The Women4Health cohort: a unique cohort to study women-specific mechanisms of cardio-metabolic regulation

Fabio Busonero ^{1,†}, Stefania Lenarduzzi ^{2,†}, Francesca Crobu ¹,
Roberta Marie Gentile³, Andrea Carta ⁴, Francesco Cracco³, Andrea Maschio ¹,
Silvia Camarda³, Michele Marongiu ¹, Daniela Zanetti ¹,
Claudio Conversano ^{1,4}, Giovanni Di Lorenzo ², Daniela Mazzà ^{2,*},
Francesco De Seta ², Giorgia Giroto ^{2,3,‡}, and Serena Sanna ^{1,5,*‡}

¹Institute of Genetic and Biomedical Research (IRGB), National Research Council (CNR), c/o Cittadella Universitaria di Monserrato, SS554 Km 4500, Monserrato, 09042, CA, Italy; ²Institute for Maternal and Child Health—IRCCS ‘Burlo Garofolo’, Via dell’Istria 65/1, Trieste, 34137, TS, Italy; ³Department of Medicine, Surgery and Health Sciences, University of Trieste, Piazzale Europa 1, Trieste, 34137, TS, Italy; ⁴Department of Business and Economics, University of Cagliari, via Università 40, 09124, Cagliari, CA, Italy; and ⁵Department of Genetics, University Medical Center Groningen, Hanzeplein 1, 97123 GZ, Groningen, The Netherlands

Received 6 December 2023; revised 22 February 2024; accepted 26 February 2024; online publish-ahead-of-print 27 February 2024

Handling Editor: Salvatore De Rosa

Aims

Epidemiological research has shown relevant differences between sexes in clinical manifestations, severity, and progression of cardiovascular and metabolic disorders. To date, the mechanisms underlying these differences remain unknown. Given the rising incidence of such diseases, gender-specific research on established and emerging risk factors, such as dysfunction of glycaemic and/or lipid metabolism, of sex hormones and of gut microbiome, is of paramount importance. The relationships between sex hormones, gut microbiome, and host glycaemic and/or lipid metabolism are largely unknown even in the homeostasis status. Yet this knowledge gap would be pivotal to pinpoint to key mechanisms that are likely to be disrupted in disease context.

Methods and results

Here we present the Women4Health (W4H) cohort, a unique cohort comprising up to 300 healthy women followed up during a natural menstrual cycle, set up with the primary goal to investigate the combined role of sex hormones and gut microbiota variations in regulating host lipid and glucose metabolism during homeostasis, using a multi-omics strategy. Additionally, the W4H cohort will take into consideration another ecosystem that is unique to women, the vaginal microbiome, investigating its interaction with gut microbiome and exploring—for the first time—its role in cardiometabolic disorders.

Conclusion

The W4H cohort study lays a foundation for improving current knowledge of women-specific mechanisms in cardiometabolic regulation. It aspires to transform insights on host–microbiota interactions into prevention and therapeutic approaches for personalized health care.

Keywords

Women • Cardiovascular diseases • Lipids • Sex hormones • Microbiome • Multi-omics study

* Corresponding authors. Tel: +39 0403785538, Email: daniela.mazza@burlo.trieste.it; Tel: +39 0706754593, Email: serena.sanna@irgb.cnr.it

† Jointly contributed to this work.

‡ Jointly supervised the study.

© The Author(s) 2024. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Cardiometabolic disorders (CMD), a group of conditions that include cardiovascular diseases (CVD), type 2 diabetes (T2D) and metabolic syndrome, are characterized by dysfunction of glycaemic and/or lipid metabolism. They are a leading cause of morbidity and mortality worldwide, with a death rate that is constantly increasing. The complexity of CMD is amplified by relevant differences in clinical manifestations, severity, and progression between sexes,^{1–3} the underlying mechanisms of which are largely unknown. Prognosis is poorer in women—for example, despite the higher age-standardized incidence of CVD observed in men than women, in Europe, CVD accounted for 39% of all deaths in men and 46% of all deaths in women.⁴ Therefore, there is a global, urgent need for better women's tailored prevention and treatment strategies.⁵

Genetic, intrinsic and extrinsic factors contribute to CMD, nonetheless female-specific mechanisms of CMD and their risk factors are only minimally explained by genetics.^{6–9} Among the extrinsic factors associated with CMD and their risk factor, particular attention has been given recently to the gut microbiota, the community of microbes inhabiting the gastrointestinal tract. In both humans and mice, significant imbalance of the gut microbiome composition and function (dysbiosis), and in gut microbiota-produced metabolites measured in faeces, has been associated with CMD and increased CMD-related risk factors,^{10–19} including insulin sensitivity and T2D,^{20,21} cholesterol levels,^{12,22,23} and atherosclerosis.¹³ There is also emerging evidence that female sex hormones could directly influence the gut microbiome: studies in humans and mice showed that large hormonal changes, such as those occurring during menopause and pregnancy, are associated with reduced species richness and diversity of the gut microbiome.^{24–26} It is thus plausible to think that large variations in female's sex hormones, impacting on the gut microbiome, could explain the sex dimorphism in CMD.

Conversely, it is understudied if small variation in sex hormones during homeostasis, e.g. those occurring regularly during the menstrual cycle, could influence gut microbiome. A recent study carried out in 20 women found nominal evidence for a role of progesterone in influencing temporal microbiota variation.²⁷ While the association was not significant after multiple testing, the observation remained inconclusive not only for the small sample size but also for the reduced precision on sex hormones values, which were inferred based on the day of the menstrual cycle rather than being directly measured. Previous studies on female's health have shown that variations in sex hormones during the menstrual cycle are significantly associated with changes in CMD-risk factors, specifically lipid levels, insulin resistance, and transaminases.^{28–30} Therefore, the menstrual cycle offers a natural endeavour to study simultaneous variation of sex hormones, gut microbiome, and CMD-related metabolism, without the need of any intervention to modulate sex hormones.

Here we present the Women4Health (W4H) cohort, an observational study of 300 women during a natural menstrual cycle. The study aims to investigate for the first time the hypothesis that the cross-talk between sex hormones and gut microbiome results in differential host metabolism regulation in females also during homeostasis and in the reproductive age.

A multi-omics strategy, including genomic, metagenomic, metabolomic, and proteomic approaches, will be used to characterize biological and physiological mechanisms.

The other unique feature of the W4H study is the collection of bio-samples to investigate the vaginal microbiome in the context of cardiometabolic regulation. While vaginal dysbiosis has been associated with gynaecological disorders (such as bacterial vaginosis, higher predisposition to local infections, miscarriages, and preterm births),^{31–33} it has never been explored whether the resulting local inflammation can perturbate nearby organs, including the gut, thus potentially affecting gut microbiome composition and function and its role in regulating host metabolism.

Here we describe the W4H cohort's study protocol and its unique (bio)data collection.

Methods

Study organization

Women4Health is a study cohort created by a joint collaborative effort between the women's and children's hospital IRCCS Burlo-Garofolo in Trieste, Italy, (P.I.: Prof. Giorgia Girotto) and the Institute for Genetics and Biomedical Research of the National Research Council (IRGB-CNR) in Cagliari, Italy (P.I.: Dr. Serena Sanna). The study was designed by Dr. Serena Sanna (IRGB-CNR), and fundings were gathered by her and by Prof. Giorgia Girotto (see 'Funding' statement). The protocol was designed by researchers from both Institutes (see 'Authors contribution' section). Daily management of volunteers' recruitment, collection of informed consent forms and biological specimens, measurement of sex hormones, and blood biochemistry parameters is performed by the IRCC Burlo-Garofolo team. All biological samples are then shipped to IRGB-CNR for long-term storage and processing. Database management and statistical data analysis are also performed by IRGB-CNR researchers. Project advertising and communication activities are carried out by several members of both research groups.

Study design

The Women4Health study aims to recruit 300 women; recruitment is currently carried out at the women's and children's hospital IRCCS Burlo-Garofolo in Trieste (Italy), according to the protocol described below and depicted in [Figure 1](#).

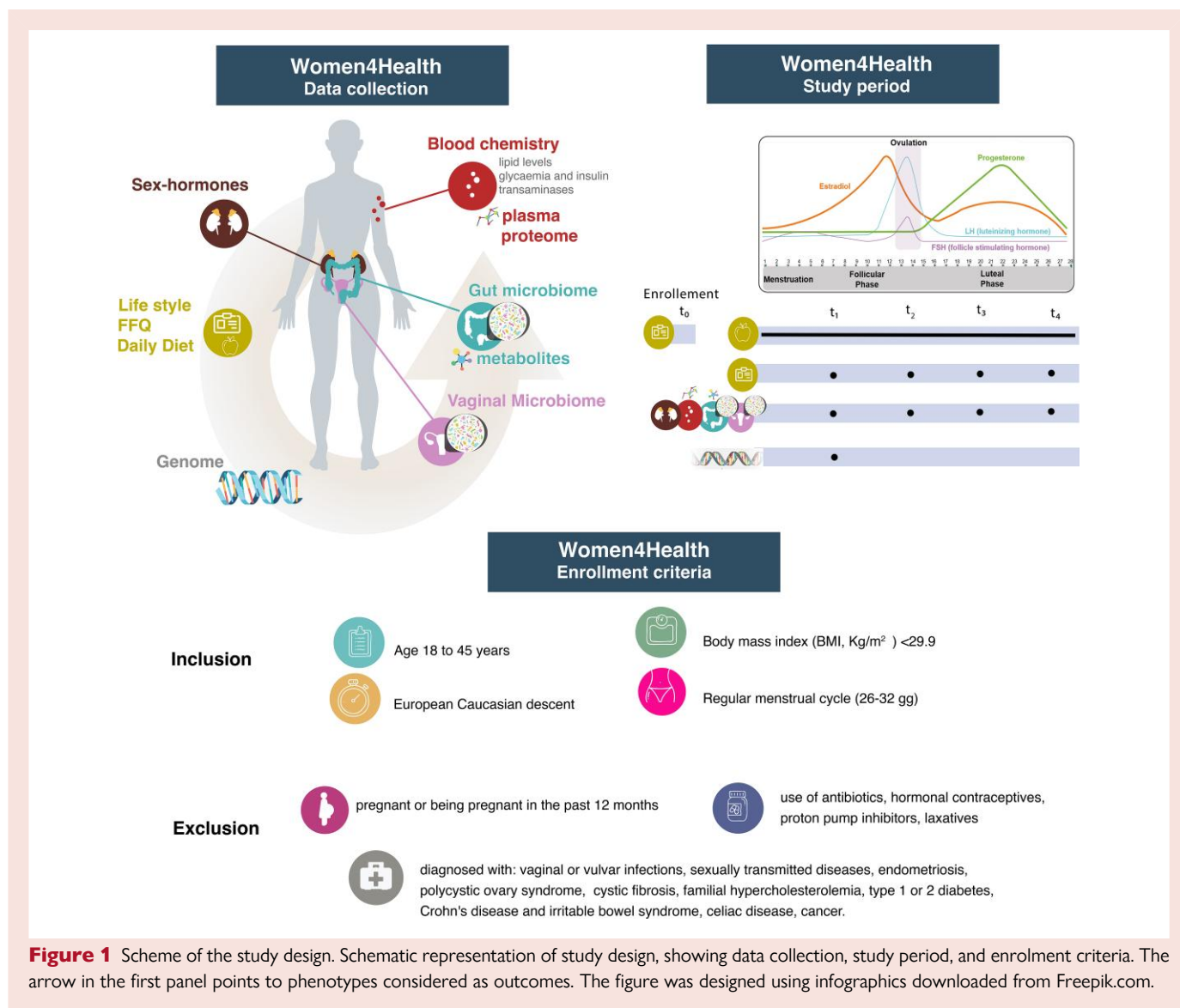
Upon inclusion and signature of informed consent, volunteers are enrolled in the study and information on family history, clinical and anthropometric data, lifestyle, and diet information are collected via questionnaires. Volunteers are asked to record their daily diet and sport at the first day of the menstrual cycle and for up to 28 days or until the next menstrual cycle. They are then scheduled for follow-up at four different time points: the 1st at the follicular phase (between days 6 and 9), the 2nd at the ovulation phase (between Days 12 and 16), the 3rd at the early luteal phase (between Days 18 and 22), and the 4th at late luteal phase (between Days 24 and 28). Depending on the day at enrolment, volunteers can start at any of the follow-up visits and continue with the remaining points of observations at the next menstrual cycle. At these time points, two stool and one vaginal swab specimens are self-collected by volunteers at home, then handed at the Obstetric and Gynaecologic (ObG) Clinic of the IRCCS Burlo-Garofolo no later than 24 h after collection. On the day biospecimens are handed to the ObG Clinic, women undergo a blood withdrawal after overnight fasting and fill in questionnaires to report health status in the week preceding collection and information on collected biological samples. Furthermore, a saliva sample for DNA extraction is collected (on the first visit only).

Consent

The study protocol—including informed consent forms—was approved by the Friuli-Venezia Giulia Ethical Committee on 09/2021 with prot. N. 0034184/P/GEN/ARCS and modifications amended on 05/09/2023.

Eligibility criteria

The W4H study enrolls women aged between 18 and 45 years old, an age range selected to avoid enrolling women in their pubertal or pre-menopausal phase. To avoid the introduction of confounding factors due to diseases or ongoing therapies, the cohort has been designed as follows: non-obese (BMI 17–29.9) women with a regular menstrual cycle (25–34 days), who have not made use in the previous 30 days of systemic or local hormonal contraceptives, antibiotics, proton-pump inhibitors, metformin, statin, and laxatives, and who do not have ongoing vaginal/vulvar infections or sexually transmitted diseases. Additionally, they must not have been diagnosed with CMD, gastrointestinal diseases, inherited metabolic disorders (cystic fibrosis), diabetes, familial hypercholesterolaemia and endometriosis or polycystic ovary syndrome, or have not been subjected to cancer therapy. Finally, we also exclude women who have been pregnant in the last 12 months. To assure genetic homogeneity across participants, we focus on women of European ancestry only, who are the majority of women living in the area of Trieste.



Recruitment protocol

Potential participants are invited in the study via the ObG Clinic, via advertisement within the hospitals and research centres in Trieste, via announcements on the Women4Health social media pages (Facebook and Instagram), via local radio and press communications, and other outreach activities. A leaflet containing details about W4H study and the abovementioned criterion for enrolment is posted in our social media pages and distributed in the hospital halls, universities, and aggregation centres.

Women who contact the recruiting team via email are invited for a meeting at the ObG clinic of the IRCCS Burlo-Garofolo and informed about the purposes of the research project and protocol of the study. Those interested in study participation and who meet the recruitment criteria are asked to sign an informed consent form to (i) agree to participate in the study and (ii) for the collection of data and biological samples to be used for research purposes only. Each enrolled volunteer is assigned a numerical code and any pseudonymized participant's information entered into the *ad hoc* created database. Participation is on a voluntary basis without any financial remuneration.

Data collection and management

Data collection started in February 2022, after infection from COVID-19 started to decrease in incidence in Italy due to the vaccination campaign.

At enrolment, women are asked to fill in an electronic questionnaire containing information about demographic and anthropometric measurements, detailed medical and gynaecological history, family disease history, direct or indirect exposure to smoke, use of pre-/pro-biotics, and dietary habits using a frequency food questionnaire validated in the Italian population.³⁴ During each follow-up visit, another questionnaire is administered to collect information regarding weekly bowel movements and stool consistency using the Bristol scale charts,³¹ history of COVID-19 vaccination or infection, use of medications or supplements, and sexual activity in the week preceding the appointment. Questionnaires are filled in by volunteers on an electronic form created and managed using the Research Electronic Data Capture (RedCap) tools hosted at IRGB-CNR.^{35,36} REDCap is a secure, web-based software platform designed to support data collection for research studies, providing (1) an intuitive interface for data entry; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for data integration and interoperability with external sources.

Daily diet and physical activity information are recorded by the volunteers themselves using MyFitnessPal, a smartphone app validated for scientific research studies.^{27,37} Daily diaries are shared by volunteers via email at the end of the observation period as PDF files.

A Data Management Plan has been set up and will be updated every 6 months.

Collection, process, and storage of biological samples

Kits for self-collection of saliva, stool, and vaginal swabs are distributed to volunteers at enrolment along with proper instructions. To facilitate self-collection and assure robustness to room temperature, we use the following collection kits from DNA Genotek®: Oragene-DNA (OG-500) for saliva, OMNImet-GUT (ME-200) and OMNIgene-GUT (OM-200) bundle for faces, and OMNIgene-VAGINAL (OMR-130) for vaginal swabs. After collection, volunteers store and hand those at room temperature to the ObG Clinic within 24 h; collection kits are then promptly stored at the appropriate temperature, as recommended by the manufacturer. Blood is collected from venepunctures in silica gel-Vacutainer, and allowed to clot by leaving it undisturbed for 1 h at room temperature (RT), then 0.5 mL of serum, obtained by centrifugation at 4000 r.p.m. (equivalent to 2990 g) for 5 min at RT, is immediately transferred into two sterile safe-lock polypropylene tubes and stored at -20°C . Plasma is separated from whole blood and collected into citrate-treated tubes by centrifugation at 3000 r.p.m. (equivalent to 1680 g) for 10 min at RT. A total of 0.5 mL of plasma is then transferred into two sterile safe-lock polypropylene tubes and stored at -80°C . Serum is used to assess sexual hormones [progesterone, 17beta-oestradiol (17BE), follicle-stimulating hormone (FSH), luteinizing hormone (LH) and prolactin, fasting insulin, thyroid-stimulating hormone TSH, free thyroxine (FT4), and testosterone (last three only at first visit) using ADVIA Centaur CP Immunoassay System (Siemens Healthineers AG)]. An aliquot of serum is used to evaluate cardiometabolic profile [glycaemia, total-, high-density lipoprotein-, and low-density lipoprotein cholesterol (TC, HDL-C, and LDL-C, respectively), and triglycerides, as well as alanine aminotransferase (ALT) and aspartate aminotransferase (AST)] using VChemistry S (Voden Medical Instrument SpA) and stored at IRCCS Burlo-Garofolo. All other biological samples (plasma, saliva, stool samples, and vaginal swabs) are shipped to the Institute for Genetic and Biomedical Research—National Research Council (IRGB-CNR) on a regular basis (every 21 days) in dry ice (stool in tube for metabolomics analyses and plasma aliquots), and at RT (saliva, vaginal samples, and stool in tube for DNA analyses). Once biological samples are received, human genomic DNA is extracted from saliva and stored at -20°C . Microbial DNA is isolated from stool (tubes for DNA analyses) and vaginal swabs, using standard protocols and DNA stored at -80°C .^{38,39} Aliquots of plasma and stool samples (tubes for metabolomic analysis) are instead directly stored at -80°C .

Plans for omics data generation

Genomic, metagenomic, metabolomic, and proteomic approaches will be used to characterize biological and physiological mechanisms.

For DNA extracted from saliva samples, we will use whole-genome genotyping arrays coupled with genotype imputation to characterize the human genome.⁴⁰ For DNA extracted from stool samples, whole-metagenomic shot-gun sequencing will be employed to achieve a taxonomic resolution down to species- and strains-level and to obtain functional information through microbial pathways definition. Furthermore, faecal metabolome will be derived using high-throughput technologies; the specific platform has not yet been identified. For DNA extracted from vaginal samples, 16S rRNA plus ITS amplicon sequencing will be instead used to determine taxonomic profiles (bacteria, archaea, and fungi composition) down to at the genus level.

Finally, the plasma proteome—including hundreds of proteins related to cardiovascular and metabolic functions—will be characterized using the Olink Proteomics platform.

Statistical analysis

Descriptive statistics presented in this manuscript were derived using the R software for statistical computing (version 4.2.1).

For each sex hormone, we assessed the significance of overall changes across the four time points of collection using the non-parametric Friedman test for paired samples. Pairwise differences in trait means were assessed using the non-parametric Wilcoxon test for paired samples and derived Bonferroni adjusted *P*-values. In addition to the measured blood-chemistry traits, we also derived parameters of insulin resistance and beta-cell function using the

Table 1 Descriptive statistics

Phenotype (unit)	Min-max	Mean (SD)
Age (years)	21–40	28.17 (5.59)
Average duration of menstruations (days)	3–7	4.95 (0.80)
Average duration of menstrual cycle (days)	23–34	28.26 (1.88)
Height (cm)	154–183	167.20 (6.30)
Weight (kg)	48.00–82.50	61.29 (9.44)
BMI (kg/m ²)	17.71–29.23	21.09 (2.9)
TSH (μLU/mL)	0.07–6.33	1.28 (0.94)
FT4 (ng/dL)	0.54–1.78	1.01 (0.21)
Testosterone (ng/dL)	7.93–29.41	14.95 (5.10)
Glucose (mg/dL)	72–121	90.78 (7.77)
Insulin (mU/L)	0.8–14	4.40 (2.45)
HOMA-IR	0.17–3.30	1.01 (0.60)
HOMA-B	0.61–10.78	3.45 (2.15)
Total cholesterol (mg/dL)	137–268	188.8 (30.53)
HDL-C (mg/dL)	35–99	64.9 (13.07)
LDL-C (mg/dL)	52–191	109.1 (13.07)
Triglycerides (mg/dL)	33–145	75.14 (24.17)
AST (U/L)	12–43	18.66 (5.11)
ALT (U/L)	6–31	15.29 (5.14)

Descriptive statistics at enrolment for the 59 volunteers for whom at least one blood withdrawal is available. Statistics for phenotypes measured in blood are given only for the first visit.

BMI, body mass index; TSH, thyroid-stimulating hormone; FT4, free thyroxine; HDL-C, high-density lipoprotein cholesterol (mg/dL); LDL-C, low-density lipoprotein cholesterol; ALT, alanine transaminase (U/L); HOMA-IR, insulin resistance; HOMA-B, insulin sensitivity; SD, standard deviation.

Homoeostasis Model Assessment, as follows: $\text{HOMA-IR} = [\text{fasting insulin } (\mu\text{U/mL}) \times \text{fasting plasma glucose } (\text{mg/dL})] / 405$; $\text{HOMA-B} = 20 \times \text{fasting insulin } (\mu\text{U/mL}) / [\text{fasting plasma glucose } (\text{mg/dL}) - 63]$.⁴¹

Code used for data clean-up, statistical analyses, and visualization of results is posted as soon as been validated in a freely available GitHub repository, at the following link: <https://github.com/Sanna-s-LAB/Women4Health/>.

Recruitment, adherence rates, and basic descriptive statistics

From February 2022–August 2023, we enrolled 86 volunteers with an average of 4.7 volunteers per month. Among these, 3 have withdrawn their consent to the study before any follow-up visit and 59 engaged for at least one follow-up visit, resulting in an adherence rate of 69%. No difference in mean age was observed between volunteers that did not engage in any follow-up visits and those who returned for at least one (two-sided Wilcoxon-rank $P = 0.93$).

In total, 47 volunteers completed all four visits, while the remaining completed only three or less visits for different reasons: COVID-19 infection (4), menstruation arrived earlier (3), work-related issues (3), car accident (1), and issues on recording diet information (1). Therefore, sample size is different at each follow-up visit (59 samples at visit 1, 58 samples at visit 2, 53 samples at visit 3, 47 samples at visit 4).

Volunteers who engaged for at least one visit were aged 28 years on average (age range was from 21–40 years), and the average length of the menstrual cycle and menses were 28 days and 4.9 days, respectively (Table 1). All volunteers had sex hormonal levels within the expected range during reproductive age, and none showed evidence of endocrine diseases based on TSH, FT4, and testosterone levels measured at first visit.

Distribution of sex hormones across time points reflect the expected variation during the menstrual cycle phases, with 17beta-oestradiol showing

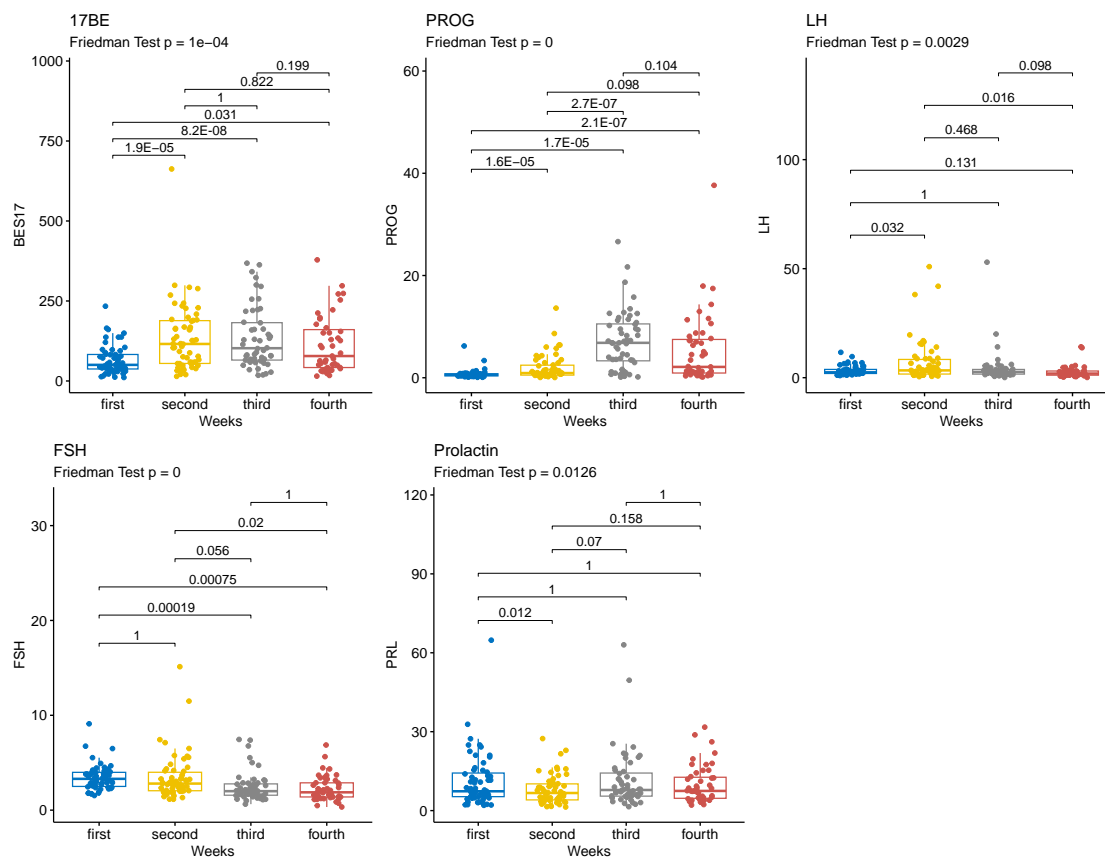


Figure 2 Variation of sex hormones by week of collection. Box plots describing the distribution of sex hormones at each time point. In each box plot, the tick line represents the median while the edges of the box represent the 25% and 75% of the distribution; whiskers represent 5% and 95% percentiles. For each trait, *P*-values of both the Friedman test for overall differences and the Wilcoxon test for pairwise differences are reported. 17BE, 17beta-oestradiol (pg/mL); PROG, progesterone (ng/mL); LH, luteinizing hormone (mIU/mL); FSH, follicular-stimulating hormone (mIU/mL); PRL, prolactin (ng/mL).

a sharp significant increase in the ovulation phases (second time point) and progesterone in the mid luteal phase (Figure 2).

Discussion

The Women4Health is the first longitudinal study that focuses on the menstrual cycle to evaluate the impact of female sex hormonal variations on commensal bacteria and CMD-related markers. Mimicking a perturbation study by following up physiological hormonal changes during the menstrual cycle, invasive hormonal treatments necessary for an intervention study will be avoided. Compared to a cross-sectional study, the longitudinal design adds the unique opportunity to understand the dynamics and derive causal relationships between components under study. The longitudinal component is in fact extremely relevant in the metagenomic field to understand the causal role of microbiome in phenotypic changes, as well as disease onset and progression.^{26,42,43} Moreover, the W4H study has a short longitudinal observation time restricted to a maximum of 30 days per person, allowing for proper control in analyses of major confounders of the microbiome, such as diet and medication use. In contrast, since menopause is a long process lasting several years, a longitudinal study targeting hormonal changes in menopause would need to follow women for up to several years, making the collection and adjustment for confounding factors unfeasible.

Furthermore, by studying females in the homeostasis status *bias from reverse causality* (diseases caused mechanisms) and *therapy-induced effects* are absent in the Women4Health cohort, biases that would instead be present in a *case-control study* of women in menopause with and without a diagnosis of CMD. Notably, studying the healthy population to learn potential mechanisms of diseases has been a successful strategy in genome-wide association studies. In fact, these have shown that genes associated with variation of quantitative traits in the general population harbour variants associated with both complex diseases and Mendelian disorders.^{44–47}

Results from the W4H study fill in the current knowledge gap on female-specific mechanisms underlying cardiovascular and metabolic functions during homeostasis and suggest putative targets to follow-up in designed cohorts of CMD patients. The identification of specific vaginal and gut microbiome features, associated with changes in host metabolism during homeostasis and disease, will provide potential unexplored routes for prevention of CMD in women through screening of microbiome signatures and treatments through microbiome modulation. This will be of great worth considering that, despite the high mortality, CMD in women remains understudied, under-recognized, underdiagnosed, and undertreated and research studies oriented on the identification of sex-specific mechanisms and sex-specific risk factors underlying the pathophysiology of these disorders are recognized as a priority of the medical research field.⁵

Noteworthy, the W4H study has high potential to boost future female-specific research studies for other complex diseases. In fact, it will provide the first comprehensive evaluation of gut and vaginal microbiota variation within a menstrual cycle, filling an existing knowledge gap on sex hormones–microbiome interaction that is essential for understanding not only CMD, but potentially many other diseases with a clear sex dimorphism and for which a putative role of sex hormones and gut microbiome has been suggested.

The Women4Health cohort study also has challenges and limitations. The most compelling challenge is the recruitment of women who do not use local or systemic hormonal contraceptives. While in Italy the percentage of females using these medications is very low (20%) compared to other European countries (50% in North Europe)⁴⁸ still this inclusion criteria needs to be added to others, such as absence of other diseases, pregnancy, as well as the determination of the volunteers to comply with a strict follow-up protocol of 30 days. The cohort is limited towards the characterization of cardiovascular biomarkers. While we measure lipids as the main risk factor for cardiovascular diseases, and hundreds of circulating proteins relevant for the cardiometabolic system, we do not have information on other relevant measurements of CVD risk, such as body fat composition, or measurements of cardiac health, such those that could be derived with an electrocardiogram. Future collaborative expansions of the Women4Health cohort will take this into consideration.

In conclusion, the unique design of Women4Health cohort study establishes the groundwork for improving current knowledge of women-specific mechanisms involved in cardiometabolic regulation. These insights have the potential to inform future translational applications aiming to convert gained knowledge on host–microbiota interactions into approaches for prevention and therapy through microbiota modulation, ultimately contributing to the development of personalized strategies.

Lead author biography



Dr Serena Sanna is Director of Research and Principal Investigator at the Institute of Genetic and Biomedical Research of the CNR, Italy. She is an internationally recognized researcher in the field of genetic epidemiology of complex traits and diseases. With her group, her research focuses on understanding the role of the genome and human microbiome interaction in the regulation of health and diseases. Key components of her studies are special cohorts, cutting edge technologies, and state-of-the art statistical methods.

Authors' contribution

All authors have read and approved the final version of the manuscript. Writing manuscript draft: F.B., F.Cro., A.M., A.C., S.S.; revision and edits: S.L., S.C., D.Z., C.C., F.D.S., D.M.; study conceptualization: S.S.; protocol definition: F.B., S.L., F.Cro., A.M., F.D.S., D.M., G.G., S.S.; data collection: S.L., R.M.G., F.Cra., G.D.L., F.D.S., D.M., G.G., S.S.; data management and analyses: A.C., M.M., S.S.; process of biological samples: F.B., S.L., F.Cro., A.M., S.C., D.M.; advertising and outreach: F.B., S.L., S.C., F.Cro., D.M., G.G., S.S.; study supervision: F.D.S., G.G., S.S.; funding: G.G., S.S.

Acknowledgements

We are grateful to all volunteers for their effort in participating in the study. We thank physicians Elena Stefani, Michele Stracquadini, Chiara

Rizzardi, Matilde Battistella, and Alma Kadric of the University of Trieste for their contribution to the recruitment phase, the Directors of I.R.C.C.S. Burlo-Garofolo and of IRGB-CNR for logistic support, the Agenzia Regionale Sardegna Ricerche for providing access to laboratories, and technologist Marco Masala at IRGB-CNR for IT support.

Funding

This study is funded by the European Union (ERC Stg 2022 to S.S., acronym SEMICYCLE, GA n.101075624) and by the Italian Ministry of Health through the contribution given to the Institute for Maternal and Child Health IRCCS Burlo Garofolo, Trieste - Italy (SD 02/21 to G.G.). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. In addition, we are pleased to mention other funds that indirectly made this study possible. In particular, we acknowledge: Marie-Curie Fellowship to D.Z., acronym 'Sex Dimorphism', GA n. 101066678; Next Generation EU grant Project GRINS (PE00000018-CUP F53C22000760007) to C.C., D70-REAR GIROTTO grant to G.G., Next Generation EU grant Project Age-It (PE0000015-CUP B83C22004880006) to S.S., and NutrAGE grant (CNR Project FOE-2021 DBA.AD005.225) to S.S.

Conflict of interest: None declared.

Data availability

We are open for collaboration with research groups that would like to join this effort as an add-on using the infrastructure at IRGB-CNR and IRCCS Burlo-Garofolo or as a collaborative effort. Detailed information can be requested by email to the corresponding authors of this manuscript. Data for this study will be available after 2 years of the conclusion of the ERC project SEMICYCLE, which funds this effort (see 'Funding' statement).

References

- Gao Z, Chen Z, Sun A, Deng X. Gender differences in cardiovascular disease. *Med Nov Technol Devices* 2019;**4**:100025.
- EUGenMed Cardiovascular Clinical Study Group; Regitz-Zagrosek V, Oertelt-Prigione S, Prescott E, Franconi F, Gerds E, Foryst-Ludwig A, Maas AH, Kautzky-Willer A, Knappe-Wegner D, Kintscher U, Ladwig KH, Schenck-Gustafsson K, Stangl V. Gender in cardiovascular diseases: impact on clinical manifestations, management, and outcomes. *Eur Heart J* 2016;**37**:24–34.
- Westergaard D, Moseley P, Sorup FKH, Baldi P, Brunak S. Population-wide analysis of differences in disease progression patterns in men and women. *Nat Commun* 2019;**10**:666.
- Townsend N, Kazakiewicz D, Lucy Wright F, Timmis A, Huculeci R, Torbica A, Gale CP, Achenbach S, Weidinger F, Vardas P. Epidemiology of cardiovascular disease in Europe. *Nat Rev Cardiol* 2022;**19**:133–143.
- Vogel B, Acevedo M, Appelman Y, Bairey Merz CN, Chieffo A, Figtree GA, Guerrero M, Kunadian V, Lam CSP, Maas AHEM, Mihailidou AS, Olszanecka A, Poole JE, Saldarriaga C, Saw J, Zühlke L, Mehran R. The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. *Lancet* 2021;**397**:2385–2438.
- Khramtsova EA, Davis LK, Stranger BE. The role of sex in the genomics of human complex traits. *Nat Rev Genet* 2019;**20**:173–190.
- Lagou V, Mägi R, Hottenga JJ, Grallert H, Perry JRB, Bouatia-Naji N, Marullo L, Rybin D, Jansen R, Min JL, Dimas AS, Ulrich A, Zudina L, Gädin JR, Jiang L, Faggian A, Bonnefond A, Fadista J, Stathopoulou MG, Isaacs A, Willems SM, Navarro P, Tanaka T, Jackson AU, Montasser ME, O'Connell JR, Bielak LF, Webster RJ, Saxena R, Stafford JM, Pourcain BS, Timpson NJ, Salo P, Shin SY, Amin N, Smith AV, Li G, Verweij N, Goel A, Ford I, Johnson PCD, Johnson T, Kapur K, Thorleifsson G, Strawbridge RJ, Rasmussen-Torvik LJ, Esko T, Mihailov E, Fall T, Fraser RM, Mahajan A, Kanoni S, Giedraitis V, Kleber ME, Silbernagel G, Meyer J, Müller-Nurasyid M, Ganna A, Sarin AP, Yengo L, Shungin D, Luan J, Horikoshi M, An P, Sanna S, Boettcher Y, Rayner NW, Nolte IM, Zemunik T, Iperen EV, Kovacs P, Hastie ND, Wild SH, McLachlan S, Campbell S, Polasek O, Carlson O, Egan J, Kiess W, Willemssen G, Kuusisto J, Laakso M, Dimitriou M, Hicks AA, Rauramaa R, Bandinelli S, Thorand B, Liu Y, Miljkovic I, Lind L, Doney A, Perola M, Hingorani A, Kivimaki M, Kumari M, Bennett AJ, Groves CJ, Herder C, Koistinen HA, Kinnunen L, Faire U, Bakker SJL, Uusitupa M, Palmer CNA, Jukema JW, Sattar N, Pouta A, Snieder H. Sex-dimorphic genetic effects and novel loci for fasting glucose and insulin variability. *Nat Commun* 2021;**12**:24.

8. Bernabeu E, Canela-Xandri O, Rawlik K, Talenti A, Prendergast J, Tenesa A. Sex differences in genetic architecture in the UK Biobank. *Nat Genet* 2021;**53**:1283–1289.
9. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, Payne AJ, Steinthorsdottir V, Scott RA, Grarup N, Cook JP, Schmidt EM, Wuttke M, Sarnowski C, Mägi R, Nango J, Gieger C, Trompet S, Leccoer C, Preuss MH, Prins BP, Guo X, Bielak LF, Below JE, Bowden DW, Chambers JC, Kim YJ, Ng MCY, Petty LE, Sim X, Zhang W, Bennett AJ, Bork-Jensen J, Brummett CM, Canouil M, Ec Kardt KU, Fischer K, Kardia SLR, Kronenberg F, Läll K, Liu CT, Locke AE, Luan J, Ntalla I, Nylander V, Schönherr S, Schurmann C, Yengo L, Bottinger EP, Brandslund I, Christensen C, Dedoussis G, Florez JC, Ford I, Franco OH, Frayling TM, Giedraitis V, Hackinger S, Hattersley AT, Herder C, Ikram MA, Ingelsson M, Jørgensen ME, Jørgensen T, Kriebel J, Kuusisto J, Ligthart S, Lindgren CM, Linneberg A, Lyssenko V, Mamakou V, Meitinger T, Mohlke KL, Morris AD, Nadkarni G, Pankow JS, Peters A, Sattar N, Stančáková A, Strauch K, Taylor KD, Thorand B, Thorleifsson G, Thorsteinsdottir U, Tuomilehto J, Witte DR, Dupuis J, Peyser PA, Zeggini E, Loos RJF, Froguel P, Ingelsson E, Lind L, Groop L, Laakso M, Collins FS, Jukema JW, Palmer CNA, Grallert H, Metspalu A, Dehghan A, Köttgen A, Abecasis GR, Meigs JB, Rotter J. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;**50**:1505–1513.
10. Ding Q-Y, Tian JX, Li M, Lian FM, Zhao LH, Wei XX, Han L, Zheng YJ, Gao ZZ, Yang HY, Fang XY, Tong XL. Interactions between therapeutics for metabolic disease, cardiovascular risk factors, and gut microbiota. *Front Cell Infect Microbiol* 2020;**10**:530160.
11. Hu X-F, Zhang WY, Wen Q, Chen WJ, Wang ZM, Chen J, Zhu F, Liu K, Cheng LX, Yang J, Shu YW. Fecal microbiota transplantation alleviates myocardial damage in myocarditis by restoring the microbiota composition. *Pharmacol Res* 2019;**139**:412–421.
12. Wang DD, Nguyen LH, Li Y, Yan Y, Ma W, Rinott E, Ivey KL, Shai I, Willett WC, Hu FB, Rimm EB, Stampfer MJ, Chan AT, Huttenhower C. The gut microbiome modulates the protective association between a Mediterranean diet and cardiometabolic disease risk. *Nat Med* 2021;**27**:333–343.
13. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung YM, Wu Y, Schauer P, Smith JD, Allayee H, Tang WH, DiDonato JA, Lusis AJ, Hazen SL. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 2011;**472**:57–63.
14. Canfora EE, Meex RCR, Venema K, Blaak EE. Gut microbial metabolites in obesity, NAFLD and T2DM. *Nat Rev Endocrinol* 2019;**15**:261–273.
15. Larsen N, Vogensen FK, van den Berg FW, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sørensen SJ, Hansen LH, Jakobsen M. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One* 2010;**5**:e9085.
16. Pingitore A, Chambers ES, Hill T, Maldonado IR, Liu B, Bewick G, Morrison DJ, Preston T, Wallis GA, Tedford C, Castañera González R, Huang GC, Choudhary P, Frost G, Persaud SJ. The diet-derived short chain fatty acid propionate improves beta-cell function in humans and stimulates insulin secretion from human islets in vitro. *Diabetes Obes Metab* 2017;**19**:257–265.
17. Karlsson FH, Tremaroli V, Nookaew I, Bergström G, Behre CJ, Fagerberg B, Nielsen J, Bäckhed F. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 2013;**498**:99–103.
18. Witkowski M, Weeks TL, Hazen SL. Gut microbiota and cardiovascular disease. *Circ Res* 2020;**127**:553–570.
19. Nemet I, Saha PP, Gupta N, Zhu W, Romano KA, Skye SM, Cajka T, Mohan ML, Li L, Wu Y, Funabashi M, Ramer-Tait AE, Naga Prasad SV, Fiehn O, Rey FE, Tang WHW, Fischbach MA, DiDonato JA, Hazen SL. A cardiovascular disease-linked gut microbial metabolite acts via adrenergic receptors. *Cell* 2020;**180**:862–877.e22.
20. Kreznar JH, Keller MP, Traeger LL, Rabaglia ME, Schueler KL, Stapleton DS, Zhao W, Vivas EI, Yandell BS, Broman AT, Hagenbuch B, Attie AD, Rey FE. Host genotype and gut microbiome modulate insulin secretion and diet-induced metabolic phenotypes. *Cell Rep* 2017;**18**:1739–1750.
21. Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Vösa U, Mujagic Z, Masclee AAM, Jonkers DMAE, Oosting M, Joosten LAB, Netea MG, Franke L, Zhernakova A, Fu J, Wijmenga C, McCarthy MI. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nat Genet* 2019;**51**:600–605.
22. Kenny DJ, Plichta DR, Shungin D, Koppel N, Hall AB, Fu B, Vasani RS, Shaw SY, Vlamakis H, Balskus EP, Xavier RJ. Cholesterol metabolism by uncultured human gut bacteria influences host cholesterol level. *Cell Host Microbe* 2020;**28**:245–257.e6.
23. Fu J, Bondar MJ, Cenic MC, Tigchelaar EF, Maatman A, Dekens JA, Brandsma E, Marczyńska J, Imhann F, Weersma RK, Franke L, Poon TW, Xavier RJ, Gevers D, Hofker MH, Wijmenga C, Zhernakova A. The gut microbiome contributes to a substantial proportion of the variation in blood lipids. *Circ Res* 2015;**117**:817–824.
24. McDonald B, McCoy KD. Maternal microbiota in pregnancy and early life. *Science* 2019;**365**:984–985.
25. Nuriel-Ohayon M, Neuman H, Ziv O, Belogolovski A, Barsheset Y, Bloch N, Uzan A, Lahav R, Peretz A, Frishman S, Hod M, Hadar E, Louzoun Y, Avni O, Koren O. Progesterone increases bifidobacterium relative abundance during late pregnancy. *Cell Rep* 2019;**27**:730–736.e3.
26. Vieira AT, Castelo PM, Ribeiro DA, Ferreira CM. Influence of oral and gut microbiota in the health of menopausal women. *Front Microbiol* 2017;**8**:1884.
27. Vandeputte D, De Commer L, Tito RY, Kathagen G, Sabino J, Vermeire S, Faust K, Raes J. Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat Commun* 2021;**12**:6740.
28. Lim GB. Role of sex hormones in cardiovascular diseases. *Nat Rev Cardiol* 2021;**18**:385.
29. Mumford SL, Schisterman EF, Siega-Riz AM, Browne RW, Gaskins AJ, Trevisan M, Steiner AZ, Daniels JL, Zhang C, Perkins NJ, Wactawski-Wende J. A longitudinal study of serum lipoproteins in relation to endogenous reproductive hormones during the menstrual cycle: findings from the BioCycle study. *J Clin Endocrinol Metab* 2010;**95**:E80–E85.
30. Yeung EH, Zhang C, Mumford SL, Ye A, Trevisan M, Chen L, Browne RW, Wactawski-Wende J, Schisterman EF. Longitudinal study of insulin resistance and sex hormones over the menstrual cycle: the BioCycle study. *J Clin Endocrinol Metab* 2010;**95**:5435–5442.
31. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol* 1997;**32**:920–924.
32. Vitali D, Wessels JM, Kaushic C. Role of sex hormones and the vaginal microbiome in susceptibility and mucosal immunity to HIV-1 in the female genital tract. *AIDS Res Ther* 2017;**14**:39.
33. Wessels JM, Felker AM, Dupont HA, Kaushic C. The relationship between sex hormones, the vaginal microbiome and immunity in HIV-1 susceptibility in women. *Dis Model Mech* 2018;**11**:dmm035147.
34. Buscemi S, Rosafio G, Vasto S, Massenti FM, Grosso G, Galvano F, Rini N, Barile AM, Maniaci V, Cosentino L, Verga S. Validation of a food frequency questionnaire for use in Italian adults living in Sicily. *Int J Food Sci Nutr* 2015;**66**:426–438.
35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;**42**:377–381.
36. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN; REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019;**95**:103208.
37. Evenepoel C, Clevers E, Deroover L, Van Loo W, Matthys C, Verbeke K. Accuracy of nutrient calculations using the consumer-focused online app MyFitnessPal: validation study. *J Med Internet Res* 2020;**22**:e18237.
38. Mattei V, Murugesan S, Al Hashmi M, Mathew R, James N, Singh P, Kumar M, Lakshmanan AP, Terranegra A, Al Khodor S. Evaluation of methods for the extraction of microbial DNA from vaginal swabs used for microbiome studies presentation_2.PPTX. *Frontiers in Cellular and Infection Microbiology* 2019;**9**. <https://doi.org/10.3389/fcimb.2019.00197>
39. Gacesa R, Kurilshikov A, Vich Vila A, Sinha T, Klaassen MAY, Bolte LA, Andreu-Sánchez S, Chen L, Collij V, Hu S. Environmental factors shaping the gut microbiome in a Dutch population. *Nature* 2022;**604**:732–739. <https://doi.org/10.1038/s41586-022-04567-7>
40. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet* 2009;**10**:387–406.
41. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 1985;**28**:412–419.
42. Warmink-Perdijk WDB, Peters LL, Tigchelaar EF, Dekens JAM, Jankipersading SA, Zhernakova A, Bossers WJR, Sikkema J, de Jonge A, Reijneveld SA, Verkade HJ, Koppelman GH, Wijmenga C, Kuipers F, Scherjon SA. Lifelines NEXT: a prospective birth cohort adding the next generation to the three-generation lifelines cohort study. *Eur J Epidemiol* 2020;**35**:157–168.
43. Imhann F, Van der Velde KJ, Barbieri R, Alberts R, Voskuil MD, Vich Vila A, Collij V, Spekhorst LM, Van der Sloot KWJ, Peters V, Van Dullemen HM, Visschedijk MC, Festen EAM, Swertz MA, Dijkstra G, Weersma RK. The 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; data release 1. *BMC Gastroenterol* 2019;**19**:5.
44. Plomin R, Haworth CMA, Davis OSP. Common disorders are quantitative traits. *Nat Rev Genet* 2009;**10**:872–878.
45. Pfeufer A, Sanna S, Arking DE, Müller M, Gateva V, Fuchsberger C, Ehret GB, Orrù M, Pattaro C, Köttgen A, Perz S, Usala G, Barbalic M, Li M, Pütz B, Scuteri A, Prineas RJ, Sinner MF, Gieger C, Najjar SS, Kao WH, Mühleisen TW, Dei M, Happle C, Möhlenkamp S, Crisponi L, Erbel R, Jöckel KH, Naitza S, Steinbeck G, Marroni F, Hicks AA, Lakatta E, Müller-Myhok B, Pramstaller PP, Wichmann HE, Schlessinger D, Boerwinkle E, Meitinger T, Uda M, Coresh J, Kääb S, Abecasis GR, Chakravarti A. Common variants at ten loci modulate the QT interval duration in the QTSCD study. *Nat Genet* 2009;**41**:407–414.
46. Uda M, Galanello R, Sanna S, Lettore G, Sankaran VG, Chen W, Usala G, Busonero F, Maschio A, Albai G, Piras MG, Sestu N, Lai S, Dei M, Mulas A, Crisponi L, Naitza S, Asunis I, Deiana M, Nagaraja R, Perseu L, Satta S, Cipollina MD, Sollaino C, Moi P, Hirschhorn JN, Orkin SH, Abecasis GR, Schlessinger D, Cao A. Genome-wide association study shows *BCL11A* associated with persistent fetal hemoglobin and amelioration of the phenotype of β -thalassaemia. *Proc Natl Acad Sci* 2008;**105**:1620–1625.
47. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A,

- Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee JY, Park T, Kim K, Sim X, Twee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X, Luan J, Lamina C, Ziegler A, Zhang W, Zee RY, Wright AF, Witteman JC, Wilson JF, Willemsen G, Wichmann HE, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF, Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM, Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJ, Scuteri A, Scott J, Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruukonen A, Rudan I, Rose LM, Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BW, Pedersen NL, Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery GW, Meitinger T, McPherson R, McCarthy MI, McArdle W, Masson D, Martin NG, Marroni F. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;**466**:707–713.
48. *Contraceptive use by method 2019: data booklet*. New York, NY: United Nations; 2019.