Project Details	
Project Code	PHS27Br Millard
Title	Cycles, signals and causality: Exploring methods for robust analysis of digital health data across menstrual cycles
Research Theme	Population Health Sciences
Project Type	Dry lab
Summary	Many symptoms, behaviours and adverse outcomes vary across the menstrual cycle, but to-date high-quality data to comprehensively characterise and understand these patterns has been lacking. CycleTrack is a ground-breaking research study that is currently collecting high-quality data across menstrual cycles. However, these data are complex to analyse, with multiple time-series and challenges such as missing data and sample heterogeneity. This project seeks to assess potential bias in analyses using these data and explore methods to mitigate these biases. You will advance your skills in data science, statistics, epidemiology and digital health, all highly valued by the FemTech industry and beyond.
Description	Background: Menstrual cycles are characterised by dynamic physiological changes due to variations in hormone levels. Across the cycle, individuals might experience changes in symptoms, behaviours such as sleep and physical activity, and adverse events such as mental health episodes. However, a lack of good quality data for menstrual health research means we do not currently have a good understanding of what a typical cycle looks like and how this varies across demographic subgroups. We also do not know what causes variations across the cycle, for example whether poor sleep during menstruation is due to heavy bleeding or pain. It may also be possible to develop predictions models to anticipate when an adverse events might occur, in order that an intervention can be given, such as predicting that an individual may be at risk of become iron deficient in their current cycle. A ground-breaking research study – CycleTrack [https://cycle-track-study.com/] – is currently being run by researchers at Exeter and Bristol, led by study PI Professor Gemma Sharp (co-supervisor on this PhD). CycleTrack is creating a world-leading resource for menstrual health research, collecting data across three menstrual cycles in a subsample of participants in two UK cohorts – ALSPAC [https://www.bristol.ac.uk/alspac/] and BiB [https://borninbradford.nhs.uk/]. A wealth of high-quality data is being collected, including self-reported measures using Ecological Momentary Assessment (EMA), passive measurement of activity, sleep and heart rate using a smart watch, and menstrual fluid samples. These data are complex, with multiple time-series and measures collected with different frequencies (e.g. daily, weekly or only once), and potential challenges such as missing data. These properties need to be carefully considered in order to use these data appropriately, both understanding the potential for bias and assessing methods to mitigate such bias. The PhD student who undertakes this project can explore methodological research questions in a range of areas including prediction and causal inference. The student will (depending on the projects chosen) gain important skills in statistics, data science, prediction (including AI), and time-series data analysis. They will also

	learn important general research skills in research conduct, open science and reproducibility. Aim: To assess potential biases using digital health data for menstrual health research and investigate methods to appropriately analyse these data. Objectives: The PhD is flexible and the student can develop their own objectives depending on their own interests, background, and training development preferences. You will choose/develop 3-4 projects that will each consider a particular applied research question and investigate challenges and potential biases with the CycleTrack digital health data and approaches to mitigating these challenges. Broadly, objectives can be in the following four general areas, that each offer distinct training opportunities: 1) Characterisation of variability across the menstrual cycle, such as how sleep or mental health symptoms (e.g. anxiety, depression, wellbeing) fluctuate across the menstrual cycle. 2) Prediction using statistical and machine learning approaches using data from a smart watch alongside cycle information, such as forecasting iron deficiency or mental health episodes, or diagnostic modelling, for example using data from the smart watch (e.g. changes in daily step counts) to identify iron deficiency. 3) Estimation of causal effects, such as whether menstrual symptoms like heavy bleeding/pain affect sleep duration and quality. 4) N-of-1 study, where instead of analysing data across a population subsample, a single participant (or sometimes extended to a small number) is studied over a longer period of time. This study design could be used to evaluate a particular causal effect identified in (3) by assessing, within an individual, whether a given exposure determines a subsequent outcome such as sleep duration. It may also be used to assess whether a personalised prediction model can be developed that improves prediction for a given individual above models developed on a wider sample (as in (2)), or to test the effect of an intervention (e.g. an alert from a prediction model). Data challenges to consider when tackling applied research questions: - High levels of missing data in EMA, and the fact this missingness will likely relate to other factors such as cycle day. - Granularity of the data collected, e.g. daily, weekly or just once. - Impact of treatments on analyses, and how best to account for this when using these data. For example, our sample will include people on hormonal treatments, such as contraceptives / HRT, as well as those not on any hormonal treatments. - Sample selection, where those who have agreed to take part in CycleTrack may be different to the wider population who menstruate. Assessment of bias can be investigated using simulation studies and through comparison of approaches with contrasting assumptions (for example comparison of approaches to dealing with missing data).
Can be completed part time?	Yes

Supervisory Team	
Lead Supervisor	
Name	Dr Louise Millard
Affiliation	Bristol
College/Faculty	Faculty of Health and Life Sciences
Department/School	Population Health Sciences, Bristol Medical School
Email Address	louise.millard@bristol.ac.uk
Co-Supervisor 1	
Name	Professor Jon Heron
Affiliation	Bristol
College/Faculty	Faculty of Health and Life Sciences
Department/School	Population Health Sciences, Bristol Medical School
Co-Supervisor 2	
Name	Professor Gemma Sharp
Affiliation	Exeter
College/Faculty	Faculty of Health and Life Sciences
Department/School	School of Psychology
Co-Supervisor 3	
Name	
Affiliation	
College/Faculty	
Department/School	