Nightingale Health Metabolic Biomarkers: Phase 1 Release

Introduction

Nightingale Health Ltd. is performing metabolic biomarker profiling of participants from the UK Biobank. This first release covers biomarker data from approximately 118,000 EDTA plasma samples from baseline recruitment and 5,000 samples from repeat assessment (with 1,500 participants having both baseline and repeat-visit sample in the first data release). These samples were measured between June 2019 and April 2020 in Nightingale Health's laboratories in Finland.

Nightingale Health's metabolic biomarker platform is based on high-throughput nuclear magnetic resonance (NMR). 249 metabolic measures, 168 in absolute levels and 81 ratio measures, are quantified per EDTA plasma sample, covering both routine biomarkers and emerging biomarkers with medical relevance. The biomarkers include detailed measures of cholesterol metabolism, fatty acid compositions, and various low-molecular weight metabolites, such as amino acids, ketones and glycolysis metabolites. For 14 lipoprotein subclasses, the lipid concentrations and composition are measured in terms of triglycerides, phospholipids, total cholesterol, cholesterol esters, and free cholesterol, and total lipid concentration within each subclass. The majority of the biomarkers are measured in absolute concentration units (mmol/L).

The present document provides an outline of the methodology and the metabolic biomarker data in the UK Biobank resource. A more extensive companion publication detailing the biomarker data and quality control procedures is expected in second quarter of 2021.

Biomarker measurements by NMR

EDTA plasma samples from aliquot 3 were measured using Nightingale Health's NMR-based metabolic biomarker profiling platform. If comparisons are done to the clinical biochemistry markers available in the UK Biobank, please note that those were measured from serum samples, primarily from aliquot 1. The samples for the first data release are a random subset of the full cohort.

The samples were prepared directly in 96 well-plates by UK Biobank. At least 85 µL plasma was aliquoted in each well using TECAN freedom EVO 150 robotic liquid handlers, which have coefficients of variation in pipetting volume at <0.75% across 8 tips. Plasma samples were shipped to Nightingale Health's laboratories in the 96-well plates on dry ice in sample batches of ~5,000-20,000.

Details of the metabolic biomarker profiling platform and experimentation have been described previously (1, 2). In brief, EDTA plasma samples were stored in a freezer at -80°C. Before preparation, frozen samples were slowly thawed at +4°C overnight, and then mixed gently and centrifuged (3 min, 3400´*g*, +4°C) to remove possible precipitate. Aliquots of each sample were transferred into 3-mm outer-diameter NMR tubes and mixed in 1:1 ratio with a phosphate buffer (75mM Na_2HPO_4 in 80%/20% H_2O/D_2O, pH 7.4, including also 0.08% sodium 3-(trimethylsilyl) propionate-2,2,3,3-d_4 and 0.04% sodium azide) automatically with an automated liquid handler (PerkinElmer Janus Automated Workstation).

The prepared samples were loaded onto a cooled sample changer, which maintains the temperature of samples waiting to be measured at +6ºC. Two NMR spectra were recorded for each plasma sample using a 500 MHz NMR spectrometer (Bruker AVANCE IIIHD). The first spectrum is a presaturated proton NMR spectrum, which features resonances arising mainly from proteins and lipids within various lipoprotein particles. The other spectrum is a T2-relaxation-filtered spectrum where most of the broad macromolecule and lipoprotein lipid signals are suppressed, leading to enhanced detection of low-molecular-weight metabolites. Automated quality control of the spectral data was performed. The metabolic biomarkers were quantified using Nightingale Health's proprietary software (Nightingale Health biomarker quantification library 2020). Appendix 1 shows biomarker distributions for the five consecutively measured sample batches of ~25,000.

As an indication of the appropriate laboratory quality management, Nightingale Health's metabolic biomarker profiling platform has qualified for CE-marking according to the required laboratory standards for clinical use (IVD 98/79/EC; Directive on In Vitro Diagnostic Medical Devices). As a result, our laboratory information management system is accredited according to ISO 13485. Comparison of Nightingale Health's biomarker measurements with clinical biochemistry in cohort studies also indicate high congruence, with correlation coefficients commonly r>0.9, see e.g. (3). The congruence in the UK Biobank is reduced a bit due to sample dilution (see Section *Sample dilution issue* below) as well as difference in the aliquot and sample type (serum vs. EDTA plasma) between the clinical biochemistry panel and Nightingale Health assays.

Sample dilution issue

UK Biobank plasma samples from aliquot 3 analysed by Nightingale Health are expected to be 5-10% diluted. The dilution is believed to come from mixing of participant sample with water due to seals that failed to hold a system vacuum in the automated liquid handling systems. While this issue may have an impact on some of the biomarker concentration values, it is expected to have limited impact on the epidemiological analyses.

Quality control assurance

Metabolic biomarker profiling by Nightingale Health's NMR platform provides consistent results over time and across spectrometers. Furthermore, the sample preparation is minimal in the Nightingale Health's metabolic biomarker platform, circumventing all extraction steps. These aspects result in highly repeatable biomarker measurements. Pre-specified quality metrics were agreed between UK Biobank and Nightingale Health to ensure consistent results across the samples, and pilot measurements were conducted.

Nightingale Health performed real-time monitoring of the measurement consistency within and between spectrometers throughout the UK Biobank samples. Two control samples provided by Nightingale Health were included in each 96-well plate for tracking the consistency across multiple spectrometers. Furthermore, two blind duplicate samples provided by the UK Biobank were included in each well plate, with the position information unlocked only after results delivery. Coefficient of variation (CV) targets across the metabolic biomarker profile were pre-specified for both Nightingale Health's internal control samples and UK Biobank's blind duplicates. The targets were met for each consecutively measured batch of ~25,000 samples. CVs across spectrometers based on Nightingale Health's internal control samples and blind duplicates measured as part of the phase 1 data

release are summarized in Figure 1. For the majority of the metabolic biomarkers the CVs are below 5%. Consistency of the blind duplicate measures for individual biomarkers is shown in Appendix 2. In addition to technical repeatability, biological stability of the biomarker measures over time between the samples from baseline and repeat assessment center visit is presented in Appendix 3.

Quality control flags

Nightingale Health's integrated quality procedures verify the sample quality by reporting signs of degradation and contamination issues. These are reported as flags along with the result data. Issues affecting the whole sample are reported as sample-level flags; issues affecting only certain biomarkers are reported as biomarker-level flags, provided as a separate data field for each biomarker. In general, if a biomarker has a flag but the biomarker value is still provided, it indicates that the presence of the interfering substance is low and deemed not to interfere with the quantification of the biomarker (i.e., the value can be trusted). Appendix 4 provides detailed information on the available quality flags. There is no need to a priori remove any biomarker values based on the flags; however, researchers may consider performing sensitivity analyses as described below.

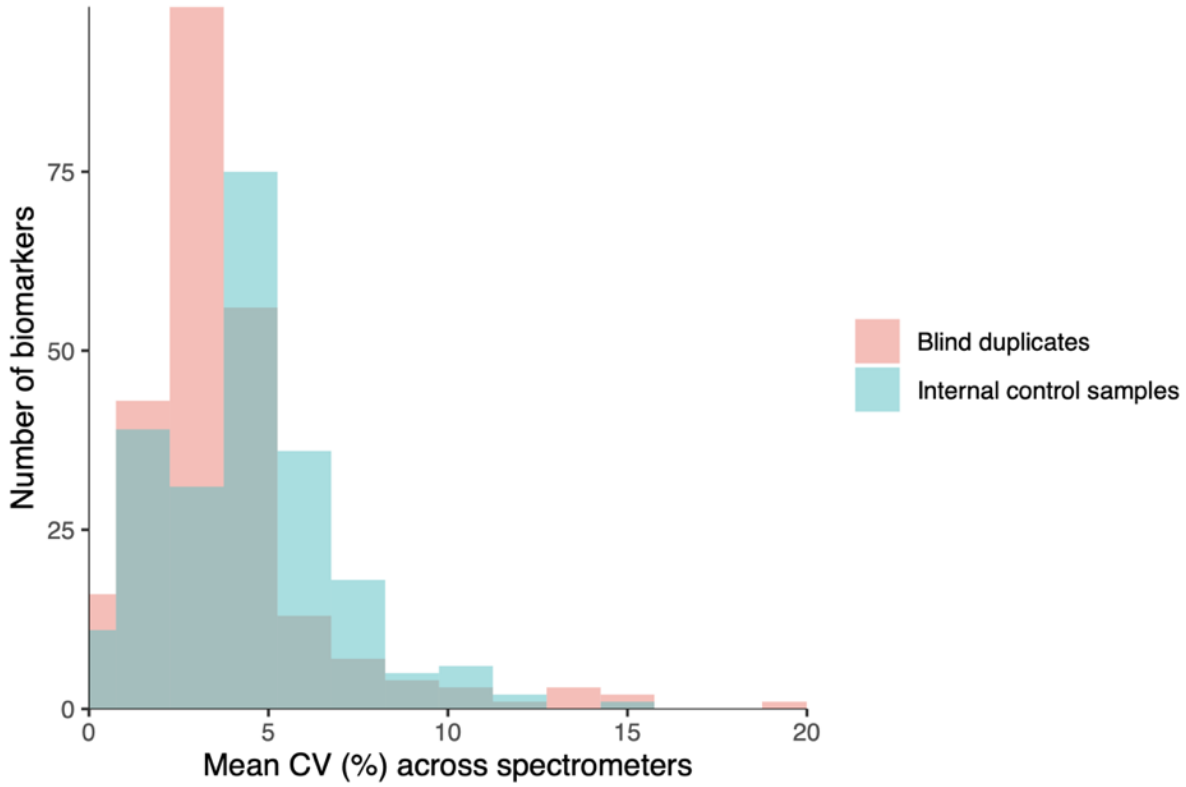

Figure 1. Distributions of coefficients of variation (CV) for the 249 metabolic measures for UK Biobank's blind duplicate samples (red) and Nightingale Health's internal control samples (blue). The CVs are across the six NMR spectrometers used for the measurements. CVs of individual biomarkers are available in Appendix 2.

Recommended approaches for data pre-processing and analysis

Data pre-processing and considerations for epidemiological analyses

Nightingale Health's metabolic biomarker profiling platform is known for high repeatability over time and absence of batch effects. This can be seen in Appendix 1, showing the biomarker distributions from five consecutively measured sample batches. The metabolic biomarker data can generally be used for epidemiological analyses without any pre-processing, and can be analysed in the same manners as the clinical biochemistry data available in the UK Biobank.

Minor differences in biomarker concentrations may arise from the different spectrometers used for the biomarker measurements. As shown in Appendix 5, the differences in the biomarker distributions between spectrometers are generally minor. However, researchers may consider correcting or adjusting for the spectrometer in epidemiological analyses.

Quality control flags

Biomarker values substantially affected by interfering substances have been removed during the quality control procedures. However, the researcher may consider performing sensitivity analyses by excluding samples flagged with "Low protein", which may indicate more severe sample dilution. Biomarker values flagged with "Below limit of quantification" may also be omitted in sensitivity analyses, since this flag indicates that the concentration of the given biomarker is smaller than the range where the Nightingale Health's metabolic biomarker profiling platform considers quantification highly accurate.

GWAS summary statistics

GWAS summary statistics of metabolic biomarkers measured by Nightingale Health will be available through MR-Base database, developed by the MRC Integrative Epidemiology Unit at the University of Bristol, from March 2021. Additional summary data from genetic analyses by University of Cambridge and Stanford University are expected in second quarter of 2021.

References

1. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circulation: Cardiovascular Genetics 2015; 8(1):192-206.
2. Ahola-Olli AV, Mustelin L, Kalimeri M, Kettunen J, Jokelainen J, Auvinen J, Puukka K, Havulinna AS, Lehtimäki T, Kähönen M, Juonala M. Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. Diabetologia 2019; 62(12):2298-2309.
3. Tikkanen E, Jägerroos V, Rodosthenous R, Holmes M, Sattar N, Ala-Korpela M, Jousilahti P, Lundqvist A, Perola M, Salomaa V, Würtz P. Metabolic biomarkers for peripheral artery disease compared with coronary artery disease: Lipoprotein and metabolite profiling of 31,657 individuals from five prospective cohorts. medRxiv 2020; doi:10.1101/2020.07.24.20158675.