Metabolomics

Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics

Pasi Soininen, PhD*; Antti J. Kangas, MSc*; Peter Würtz, PhD; Teemu Suna, MSc; Mika Ala-Korpela, PhD

Abstract—Metabolomics is becoming common in epidemiology due to recent developments in quantitative profiling technologies and appealing results from their applications for understanding health and disease. Our team has developed an automated high-throughput serum NMR metabolomics platform that provides quantitative molecular data on 14 lipoprotein subclasses, their lipid concentrations and composition, apolipoprotein A-I and B, multiple cholesterol and triglyceride measures, albumin, various fatty acids as well as on numerous low-molecular-weight metabolites, including amino acids, glycolysis related measures and ketone bodies. The molar concentrations of these measures are obtained from a single serum sample with costs comparable to standard lipid measurements. We have analyzed almost 250,000 samples from around 100 epidemiological cohorts and biobanks and the new international set-up of multiple platforms will allow an annual throughput of more than 250,000 samples. The molecular data have been used to study type 1 and type 2 diabetes etiology as well as to characterize the molecular reflections of the metabolic syndrome, long-term physical activity, diet and lipoprotein metabolism. The results have revealed new biomarkers for early atherosclerosis, type 2 diabetes, diabetic nephropathy, cardiovascular disease and all-cause mortality. We have also combined genomics and metabolomics in diverse studies. We envision that quantitative high-throughput NMR metabolomics will be incorporated as a routine in large biobanks; this would make perfect sense both from the biological research and cost point of view – the standard output of over 200 molecular measures would vastly extend the relevance of the sample collections and make many separate clinical chemistry assays redundant.

Epidemiology is Changing

Applications of comprehensive metabolic profiling, broadly termed metabolomics, are increasingly common in epidemiology and genetics. This is because of recent developments in quantitative methodologies and various appealing results from their applications on understanding life-course health and disease causes. Most epidemiological studies routinely analyze blood biomarkers, for example, total cholesterol and glucose. Our team has developed a high-throughput serum nuclear magnetic resonance (NMR) metabolomics platform that measures the concentrations of standard biomarkers, such as various cholesterol measures, triglycerides, and creatinine, but in the same experiment provides also quantitative molecular data on lipoprotein subclasses, such as lipids, fatty acids, and apolipoproteins as well as on various low-molecular-weight metabolites, including amino acids, glycolysis-related metabolites, and ketone bodies. All these metabolic measures are obtained from a single serum or plasma sample with costs comparable with standard lipid measurements. To date (end of 2014) we have used the platform to analyze almost 250,000 samples (in about 6 years) in around 100 epidemiological and clinical cohorts and the studies have revealed new biomarkers for early atherosclerosis, type 2 diabetes mellitus, diabetic nephropathy, cardiovascular disease and all-cause mortality. The molecular data have been extensively used in studies of type 1 and type 2 diabetes mellitus pathogenesis as well as to characterize the molecular reflections of the metabolic syndrome and lipoprotein metabolism. We have also examined the genetic determinants and heritability of a detailed molecular snapshot of systemic metabolism and taken various steps toward multomics systems epidemiology, for example, to understand liver function, to identify causal networks of gene expression modules, and to design a metabolite-cluster genomewide approach to identify genes for atherosclerosis. Moreover, we have interrogated the metabolic effects of long-term physical activity as well as aging and menopause and illustrated that the baseline metabolic profile affects the systemic effects of dietary interventions. The increasing amounts of quantitative data on systemic metabolism are revealing a plethora of novel molecular biomarkers suggesting that quantitative metabolomics will eventually reshape the way how epidemiology and genetics are practiced in the near future.

This article focuses on the applications of quantitative NMR metabolomics to study metabolic diseases and...
<table>
<thead>
<tr>
<th>Focus</th>
<th>Number of Study Participants and Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metabolic risk factor characterization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-throughput serum NMR metabolomics platform</td>
<td>n=4407; illustration of the epidemiological relevance of quantitative NMR spectroscopy data (metabolic syndrome); basis for the further development of the automated quantitative platform; serum</td>
<td>Soininen et al²⁵⁴</td>
</tr>
<tr>
<td>Subclinical atherosclerosis</td>
<td>n=4309; young adults with systemic metabolic data and carotid intima media thickness measurements; self-organizing map analyses for metabolic phenotyping to illustrate the heterogeneity of metabolic risk; serum</td>
<td>Würtz et al³⁴</td>
</tr>
<tr>
<td>Long-term physical activity</td>
<td>16 twin-pairs with &gt;30-year discordance for physical activity and 1037 age- and sex-matched pairs from 3 population cohorts; numerous metabolic differences found between persistently physically active and inactive individuals and indications for better metabolic health in the physically active ones; serum</td>
<td>Kujala et al²⁶³</td>
</tr>
<tr>
<td>Insulin resistance</td>
<td>n=7098 in 2 population cohorts; cross-sectional metabolic signatures of insulin resistance in young healthy adults; sex-specific fingerprints; serum</td>
<td>Würtz et al²³⁰</td>
</tr>
<tr>
<td>Fasting and postload glycemia</td>
<td>n=1873 middle-aged individuals from 2 population cohorts; cross-sectional associations of metabolite panel and 6-year prospective associations in 1 of the cohorts; indications that alterations in amino acid metabolism precede perturbations in glucose homeostasis; serum</td>
<td>Würtz et al²³</td>
</tr>
<tr>
<td>Insulin resistance, glycemia, and type 2 diabetes mellitus</td>
<td>Up to n=9398 Finnish men; cross-sectional and prospective associations of amino acids, ketone bodies, fatty acids, and lipoprotein subclasses with insulin resistance, the continuum of glycemia and the risk for type 2 diabetes mellitus; metabolite associations with known genetic risk loci for glycemia and diabetes mellitus; multiple results indicating a key role for comprehensive metabolic phenotyping in understanding the pathogenesis of type 2 diabetes mellitus; serum</td>
<td>Stančáková et al¹⁴²</td>
</tr>
<tr>
<td>Hypertension</td>
<td>n=4630 from 17 population samples worldwide; spectral analyses of metabolic phenotype diversity and quantification of 4 discriminatory metabolites for association with blood pressure; cross-population metabolic differences illustrated; urine Black (n=369) compared with non-Hispanic white Americans (n=1190) aged 40 to 59 yr from 8 population samples; dietary and urinary metabolic factors associating with blood pressure; urine</td>
<td>Holmes et al³¹</td>
</tr>
<tr>
<td>Presence of drug metabolites</td>
<td>n=4680 from multiple population samples worldwide; screening strategy for analgesic usage, validation with self-reported medication data; urine</td>
<td>Loo et al⁴⁶</td>
</tr>
<tr>
<td>Aging</td>
<td>n=26065 from 8 population studies; metabolic trends of aging and menopause; serum and plasma</td>
<td>Auro et al²³</td>
</tr>
<tr>
<td>Obesity</td>
<td>n=7255 in 2 population studies; genome metabolome integrated network analysis; serum</td>
<td>Valcárcel et al¹⁹⁰</td>
</tr>
<tr>
<td>Fatty acids and lipoprotein subclasses</td>
<td>n=1269 individual twins, including 561 complete pairs; genetic and environmental cause of the associations of serum fatty acids with lipoprotein profile; serum</td>
<td>Jelenkovic et al⁴¹</td>
</tr>
<tr>
<td><strong>Biomarkers and risk assessment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subclinical atherosclerosis</td>
<td>n=1595 young adults; circulating biomarkers for 6-year high carotid intima media thickness, new systemic biomarkers with improved risk stratification for subclinical atherosclerosis in comparison with conventional lipids; serum</td>
<td>Würtz et al²³⁰</td>
</tr>
<tr>
<td>Type 1 diabetes mellitus and kidney disease</td>
<td>Up to n=3544 patients with type 1 diabetes mellitus; cross-sectional and prospective associations of various systemic metabolites and lipoprotein subclass measures with the severity of diabetic kidney disease and mortality; introducing multiparametric risk assessment of diabetic nephropathy; serum</td>
<td>Makinen et al⁴²⁴</td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>n=17345 from 2 general population cohorts; 4 circulating biomarkers for 5-year risk of death; biomarker associations with multiple causes of death suggest novel systemic connectivities across seemingly disparate morbidities; improved prediction of the short-term risk of death from all causes above established risk factors; serum and plasma</td>
<td>Fischer et al²⁶</td>
</tr>
<tr>
<td><strong>Genetics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic determinants of circulating metabolites and lipids</td>
<td>n=8330 from 5 population-based cohorts; genomewide association study of 99 metabolite traits and 117 ratios, and heritability estimates from 561 twin pairs; serum</td>
<td>Kettunen et al⁴³⁳</td>
</tr>
<tr>
<td>Genetic determinants of urine metabolites</td>
<td>n=6608 from 2 population cohorts of young adults; genomewide association study of 11 metabolite networks; multtissue gene expression; candidate genes for atherosclerosis; illustration of the statistical power of multiphenotype genomewide association study; serum</td>
<td>Inouye et al²²²</td>
</tr>
<tr>
<td>Metabolic characterization of lipid genes</td>
<td>n=8330 from 5 population-based cohorts; metabolic characterization and fine mapping of 95 known lipid gene loci; demonstration of the advantages of combining detailed genotyping and a wide metabolite panel; serum</td>
<td>Tukiainen et al⁴¹⁸</td>
</tr>
<tr>
<td></td>
<td>n=10 547 from 3 population-based cohorts; assessment of pleiotropy in the genetic variants of lipid risk factors for coronary heart disease; serum</td>
<td>Würtz et al³⁷</td>
</tr>
</tbody>
</table>

(Continued)
circulating biomarkers. With metabolomics we refer to comprehensive metabolic profiling of multiple molecular pathways, specific methods on individual molecules or single molecular groups are not incorporated. Surprisingly few NMR-based metabolomics studies meet the criteria of absolute metabolite quantification and large sample size appropriate for epidemiology and genetics. We chose a threshold of a 1000 individuals for the inclusion on the basis of recent power calculations, which revealed that sample sizes of a few thousand should offer sufficient statistical precision to detect proton NMR-based biomarkers quantifying predisposition to disease. Table 1 summarizes epidemiological and genetic studies that include >1000 individuals and have applied NMR metabolomics with absolute quantification of metabolites.

NMR Makes Metabolomics Available for the Masses

Cost effectiveness via high throughput, along with absolute quantification of molecular measures, is a prerequisite for metabolomics methodologies in epidemiology. NMR spectroscopy and mass spectrometry (MS) have become the key complementary technologies in the metabolomics field. Owing to fundamental methodological differences, the metabolomics applications of NMR and MS are divergent; MS being broadly applied for detailed characterization of (patho)physiology and metabolic individuality while NMR is the methodology that paves the way of metabolomics into large epidemiological and genetic studies. NMR detects molecule-specific signals of rather high-concentration substances and thereby inherently filters out adverse complications in the data. MS separates signals from individual molecular identities with a mass difference and thereby naturally leads to rich and complicated spectroscopic data. However, analysis of the quantitative MS data, particularly in an automated and robust fashion, poses currently considerable challenges hampering the use of quantitative MS in large-scale epidemiology. Also, the per-sample costs in quantitative MS tend to be high, partly because of the necessity of reference substances for quantifications in each molecular category. An additional drawback in MS is that it cannot analyze lipoproteins. Therefore, NMR is currently the only methodology capable of offering reproducible high-throughput metabolite quantifications in a cost-effective manner.

NMR spectroscopy has gained a lot of attention as a method inherently optimal for obtaining quantitative lipoprotein data. Thus, even though the focus of this work is on NMR-based metabolomics as defined above, we would like to refer to a recent review by Mallol et al that gives an extensive analytic and application update of the field. As summarized in the review, the feasibility of NMR spectroscopy for lipoprotein quantification has attracted several research teams to test and apply the methodology, however, mainly with small numbers of study subjects. Also a few commercial companies are offering lipoprotein analysis via NMR spectroscopy, for example, the spectrometer manufacturer Bruker BioSpin and Lipofit’s numares HEALTH (www.numares-health.com). We are not aware of any scientific publications based on the Bruker methodology, and the published data on the Lipofit method suggest quantitative challenges even at the level of total lipoprotein measures, low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol. In contrast, the commercial NMR-based lipoprotein quantification methodology of LipoScience (www.liposence.com) has been widely applied in biomedical applications. The company is advocating the role of lipoprotein particle concentrations (instead of lipoprotein lipid concentrations) in clinical risk assessment. A recent systematic review evaluating the effects of lipid-lowering therapy on lipoprotein and lipid values suggested that the use of LDL particle number for monitoring lipid-lowering therapy, particularly for statins, can provide a more accurate assessment of residual cardiovascular risk than LDL cholesterol. A recent study has also demonstrated a more consistent inverse association between cardiovascular end points and NMR-derived HDL particle number compared with HDL cholesterol. Regarding clinical applications of LipoScience’s lipoprotein panel, it is notable that their NMR-based method has been approved by the US Food and Drug Administration to directly quantify LDL particles. The company is also introducing fully automated, high-throughput NMR spectroscopy for lipoprotein analyses in a clinical laboratory setting. The foundation of the NMR-based technology for lipoprotein quantification was independently developed by Otvos et al and us during early 1990s. The method by Otvos et al was taken as the basis for the LipoScience application which has been applied in many clinical research studies and one of the largest studies published is a study on comparing lipoprotein particle profiles by NMR with standard lipids and apolipoproteins in predicting incident cardiovascular
disease in women (n=27,673). The long-term application of NMR-based lipoprotein analysis in medical research is obviously an encouraging example of the epidemiological and clinical prospects of NMR-based technologies. Predicting cardiovascular disease with LipoScience’s NMR-based lipoprotein measures was found comparable with standard lipids or immunoassay-measured apolipoproteins. These results together with recently established new biomarkers for cardiovascular disease and diabetes mellitus advocate the need to extend beyond lipoprotein lipids in disease risk assessment. Nevertheless, incorporating the traditional clinically used lipid and apolipoprotein information as part of the metabolomics panel, as in our platform, seems sensible.

Only Quantitative Metabolomics Aids Epidemiology

Detailed quantitative molecular profiling can be a treasure trove to study the (patho)physiology of any condition of biomedical interest. The quantitative metabolite data can readily be used to address various biological questions without any complications and limitations of spectral-based multivariate data analysis. Importantly, the full arsenal of statistical epidemiology tools are directly applicable to quantitative data; this makes it straightforward, for example, to combine data from various sources (such as metabolomics and standard clinical chemistry assays), to compare association magnitudes, to account for multiple testing, to adjust for confounding factors, and to perform independent replication and meta-analyses. These are prerequisites for good epidemiological research but are not met in customary metabolomics-driven spectroscopy-based studies, which are typically cross-sectional and limited with respect to the number of study subjects. They also often apply case-control settings for surprisingly single-minded search for nonexistent binary disease states. It is of particular note that sound study design in epidemiological metabolomics follows similar principles as standard epidemiology. Metabolomics technologies can also be directly applied in already existing large epidemiological sample collections. There is no need for a particular study design to apply quantitative metabolomics. Notably, quantitative metabolomics data are in fact identical to an extensive collection of molecular concentrations.

Here, we describe the premise of quantitative NMR metabolomics of blood specimens for applications in epidemiology and genetics. We exemplify a few applications of our serum NMR metabolomics platform to study metabolic health and diseases, to identify new biomarkers and to combine various omics data to elucidate lipoprotein metabolism and pinpoint genetic pleiotropy. We also summarize the current status of quantitative NMR metabolomics in epidemiology and aim to envision the road ahead in this field.

NMR Provides the Gold Standard for Molecular Identification and Quantification

NMR spectroscopy is one of the most important methods in chemistry, particularly for molecular identification and absolute quantification of soluble compounds. In proton, (H) NMR spectroscopy each molecule with hydrogen atoms gives a characteristic signal, the shape of which is quantum mechanically distinctive and the area of which is proportional to the concentration of the molecule. These fundamental principles make NMR the gold standard for metabolite quantification of low-molecular-weight metabolites and lipid molecules in liquid phase. Although NMR is not a sensitive method for metabolite detection (eg, in comparison with MS), only a few minutes measurement time is enough to quantitatively capture a comprehensive molecular signature from a serum sample. The molecular variety and multiple chemical environments in a biosample often cause the individual molecular signals to overlap. However, current linefitting methods that rely on molecule-specific model lineshapes and regression modeling can robustly handle the quantitative analyses of overlapping information. It is also well established that NMR is inherently suitable for detecting and quantifying lipoproteins because the physical lipoprotein particle structure fundamentally leads to a relation between the particle size and the NMR chemical shift. A recent systematic analytic approach verified that NMR spectra would allow statistically significant modeling of 13 lipoprotein subclasses; this number is identical to what we apply in our platform in the same lipoprotein particle size range. In addition, we quantify a category of larger very-low-density lipoprotein and chylomicron particles.

Our serum NMR metabolomics platform, illustrated in Figure 1, takes advantage of the abovementioned fundamental characteristics and possibilities of proton NMR spectroscopy. The platform is highly automated with respect to experimentation as well as spectral data analysis and provides quantitative data on the molar concentrations of >200 metabolic measures as summarized in Figure 2. Hundreds of biologically relevant derived measures (eg, metabolite ratios as proxies for enzymatic activity) can also be calculated. Of course, with respect to the entire serum metabolome of thousands of compounds related to systemic metabolism, the capture by NMR is still sparse. Yet, from the epidemiological perspective, it is a noteworthy addition and has already produced a wide variety of new scientific findings.

Optimal Quantitative Platform for Large-Scale Epidemiology

The levels of systemic metabolites in both the fasting and nonfasting state arise from a broad combination of genetic and lifestyle factors, including the level of adiposity, diet, and prevalent diseases, similarly as standard cardiovascular risk factors. The serum NMR metabolomics platform provides a comprehensive snapshot of the individual’s physiological status as reflected in the systemic metabolism. In normal conditions, the metabolite profile is stable and tracks well over long periods of time although published data are sparse. These metabolic measures reflecting systemic metabolism can, therefore, be used as markers of health and disease with the premise that they extend our information on physiological processes and disease pathogenesis, as well as potentially allow for improved risk prediction beyond conventional risk factors. A key point to note in relation to Figure 2 is that the metabolite data based on the platform (Figure 1) are in absolute molar units, for
Figure 1. The automated high-throughput serum nuclear magnetic resonance (NMR) metabolomics platform—process from the population to the individual comprehensive quantitative molecular data. Any routine serum (or ethylenediaminetetraacetic acid plasma) samples can be profiled. For integrity and biological stability, long-term (>1 month) storage must be at −80°C and the samples must stay frozen during the transport. The samples are handled in 96-well plates, every plate containing 2 quality control samples—a serum mimic and a mixture of 2 low-molecular-weight metabolites. The former is used to monitor the consistency of quantifications, whereas the latter is a technical reference to monitor the performance of the automated liquid handler and the spectrometer. Barcoding is preferred for sample identification. Before the NMR measurements, 260 μL of serum and 260 μL of a sodium phosphate buffer (75 mmol/L Na2HPO4 in 80%/20% H2O/D2O, pH 7.4; including also 0.08% sodium 3-(trimethylsilyl)propionate-2,2,3,3-d4 and 0.04% sodium azide) is carefully mixed and moved to the NMR tubes. All the liquid handling steps for native serum samples are done with a PerkinElmer JANUS Automated Workstation equipped with an 8-tip dispense arm with Varispan. The automated liquid handler can prepare 96 samples in ~30 minutes. Our current NMR metabolomics laboratory setup combines a Bruker AVANCE III 500 MHz and Bruker AVANCE III HD.
example, in mmol/L. This is absolutely crucial for the metabolomics data to be of value in epidemiology and genetics. 3,75 It is now broadly recognized that the 2 key metabolomics technologies, NMR and MS, are highly complementary. 7,26 However, the serum NMR platform offers robust quantification (for a very extensive metabolic profile from the traditional epidemiology perspective) at a fraction of the costs related to MS. Thus, even though epidemiological studies would benefit from as extensive metabolite coverage as possible, the current choice of profiling platform essentially boils down to a trade-off between a lot-more (metabolite coverage) for a lot-more (money) with MS and less for a lot-less with NMR. This dilemma is often amplified because the fundamental study designs (with multigenic continuous conditions with multiple confounding factors) call more importantly for large numbers of study samples than the most extensive metabolic coverage. The current state of metabolomics may be likened to the candidate gene era in genetics; 77 once robust statistics and replication became prerequisites, it led to a leap in sample size and attention to statistics and replication in the field of metabolomics to strengthen the scientific value of the work, irrespective of the analytic platform used. Toward these goals, our long-term focus in the case of the serum NMR metabolomics platform has been in full automation, utmost cost effectiveness via high throughput and reasonably extensive metabolic coverage.

The quantification limit with the automated operation of the platform is ≈10 μmol/L, however, the exact limit depending on the metabolite characteristics. Collecting the NMR data for remarkably larger times or analyzing the spectra (semi)manually, could lead to quantification of maybe a dozen metabolites more—with the disadvantage of reducing the experimental throughput to only one fourth and multiplying the costs. We have deliberately chosen what we think is an optimal balance between throughput and metabolic coverage. The current platform allows annual analyses of ≈80,000 serum samples in full automation in a laboratory equipped with 2 NMR spectrometers and the necessary liquid handling instruments (Figure 1). This setup enables quantification of an extensive metabolite profile with >200 measures for costs comparable with standard lipid measurements.

**Standard Tools of Epidemiology Suit Quantitative Metabolomics—and Vice Versa**

Risk associations in different populations, lifestyle determinants, and the genetic underpinnings for standard blood markers used in cardiovascular risk assessment have been studied for decades. Because quantitative metabolomics provides metabolite concentrations in physiological units, the data can be analyzed using the same statistical methods as any other biomarker assay data. As a corollary, no special knowledge on spectroscopy is required for the epidemiological analyses because concentrations are determined for identified metabolites as illustrated in Figure 2. As a matter of fact, the actual experimental methodology is of limited interest in epidemiological studies. Whether an amino acid concentration is quantified by NMR, MS, or chromatography does not make the molecular measure any different—only the measurement accuracy (detection limit and bias) may be affected.

The use of standard medical statistics, such as linear models, is often suitable for the initial biomarker assessment and validation in independent studies. This approach greatly facilitates the interpretation—if the metabolites have not previously been associated with the risk factor or disease outcome, it seems most intuitive to study association magnitudes in terms of, for example, regression coefficients and odds ratios, which are adjustable for potential confounders. The effect estimates may then be compared with those of standard risk factors and replicated in independent studies. 17,19,28

**Hypothesis and Data-Driven Science Hand in Hand**

Personalized medicine has been hyped for many years. 79,80 However, the success of this potentially useful concept to base health and disease assessment on individual molecular data is intrinsically hampered in the area of common metabolic disorders. This is because of genetic and metabolic complexity, manifested by continuous biological processes and phenomena that neither allow for clear classifications nor accurate predictions with currently available information. 3,81 However, a recent

---

**Figure 1. (Continued)** 600 MHz spectrometers, both with the SampleJet robotic sample changer. The SampleJet can hold 480 samples (5 positions for NMR tubes in the standard 96-well plate array) in a cooled (+6°C) temperature. The cooled sample changer is a prerequisite when working with biological samples to prevent degradation. The 500 MHz spectrometer is equipped with a selective inverse room temperature probe head, whereas the 600 MHz system has a cryogenically cooled triple resonance probe head (CryoProbe Prodigy TCI). For the native serum samples the lipoprotein (LPO) and low-molecular-weight metabolites (LMWM) data can be automatically collected either with the 500 MHz or the 600 MHz spectrometer. Standardized parameters are used for data acquisitions. The entire time required for the sample handling and measurements is ≈8 and 5 minutes at 500 and 600 MHz, respectively. After these measurements, the same samples go through a standardized lipid extraction procedure based on multiple extraction steps containing saturated sodium chloride solution, methanol, dichloromethane, and deuteriochloroform. The extraction procedure is done manually with an Integra Biosciences VIAFLO 96 channel electronic pipette. These lipid extracts are then moved into the NMR tubes and the extracted lipid (LIPID) data are collected in full automation with the 600 MHz instrument using a standard parameter set. The time required for the sample handling and the LIPID measurement is ≈5 minutes. The depicted setup of a 500 MHz and a 600 MHz platform allows ≈80,000 samples to be analyzed annually. The initial data processing, including the Fourier transformations to NMR spectra and automated phasing are done using the computers that control the spectrometers; the spectra are then automatically transferred to a centralized server, which performs various further automated spectral processing steps, including overall signal check for missing/extra peaks, background control, baseline removal and spectral area-specific signal alignments. The spectral information of the actual sample also undergoes various comparisons with the spectra of the 2 quality control samples; the data for which is also followed and compared in a consecutive manner. For those spectral areas that pass all the quality control steps, regression modeling products are performed to produce the quantified molecular data. Also the individual metabolic measures undergo various statistical quality control steps and are also checked against an extensive database of quantitative molecular data. The metabolic measures that pass all quality control steps are stored in the database and are ready for various epidemiological data analysis. FA indicates fatty acid; MUFA, monounsaturated fatty acids; and PUFA, polyunsaturated fatty acids.
**METABOLIC MEASURES**

<table>
<thead>
<tr>
<th>Ketone bodies (mmol/l)</th>
<th>Amino acids (mmol/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Acetate</td>
<td>• Alanine</td>
</tr>
<tr>
<td>• Acetoacetate</td>
<td>• Glutamine</td>
</tr>
<tr>
<td>• 3-hydroxybutyrate</td>
<td>• Glycine</td>
</tr>
<tr>
<td></td>
<td>• Histidine</td>
</tr>
</tbody>
</table>

**Glycolysis related metabolites (mmol/l)**

- Glucose
- Lactate
- Pyruvate
- Citrate
- Glycerol

**Inflammation (mmol/l)**

- Glycoprotein acetyls, mainly α-1-acid glycoprotein

**Fatty acids and saturation**

- Total fatty acids
- Estimated fatty acid chain length
- Estimated degree of unsaturation

**Fatty acids (mmol/l and % of total FAs)**

- Omega-3 fatty acids
- Omega-6 fatty acids
- Polyunsaturated fatty acids
- Monounsaturated fatty acids; 16:1, 18:1
- Saturated fatty acids
- Docosahexaenoic acid; 22:6
- Linoleic acid; 18:2
- Conjugated linoleic acid

**Cholesterol (mmol/l)**

- VLDL cholesterol
- LDL cholesterol
- HDL cholesterol
- HDL₃ cholesterol
- HDL₄ cholesterol
- Cholesterol
- Free cholesterol
- Esterified cholesterol
- Remnant cholesterol

**Apolipoproteins (g/l)**

- ApoA-I
- ApoB
- ApoB/ApoA-I

**Lipoprotein particle size (nm)**

- Mean diameter of VLDL particles
- Mean diameter of LDL particles
- Mean diameter of HDL particles

**Lipid measures for each subclass**

- Esterified cholesterol (mmol/l and % of total lipids)
- Free cholesterol (mmol/l and % of total lipids)
- Triglycerides (mmol/l and % of total lipids)
- Phospholipids (mmol/l and % of total lipids)

**Fluid balance**

- Creatinine (mmol/l)
- Albumin (g/l)

**Glycerides & phospholipids (mmol/l)**

- VLDL triglycerides
- LDL triglycerides
- HDL triglycerides
- Triglycerides
- Diglycerides
- Phosphoglycerides
- Ratio of diglycerides to triglycerides
- Ratio of triglycerides to phosphoglycerides
- Phosphatidylcholine and other cholesterols
- Sphingomyelins
- Total cholestrols

**Remnant cholesterol** is defined as (non-HDL and non-LDL)-cholesterol.²⁷

**14 LIPOPROTEIN SUBCLASSES**

**Average particle size (diameter in nm)**

- >75
- 53.6
- 44.5
- 36.8
- 31.3
- 28.6
- 25.5
- 23.0
- 18.7
- 14.3
- 12.1
- 10.9
- 8.7

**Average lip composition (%)**

- 74%
- 64%
- 64%
- 56%
- 37%
- 20%
- 45%
- 41%
- 40%
- 37%
- 45%
- 45%
- 39%
- 40%
- 25%

**Figure 2.** The quantitative molecular output from the serum nuclear magnetic resonance metabolomics platform depicted in Figure 1. The colored circles depict the spectral origin of the metabolic measures according to Figure 1. The metabolic measures are collected under headers that link them to a molecular group or a biological rationale. Each filled bullet denotes a primary measure and open bullet a derived measure—for instance, for the fatty acids it often makes biological sense to use the percentage of a fatty acid of the total fatty acids. As an example, characteristic population distributions are illustrated for omega-6 fatty acids both in mmol/L and as percentage of omega-6 fatty acids relative to total serum fatty acids. If not otherwise specified, each measure refers to total concentration in serum as measured by the platform as illustrated in Figure 1. For example, cholesterol means serum total cholesterol and triglycerides refer to serum total triglycerides. Remnant cholesterol is defined as (non-HDL and non-LDL)-cholesterol.²⁷ The platform provides data on 14 lipoprotein subclasses that are characterized by particle size as indicated on the bottom of the figure. For each subclass esterified and free cholesterol, triglycerides and phospholipids are quantified, allowing calculation of the average lipid compositions of the lipoprotein subclass particles as pictured in the figure. The number of metabolic measures shown in the figure is 233. The platform development is still ongoing and some additional metabolite measures will become available; these will be based on improved automated data analyses and will, therefore, be available retrospectively for all measured samples. apoA-I indicates apolipoprotein A-I; apoB, apolipoprotein B; FA, fatty acid; HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; and VLDL, very-low-density lipoprotein.
longitudinal work in the case of a single individual has demonstrated the use of extensive integrative personal omics profiles to interpret biological pathways and infer disease risks. Nevertheless, deep multomics phenotyping of an individual is distant from current public health care and can be envisioned to remain unreasonably costly (and impractical to implement) for many years to come. Notwithstanding, the principle of systemic phenotypes and multivariate data analyses may well not be that many years away from customary applications in population epidemiology and biomedicine. We have demonstrated earlier that multivariate phenotyping based on serum NMR can be applied to classify individuals with type 1 diabetes mellitus to metabolically distinct risk groups. The phenotypes were identified in a hypothesis-free data-driven analysis based on comprehensive metabolomics data. We have applied a similar approach more recently, for example, to study the disease progression and pathogenesis of diabetic kidney disease and to show that individual systemic characteristics affect the metabolic outcomes of dietary interventions. The results demonstrate that the entire metabolic profile, rather than individual metabolites, is beneficial for capturing (patho)physiological processes.

It is important to note that even though it is now possible to identify characteristic metabolic phenotypes based on comprehensive systemic data, we are still bound to population relationships in inferring individual health prospects. This is because the biological heterogeneity and overlap of the current phenotypes are too large for statistically pertinent individual assessment. This limitation is likely to change as comprehensive metabolic data becomes available in such numbers that allow in-depth biological stratification. It is difficult to rationalize what the number of people would need to be to assess this, but we estimate ≈100,000 being a minimum for such rigorous population level stratifications that would allow substantial improvements in risk prediction models. This would require that each metabolic phenotype would be homogeneous enough to capture preponderance of multiple biological pathways and convey partially non-overlapping multidimensional metabolic data. To reach this goal, it is essential not only to focus on disease diagnostics but also put more effort on continuous definitions of health and systemic function (and dysfunction) over the life course. In addition, for a truly global understanding of systemic metabolism, we would need extensive data on metabolic biomarkers in different age groups and ethnicities. It will also be important to get a representative experience of lifestyle and environmental effects and to collect longitudinal data in population settings.

Quantitative metabolomics allows both testing clear hypotheses and hypotheses-free assessment of biological questions. In the case of extensive epidemiological data, we do advocate hypothesis-free exploratory analyses. Our knowledge on many new metabolic measures, now becoming available via the applications of new technologies, is limited and recent works have illustrated intriguing new findings based on comprehensive metabolomics data in relation to metabolic disease outcomes. Importantly, integration of quantitative metabolomics data in hypothesis as well as in data-driven research is straightforward and research versatility can be maintained.

### Physiological Applications

Large-scale quantitative metabolomics data allow a choice of a focused or a holistic research approach. This choice is illustrated in Figure 3: the comprehensive metabolic data obtained from the NMR metabolomics platform makes it equally possible to conduct a focused examination of amino acids in diabetes mellitus or a holistic assessment on the effects of long-term physical activity. In the following sections, we highlight some examples of quantitative metabolomics studies that characterize metabolic risk factors and discover novel biomarkers.

### Systemic Metabolic Effects of Physical Activity

It is commonly accepted that exercise is good for us and many guidelines recommend regular physical activity. Long-term physical activity is also known to reduce metabolic risks, although molecular information on its effects on the systemic metabolism is sparse because of challenges from multiple interconnected pathways. In a situation like this, the serum NMR metabolomics platform offers excellent opportunities to concurrently probe various metabolic pathways. This holistic approach is illustrated in Figure 3 for our recent study on the long-term leisure-time physical activity and serum metabolome. Here, we profiled 16 twin pairs who were discordant for physical activity for >30 years and >1000 age- and sex-matched pairs of unrelated adults from 3 population-based cohorts who were persistently (≥25 years) active or inactive. This study is a good example how quantitative metabolomics inherently allows for an elegant epidemiological study setup with appropriate numbers and independent biological replication. The holistic approach revealed various biological findings: the active individuals compared with inactive individuals had better lipoprotein profiles and higher levels of polyunsaturated relative to total fatty acids. These findings pinpointing better lipid profiles for those with long-term physical activity were expected and provided good support for the new methodology (as well as for the holistic research approach). This also substantiates that the experimental methodology is not central in epidemiological studies when concentrations of specific molecular identities are measured. On the top of the expected findings on serum lipoprotein and lipid profiles, we also made a few new findings, for example observations that lower isoleucine (a branched-chain amino acid [BCAA]) and lower glycoprotein (an acute-phase reactant) concentrations relate to persistent physical activity.

The comprehensive metabolic characterization highlighted consistent associations with physical activity across multiple metabolic pathways, including specific lipids, amino acids, and inflammation markers, and findings persisted after adjustment for body mass index. The widespread metabolic differences demonstrated better metabolic health in physically active than in inactive individuals. The great importance of physical activity for cardiovascular prevention thus arises from multiple metabolic mechanisms rather than pertain to a single process. In fact, most risk factors exert systemic influences on numerous metabolites rather than a few measures. Detailed quantitative metabolite profiling, therefore, provides a valuable opportunity to characterize such overall systemic effects. And, as suggested...
in the related editorial by Cheng, the individual serum metabolome could allow not only risk assessment of disease but also tailored prescription for intervention.

**Amino Acids and Type 2 Diabetes Mellitus**

The holistic use of the comprehensive systemic data is well suited to simultaneously explore multiple metabolic pathways. However, the individual quantitative molecular measures can also be used in a focused way. This is illustrated in Figure 3 for the BCAAs (isoleucine, leucine, and valine) with respect to their potential role in type 2 diabetes mellitus. The circulating levels of BCAAs have long been known to associate with obesity and insulin resistance. Recent prospective metabolomics studies have revealed that BCAAs (and tyrosine and phenylalanine) also reflect the risk for developing type 2 diabetes mellitus. These initial results by Wang et al., based on plasma MS profiling using a nested, case-control design in 2 population cohorts, were corroborated by results from our serum NMR metabolomics platform for >9000 Finnish men.

To clarify the role of BCAAs in insulin resistance and hyperglycemia, we and our collaborators went on to analyze the amino acid data in relation to various stages in the pathogenesis of type 2 diabetes mellitus. As a summary, the BCAAs were found to strongly associate with insulin resistance already in young adults, even when adjusting for the risk factor components of the metabolic syndrome. The associations were stronger for men, and no association were found for women with low waist circumference; specific findings that also exemplify the importance of clinical stratification allowed by large sample size and quantitative data. In addition, the BCAA concentrations were found to gradually increase from normal to the prediabetic range, and further along the extent of hyperglycemia. In prospective analyses, BCAAs predicted the levels of insulin resistance in young adults at 6-year follow-up, but it required an older study population to find prospective associations with 6-year glucose levels. The BCAA concentrations were also associated with the risk for onset of diabetes mellitus during a 5-year period among men with normal glycemic control. However, associations prevailed when adjusting for baseline insulin secretion; however, the associations were diminished when adjusting for insulin resistance. These results are indications that insulin resistance plays a mediating role in the relation between BCAAs and type 2 diabetes mellitus. However,
initial assessment of genetic variants regulating circulating BCAA levels did not suggest causal associations with insulin resistance. Overall, these results from several large population-based cohorts are a good exemplar of the focused molecular use of quantitative metabolic data to understand disease pathogenesis.

Disease Biomarkers
For decades the repertoire of blood biomarkers routinely used in epidemiological studies of disease pathogenesis and prediction has been surprisingly limited, and these same key measures, for example, cholesterol and glucose, are still often the cornerstone. The considerable complexity of cardiovascular and metabolic diseases would suggest that this small group of molecular measures does only constitute the beginning of understanding these diseases. Thus, several new biomarker findings based on comprehensive quantitative metabolomics are not unexpected. The serum NMR metabolomics platform can aid biomarker discovery in a hypothesis-free manner by providing metabolite quantifications across multiple pathways: all metabolic measures can then be separately tested for the potential disease association or incidence. This should be followed by appropriate independent replication of the candidate biomarkers identified in the discovery cohort. Unfortunately, the promise of metabolomics in biomarker discovery has not been realized and although various articles have been published, there is little consistency and relatively little rigor in the metabolomics works in this area as recently pointed out by Wishart et al.

Our experience indicates that many common systemic metabolites, such as amino acids and polyunsaturated fatty acids, show consistent associations in properly powered epidemiological studies. For instance, we have demonstrated prospective associations of aromatic amino acids with carotid intima-media thickness, a subclinical measure of atherosclerosis, in young adults. These findings have been corroborated by MS-based profiling and we have recently extended their role to later stages of cardiovascular disease in large population cohorts (Würtz et al, 2015, doi: 10.1161/CIRCULATIONAHA.114.013116). The association strengths of aromatic amino acids are comparable with those of established risk factors, and they remain predictive for the disease even when adjusted for standard lipids and glycemic traits. We have also indicated that serum concentrations of polyunsaturated fatty acids, including docosahexaenoic acid, are independent predictors of subclinical atherosclerosis and cardiovascular risk. Our findings indicate that aromatic amino acids, glycolysis substrates, and polyunsaturated fatty acids are reflective of cardiovascular disease risk in general population settings. Extensive work in this area is ongoing and it is too early to conclude if these biomarkers will be helpful in refining risk predictions with a single prediction model for the whole population. However, we already know that these biomarkers highlighted by quantitative metabolomics seem consistent in multiple populations and as assayed by NMR and MS profiling technologies.

A lot of work is ongoing in the area of cardiovascular disease and metabolic biomarkers. There are also other areas that would benefit of a general metabolic risk assessment. In our hypothesis-free analyses, we recently encountered an intriguing finding: among the metabolic measures quantified by the serum NMR metabolomics platform, we found 4 biomarkers that are predictive for all-cause mortality. Early and accurate identification of high-risk individuals, who seem healthy, but in fact have an underlying serious illness, would provide a rationale for further medical assessments and valuable opportunities for preventive treatments. Currently, there is no such test that could accurately assess whether a person is at risk of ill health generally, or likely to die soon from a disease. Thereby our unforeseen finding of 4 biomarkers that predict the risk of short-term death (within 5 years) among a general population, rather than within people already known to be ill, is of high clinical importance. The finding was based on 2 large population-based cohorts, a discovery cohort of almost 10,000 Estonian individuals and a replication cohort of ≈7500 Finnish people. The 4 biomarkers—albumin, glycoprotein, very-low-density lipoprotein particle size, and citrate—are implicated in various pathophysiological mechanisms, including fluid imbalance, inflammation, lipoprotein metabolism, and metabolic homeostasis. Although these findings raise more questions than they provide answers, they are likely to open a new path to unravel novel relationships between systemic biomarkers and diverse morbidities.

Causality, Pleiotropy, and Functional Genetics
An impactful direction in epidemiology is assessment of causal relationships; a widespread tool for this is the Mendelian randomization approach. Relying on specific genetic instruments Mendelian randomization enables inference of causal pathways, for example, whether obesity is causally affecting systemic metabolic concentrations. Mendelian randomization analysis is an elegant way to diminish the effects of confounders in epidemiological studies but it entails specificity of the genetic instruments and calls for extensive data sets. Quantitative serum NMR metabolomics is assisting Mendelian randomization studies with respect to both of these challenges. In the assessment of the validity of a genetic instrument, for example, if a single-nucleotide polymorphism is pleiotropic, the comprehensive metabolic coverage offered by the platform is extending the evaluation of the gene-metabolite associations to a lot more refined situation than possible with only a few conventional epidemiological markers. We have recently illustrated that many single-nucleotide polymorphisms in various lipid genes, thought to be good genetic instruments based on evaluations only with traditional lipid measures, are actually highly pleiotropic. An example hereof is given in Figure 4 that illustrates the associations of PCSK9, a gene previously interpreted to be specific for LDL cholesterol, and an extensive set of serum metabolic measures. The association pattern of the genetic variant in PCSK9 shows extensive pleiotropy, extending from small LDL over intermediate-density lipoprotein and small very-low-density lipoprotein lipids, and additionally showing associations with, for example, omega-6 fatty acids and sphingomyelin.

Mendelian randomization is likely to be used more and more in epidemiology and also as a substitute for randomized controlled trials. Genomewide association studies have identified many variants that each affects multiple traits, hereby suggesting that pleiotropic effects on human complex traits are
The reliability of Mendelian randomization studies highly depends on genuine genetic instruments and thereby the ability to assess the validity of the genetic instruments via detailed metabolic phenotyping is of immense value.

The genomewide association study era has improved our understanding of disease pathogenesis by identifying genetic variants associated with complex human traits and disease phenotypes. Despite the massive association information on genes and multiple traits, the functional information on the implicated genes is limited and the genetic data are still typically able to explain only a small portion of the trait variance. Detailed quantitative metabolic data can assist in both of these challenges. The effect size between a genetic variant and a trait depends on the fundamental definition of the trait—the relations are likely to get stronger as 1 moves closer to the molecular source. The missing heritability is, therefore, not simply a reflection of what cannot be found by a common genetic variant association; it relates fundamentally to the biological and molecular rationale of the trait. Total cholesterol is a good exemplar of wide-ranging biological heterogeneity intrinsically incorporated into a single measure; total cholesterol sums up all the cholesterol molecules in circulation and thereby has no distinction for the lipoprotein particle it is carried with. The lipoprotein metabolism is complex and involves various particle subclasses that have a different and even opposite biological roles—LDL and HDL particles being a well-known example. We have recently demonstrated this in a genetic and metabolic fine mapping study on 8330 Finnish individuals, including data on 14 lipoprotein subclasses, quantified by the serum NMR metabolomics platform. The results illustrate the immense value of detailed molecular profiling to address the validity of genetic instruments used in Mendelian randomization. Remnant cholesterol is defined as (non-HDL and non-LDL)-cholesterol. ApoA-I indicates apolipoprotein A-I; apoB, apolipoprotein B; C, cholesterol; CE, cholesterol esters; DHA, docosahexaenoic acid; FC, free cholesterol; HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; L, large; L, total lipids; LA, linoleic acid; LDL, low-density lipoprotein; M, medium; MUFA, monounsaturated fatty acids; PL, phospholipids; PUFA, polyunsaturated fatty acids; S, small; Sat FA, saturated fatty acids; TG, triglycerides; XL, very large; XS, very small; XXL, extremely large; and VLDL, very-low-density lipoprotein.

Figure 4. Associations of rs11591147 in PCSK9 with various lipoprotein subclass and metabolite measures from the serum nuclear magnetic resonance metabolomics platform. The colored circles depict the spectral origin of the metabolic measures according to Figures 1 and 2. Color-coding denotes association magnitude as assessed by linear regression analysis in units of standardized metabolite concentration per allele. Statistical significance is indicated by dots and stars for nominally significant (P<0.05) and Bonferroni-corrected (P<0.0001) associations, respectively. The associations were meta-analyzed for >10500 Finnish individuals as previously disclosed. The association pattern of PCSK9 pinpoints extensive pleiotropy, extending from small LDL over IDL as well as small and medium VLDL lipids, and additionally showing associations with, for example, omega-6 fatty acids and sphingomyelin. These results illustrate the immense value of detailed molecular profiling to address the validity of genetic instruments used in Mendelian randomization. Remnant cholesterol is defined as (non-HDL and non-LDL)-cholesterol. ApoA-I indicates apolipoprotein A-I; apoB, apolipoprotein B; C, cholesterol; CE, cholesterol esters; DHA, docosahexaenoic acid; FC, free cholesterol; HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; L, large; L, total lipids; LA, linoleic acid; LDL, low-density lipoprotein; M, medium; MUFA, monounsaturated fatty acids; PL, phospholipids; PUFA, polyunsaturated fatty acids; S, small; Sat FA, saturated fatty acids; TG, triglycerides; XL, very large; XS, very small; XXL, extremely large; and VLDL, very-low-density lipoprotein.

The finding indicates distinct metabolic characteristics for small and large HDL particles as also previously indicated by the gene coexpression patterns in circulating leukocytes. The fine-grained association patterns provide better understanding of the biological processes modifying lipid levels and could potentially clarify the elusive role of HDL in cardiovascular disease. Using similar metabolic fine mapping, our colleagues...
have recently started to use the detailed serum metabolomics data and moved toward more functional studies, for example, elucidating how certain genes link with metabolic intermediates in the case of glycemia and type 2 diabetes mellitus, blood pressure, and liver enzymes. It is evident from these pioneering works that the comprehensive quantitative data based on the serum NMR metabolomics platform will have a substantial role also in characterizing gene functions.

Past, Present, and the Future of NMR-Based Metabolomics

NMR-metabolomics has its roots in biomedical NMR spectroscopy. Some key works in the early 1980s spotlighted the emerging flexibility of proton NMR to study biofluids and infer various molecular information of potential medical relevance. It soon became evident that proton NMR can also be used to quantify lipoprotein lipids and the term...
metabolomics was coined by Jeremy K. Nicholson in the late 1990s to mean understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Since then, metabolomics has become the generally accepted broad concept for the science studying chemical and biomedi- cal processes related to multiple metabolites; by the end of 2013 almost 7000 items had accumulated into PubMed with either metabolomics or metabolomics. There is no doubt that quantitative metabolomics is providing (and will continue to provide) original data in large epidemiological studies for molecular insights in physiology and disease pathogenesis. When looking at the road ahead, we envision a few periods of development from the current epidemiological research toward data-driven health care; these eras and their characteristic features are outlined in Figure 5.

Although metabolomics is rapidly becoming an integral part of epidemiological studies, it seems that our platform is the only 1 used in publications with comprehensive serum NMR metabolomics data for >1000 individuals. To date, there are only 3 studies published in which the number of individuals exceeds 10000 (Table 1). Evidently, present metabolomics studies have only started to scratch the surface in terms of potential applications in epidemiology. That said, the published works do not represent the current experimental situation; we have now analyzed almost 250000 samples (from around 100 epidemiological and clinical cohorts) with the serum NMR metabolomics platform and are routinely conducting studies encompassing >25000 people. We expect the first combined studies >100000 individuals to be completed during 2015. All the current data have been acquired using 2 platforms, 1 operational since January 2009 and the other since February 2013. On the basis of the present funding and strategic decisions, 26 additional serum NMR metabolomics platforms are expected to be up and running in Finland and in the United Kingdom by the end of 2015, which increases the capacity of the platform analyses to >250000 samples per year.

We are phasing an era when extensive sets of samples can be profiled by NMR metabolomics in a short time scale. The resulting coherent quantitative metabolite data can be meta-analyzed across multiple cohorts in standard manners. Interestingly, in comparison with the current expenditure on genomewide arrays, the cost of the quantitative serum NMR metabolomics is only around one third. It is, therefore, no longer unrealistic to envision that quantitative metabolomics would be incorporated as a routine for all serum samples in large biobanks; this would make perfect sense both from the biological research and cost point of view—the standard output of >200 molecular measures would vastly extend the biomedical relevance of the sample collections. Moreover, keeping in mind that many routine clinical and epidemiological markers, including various cholesterol measures, triglycerides, apolipoproteins A-I and B, creatinine, albumin, and glucose are also included in the platform output, the saving from avoiding many separate clinical chemistry assays would simply cover the cost of the metabolomics.

Sources of Funding

Our work in the area of quantitative serum NMR metabolomics has been supported by the Academy of Finland, TEKES—the Finnish Funding Agency for Technology and Innovation, the Sigrid Juselius Foundation, the Novo Nordisk Foundation, the Finnish Diabetes Research Foundation, and the strategic and infrastructural research funding from the University of Oulu, Finland. In addition, our current endeavors toward large-scale systems epidemiology are funded by the British Heart Foundation, the UK Wellcome Trust, and the UK Medical Research Council via the University of Bristol Integrative Epidemiology Unit.

Disclosures

All authors are shareholders of Brainshake Ltd. (www.brainshake.fi), a startup company offering NMR-based metabolite profiling.

References


Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics
Pasi Soininen, Antti J. Kangas, Peter Würtz, Teemu Suna and Mika Ala-Korpela

Circ Cardiovasc Genet. 2015;8:192-206
doi: 10.1161/CIRCGENETICS.114.000216

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/8/1/192

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org//subscriptions/