AHA Scientific Statement

Potential Impact and Study Considerations of Metabolomics in Cardiovascular Health and Disease
A Scientific Statement From the American Heart Association

Susan Cheng, MD, MMSc, MPH, Chair; Svati H. Shah, MD, MHS, FAHA, Co-Chair;
Elizabeth J. Corwin, PhD; Oliver Fiehn, MD; Robert L. Fitzgerald, PhD;
Robert E. Gerszten, MD, FAHA; Thomas Illig, PhD; Eugene P. Rhee, MD;
Pothur R. Srinivas, PhD, MPH; Thomas J. Wang, MD, FAHA; Mohit Jain, MD, PhD;
on behalf of the American Heart Association Council on Functional Genomics and Translational Biology;
Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; and Stroke Council

Abstract—Through the measure of thousands of small-molecule metabolites in diverse biological systems, metabolomics now offers the potential for new insights into the factors that contribute to complex human diseases such as cardiovascular disease. Targeted metabolomics methods have already identified new molecular markers and metabolomic signatures of cardiovascular disease risk (including branched-chain amino acids, select unsaturated lipid species, and trimethylamine-*N*-oxide), thus in effect linking diverse exposures such as those from dietary intake and the microbiota with cardiometabolic traits. As technologies for metabolomics continue to evolve, the depth and breadth of small-molecule metabolite profiling in complex systems continue to advance rapidly, along with prospects for ongoing discovery. Current challenges facing the field of metabolomics include scaling throughput and technical capacity for metabolomics approaches, bioinformatic and chemoinformatic tools for handling large-scale metabolomics data, methods for elucidating the biochemical structure and function of novel metabolites, and strategies for determining the true clinical relevance of metabolites observed in association with cardiovascular disease outcomes. Progress made in addressing these challenges will allow metabolomics the potential to substantially affect diagnostics and therapeutics in cardiovascular medicine. (***Circ Cardiovasc Genet***. 2017;10:e000032. DOI: 10.1161/HCG.0000000000000032.)

    Key Words: AHA Scientific Statements ■ cardiovascular diseases ■ metabolome ■ metabolomics

Substantial technical advances have enabled robust and comprehensive profiling of molecular markers in the study of human health and disease. These advances have led to refinement in the application of genomic, epigenetic, transcriptomic, proteomic, and metabolomic tools in clinical studies designed to elucidate disease pathobiology, to identify clinically relevant biomarkers, and to guide therapy. Metabolomics, among the most recently emerged "omics" fields, involves the systematic measure and study of small-molecule metabolites across biological systems using biosamples that include plasma, urine, saliva, and tissues. Metabolomics serves as a reflection of gene and protein functional activity and, in doing so, captures information that is proximal to a given disease phenotype while spanning the spectrum from genetic sequence to cellular physiology. Therefore, metabolomics offers the potential to complement other cellular measures, including the vast genetic information derived from DNA sequencing in humans. Notably, whereas

Circ Cardiovasc Genet **is available at http://circgenetics.ahajournals.org**   **DOI: 10.1161/HCG.0000000000000032**

the genetic sequence is largely static from birth, metabolomic measures are relatively dynamic, representing cellular activity and the effects of extrinsic exposures, including dietary intake, physical fitness, microbiota variation, and toxicant exposures. The ability to integrate measures of both intrinsic and extrinsic origin, as well as their interplay, may effectively bridge investigations of both gene and environment and will be critical for understanding complex, multifaceted human diseases such as cardiovascular disease (CVD). Accordingly, the rich small-molecule metabolite data represented by an individual's metabolome, along with the chemical diversity of population-scale metabolomes, could lead to key pathobiological insights and offer powerful tools for personalizing risk assessment and disease prevention. In this American Heart Association scientific statement, we aim to provide an overview of the rapidly evolving field of metabolomics, its applications to date for understanding CVD, and its future role in biomedicine. Instead of comprehensively reviewing technical metabolomics, which has been offered by prior reports, we sought herein to highlight for the scientific community the key issues and challenges relevant to the application of metabolomics, including those related to metabolite standards, databases and identification, study design, bioinformatics, and clinical translation.

Metabolomics Approaches

A range of technical approaches have been used for the analysis of metabolites in biospecimens, including nuclear magnetic resonance (NMR) and mass spectrometry (MS). These analytical tools have been extensively reviewed elsewhere.[1,2] Although each tool has its own strengths and detractions, both are capable of assaying metabolites in a number of biofluids and tissues. NMR typically requires minimal sample processing and provides absolute quantitation of analytes, but it is limited by sensitivity and thus generally allows measurement of only the most abundant of metabolites. By contrast, MS-based metabolomics can involve direct infusion into a mass spectrometer or, more routinely, coupling to upfront chromatography, either liquid chromatography (LC) or gas chromatography, to reduce sample complexity before mass analysis. Additionally, both LC-MS and gas chromatography–MS require sample preparation with extraction of metabolites from complex specimens and, in many cases, even chemical processing before assay. Necessary preprofiling procedures notwithstanding, MS-based approaches offer a much greater range of detection with high sensitivity, allowing the measure of up to thousands of small molecules at a given time. Such MS-based approaches may be operated either in a "targeted" fashion, with detection of prespecified molecules (typically up to 100–200 metabolites at a time), or as part of an "untargeted" full scan, in discovery mode without prespecification of molecules (typically measuring >1000 molecules at a time). Thus, advancing from NMR to targeted MS to untargeted MS approaches provides increasing analytical coverage and hence greater discovery potential. However, with a greater number of measures come additional analytical and computational hurdles. These challenges include those pertaining to alignment of data to allow comparison of specific metabolites across hundreds to thousands of samples, handling and extraction of large spectral data files, and ultimately the definitive identification of metabolites of interest (discussed below in Current Applications).

Metabolomics Discoveries in CVD

Although still early in its technical and scientific evolution, metabolomics and its application to the study of clinical CVD have already demonstrated substantial potential for discovery and insight.

Metabolite Correlates of CVD

An overview of clinical studies of metabolomics in CVD is provided in Table 1. Early metabolomics studies made use of NMR-based approaches to distinguish individuals with multivessel coronary artery disease (CAD) from individuals without angiographic CAD using spectral fingerprint patterns.[3] Although initially promising with reportedly >90% predictive power for identifying patients with CAD from control subjects, spectral fingerprint data were in hindsight more reflective of sex and statin medication use than independent correlates of CAD.[4] These results underscore the need for comprehensive follow-up to identify signatures of a potential cause rather than a consequence of disease, careful chemistry to annotate metabolite peaks, and statistical evaluation of potential confounders that can influence dynamic metabolomic signals. A later study used a targeted tandem MS (MS/MS) approach to identify and quantify 69 metabolites in individuals with angiographic CAD, revealing higher levels of branched-chain amino acids (BCAAs) in association with CAD even after adjustment for cardiovascular risk factors.[7,13] Subsequently, high-throughput NMR measures of 68 abundant plasma metabolites, performed in a multicohort epidemiological study (including the National Finnish FINRISK study, Southall and Brent Revisited Study, and British Women's Health and Heart Study), showed that higher phenylalanine and monounsaturated fatty acid levels were associated with increased CVD risk, whereas higher concentrations of omega-6 fatty acids and docosahexaenoic acid were associated with lower risk of CVD over a follow-up period spanning decades. Importantly, these observations were validated not only in multiple independent cohorts but also with orthogonal targeted MS approaches applied to an additional 2000 participants from the Framingham Offspring Study, demonstrating the potential for generalizability and validity of metabolomic observations.[16]

Untargeted, also known as nontargeted, MS-based analyses have now extended the measure from several dozen to several thousand metabolites in a single sample. Recent applications of LC-MS have identified 4 unsaturated lipids (lysophosphatidylcholine 18:1, lysophosphatidylcholine 18:2, monoglyceride 18:2, and sphingomyelin 28:1) in association with CVD independently of traditional risk factors in the TwinGene cohort; these findings were then validated in the ULSAM (Uppsala Longitudinal Study of Adult Men) and PIVUS (Prospective Investigation of the Vasculature in Uppsala Seniors) cohorts.[14]

Table 1.   Major Findings to Date From Selected Clinical Studies of Metabolomics and Cardiovascular Disease

Study	Study Design and Size	Method	Major Findings
Brindle et al[3] (2002)	Case-control: n=36 patients with triple-vessel disease compared with n=30 patients with normal coronaries	NMR	Case-control: PLS-DA (80% training set, 20% test set) demonstrated the most discriminating analytes to be lipids, mostly VLDL, LDL, HDL, and choline
	Cross-sectional: n=76 patients with CAD with 1-, 2-, or 3-vessel disease		Cross-sectional: NMR chemical shifts of δ1.30 and δ3.22 (representing LDL variants and other lipid elements) discriminated CAD severity, whereas traditional clinical risk factors did not
Kirschenlohr et al[4] (2006)	Cross-sectional: n=322 patients with normal coronaries or 1-, 2-, or 3-vessel CAD	NMR	Predictions for CAD presence were only 80% correct for patients not treated with statins and 61% for treated patients
Lewis et al[5] (2008)	Experimental: n=36 planned MI patients	Targeted LC-MS	Experimental: metabolites involved in pyrimidine metabolism, TCA cycle, and the pentose phosphate pathway were altered after planned MI
	Cross-sectional: 12 acute MI patients; 9 non-CAD patients		Cross-sectional: aconitic acid, hypoxanthine, TMAO, and threonine distinguished presence vs absence of spontaneous MI
Turer et al[6] (2009)	Experimental: n=37 patients undergoing CABG	Targeted LC-MS	Lower extraction rates of acetylcarnitine and 3-hydroxybutyryl-carnitine by the myocardium were observed after intraoperative reperfusion
Shah et al[7] (2010)	Cross-sectional and prospective: n=757 patients total (3 groups of CAD case-control pairs)	Targeted LC-MS	Cross-sectional: BCAAs and urea cycle metabolites were associated with presence of CAD in derivation and validation cohorts
			Prospective: dicarboxylacylcarnitines predicted incident CAD events
Wang et al[8] (2011)	Cross-sectional: n=1876 patients undergoing elective cardiac catheterization	Targeted LC-MS	Higher levels of TMAO, choline, and betaine were associated with presence of CVD
Shah et al[9] (2012)	Prospective: n=2023 patients undergoing elective cardiac catheterization	Targeted LC-MS	Long-chain dicarboxylacylcarnitines, BCAAs, and fatty acids were associated with death or MI
			Medium-chain acylcarnitines were also associated with death
Koeth et al[10] (2013)	Cross-sectional and prospective: n=2595	Targeted LC-MS	Higher L-carnitine in plasma was associated with prevalent CVD and incident coronary, stroke, and mortality events in individuals with concurrently high TMAO levels
Tang et al[11] (2013)	Prospective: n=4007 patients undergoing elective cardiac catheterization	Targeted LC-MS	Higher baseline TMAO was associated with major adverse cardiovascular event (death, MI, or stroke)
Zheng et al[12] (2013)	Prospective: n=1744 blacks	Targeted LC-MS	Two unknown novel metabolites were associated with incident heart failure in multivariable models: dihydroxy-docosatrienoic acid and an isoform of either hydroxyleucine or hydroxyisoleucine
Bhattacharya et al[13] (2014)	Cross-sectional: n=1983 patients with no vs significant CAD	Targeted LC-MS	Two distinct PCA-derived metabolite factors were associated with both CAD presence and CAD severity in adjusted models: 1 factor was composed of BCAA and the other was composed of short-chain acylcarnitines
Ganna et al[14] (2014)	Prospective: n=3668 from 3 separate derivation and replication cohorts	Untargeted MS	Four metabolites were associated with incident CAD: lysophosphatidylcholine 18:1, lysophosphatidylcholine 18:2, monoglyceride 18:2, and sphingomyelin 28:1
			Monoglyceride 18:2 showed association with an SNP in the ZNF259/APOA5 region with a weak but positive causal effect in mendelian randomization analysis
Jové et al[15] (2015)	Prospective and cross-sectional: n=131 patients with TIA in derivation and n=162 in validation cohorts	Targeted LC-MS	Prospective: low lysophosphatidylcholine (16:0) was associated with increased stroke risk
			Cross-sectional: lower levels of lysophosphatidylcholine (20:4) and lysophosphatidylcholine (22:6) also demonstrated potential risk for stroke or atherosclerosis
Würtz et al[16] (2015)	Prospective multicohort: n=7256 derivation; n=2622 and n=3536 replication	NMR and targeted LC-MS	Four metabolites were associated with incident CVD in meta-analyses: higher phenylalanine and monounsaturated fatty acids, lower omega-6 fatty acids, and docosahexaenoic acids

BCAA indicates branched-chain amino acid; CABG, coronary artery bypass graft; CAD, coronary artery disease; CVD, cardiovascular disease; HDL, high-density lipoprotein; LC, liquid chromatography; LDL, low-density lipoprotein; MI, myocardial infarction; MS, mass spectrometry; NMR, nuclear magnetic resonance; PCA, principal components analysis; PLS-DA, partial least-squares discriminant analysis; SNP, single-nucleotide polymorphism; TCA, tricarboxylic acid; TIA, transient ischemic attack; TMAO, trimethylamine-N-oxide; and VLDL, very low-density lipoprotein.

Using alternative untargeted LC-MS approaches, others have assessed thousands of metabolite features in a small cohort of CVD cases and controls; the metabolite trimethylamine-N-oxide (TMAO) was observed to be highly associated with CAD, and this finding was validated in multiple large cohorts in models adjusting for clinical risk factors.[8,11] Interestingly, TMAO was found to be produced by the interaction of gut bacteria with dietary intake of phosphatidylcholine and carnitine, components typically overrepresented in a meat-based diet.[8,10,17] Furthermore, supplementation with TMAO or intermediary metabolites involved in the production of TMAO, including γ-butyrobetaine, was found to promote atherosclerosis formation in mice.[8,18] Although the exact mechanisms underlying the atherosclerosis-promoting effects of TMAO and related metabolites have yet to be defined, initial work has suggested that TMAO may regulate sterol metabolism and promote macrophage foam cell formation.[8]

Collectively, the breadth and depth of metabolomics studies conducted over the past decade have begun to reveal previously unappreciated factors that may contribute to the pathogenesis of CVD, including dietary factors and gut microbiota variation. As the number of human investigations with metabolomics profiling data available has expanded, studies have begun to associate circulating metabolites not only with global CVD outcomes[7,9] but also with specific CVD subtypes, including myocardial ischemia and infarction,[5] congestive heart failure,[6,12] and stroke.[15]

Metabolite Correlates of Cardiometabolic Risk Factors

Overall, CVD represents a complex collection of multifactorial disease states and risk factors. Thus, to complement the study of the end-disease phenotypes, extensive complementary efforts have been made to understand the pathophysiology of more well-circumscribed risk factors. As seen in other "omics" fields, research focused on individual CVD risk factors can yield more specific information on potential pathways to disease progression. For instance, metabolomics studies of traditional risk factors have identified both plasma and urinary metabolite correlates of blood pressure and hypertension. These include amino acids such as alanine,[19,20] hippuric acid derivatives of gut microbial activity,[19,21] and hexadecanedioate, a dicarboxylic acid that has also been associated with mortality.[22] Given the established importance of lipids (including triacylglycerides and cholesterol-based species) to CVD pathobiology, lipidomic correlates of conventional cholesterol subfractions (eg, high-density lipoprotein, low-density lipoprotein, and very low-density lipoprotein) have also been the subject of multiple investigations.[23–26] Studies to date have demonstrated varying degrees of intercorrelation among lipid species, underscoring the potential of lipidomic variation to shed important insights into cardiovascular risk that are not reflected by traditional lipid measures alone.[27,28]

Perhaps the most consistent and developed finding with respect to cardiometabolic risk traits has been the link between BCAAs with insulin resistance and type 2 diabetes mellitus. BCAAs and related metabolites have been associated with insulin resistance in several cohorts,[29–31] and an investigation in the Framingham Heart Study observed that baseline levels of circulating BCAAs predicted the development of diabetes mellitus up to 12 years later.[32] In another study, BCAA levels were found to decrease more after bariatric surgery than after behavioral weight loss, even for an equivalent amount of weight lost,[33] paralleling the greater impact of surgical weight loss on glucose homeostasis. Furthermore, BCAAs appear lower in metabolically healthy versus metabolically unwell overweight/obese individuals[34] and can predict improvement in insulin resistance with weight loss.[35] Studies in rats suggest that dietary supplementation of high-fat diets with BCAAs can promote insulin resistance regardless of weight gain.[29] In humans, changes in circulating BCAA levels have been independently associated with CVD, adding further to the growing body of data pointing to common metabolic disturbances giving rise to both diabetes mellitus and CVD.[7,13] Although the mechanistic underpinnings of these findings to date remain unclear, recent work has raised the intriguing possibility that elevated BCAAs may originate from gut microbiota rather than from endogenous sources. In identical twins discordant for obesity,[36] fecal microbiota transplantation from the obese twin into germ-free mice led to increased total body and fat mass and the development of obesity-associated metabolic phenotypes; these results stood in contrast to those in mice transplanted with microbiota from the lean twin.[36] Transplantation of microbiota from obese twins was also associated with increased expression of genes involved in metabolism of BCAAs and higher BCAA levels in serum, again offering support for a possible unifying source of risk for obesity, diabetes mellitus, and CVD.

Metabolomics research has also focused on lifestyle behaviors that can alter the natural history of CVD, including exercise- and diet-based interventions, given their pleiotropic role in attenuating the effects of multiple traditional risk factors. Anticipating that metabolite profiles will differ on the basis of the timing and duration of physical activity, both short-term and long-term studies have sought to characterize the biochemical response to exercise. Early studies by Lewis et al have shown that the acute response to an exercise stress test involves an increase in plasma markers of glycogenolysis, lipolysis, and adenine nucleotide catabolism, in addition to an increase specifically in concentrations of amino acids, span 2 tricarboxylic acid cycle intermediates, and niacinamide, a modulator of insulin release and glycemic control.[37] Metabolite changes after prolonged exercise (ie, marathon running) have been correlated with a further increase in markers of lipolysis, an increase in products of ketogenesis, and a marked downturn in concentrations of most amino acids. Huffman et al[38] studied previously sedentary individuals undergoing 6 months of aerobic exercise and observed that inactive overweight to obese adults who underwent exercise training had improved insulin sensitivity that was associated with increased plasma glycine, proline, and alanine levels, as well as increased concentrations of free fatty acids and products of fatty acid oxidation. More recently, studies of adult twin pairs and age- and sex-matched pairs of unrelated adults have found that more active individuals have better lipoprotein cholesterol profiles and higher levels of polyunsaturated relative to saturated fatty acids.[39] Corroborating the findings from studies of cardiometabolic disease and

diabetes mellitus, Kujala et al[39] also found that the BCAA isoleucine was lower in active than inactive individuals, with similar findings for valine, tyrosine, and phenylalanine. With respect to diet, comparatively fewer investigations have been reported, in part because of the relatively greater challenge of analytically disentangling nutrient from nonnutrient signals detectable in biofluids.[40] Nonetheless, a recent experimental study demonstrated the potential of a high-choline or L-carnitine diet in mice to inhibit intestinal microbial production of trimethylamine,[41] a previously reported correlate of cardiometabolic risk. Follow-up preclinical and human studies are needed to further clarify the potential for dietary changes to induce metabolomic variation that, in turn, affects clinical outcomes.

Moving forward, it will be important for all study results to be validated by independent groups in different clinical populations, to delineate causality of CVD risk-related metabolites in experimental models, and to define the biological mechanisms by which these metabolites either promote or attenuate atherogenesis. For metabolites meeting this high bar, it will be essential to determine whether their dietary or pharmacological modulation alters both metabolomic signatures and the associated natural history of CVD in humans. In addition to shedding light on fundamental mechanisms of disease, changes in the human metabolome may serve as indicators of how efficacious a given cardiovascular intervention may be and could even elucidate the molecular basis for such efficacy (ie, pharmacometabolomics).[42] Kitzmiller et al[43] demonstrated that multiple metabolic pathways, including those involving the microbiome, potentially contribute to the low-density lipoprotein cholesterol lowering efficacy and the pleiotropic effects of simvastatin. Another study investigating the differential response of hypertensive patients to the β-blocker atenolol demonstrated that the variability in metabolic pathways is a function of race and genotype.[44] Although still early in its infancy, pharmacometabolomics could prove important for personalizing CVD therapy.

Current Applications

Although broad-scale metabolomics assays have not been approved by the US Food and Drug Administration for clinical purposes, MS is routinely used under US Food and Drug Administration approval to measure a variety of select compounds or metabolites in clinical laboratories, as shown in Table 2. The main criteria required to deploy an assay in clinical practice are safety and efficacy. A review of the US Food and Drug Administration's perspective on clinical MS provides several examples of pathways leading to approved diagnostic tests.[45] In the research setting, metabolomics data have been used primarily to classify different populations with respect to presence or absence of disease or with respect to risk for disease. For metabolomics to be used in the clinical setting for individual patients, substantial evidence will be required to demonstrate that metabolomics data are indeed effective for providing either a diagnosis or information that will optimize therapy. Some of the potential requirements for transitioning metabolomics from a research to a clinical practice application are shown in the Appendix.

A key issue concerning the possible future translation of metabolomics from research to clinical practice is the fact that most laboratory data used for patient care involve quantification of results expressed as concentration units (absolution quantification); most metabolomics data are currently expressed as a percent increase/decrease relative to a reference population (relative quantification). In the clinical laboratory, quantitative measurements are very tightly controlled. By contrast, metabolomics data are collated, analyzed, and reported from different laboratories using a variety of methods. Accordingly, quality control systems for relative quantification need to be developed before metabolomics can transition to routine clinical practice. One approach to this challenge could be to develop quantification relative to an analyte that is present in all samples, as is currently done when quantifying the percent of glycated hemoglobin for monitoring diabetes mellitus therapy. Nonetheless, controlling data from multiple different sources also poses challenges with regard to standardized analytical practices. Thus, for metabolomics to migrate from a research application to a patient care application, major investments in infrastructure will be required. An example of this kind of structure is the Clinical Proteomic Tumor Analysis Consortium of the National Cancer Institute for developing proteomics platforms.[46]

Study Design and Technical Considerations

The challenges pertaining to translation notwithstanding, several key study design and technical considerations are also required for successful implementation of metabolomics approaches in research (Table 3). The type of study subjects, size of the study cohort, and availability of appropriate control subjects are all important factors for evaluating the extent to which there is adequate statistical power for detecting associations between metabolites and a given outcome of interest (Data Analyses and Data Reporting below). Investigators should also carefully review the quality of the collection and storage of biospecimens for proposed metabolomics profiling because some types of analytes may be at risk for at least partial degradation under certain sample handling conditions. Because variation in metabolites may be related to a range of exogenous and endogenous factors, data collection on study subjects should ideally include information about medication use, dietary patterns, and environmental exposures, as well as concomitant clinical conditions.

With respect to the technical aspects of metabolomic profiling, it should be decided at the outset whether a targeted or untargeted approach will be used, along with the most appropriate quality control methods. If a targeted approach is selected, the extent to which candidate analytes will be measured as part of a hypothesis-driven analysis should be prespecified. For example, if BCAAs and acylcarnitines are to be analyzed to test a given hypothesis, a variety of targeted analytical platforms are available, including commercial kits that can provide validated quantitative values. Given their feasibility and accessibility, targeted MS approaches have been used in the majority of CVD studies published to date. However, studies geared toward generating novel hypotheses or discovering disease-stage specific biomarkers are more likely to benefit from an untargeted approach that offers the simultaneous

Table 2. Current Routine Uses of Mass Spectrometry in the Clinical Laboratory

Type of Test	Test Description
FDA-approved in vitro diagnostic tests	LC-MS/MS for quantification of tacrolimus
	LC-MS/MS for newborn screening
	MALDI-TOF for microbial identifications
Laboratory-developed tests: routinely used	GC-MS and LC-MS/MS for drugs of abuse confirmation
	GC-MS and LC-MS/MS for monitoring pain management compliance
	LC-MS/MS for quantifying steroids
	LC-MS/MS for monitoring immunosuppressants
	LC-MS/MS for quantifying vitamin D and related metabolites
Laboratory-developed tests: specialized assays (partial list)	Amino acid analysis
	Heavy-metal analysis
	Light-chain analysis
	Metanephrine analysis
	Methylmalonic acid quantification
	Novel psychoactive substances
	Parathyroid hormone–related protein
	Quantification of thyroglobulin
	Renin
	Therapeutic drug monitoring
	Thyroid hormone

FDA indicates US Food and Drug Administration; GC, gas chromatography; LC, liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization–time of flight; MS, mass spectrometry; and MS/MS, tandem mass spectrometry.

Table 3. Metabolomics Study Design and Implementation Considerations

Study Component	Considerations
Study design	Prospective vs retrospective
	Cohort vs case-control
	Appropriate control group
Sample collection	Fasting status and length of time of fasting
	Consistency of collection protocol across the study cohort
	Time between sample collection and processing
	Time between sample collection and freezing to <−70°C
	Protocols for sample preparation, depending on type of tissue
	Type of storage tubes and caps used
	Automated vs manual sample handling methods used for dividing into aliquots
	Total storage time and conditions before profiling (eg, number of freeze-thaw cycles)
Data collection	Concomitant medication use
	Dietary history (acute and chronic)
	Environmental factors
	Clinical comorbidities and outcomes
Metabolomic profiling	Sensitivity of platform being used
	Quality control measures (repeat samples, pooled samples, standards)
	Metabolite identification methods (for nontargeted data)
	Biological and analytical reproducibility
Statistical analysis	Quality control assessment
	Missing values, values below lower limits of detection
	Outliers
	Batch effects
	Multiple comparisons
	Multidimensional data reduction approaches
	Supervised vs unsupervised analyses
	Potential clinical and measurable confounders (comorbidities, renal function, medications, fasting status)
	Assessment of independent and incremental association

measurement of all detectable metabolites in a given biospecimen. Because of the widespread availability and robustness of accurate mass spectrometers, such instruments are often used for study designs based on untargeted metabolomics. Notably, an untargeted approach does not exclusively require accurate mass analyses that provide part-per-million mass accuracy. In fact, biomarker discovery and biomarker identification can logically be separated into 2 different objectives within a given research study. For example, a multitargeted method for specific compounds could be used in conjunction with untargeted (full scan) data acquisitions with a nominal mass instrument. After identification of metabolites of potential clinical importance, an in-depth secondary analysis can be conducted with an accurate mass instrument. Furthermore, although untargeted analysis is generally considered unbiased, as is the case with virtually all scientific approaches, it too is inherently biased with respect to the types of metabolites (ie, hydrophobic versus hydrophilic) extracted from complex specimens given the use of specific organic solvents and the chemical class of molecules assayed with specific chromatography techniques. For example, if an LC reversed-phase C18 column is used, hydrophilic compounds will not be well retained. Such an untargeted method would be unable to distinguish the

presence and abundance of hexose isomers (eg, common molecules such as fructose and galactose) and hexoses that are rarely reported (eg, tagatose, allose, or idose). Alternatively, such metabolites are easily separated by gas chromatography when detected in human biofluids. Therefore, comprehensive and unbiased interrogation of the human metabolome will likely require multiple complementary methods.

Bioinformatics and Statistical Considerations

As seen with genomics, the promise of any profiling technology is fully realized only in the setting of a computational infrastructure that is capable of managing the vast amount of high-dimensional data generated. Robust approaches for storage, integration, statistical analyses, and visualization of metabolomics data remain in the early stages of development. Thus, the need for data tools is of particular importance in dealing with untargeted metabolomics, in which both the volume and complexity of data can scale by multiple orders of magnitude, especially when applied to large population-level studies.

Data Processing

Central to metabolomics is the actual collection and handling of either MS or NMR data. Particularly with mass spectra, data processing can involve a number of critically important steps, including filtering background noise, detecting and defining thousands of spectral peaks from a single sample, and aligning spectra across multiple samples to allow comparison of metabolite features across sample groups. Although dozens of open-source and commercial tools have been used to process and report the results of MS- or NMR-based metabolomics data,[47] there remains little uniformity or standardization from platform to platform, often complicating efforts to perform independent replication of metabolomics findings.

Data processing and robust analyses are much more difficult for untargeted metabolomics than for targeted metabolomics. Existing widely used software packages have recognized limitations, including the most recent version of MS-DIAL, which can be used for data-independent mass fragmentations (SWATH-type data acquisitions), and classic software such as MZmine or XCMS. Although MS-DIAL excels in combining the different adducts and potential in-source fragments of all compounds on the level of intact molecules, in contrast to simple mass-to-charge (m/z) and retention time (rt) feature-finding software such as XCMS, MS-DIAL has not yet been benchmarked against alternative software (eg, MZmine) with respect to achieving lower rates of false-negative and false-positive peak finding results. Accordingly, caution is advised in the interpretation of reports of thousands of distinct metabolites present in a single LC-MS run, which may result from poor software performance and a high number of chemical noise peaks that are not derived from a truly biological background. As the field of metabolomics continues to advance, so too will data processing approaches as part of ongoing collective efforts to ensure the robustness and reproducibility of findings across laboratories.

Data Analyses

Similar to issues in data processing, there currently exists little uniformity in the biostatistical management and analysis of metabolomics data. Statistical challenges for handling and analyzing metabolomics data are similar to those for other large-scale profiling data but also include several unique features. Metabolomics measures are influenced by time-dependent factors, including short-term and long-term perturbations or exposures that will influence detectable levels of metabolites to varying degrees. Time-dependent factors may or may not contribute to variable missingness of values. Statistical approaches to handling missing data, especially for measurements obtained from a broad untargeted metabolomics method, need to consider technical and biological contributions to missingness. Measurement values for individual analytes can be missing because they are below the lower limit of detection of a given platform, because of technical issues not specifically related to detection limits, or because they are truly biologically absent in a given tissue sample. Accordingly, the rate of missingness for an analyte profiled across a human cohort is often, but not always, correlated with the relative abundance of that analyte in samples obtained from that cohort. How to best handle metabolomics data depends on the nature and distribution of measurement values, including rates and types of missingness (eg, missing not at random). Thus, imputation approaches to handling missingness may be appropriate for some situations and not for others.

Beyond issues related to how metabolite variables are handled, a number of different approaches have been used to identify potentially biological meaningful relationships between metabolites and clinical outcomes of interest in univariate and multivariable analyses. A traditional approach that is commonly applied involves analyzing each metabolite separately and adjusting for multiple comparisons across the number of metabolites studied with Bonferroni or false discovery rate thresholds. Given collinearity between metabolites, reflecting shared biological pathways and shared environmental perturbations, data reduction approaches are often additionally applied to reduce the statistical burden of multiple comparisons. An advantage of data reduction approaches is that the identification of interrelated groups of metabolites can highlight underlying biology in ways that prioritization of single metabolite analysis cannot. Such approaches include but are not limited to principal components analysis and hierarchical cluster analyses. Complementary methods that are optimized for feature selection or prediction and are particularly suitable for high-dimensional data include discriminant analysis, partial least-squares analysis, least absolute shrinkage and selection operator and related analyses, rule-based approaches (eg, random forests), and instance-based approaches (eg, support vector machines). The relative advantages and disadvantages of different statistical approaches can be understood on the basis of several features: (1) the extent to which a dimension reduction step is either required or not required before association analyses, (2) the extent to which metabolite associations are assumed to be linear or nonlinear, (3) the extent to which prediction of an outcome versus variable selection is favored, and (4) the desire for sparseness in the predictive model versus identification of the larger set of correlated analytes for biological interpretation. Not all statistical approaches are efficient to run with standard computing resources, particularly those that involve recursive modeling, and so practical

considerations are also involved in the selection of the most appropriate approach for a given dataset. Notably, commonly used methods have yet to be formally evaluated in comparison to each other for analyzing high-dimensional metabolomics data acquired from either large or small cohorts. It is possible that a combination of statistical methods (eg, ensemble approach) may be desirable for some data sets. For approaches that require a dimension reduction step, the extent to which this initial stage of analysis should be unsupervised or supervised to account for intercorrelations between metabolite variables is not entirely clear and depends on the purpose of the study. The next generation of discoveries in metabolomics will rely heavily on progress made to establish robust statistical methods for analyzing these high-dimensional data and improvements in the annotation of metabolite-related biochemical pathways to enable in silico pathway analyses.

Data Reporting

With respect to reporting metabolomics data and results, standardization is also greatly needed. Metabolomics data reporting procedures should ideally include a universally accepted set of minimal criteria to be used in all published reports, including detailed methods and quality control metrics. Additionally, efforts should make use of standard pooled quality control samples. These control samples could include the National Institute of Standards & Technology Standard Reference Material 1950 standard plasma, which would also enable researchers to compare results, including use of untargeted metabolomics methods for which abundance of small molecules could be reported in relation to abundances found in that standard plasma. In addition, publication requirements should ensure the deposit of all metabolomics data, including raw and processed data, in a central publicly accessible repository, as is currently required for other "omics" data. In the absence of raw data, verification of metabolite identifications cannot be performed by independent reviewers, and scientific claims may not be reproducible. Additionally, a centralized storage of metabolomics data provides opportunities for secondary analysis and data mining, ensuring the continued value of generated data sets. Such repositories will be essential for uncovering new metabolite correlates and detailing the dynamic range of metabolites across populations and samples. It should be noted that achieving this priority may be complicated by the dependence of some groups on third-party clinical research organizations from which raw spectral data are generally not available, making independent verification impossible. In anticipation of the need for a central data repository, the National Institutes of Health has implemented a metabolomics repository for Common Funds projects that is open also for general purpose metabolomics[48] and is highly endorsed for National Institutes of Health–funded studies. A similar resource has also been established in Europe.[48a]

Moving forward, efforts must be made across the field to include distinct database identifiers in metabolomics reports, at a minimum in supplemental tables. It is a common misconception that metabolite names are unique. For some compounds, up to hundreds of names exist; indeed, for most compounds, at least a dozen different names have been reported, often with overlapping use. A continued practice of using names as chemical identifiers leads to ambiguity and also severely hampers electronic retrieval of findings such as from PubMed searches. Instead, the National Institute of Standards & Technology and International Union of Pure and Applied Chemistry have defined a standardized, open-source way to encode chemical structures into a simple string of letters, the International Chemical Identifier code. This code is accompanied by a very short International Chemical Identifier key that fits easily into tables and documents to facilitate automated searches.

Current Challenges

To ensure that metabolomics can realize its potential to substantially affect research and practice, experts in the field are currently addressing key challenges in implementation and application, including high-throughput approaches to extend profiling capacity; strategies for determining the biochemical identity and, in turn, the functional role of novel analytes; and methods for determining the true clinical importance of analytes observed in association with cardiovascular phenotypes.

High-Throughput Metabolomics

As tools for applying metabolomics to larger-scale cohorts are developed, so too will the utility of metabolomics and metabolic markers in clinical practice. This is especially true of untargeted metabolomics applied to population-based studies, an essential step for defining the human metabolome and providing new knowledge of the biochemical landscape of health and disease. Because untargeted metabolomics allows monitoring thousands of metabolite features and assuming that metabolite-disease effect sizes will be comparable to those observed in targeted metabolomics studies, application across much larger cohorts than those previously studied will be needed to limit false-negative errors. In turn, the ability to profile multiple separate cohorts will be needed to validate findings and to limit false-positive errors. Given these imperatives, emerging high-throughput technologies must leverage higher-selectivity mass analysis and maintain acceptable levels of reproducibility and sensitivity while scaling capacity.

Until recently, high-scale sample throughput was achievable only with NMR-based approaches, which are technically robust, are amendable to automation, and do not require extensive upfront sample preparation.[49] However, NMR approaches have limited sensitivity for detecting low-abundancy metabolites. Thus, extending from NMR to more comprehensive measures of the metabolome with MS-based techniques represents the next challenge in throughput technology. Typically, MS-based approaches require coupling to chromatographic methods to reduce sample complexity before introduction into a mass spectrometer. Although very effective, chromatographic methods are often time consuming, especially when serial chromatography is applied to each sample to capture representative molecules across very different chemical classes in a discovery mode (ie, subplatforms). Given the costs associated with obtaining and operating an MS infrastructure and the need for trained personnel to manage such instrumentation, extended analytic times can translate into substantial cost per sample. Time- and cost-related

limitations will, in part, be overcome with the recently established National Institutes of Health Regional Comprehensive Metabolomics Resource Cores and the development of similar academic and commercial facilities capable of multiplexing metabolomics studies. As seen over the past decade in the field of genomics, the currently high cost per sample of metabolomics profiling is expected to decline over the next several years as throughput technologies continue to develop.

Technical advances, including the integration of ultrathroughput high-resolution MS with automated sample handling and computational identification of mass spectral features, are facilitating the transition to an era when thousands of samples can be analyzed in a short time, with the measure of thousands of metabolites per specimen. Such throughput applied to large-scale studies will allow characterization of chemical diversity across multiple racial/ethnic populations in relation to prospective disease phenotypes and in conjunction with integrated data on genomic variation, epigenetic markers, mRNA transcription, and other biomarkers to enable a true systems-level analysis of human physiology and disease.

Biochemical Identity and Role of Novel Metabolites

Early chemoinformatics approaches have been developed to aid in metabolite identification and data visualization, although existing tools remain relatively rudimentary. Thus, improvement in experimental and computational tools designed to clarify the identity and role of novel metabolites represents an important future challenge for the field. With untargeted analysis of any complex human biospecimen, only a minority fraction of the metabolites is definitively identifiable by current standards, with the majority of molecules remaining unknown with respect to structure and origin (ie, endogenous versus exogenous). Future technical advances in MS may aid in molecular identification, including differential mobility spectrometers that could be placed between an LC and an MS instrument that could further extend metabolomic coverage and theoretically assist with compound identification. At present, a number of spectral databases currently exist[47] and serve as invaluable resources. Expanding these databases, along with computational methods for predicting metabolite identification, will be essential for progress in the field. To guide database and computational methods development, the Metabolomics Standards Initiative has put forward criteria for reporting of metabolites that importantly include the level of rigor used for identification of metabolites.[49a] Because the practice of metabolomics extends from the analytical chemist to the mainstream biologist, additional data analysis and visualization tools will be needed to facilitate scientific discovery as well as communication, including those tools allowing pathway analysis or metabolite set enrichment analysis, as well as correlation-based network analysis. Sophisticated methods for mapping metabolomics data to pathways will be critical for interrogating not only interanalyte relationships but also metabolite relationships with genomic, epigenetic, transcriptomic, and proteomic factors, as well as clinical phenotypes.

Clinical Relevance

Metabolomics studies conducted in population-based cohorts will allow studies of how various factors, including diet and

medical comorbidities, affect the metabolome. As has been demonstrated with metabolomic studies of oral glucose challenge, acute caloric intake, and the attendant stimulation of insulin signaling, select metabolic interventions have a profound impact on circulating metabolites. Much less is known about how differences in dietary composition modulate the metabolome chronically and acutely. Indeed, the majority of metabolomics biomarker studies to date have examined plasma obtained from fasting individuals and thus are agnostic to effects of metabolite variation in the fed or challenged state. Disentangling the relative contribution of various medical comorbidities to circulating metabolite levels is also important in assessment of the potential value of select metabolites as markers of CVD risk. In addition to having a causal role in CVD pathogenesis, diabetes mellitus and obesity have a substantial impact on the metabolome that can confound metabolomics studies of CVD prediction. Perhaps less appreciated is the substantial impact that kidney disease can have on both CVD risk and circulating metabolite levels. Although recognized for their ability to excrete select nitrogenous metabolites such as creatinine and urea, the kidneys also exert broad and heterogeneous effects, depending on the extent to which metabolites of interest are filtered, absorbed, secreted, catabolized, or even synthesized by the kidney.[50] Ultimately, a full accounting of diet, medical comorbidities, and other potential confounders in metabolomics studies of CVD will require a variety of approaches. First, metabolomics data acquired from large, well-phenotyped cohorts are required to establish the matrix of correlations between metabolite levels and clinical variables known to be associated with cardiometabolic disease (eg, age, race, sex, body size, insulin resistance, blood pressure, and renal function). Second, more in-depth studies of individuals over time are required to clarify how fasting versus fed status, dietary patterns, circadian cycles, other time-dependent factors, and random variation can affect circulating metabolites. Third, although less important from a strict biomarker sense, studies that use invasive catheterization of select vascular beds are important to establish the potential organ specificity of select peripheral venous metabolite signatures. Finally, understanding the genetic determinants of plasma metabolite levels and cross-referencing with prior genetics studies of CVD offer the potential to implicate select metabolite biomarkers in causal pathways.

Future Directions

With continued advancement in metabolomic technologies, several possible applications in cardiovascular medicine may be on the horizon. With respect to understanding the pathophysiology of disease, integration of genomic and metabolomic data could allow clinicians and researchers to better explain why an individual develops CVD in response to a given exposure history such as tobacco or chronic stress. For example, sources of interindividual variation in response to risk exposures could include the possibility that a specific metabolite is not produced as a result of exposure to the stimulus or is produced but degraded at a different rate. With respect to risk prediction, the ability to one day perform actual or nearly real-time monitoring of blood or urine metabolites could allow clinicians to detect small-molecule biomarkers associated with worsening or improving clinical trajectories

and to target early interventions. Similarly, changes in metabolite profiles over time, with aging, or after administration of a drug could be used to define an individual's predisposition for disease and response to therapy. Ultimately, the ability to efficiently and effectively use metabolomics tools to conduct molecular phenotyping could serve to substantially advance the goals of precision medicine.

The promise that metabolomics offers for improving medical diagnostics and therapeutics notwithstanding, advances in the field will depend on overcoming certain intrinsic challenges. The tractable number of metabolites potentially measurable in the metabolome is a potential analytic advantage over other detailed profiling assays such as genomics and proteomics.[51] On the other hand, the complexity of metabolite profiling exceeds that of other profiling techniques such as genomics or transcriptomics, given that the basic chemical constituents of DNA or mRNA are far less diverse. The inability of basic investigators to transparently convey the challenges intrinsic to the field, combined with the clinical community's desire for rapid integration of new technologies, has led to an initial mismatch of expectations and "metabolomic deliverables." However, this divide between expectation and deliverables is beginning to narrow. Protocols for evaluating major classes of compounds have become increasingly more standardized. Authentic chemical standards for providing absolute quantitation of analytes have been integrated into these protocols, throughput technologies have become more powerful, and databases for compound identification have become more robust. In turn, standardized procedures are increasingly being applied to large clinical cohorts, including individuals representing the spectrum of cardiovascular and metabolic disease risk. As a result, the scientific community is fortunately beginning to see consistency in the findings reported from across different metabolomics studies. Beyond feasibility and reliability, however, the ultimate challenge pertains to exactly how much new information metabolomics will add to established diagnostic approaches and tests, the critical bar for any nascent technology.

The next generation of scientific discoveries will emerge from an improved understanding of the mechanisms underlying disease risk and phenotypes at the individual level.[52] Given the potential of metabolomics to integrate detailed biochemical information with data on both endogenous and exogenous exposures, metabolomics also offers the opportunity to bridge the methodological and informational gap between the fields of genomics and environmental science in particular.[53,54] Such an integrated approach to understanding and targeting the links between genetic predisposition and external risk exposures is especially needed for complex and multifactorial disease entities such as CVD.

Summary

Having evolved considerably from traditional analytical chemistry, metabolomics now provides extremely detailed molecular profiling of biospecimens and offers an integrated approach to investigating both the intrinsic and extrinsic factors that contribute to CVD risk. Early work in the field has led to the discovery of novel molecular markers of CVD risk, including metabolites related to dietary patterns and gut

microbiome activity. Future work will leverage ongoing technical advances, which are widening the scope and throughput with which small-molecule profiling can be conducted. Therefore, the prospects for further discovery are rapidly growing. Nonetheless, the extent to which metabolomics work will yield a next generation of important discoveries in cardiovascular science will depend on the success with which the field can address key challenges, including bioinformatics approaches for handling high-throughput untargeted metabolomics data, strategies for identifying the biochemical structure and functional role of novel metabolites, and methods for determining the true relevance of metabolites observed in association with clinical outcomes. Progress made in addressing these challenges will allow the potential for metabolomics approaches to substantially affect diagnostics and therapeutics in cardiovascular medicine.

Appendix

Appendix. Requirements Necessary to Transition Metabolomics From Research Applications to Clinical Diagnostics

Component	Considerations
Guidance documents	Structure for good laboratory practices needs to be established
	Input needed from metabolomics experts
	Input needed from clinical laboratories
	Input needed from regulatory agencies
Regulatory structure	Laboratory standards need to be established
	Laboratory inspection protocols need to be developed
	Proficiency testing materials need to be developed
	Laboratory inspectors need to be trained
	Reimbursement needs to be defined
Clinical concerns	Patients who will benefit from metabolic profiling need to be identified
	The clinical utility of testing needs to be defined
	The safety and efficacy of the testing for individual patients need to be demonstrated
Analytical considerations	Patient preparation needs to be defined
	Protocols need to be standardized
	Performance of tests needs to be compared with established techniques
	Standard reference materials need to be developed for harmonization (eg, National Institute of Standards & Technology)
	Proficiency materials need to be developed
	Quality control metrics need to be developed
	Expected ranges for different patient populations need to be defined
Education	Physicians need to be trained on clinical utility
	Laboratories need to gain expertise in analysis
	Insurers (payers) need to understand clinical value of testing
Quality control	Consensus documents on degree of control are needed (eg, with untargeted metabolomics, how to monitor recovery and relative concentration for unknown metabolites)

Disclosures

Writing Group Disclosures

Writing Group Member	Employment	Research Grant	Other Research Support	Speakers' Bureau/ Honoraria	Expert Witness	Ownership Interest	Consultant/ Advisory Board	Other
Susan Cheng	Brigham and Women's Hospital Department of Medicine	AHA†; NIH/NHLBI†	None	None	None	None	None	None
Svati H. Shah	Duke University Medicine	NIH†	None	None	None	Stock ownership/patent*	None	None
Elizabeth J. Corwin	Emory University School of Nursing	None	None	None	None	None	None	None
Oliver Fiehn	University of California, Davis Genome Center	NIH†; Agilent*; Sciex*; ThermoScientific*	None	None	None	None	None	None
Robert L. Fitzgerald	UCSD School of Medicine Pathology	None	None	None	None	None	None	None
Robert E. Gerszten	Beth Israel Deaconess Medical Center	None	None	None	None	None	None	None
Thomas Illig	Hannover Medical School	None	None	None	None	None	None	None
Mohit Jain	University of California at San Diego	NIH†; Doris Duke Charitable Foundation*	None	None	None	None	None	None
Eugene P. Rhee	Massachusetts General Hospital Department of Medicine	None	None	None	None	None	None	None
Pothur R. Srinivas	NIH Cardiovascular Diseases	None	None	None	None	None	None	None
Thomas J. Wang	Vanderbilt University School of Medicine	NIH†	None	None	None	None	None	None

This table represents the relationships of writing group members that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all members of the writing group are required to complete and submit. A relationship is considered to be "significant" if (a) the person receives $10 000 or more during any 12-month period, or 5% or more of the person's gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10 000 or more of the fair market value of the entity. A relationship is considered to be "modest" if it is less than "significant" under the preceding definition.

*Modest.

†Significant.

Reviewer Disclosures

Reviewer	Employment	Research Grant	Other Research Support	Speakers' Bureau/ Honoraria	Expert Witness	Ownership Interest	Consultant/ Advisory Board	Other
Stephen B. G. Abbott	University of Virginia	None	None	None	None	None	None	None
Michaela A. Aldred	Cleveland Clinic	None	None	None	None	None	None	None
Zdenka Pausova	University of Toronto (Canada)	CIHR/HSF* (ongoing operating research grants from the Canadian funding agencies to support lipidomics research in a population-based cohort of adolescents and their parents)	None	None	None	None	None	None

This table represents the relationships of reviewers that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all reviewers are required to complete and submit. A relationship is considered to be "significant" if (a) the person receives $10 000 or more during any 12-month period, or 5% or more of the person's gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10 000 or more of the fair market value of the entity. A relationship is considered to be "modest" if it is less than "significant" under the preceding definition.

*Modest.

References

1. Rhee EP, Gerszten RE. Metabolomics and cardiovascular biomarker discovery. *Clin Chem*. 2012;58:139–147. doi: 10.1373/clinchem.2011.169573.

2. Oldiges M, Lütz S, Pflug S, Schroer K, Stein N, Wiendahl C. Metabolomics: current state and evolving methodologies and tools. *Appl Microbiol Biotechnol*. 2007;76:495–511. doi: 10.1007/s00253-007-1029-2.

3. Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, Clarke S, Schofield PM, McKilligin E, Mosedale DE, Grainger DJ. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics [published correction appears in *Nat Med*. 2003;9:477]. *Nat Med*. 2002;8:1439–1444. doi: 10.1038/nm802.

4. Kirschenlohr HL, Griffin JL, Clarke SC, Rhydwen R, Grace AA, Schofield PM, Brindle KM, Metcalfe JC. Proton NMR analysis of plasma is a weak predictor of coronary artery disease [published correction appears in *Nat Med*. 2006;12:862]. *Nat Med*. 2006;12:705–710. doi: 10.1038/nm1432.

5. Lewis GD, Wei R, Liu E, Yang E, Shi X, Martinovic M, Farrell L, Asnani A, Cyrille M, Ramanathan A, Shaham O, Berriz G, Lowry PA, Palacios IF, Taşan M, Roth FP, Min J, Baumgartner C, Keshishian H, Addona T, Mootha VK, Rosenzweig A, Carr SA, Fifer MA, Sabatine MS, Gerszten RE. Metabolite profiling of blood from individuals undergoing planned myocardial infarction reveals early markers of myocardial injury. *J Clin Invest*. 2008;118:3503–3512. doi: 10.1172/JCI35111.

6. Turer AT, Stevens RD, Bain JR, Muehlbauer MJ, van der Westhuizen J, Mathew JP, Schwinn DA, Glower DD, Newgard CB, Podgoreanu MV. Metabolomic profiling reveals distinct patterns of myocardial substrate use in humans with coronary artery disease or left ventricular dysfunction during surgical ischemia/reperfusion. *Circulation*. 2009;119:1736–1746. doi: 10.1161/CIRCULATIONAHA.108.816116.

7. Shah SH, Bain JR, Muehlbauer MJ, Stevens RD, Crosslin DR, Haynes C, Dungan J, Newby LK, Hauser ER, Ginsburg GS, Newgard CB, Kraus WE. Association of a peripheral blood metabolic profile with coronary artery disease and risk of subsequent cardiovascular events. *Circ Cardiovasc Genet*. 2010;3:207–214. doi: 10.1161/CIRCGENETICS.109.852814.

8. 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. doi: 10.1038/nature09922.

9. Shah SH, Sun JL, Stevens RD, Bain JR, Muehlbauer MJ, Pieper KS, Haynes C, Hauser ER, Kraus WE, Granger CB, Newgard CB, Califf RM, Newby LK. Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. *Am Heart J*. 2012;163:844–850.e1. doi: 10.1016/j.ahj.2012.02.005.

10. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, Britt EB, Fu X, Wu Y, Li L, Smith JD, DiDonato JA, Chen J, Li H, Wu GD, Lewis JD, Warrier M, Brown JM, Krauss RM, Tang WH, Bushman FD, Lusis AJ, Hazen SL. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med*. 2013;19:576–585. doi: 10.1038/nm.3145.

11. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y, Hazen SL. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med*. 2013;368:1575–1584. doi: 10.1056/NEJMoa1109400.

12. Zheng Y, Yu B, Alexander D, Manolio TA, Aguilar D, Coresh J, Heiss G, Boerwinkle E, Nettleton JA. Associations between metabolomic compounds and incident heart failure among African Americans: the ARIC Study. *Am J Epidemiol*. 2013;178:534–542. doi: 10.1093/aje/kwt004.

13. Bhattacharya S, Granger CB, Craig D, Haynes C, Bain J, Stevens RD, Hauser ER, Newgard CB, Kraus WE, Newby LK, Shah SH. Validation of the association between a branched chain amino acid metabolite profile and extremes of coronary artery disease in patients referred for cardiac catheterization. *Atherosclerosis*. 2014;232:191–196. doi: 10.1016/j.atherosclerosis.2013.10.036.

14. Ganna A, Salihovic S, Sundström J, Broeckling CD, Hedman AK, Magnusson PK, Pedersen NL, Larsson A, Siegbahn A, Zilmer M, Prenni J, Arnlöv J, Lind L, Fall T, Ingelsson E. Large-scale metabolomic profiling identifies novel biomarkers for incident coronary heart disease. *PLoS Genet*. 2014;10:e1004801. doi: 10.1371/journal.pgen.1004801.

15. Jové M, Mauri-Capdevila G, Suárez I, Cambray S, Sanahuja J, Quílez A, Farré J, Benabdelhak I, Pamplona R, Portero-Otín M, Purroy F. Metabolomics predicts stroke recurrence after transient ischemic attack. *Neurology*. 2015;84:36–45. doi: 10.1212/WNL.0000000000001093.

16. Würtz P, Havulinna AS, Soininen P, Tynkkynen T, Prieto-Merino D, Tillin T, Ghorbani A, Artati A, Wang Q, Tiainen M, Kangas AJ, Kettunen J, Kaikkonen J, Mikkilä V, Jula A, Kähönen M, Lehtimäki T, Lawlor DA, Gaunt TR, Hughes AD, Sattar N, Illig T, Adamski J, Wang TJ, Perola M, Ripatti S, Vasan RS, Raitakari OT, Gerszten RE, Casas JP, Chaturvedi N, Ala-Korpela M, Salomaa V. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation*. 2015;131:774–785. doi: 10.1161/CIRCULATIONAHA.114.013116.

17. Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, van der Ouderaa F, Bingham S, Cross AJ, Nicholson JK. Susceptibility of human metabolic phenotypes to dietary modulation. *J Proteome Res*. 2006;5:2780–2788. doi: 10.1021/pr060265y.

18. Koeth RA, Levison BS, Culley MK, Buffa JA, Wang Z, Gregory JC, Org E, Wu Y, Li L, Smith JD, Tang WH, DiDonato JA, Lusis AJ, Hazen SL. γ-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab*. 2014;20:799–812. doi: 10.1016/j.cmet.2014.10.006.

19. Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA, Daviglus ML, Kesteloot H, Ueshima H, Zhao L, Nicholson JK, Elliott P. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature*. 2008;453:396–400. doi: 10.1038/nature06882.

20. Cheng S, Rhee EP, Larson MG, Lewis GD, McCabe EL, Shen D, Palma MJ, Roberts LD, Dejam A, Souza AL, Deik AA, Magnusson M, Fox CS, O'Donnell CJ, Vasan RS, Melander O, Clish CB, Gerszten RE, Wang TJ. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation*. 2012;125:2222–2231. doi: 10.1161/CIRCULATIONAHA.111.067827.

21. Zheng Y, Yu B, Alexander D, Mosley TH, Heiss G, Nettleton JA, Boerwinkle E. Metabolomics and incident hypertension among blacks: the Atherosclerosis Risk in Communities Study. *Hypertension*. 2013;62:398–403. doi: 10.1161/HYPERTENSIONAHA.113.01166.

22. Menni C, Graham D, Kastenmüller G, Alharbi NH, Alsanosi SM, McBride M, Mangino M, Titcombe P, Shin SY, Psatha M, Geisendorfer T, Huber A, Peters A, Wang-Sattler R, Xu T, Brosnan MJ, Trimmer J, Reichel C, Mohney RP, Soranzo N, Edwards MH, Cooper C, Church AC, Suhre K, Gieger C, Dominiczak AF, Spector TD, Padmanabhan S, Valdes AM. Metabolomic identification of a novel pathway of blood pressure regulation involving hexadecanedioate. *Hypertension*. 2015;66:422–429. doi: 10.1161/HYPERTENSIONAHA.115.05544.

23. Stegemann C, Pechlaner R, Willeit P, Langley SR, Mangino M, Mayr U, Menni C, Moayyeri A, Santer P, Rungger G, Spector TD, Willeit J, Kiechl S, Mayr M. Lipidomics profiling and risk of cardiovascular disease in the prospective population-based Bruneck study. *Circulation*. 2014;129:1821–1831. doi: 10.1161/CIRCULATIONAHA.113.002500.

24. Kien CL, Bunn JY, Poynter ME, Stevens R, Bain J, Ikayeva O, Fukagawa NK, Champagne CM, Crain KI, Koves TR, Muoio DM. A lipidomics analysis of the relationship between dietary fatty acid composition and insulin sensitivity in young adults. *Diabetes*. 2013;62:1054–1063. doi: 10.2337/db12-0363.

25. Kulkarni H, Meikle PJ, Mamtani M, Weir JM, Barlow CK, Jowett JB, Bellis C, Dyer TD, Johnson MP, Rainwater DL, Almasy L, Mahaney MC, Comuzzie AG, Blangero J, Curran JE. Plasma lipidomic profile signature of hypertension in Mexican American families: specific role of diacylglycerols. *Hypertension*. 2013;62:621–626. doi: 10.1161/HYPERTENSIONAHA.113.01396.

26. Weir JM, Wong G, Barlow CK, Greeve MA, Kowalczyk A, Almasy L, Comuzzie AG, Mahaney MC, Jowett JB, Shaw J, Curran JE, Blangero J, Meikle PJ. Plasma lipid profiling in a large population-based cohort. *J Lipid Res*. 2013;54:2898–2908. doi: 10.1194/jlr.P035808.

27. Hinterwirth H, Stegemann C, Mayr M. Lipidomics: quest for molecular lipid biomarkers in cardiovascular disease. *Circ Cardiovasc Genet*. 2014;7:941–954. doi: 10.1161/CIRCGENETICS.114.000550.

28. Puri R, Duong M, Uno K, Kataoka Y, Nicholls SJ. The emerging role of plasma lipidomics in cardiovascular drug discovery. *Expert Opin Drug Discov*. 2012;7:63–72. doi: 10.1517/17460441.2012.644041.

29. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, Haqq AM, Shah SH, Arlotto M, Slentz CA, Rochon J, Gallup D, Ilkayeva O, Wenner BR, Yancy WS Jr, Eisenson H, Musante G, Surwit RS, Millington DS, Butler MD, Svetkey LP. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance [published correction appears in *Cell Metab*. 2009;9:565–566]. *Cell Metab*. 2009;9:311–326. doi: 10.1016/j.cmet.2009.02.002.

30. Shaham O, Wei R, Wang TJ, Ricciardi C, Lewis GD, Vasan RS, Carr SA, Thadhani R, Gerszten RE, Mootha VK. Metabolic profiling of the human

response to a glucose challenge reveals distinct axes of insulin sensitivity. *Mol Syst Biol*. 2008;4:214. doi: 10.1038/msb.2008.50.

31. Felig P, Marliss E, Cahill GF Jr. Plasma amino acid levels and insulin secretion in obesity. *N Engl J Med*. 1969;281:811–816. doi: 10.1056/NEJM196910092811503.

32. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, Lewis GD, Fox CS, Jacques PF, Fernandez C, O'Donnell CJ, Carr SA, Mootha VK, Florez JC, Souza A, Melander O, Clish CB, Gerszten RE. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17:448–453. doi: 10.1038/nm.2307.

33. Laferrère B, Reilly D, Arias S, Swerdlow N, Gorroochurn P, Bawa B, Bose M, Teixeira J, Stevens RD, Wenner BR, Bain JR, Muehlbauer MJ, Haqq A, Lien L, Shah SH, Svetkey LP, Newgard CB. Differential metabolic impact of gastric bypass surgery versus dietary intervention in obese diabetic subjects despite identical weight loss. *Sci Transl Med*. 2011;3:80re2. doi: 10.1126/scitranslmed.3002043.

34. Batch BC, Shah SH, Newgard CB, Turer CB, Haynes C, Bain JR, Muehlbauer M, Patel MJ, Stevens RD, Appel LJ, Newby LK, Svetkey LP. Branched chain amino acids are novel biomarkers for discrimination of metabolic wellness. *Metabolism*. 2013;62:961–969. doi: 10.1016/j.metabol.2013.01.007.

35. Shah SH, Crosslin DR, Haynes CS, Nelson S, Turer CB, Stevens RD, Muehlbauer MJ, Wenner BR, Bain JR, Laferrère B, Gorroochurn P, Teixeira J, Brantley PJ, Stevens VJ, Hollis JF, Appel LJ, Lien LF, Batch B, Newgard CB, Svetkey LP. Branched-chain amino acid levels are associated with improvement in insulin resistance with weight loss. *Diabetologia*. 2012;55:321–330. doi: 10.1007/s00125-011-2356-5.

36. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Van Treuren W, Walters WA, Knight R, Newgard CB, Heath AC, Gordon JI. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*. 2013;341:1241214. doi: 10.1126/science.1241214.

37. Lewis GD, Farrell L, Wood MJ, Martinovic M, Arany Z, Rowe GC, Souza A, Cheng S, McCabe EL, Yang E, Shi X, Deo R, Roth FP, Asnani A, Rhee EP, Systrom DM, Semigran MJ, Vasan RS, Carr SA, Wang TJ, Sabatine MS, Clish CB, Gerszten RE. Metabolic signatures of exercise in human plasma. *Sci Transl Med*. 2010;2:33ra37. doi: 10.1126/scitranslmed.3001006.

38. Huffman KM, Slentz CA, Bateman LA, Thompson D, Muehlbauer MJ, Bain JR, Stevens RD, Wenner BR, Kraus VB, Newgard CB, Kraus WE. Exercise-induced changes in metabolic intermediates, hormones, and inflammatory markers associated with improvements in insulin sensitivity. *Diabetes Care*. 2011;34:174–176. doi: 10.2337/dc10-0709.

39. Kujala UM, Mäkinen VP, Heinonen I, Soininen P, Kangas AJ, Leskinen TH, Rahkila P, Würtz P, Kovanen V, Cheng S, Sipilä S, Hirvensalo M, Telama R, Tammelin T, Savolainen MJ, Pouta A, O'Reilly PF, Mäntyselkä P, Viikari J, Kähönen M, Lehtimäki T, Elliott P, Vanhala MJ, Raitakari OT, Järvelin MR, Kaprio J, Kainulainen H, Ala-Korpela M. Long-term leisure-time physical activity and serum metabolome. *Circulation*. 2013;127:340–348. doi: 10.1161/CIRCULATIONAHA.112.105551.

40. Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B. Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr*. 2005;82:497–503.

41. Wang Z, Roberts AB, Buffa JA, Levison BS, Zhu W, Org E, Gu X, Huang Y, Zamanian-Daryoush M, Culley MK, DiDonato AJ, Fu X, Hazen JE, Krajcik D, DiDonato JA, Lusis AJ, Hazen SL. Non-lethal inhibition of gut microbial trimethylamine production for the treatment of atherosclerosis. *Cell*. 2015;163:1585–1595. doi: 10.1016/j.cell.2015.11.055.

42. Clayton TA, Baker D, Lindon JC, Everett JR, Nicholson JK. Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A*. 2009;106:14728–14733. doi: 10.1073/pnas.0904489106.

43. Kitzmiller JP, Luzum JA, Baldassarre D, Krauss RM, Medina MW. CYP3A4*22 and CYP3A5*3 are associated with increased levels of plasma simvastatin concentrations in the Cholesterol and Pharmacogenetics Study cohort. *Pharmacogenet Genomics*. 2014;24:486–491. doi: 10.1097/FPC.0000000000000079.

44. Wikoff WR, Frye RF, Zhu H, Gong Y, Boyle S, Churchill E, Cooper-Dehoff RM, Beitelshees AL, Chapman AB, Fiehn O, Johnson JA, Kaddurah-Daouk R; Pharmacometabolomics Research Network. Pharmacometabolomics reveals racial differences in response to atenolol treatment. *PLoS One*. 2013;8:e57639. doi: 10.1371/journal.pone.0057639.

45. Lathrop JT, Jeffery DA, Shea YR, Scholl PF, Chan MM. US Food and Drug Administration perspectives on clinical mass spectrometry. *Clin Chem*. 2016;62:41–47. doi: 10.1373/clinchem.2015.244731.

46. Ellis MJ, Gillette M, Carr SA, Paulovich AG, Smith RD, Rodland KK, Townsend RR, Kinsinger C, Mesri M, Rodriguez H, Liebler DC; Clinical Proteomic Tumor Analysis Consortium (CPTAC). Connecting genomic alterations to cancer biology with proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. *Cancer Discov*. 2013;3:1108–1112. doi: 10.1158/2159-8290.CD-13-0219.

47. Alonso A, Marsal S, Julià A. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front Bioeng Biotechnol*. 2015;3:23. doi: 10.3389/fbioe.2015.00023.

48. Metabolomics Workbench website. http://www.metabolomicsworkbench.org/data/index.php. Accessed December 15, 2016.

48a. MetaboLights website. http://www.ebi.ac.uk/metabolights/. Accessed December 15, 2016.

49. Fischer K, Kettunen J, Würtz P, Haller T, Havulinna AS, Kangas AJ, Soininen P, Esko T, Tammesoo ML, Mägi R, Smit S, Palotie A, Ripatti S, Salomaa V, Ala-Korpela M, Perola M, Metspalu A. Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons. *PLoS Med*. 2014;11:e1001606. doi: 10.1371/journal.pmed.1001606.

49a. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reily MD, Thaden JJ, Viant MR. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*. 2007;3:211–221.

50. Rhee EP, Clish CB, Ghorbani A, Larson MG, Elmariah S, McCabe E, Yang Q, Cheng S, Pierce K, Deik A, Souza AL, Farrell L, Domos C, Yeh RW, Palacios I, Rosenfield K, Vasan RS, Florez JC, Wang TJ, Fox CS, Gerszten RE. A combined epidemiologic and metabolomic approach improves CKD prediction. *J Am Soc Nephrol*. 2013;24:1330–1338. doi: 10.1681/ASN.2012101006.

51. Shah SH, Kraus WE, Newgard CB. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. *Circulation*. 2012;126:1110–1120. doi: 10.1161/CIRCULATIONAHA.111.060368.

52. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372:793–795. doi: 10.1056/NEJMp1500523.

53. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. *Nature*. 2012;491:384–392. doi: 10.1038/nature11708.

54. Shah SH, Newgard CB. Integrated metabolomics and genomics: systems approaches to biomarkers and mechanisms of cardiovascular disease. *Circ Cardiovasc Genet*. 2015;8:410–419. doi: 10.1161/CIRCGENETICS.114.000223.

Circulation
Genomic and Precision Medicine

Potential Impact and Study Considerations of Metabolomics in Cardiovascular Health and Disease: A Scientific Statement From the American Heart Association

Susan Cheng, Svati H. Shah, Elizabeth J. Corwin, Oliver Fiehn, Robert L. Fitzgerald, Robert E. Gerszten, Thomas Illig, Eugene P. Rhee, Pothur R. Srinivas, Thomas J. Wang and Mohit Jain on behalf of the American Heart Association Council on Functional Genomics and Translational Biology; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; and Stroke Council

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/10/2/e000032