Title

The Use of Metabolomics in Population-Based Research

Journal

Advances in Nutrition, 5(6)

ISSN

2161-8313

Authors

Su, L Joseph
Fiehn, Oliver
Maruvada, Padma
et al.

Publication Date

2014-11-01

DOI

10.3945/an.114.006494

Peer reviewed

The Use of Metabolomics in Population-Based Research[1–3]

L. Joseph Su,[4]* Oliver Fiehn,[6] Padma Maruvada,[7] Steven C. Moore,[5] Stephen J. O'Keefe,[8] David S. Wishart,[9] and Krista A. Zanetti[4]

[4]Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, and [5]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD; [6]Department of Molecular and Cellular Biology and Genome Center, University of California at Davis, Davis, CA; [7]Division of Digestive Diseases and Nutrition, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD; [8]Department of Medicine, University of Pittsburgh, PA; and [9]Department of Biological Sciences and Computing Science, University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

The NIH has made a significant commitment through the NIH Common Fund's Metabolomics Program to build infrastructure and capacity for metabolomics research, which should accelerate the field. Given this investment, it is the ideal time to start planning strategies to capitalize on the infrastructure being established. An obvious gap in the literature relates to the effective use of metabolomics in large-population studies. Although published reports from population-based studies are beginning to emerge, the number to date remains relatively small. Yet, there is great potential for using metabolomics in population-based studies to evaluate the effects of nutritional, pharmaceutical, and environmental exposures (the "exposome"); conduct risk assessments; predict disease development; and diagnose diseases. Currently, the majority of the metabolomics studies in human populations are in nutrition or nutrition-related fields. This symposium provided a timely venue to highlight the current state-of-science on the use of metabolomics in population-based research. This session provided a forum at which investigators with extensive experience in performing research within large initiatives, multi-investigator grants, and epidemiology consortia could stimulate discussion and ideas for population-based metabolomics research and, in turn, improve knowledge to help devise effective methods of health research. *Adv Nutr 2014;5:785–788.*

Introduction

Metabolomics is the quantitative measurement of the dynamic metabolic responses of living systems to pathophysiologic stimuli or genetic modifications (1). In a human context, metabolomics is the comprehensive characterization of molecules of both endogenous and exogenous origins in biospecimens and then applied to biochemical pathways and/or integrative metabolic physiology. Metabolomics has received more attention in recent years as an ideal methodology to unravel signals closer to the culmination of the disease process. The compounds identified through metabolomics profiling are found in a range of intermediate metabolic pathways and may serve as biomarkers of exposure or susceptibility to disease. Therefore, metabolomics is a valuable approach for deciphering metabolic outcomes with a phenotypic change. Through the sessions at the ASN Scientific Sessions and Annual Meeting at Experimental Biology 2014, there were calls for better biomarkers for dietary intakes and alternatives for existing dietary assessment tools. Some sessions discussed biomarkers for environmental, lifestyle, and dietary exposures; intermediate outcomes; and early detection of diseases. There were also calls for better integration of genomic knowledge gained with nutritional research. The session "Use of Metabolomics in Population-based Research" was well poised on the concluding day of the meeting to address a relatively new approach for many of the issues raised throughout the proceedings.

Dr. Joseph Su introduced the session by highlighting that the examination of the role of metabolites in human health

[1] This article is a summary of the symposium "The Use of Metabolomics in Population-based Research" held 30 April 2014 at the ASN Scientific Sessions and Annual Meeting at Experimental Biology 2014 in San Diego, CA. The symposium was sponsored by the American Society for Nutrition (ASN) and the ASN Nutritional Epidemiology Research Interest Section.

[2] Supported, in part, by the Breast Cancer Research Stamp Fund, awarded through competitive peer review and the Intramural Research Program of the National Cancer Institute, NIH, U.S. Department of Health and Human Services. The Shanghai Women's Health Study was supported primarily by R37CA70867, and the Shanghai Men's Health Study was supported by R01CA082729. O. F. received funding from NIH grant U24 DK097154. S. J. O. is supported by NIH grant R01 CA135379 and the NIH Research Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London (Jeremy Nicholson).

[3] Author disclosures: L. J. Su, O. Fiehn, P. Maruvada, S. C. Moore, S. J. O'Keefe, D. S. Wishart, and K. A. Zanetti, no conflicts of interest.

* To whom correspondence should be addressed. E-mail: sulj@mail.nih.gov.

is not new in the field of nutrition. However, metabolomics is a relatively new "-omics" field that followed genomics, transcriptomics, and epigenomics, which generated a renewed interest in the examination of the effect of metabolites on human health from a more integrative prospective. With the generation of high-dimensional data by using either targeted or untargeted approaches, health researchers have new opportunities to apply metabolomics analyses to population-based research. At the same time, challenges need to be addressed to effectively use this tool and to interpret the data generated from population-based research.

Dr. Padma Maruvada opened the session with a presentation on the NIH Common Fund's Metabolomics Program, which has the goal of building metabolomics capacity and bridging discovery to translation. The program has 5 cores areas of focus, including comprehensive resource cores, technology development, development of reference standards, training, and data sharing and international collaboration. To date, 6 Regional Comprehensive Metabolomics Resource Cores have been established at the University of Michigan; Research Triangle Institute; the University of California, Davis; the University of Florida; the University of Kentucky; and the Mayo Clinic (2). The resource cores provide training and workshops, fee-for-service support, and the opportunity for investigators to apply for collaborative pilot and feasibility studies. Additionally, a Metabolomics Data Repository and Coordination Center (DRCC)[10] has also been established at the University of California, San Diego, to serve as a national data repository and to coordinate data sharing and dissemination activities from national and international metabolomics centers. The DRCC develops analytical tools and data standards and has established a cloud-based data repository, which is a public access database, called Metabolomics Workbench (http://www.metabolomicsworkbench.org/) (3). Additionally, the DRCC is accepting metabolomics data from NIH grantees/projects to be deposited in Metabolomics Workbench from the larger metabolomics community as long as the minimum metadata- and raw data–sharing requirements are followed. Furthermore, the DRCC has established a Metabolite Standards Synthesis Core that supports synthesis and characterization of metabolomics standard reference materials to provide to the metabolomics community through contracts. The Common Fund Program also supports training through Mentored Development Awards (K01 grant mechanism) and development of courses and workshops through the R25 grant mechanism.

Dr. Steven Moore highlighted examples that illustrated their efforts to identify metabolomics signatures for dietary intake and BMI. In the context of epidemiologic research, 2 applications of metabolomics that are of current interest to the field of nutrition are the assessment of diet biomarkers to improve on existing self-report measures and the assessment of obesity-related biomarkers to explore biologic pathways that mediate BMI-disease associations. However, there are limited data on the breadth and classification of metabolites from current metabolomics platforms that are related to diet or BMI. Dr. Moore examined cross-sectional associations between diet, BMI, and concentrations of ~400 blood metabolites. For diet, pairwise correlations were calculated for 137 different dietary items vs. 412 metabolites in 500 men and women (4). In total, 46 diet-metabolite associations met the Bonferroni-corrected threshold of statistical significance, including biomarkers for citrus, greens, red meat, fish, peanuts, butter, alcohol, supplement intake, and coffee. Most biomarkers were constituent ingredients or downstream metabolites of the relevant dietary items. For BMI, they estimated metabolite correlations for 317 metabolites in a population of 946 men and women (5). They identified 36 BMI-metabolite associations that met the Bonferroni threshold, 15 of which were novel findings. Biomarkers included branched-chain amino acids and markers of possible clinical relevance, including mannose and lathosterol. Together, their results for diet and BMI highlight that metabolomics holds great potential as a tool for characterizing exposure in epidemiologic studies.

Dr. Oliver Fiehn introduced the NIH West Coast Metabolomics Center (WCMC), which is 1 of the Comprehensive Resource Cores funded by the NIH Common Fund Metabolomics Program. The goal of the center is to integrate metabolomics into large-scale health research of particular relevance to the objective of this symposium. The service cores at the WCMC are organized on the basis of compound-centric platforms that tackle both targeted and untargeted metabolomics analyses. The metabolites of interest for these platforms include the following: polar and neutral lipids; volatiles; primary small metabolites, such as sugars, FFAs, and amino acids; and other specific metabolites, such as oxylipids, sterols and bile acids, flavonoids, and glucuronidated and glycosylated aglycones. Dr. Fiehn highlighted 2 epidemiologic studies: The Environmental Determinants of Diabetes in the Young (TEDDY) project, which examined >12,000 samples, and Functional Cardio-Metabolomics, which examined metabolic pathways associated with atherosclerosis and cardiovascular risk factors involving 1700 samples. These examples demonstrated the capacity and approaches that the WCMC uses to successfully process high-throughput lipidomics projects. He also highlighted the capacity for and process of the data processing and normalizations using internal standards and batch correction at the WCMC. Additionally, he showcased how the WCMC integrates data from different platforms, including 30 mass spectrometers and 5 NMRs, and utilizes biochemical mapping through MetaCyc, a database of experimentally verified metabolic pathways from all domains of life (2151 pathways from >2515 different organisms), and other biochemical databases and chemical mapping through substructure comparison to cover >1200 targeted compounds. There has been strong interest among epidemiologists to apply metabolomics for an increased understanding of the etiologic and preventive factors on health outcomes. Less emphasis has been placed on the utilities of metabolomics for phenotyping the response to drug

[10] Abbreviations used: DRCC, Data Repository and Coordination Center; TMIC, The Metabolomics Innovation Centre; WCMC, West Coast Metabolomics Center.

treatment. However, Dr. Fiehn highlighted an example of a study that was directed by Dr. Rima Kaddurah-Doauk at Duke University in which a pharmacometabolomics network that consisted 5 clinical centers and 4 analytical laboratories to examine the metabolomics effect of 5 drugs (6). This study integrated metabolomics and genomics to predict which patients would benefit from specific treatments and dosages before the initiation of the treatment and alternative therapies. Most impressively, he demonstrated that metabolomics can be used to inform the single nucleotide polymorphism analysis results from these patients to guide treatment plans.

Throughout the Experimental Biology 2014 sessions, there were a number of conversations regarding the proper assessment of diet in free-living populations. The relation between diet and health outcomes, such as red wine consumption and cardiovascular disease prevention, red meat intake and risk of cancer, high-purine-content foods and gout, as well as myocardial infarction, has been identified or investigated. However, it has long been recognized that dietary questionnaires have limitations for measuring true dietary intake and an individual's response to diet. Accurate assessment of dietary intake is limited by the potential of recall bias, imprecise and inconsistent portion-size estimation, under-reporting, and incorrect reporting, among others. Furthermore, the field of nutritional epidemiology has been searching for biomarkers that quantify total dietary intake (e.g., doubly labeled water for energy), correlate to total dietary intake (e.g., plasma ascorbate for vitamin C), serve as proxies for estimating dietary intake (e.g., carotene for vegetables), indicate prior consumption, and/or assess compliance for certain restricted diets.

Dr. David Wishart presented the efforts at The Metabolomics Innovation Centre (TMIC) in developing applications of metabolomics for dietary monitoring. Metabolomics offers a quantitative method for measuring diets and dietary responses. However, systematic studies need to be performed to gather and validate key dietary biomarkers and markers of physiologic responses to dietary intake. As identified by Dr. Wishart, the key issues surrounding the use of metabolomics for diet monitoring and assessment involve the need for the following: *1*) standard lists of food components and lists of food biomarkers from dietary feeding studies, *2*) the ability to link food components with food metabolites, *3*) the ability to identify specific food biomarkers found in biofluids or tissues, *4*) the need to correlate food biomarker concentrations with food intake, and *5*) the need to recognize which food biomarkers are ideal for specific time scales (7). He showcased the progress that has been made to address these challenges at TMIC. This included a description of Phenol-Explorer (http://www.phenol-explorer.eu/) (8), a Web-based comprehensive database on polyphenol content and polyphenol concentrations in foods, that was jointly developed by TMIC and Dr. Augustin Scalbert at the French National Institute for Agricultural Research (INRA), Clermont-Ferrand. Content data, including food, food variety, commercial origin, extraction protocol, analytical methods, and references; >900 peer-reviewed papers; and 37,000 data entries on 502 polyphenols in 459 foods are searchable by food name, compound name, or polyphenol class in the system. Colorful polyphenol content graphs, structured polyphenol taxonomy, and food taxonomy can be displayed through the search of Phenol-Explorer. Additionally, TMIC is developing a comprehensive database of raw food components called FooDB (http://foodb.ca/) (9) to extend and validate the composition data and identify candidate food biomarkers. FooDB currently consists of >1000 raw foods with descriptions, food associations, and their concentrations derived through the literature and data mining, as well as >60,000 hyperlinks to 11 different databases. Last, he demonstrated the Human Metabolome Database (HMDB; http://www.hmdb.ca/) (10), a Web-accessible resource developed by TMIC that contains detailed information on 40,214 quantified, detected, and expected human metabolites, The Human Metabolome Database also captures disease associations, food metabolites, food-microbial metabolites, and food biomarkers.

The symposium session concluded with a presentation from Dr. Stephen O'Keefe in which he discussed his efforts to identify dietary factors in colon cancer prevention. Dr. O'Keefe reminded the audience that colon cancer is a westernized disease with an incidence rate in African Americans of ~65:100,000, compared with <5:100,000 in rural Africa populations. On the basis of the analysis of epidemiologic evidence from around the world, Doll and Peto concluded that ˚90% of gastroenterological cancers were caused by diet (11), and the most recent meta-analysis of the available literature by the World Cancer Research Fund/American Institute for Cancer Research concluded that dietary fiber and red processed meat convincingly affect risk (12). To explore the hypothesis that diet asserts its effect on cancer risk through its influence on colonic microbial metabolism, they examined differences in fecal and colonic contents between samples of the healthy middle-aged African American and rural South African populations by ^{1}NMR spectroscopy. The study first collected biospecimens while study subjects consumed their usual diets (i.e., high in meat and fat and low in fiber for African Americans and low in meat and fat and high in fiber for Africans) and then again after switching their diets for 2 wk. The results showed differences in colonic microbial metabolites at baseline. After the dietary switch, the preliminary results showed that microbial metabolic phenotype changed with mucosal biomarkers of cancer risk. This study was able to investigate the authors' hypothesis that the physiologic modification of the diet can be captured through metabolomics assessment of biospecimens. This study offered a significant implication for future studies of diet and cancer using metabolomics to assess dietary changes.

In conclusion, this session highlighted the current NIH investment for metabolomics as well as several studies exemplifying that the generation of metabolomics data offer new opportunities for health researchers to apply this tool in population-based research, especially in the field of nutrition. Additionally, the session highlighted the importance of metabolomics database development to support nutrition and nutrition-related research. Finally, although there are many

opportunities to move the use of metabolomics analyses into population-based studies, challenges for its effective and valid use remain and will need to be addressed.

Acknowledgments

All authors read and approved the final manuscript.

References

1. Nicholson JK, Lindon JC, Holmes E. Metabolomics: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 1999;29:1181–9.
2. National Institutes of Health. National Institutes of Health, Office of Strategic Coordination—the common fund: metabolomics: funded research. Bethesda (MD): NIH; 2014.
3. University of California San Diego. Metabolomics workbench. [Cited 2014 July 15]. Available from http://www.metabolomicsworkbench.org/. 2014.
4. Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100:208–17.
5. Moore S, Matthews C, Sampson J, Stolzenberg-Solomon R, Zheng W, Cai Q, Tan YT, Chow WH, Ji BT, Liu DK, Xiao Q, Boca SM, Leitzmann MF, Yang G, Xiang YB, Sinha R, Shu XO, Cross AJ. Human metabolic correlates of body mass index. Metabolomics 2014;10:259–69.
6. 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 reveals racial differences in response to atenolol treatment. PLoS ONE 2013;8:e57639.
7. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJ, Wishart DS. The food metabolome: a window over dietary exposure. Am J Clin Nutr 2014;99:1286–308.
8. Rothwell JA, Urpi-Sarda M, Boto-Ordoñez M, Knox C, Llorach R, Eisner R, Cruz J, Neveu V, Wishart D, Manach C, Andres-Lacueva C, Scalbert A. Phenol-Explorer 2.0: a major update of the Phenol-Explorer database integrating data on polyphenol metabolism and pharmacokinetics in humans and experimental animals. Database 2012;2012:bas031, doi:10.1093/database/bas031.
9. The Metabolomics Innovation Centre (TMIC). FooDB, version 1.0. [Cited 2014 July 15]. Available from http://foodb.ca/2014.
10. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Yallou F, Bjorndahl T, Perez-Pineiro R, Eisner R, Allen F, Neveu V, Greiner R, Scalbert A. HMDB 3.0—the Human Metabolome Database in 2013. Nucleic Acids Res 2013;41:D801–7.
11. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 1981;66:1191–308.
12. American Institute for Cancer Research; World Cancer Research Fund. Food, nutrition, physical activity and the prevention of cancer: a global perspective: a project of World Cancer Research Fund International. Washington: American Institute for Cancer Research; 2007.