# Metabolomics: The Stethoscope for the Twenty-First Century

Hutan Ashrafian[a]    Viknesh Sounderajah[a]    Robert Glen[a, b]    Timothy Ebbels[a]

Benjamin J. Blaise[a]    Dipak Kalra[c]    Kim Kultima[d]    Ola Spjuth[e]    Leonardo Tenori[f]

Reza M. Salek[g]    Namrata Kale[g]    Kenneth Haug[g]    Daniel Schober[h]    Philippe Rocca-Serra[i]

Claire O'Donovan[g]    Christoph Steinbeck[j]    Isaac Cano[k]    Pedro de Atauri[l]    Marta Cascante[l]

[a]Institute of Global Health Innovation and Department of Surgery and Cancer, Imperial College London, London, UK; [b]Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Cambridge, UK; [c]Department of Medical Informatics and Statistics, University of Ghent, Ghent, Belgium; [d]Department of Medical Sciences, Uppsala University, Uppsala, Sweden; [e]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden; [f]Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy; [g]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, UK; [h]Department of Stress and Developmental Biology, Leibniz Institute of Plant Biochemistry, Halle (Saale), Germany; [i]Department of Engineering Science, Oxford e-Research Centre, University of Oxford, Oxford, UK; [j]Institute for Inorganic and Analytical Chemistry, Friedrich-Schiller-University, Jena, Germany; [k]Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain; [l]Department of Biochemistry and Molecular Biomedicine, Faculty of Biology, Universitat de Barcelona and CIBERHD (CIBER de Enfermedades hepáticas y digestivas), Barcelona, Spain

---

### Significance of the Study

- Metabolomics encompasses the systematic identification and quantification of all metabolic products from the human body.
- This field could provide clinicians with novel sets of diagnostic biomarkers for disease states, in addition to quantifying treatment response to medications at an individualized level.
- Outstanding issues preventing widespread clinical use are scalability of data interpretation, standardization of sample handling practice, and e-infrastructure.

---

## Abstract
Metabolomics encompasses the systematic identification and quantification of all metabolic products in the human body. This field could provide clinicians with novel sets of diagnostic biomarkers for disease states in addition to quantifying treatment response to medications at an individualized level. This literature review aims to highlight the technology underpinning metabolic profiling, identify potential applications of metabolomics in clinical practice, and discuss the translational challenges that the field faces. We searched PubMed, MEDLINE, and EMBASE for primary and secondary research articles regarding clinical applications of metabolo-

Correspondence to:
Hutan Ashrafian, hutan@imperial.ac.uk

mics. Metabolic profiling can be performed using mass spectrometry and nuclear magnetic resonance-based techniques using a variety of biological samples. This is carried out in vivo or in vitro following careful sample collection, preparation, and analysis. The potential clinical applications constitute disruptive innovations in their respective specialities, particularly oncology and metabolic medicine. Outstanding issues currently preventing widespread clinical use are scalability of data interpretation, standardization of sample handling practice, and e-infrastructure. Routine utilization of metabolomics at a patient and population level will constitute an integral part of future healthcare provision.

## Introduction

In recent years, the novel field of metabolomics (synonymous with "metabonomics" and "metabolic profiling"), which included in the "omics" subclass of biological studies, has gained increasing attention from both clinical and academic health circles. Metabolomics refers to the "systematic identification and quantification of the small molecule metabolic products (the metabolome, which consists of 40,000 metabolites in humans [1]) of a biological system at a specific point in time" [2]. It has emerged as a fertile area for research and development given its inherently vast translational promise. The fundamental paradigm underpinning clinical metabolic phenotyping is that any localized metabolic, physical, or histological perturbation in the human body will result in global changes characterized in biological samples. These changes are statistically connected to both the disease process and complex gene-environment interactions. Through this, predictions regarding disease risk and treatment responses may be estimated at an individual level [2]. These benefits have tangible health and socioeconomic benefits as effective provision of care improves patient safety and increases the cost-effectiveness of delivered therapies [3].

Further to the fundamental tenants of the field, this technology embraces several concepts popular within contemporary healthcare. Metabolomics feeds into the relatively novel concept of systems medicine [4–6], which supports the importance of viewing each patient as a distinct combination of biochemical, physiological, and environmental interactions. In turn, this holistic appreciation facilitates the paradigm shift in care provision towards a "P4" approach, consisting of prediction, prevention, personalization, and participation [7]. As such,

metabolomics is a foundation exemplar of precision medicine [8] and fulfils the needs of offering the powerful strengths of mathematics and analytics to healthcare (iatromathematics) [9]. However, despite the vast translational potential of this field, there is relatively scarce knowledge of this field outside research circles, particularly in comparison to other "omic" based ventures as such the Human Genome Project [10]. As a consequence, the uptake of this potentially disruptive innovation may be hamstrung by a lack of clinician awareness of its pending utility within clinical practice. This gap in collective knowledge may subsequently result in substandard care with respect to both individual and population health as a whole. This review aims to highlight the technology underpinning metabolic profiling, identify potential applications of metabolomics in clinical practice, and discuss the translational challenges that metabolomics faces.

## Technology

The process of attaining relevant metabolomic data requires 3 steps [11]:
1. Sample collection
2. Sample preparation
3. Sample analysis
4. Sample collection

Metabolic profiling can be performed on both in vivo and in vitro samples [12]. These include a range of samples consisting of cells, fluids, or tissues. In practice, biofluids are amongst the easiest to acquire and work with. Such samples include serum, plasma, urine, saliva, or faecal content. All samples require precise handling as metabolic pathways are highly susceptible to exogenous environmental factors, which can lead to inaccurate results upon analysis. Therefore, for example, maintaining a low temperature as well as ensuring consistent sample extraction is vital for both modalities.

## Sample Preparation

The compounds within a given sample are usually of a complex nature. These sample constituents may be separated to allow for siphoning of pertinent analytes from others that cannot be resolved by the detector [13]. The preparation of the sample is highly dependent on the detection modality that is subsequently used. Of the 2 principles modalities, nuclear magnetic resonance (NMR)

Ashrafian et al.

spectroscopy and mass spectroscopy (MS) [14], NMR techniques often omit a separation phase.

Sample pretreatment depends on whether a targeted or non-targeted study is being performed [15]. Pertinently for non-targeted studies, minimal pretreatment is desired so that metabolites are not lost. There are a number of pretreatment separation modalities used with MS that hold distinct advantages in characterizing particular aspects of a metabolome.

The principle separation modalities used with MS include the following: gas chromatography allows for gas phase separation of molecules [16]. It is most useful for analysis of trace amounts of volatile compounds. High-performance liquid chromatography [17] uses chromatographic columns, which are filled with microparticles, to allow for high pressure elution of the sample, allowing for increased chromatographic separation. It is the most commonly used analytical technique given its versatility and ability to retain a number of compounds.

Capillary electrophoresis [18] allows for the electrokinetic separation of the sample. It is another versatile technique which enables separation of a wide range of analytes ranging from small inorganic ions to larger proteins.

It should be noted that there are many further pre-analytical considerations of vital importance, which include the collection of samples from either fasting or fed states, the importance of circadian rhythms, the consideration of which additives are added to collection tubes as well as the potential for sample haemolysis. These are outside the scope of this review but are covered in detail in other reports [19, 20].

### Detection Methods

As mentioned earlier, metabolic profiling is based upon 2 principle analytical modalities: NMR spectroscopy and MS. Both are able to simultaneously identify and quantify information on a wide range of molecules. Moreover, both require only a small amount of sample.

NMR spectroscopy relies upon the ability of spin active nuclei to absorb and re-emit pulsed electromagnetic radiation when placed within a magnetic field [21, 22]. The frequency pattern, a signature which is a consequence of the interaction of the nuclei with the electromagnetic field, provides information regarding the molecular structure, motion, and chemical environment. In biological samples, hydrogen is the most commonly targeted nucleus, due to its abundance in biological samples. Other atoms such as carbon and phosphorus may also be targeted. NMR is perceived to be a highly reproducible and rapid platform which crucially offers a non-destructive method of sample analysis, which may be either fluid or solid. It offers exact quantification of a wide range of chemical structures and remains the only tool that can provide atom-centred information. Practically, it is relatively low cost per sample, requires minimal sample preparation, and offers prompt throughput (15 min per sample). Moreover, samples subjected to NMR analysis may be subsequently further analysed by MS. Deficiencies of this technology are that it is insensitive, requires high user skills, and that there is an initial high start-up cost in order to acquire the instruments.

In comparison, mass spectrometry is a destructive analytical process which relies upon the formation of gas phase ions, which are subsequently separated by their mass/charge ratio. The ions then hit a detector, which accounts for the number of ions for each mass/charge ratio [23]. This is subsequently analysed and compared against available mass spectral databases in order to predict the molecular identity of the constituents. MS is a highly sensitive method of sample analysis, which can be used for targeted and non-targeted analyses. However, the sensitivity and accuracy of the detection are highly dependent upon the experimental conditions and the instrument settings.

### Sample Analysis

There is a need for rapid and accurate statistical tools that can process the complexity and volume of the vast amount of data that is generated. Different metabolomic features may be used as the input for data analysis. These include spectral peak areas, metabolite concentrations, and spectral bin areas [24]. A number of univariate and multivariate statistical approaches can be performed, which also focus upon data pre- and post-processing tasks such as signal extraction/peak detection, noise reduction, correction of run order drifts or batches, and peak fitting. As a class, they are known as chemometric methods [25].

Univariate methods analyse the metabolomic features independently [26]. They are often easier to interpret as they employ more commonly used and understood statistical approaches. However, this analysis does not consider the presence of interactions between different metabolic features. Confounding variables such as gender, diet, or BMI are not accounted for. This leads to an increased probability of incorrect results. The choice of sta-

tistical analysis is particularly important given the number of features that are simultaneously analysed, thus increasing the risk of a false-positive statistical result, known as the multiple testing problem. However, it can also be argued that situations which harbour significant confounders are best solved by careful cohort stratification from the outset.

In contrast to univariate analysis, multivariate analysis considers all the imputed metabolomic features and attempts to identify relationships between them [27]. These methods can be classified into 2 groups, supervised and unsupervised methods. Unsupervised methods are able to effectively detect data patterns with biological variables, with the most common unsupervised method being principal component analysis. Supervised methods identify patterns within variables of interest while downweighting other sources of variance. The most common supervised statistical method is partial least squares regression analysis.

### Clinical Applications

Although there are admittedly limited clinical applications for metabolomics currently, there are a burgeoning number of potential applications that could disrupt clinical practice in the near future. As noted, the conceptual appeal of the field is to provide biomarkers, which may be either predictive, diagnostic, or prognostic.

#### Oncology

Metabolomics offers a particularly broad set of oncological applications, particularly with respect to providing serum [28] or imaging-based biomarkers of cancer. Metabolic profiling will assist in the diagnosis of several tumour types. This has been prominently noted with breast cancer in particular. There have been over 30 endogenous metabolites which have been noted in breast cancer specimens, including tCHo levels (resulting from increased phosphocholine), low glycerophosphocholine, and low glucose [29]. Moreover, Bathen et al. [30] stated that a malignant phenotype can be reliably detected against normal tissue with sensitivity and sensibility between 83 and 100% for tumour size, lymph node status, hormone status, and histology. Metabolic signatures have also been mapped for ovarian cancer [31], lung cancer [32], endometrial cancer [33], and colorectal cancer.

There has also been particular focus upon harnessing metabolomics to guide oncological surgery. Inglese et al. [34] highlighted the potential of combining 3D mass spectrometry imaging with unsupervised neural network-based techniques to precisely determine the extent of malignant tissue clusters pre-operatively. Intra-operatively, rapid ionization mass spectrometry has been successfully coupled to electrosurgical tools to allow for near real-time characterization of margins during cautery-led tumour dissection [35, 36]. This device, known as the "intelligent knife" (iKnife) has been tested in vivo and has garnered promising post-operative histopathological support [37]. However, current limitations with respect to the iKnife include the costly start-up costs as well as the need for robust, histologically specific mass spectral libraries.

In addition to diagnosis, metabolic profiling can be used in prognosticating clinical outcomes. NMR-based techniques have predicted which samples of human glioma cell cultures would be drug sensitive and drug resistant prior to treatment with either chemotherapy or hormonal therapy. Glunde et al. [38] demonstrated that a decrease in tCHO signal equates to a promising response to chemotherapy and may be an early marker of therapeutic effect, which is detectable prior to changes in conventional imaging in breast, brain, or prostate cancer. Moreover, micrometastases were predicted in a study of patients with breast cancer [39], who were noted to have higher levels of plasma glucose proline, lysine, phenylalanine, N-acetylcysteine, and lower lipid levels. Furthermore, Tenori et al. [40] demonstrated that pretreatment serum samples for patients with metastatic breast cancer are predictive for overall survival, time to progression and treatment toxicity, according to serum phenylalanine, glutamate, and glucose levels respectively.

#### Endocrinology

Quantifying the individual risk of type 2 diabetes has also been a key target for the field of metabolomics [41]. Wang et al. [42] suggest that the metabolite 2-aminoadipic acid (2-AAA) is a marker of diabetes risk and a potential modulator of glucose homeostasis in humans. They performed a nested case-control study over 12 years on 188 individuals who had developed diabetes and 188 propensity-matched normoglycemic controls identified from the Framingham Heart Study. It was demonstrated that those with 2-AAA concentrations in the top quartile had a >4-fold risk of developing diabetes. It was observed to be elevated up to 12 years before the onset of a clinically appreciable disease state. Moreover, it was noted that 2-AAA concentrated were not well correlated with other metabolite biomarkers of diabetes, which suggests

that 2-AAA is from a distinct pathophysiological pathway. Furthermore, 2-AAA treatment enhanced insulin secretion from a pancreatic β cell line as well as from murine and human islets. However, no human trials with 2-AAA analogues have been conducted.

### Rheumatology

There appears to be a use for metabolomics in the production of diagnostic biomarkers of both inflammatory and non-inflammatory rheumatological conditions [43]. Ouyang et al. [44] demonstrated that serum from patients with systemic lupus erythematosus had significant reductions in valine, tyrosine, phenylalanine, lysine, isoleucine, histidine, glutamine, and alanine amongst others in comparison to patients with rheumatoid arthritis and healthy controls. Kim et al. [45] have shown that over 20 metabolites are potential biomarkers to discriminate rheumatoid arthritis from other conditions such as ankylosing spondylitis, Behçet's disease, and gout. In addition to these inflammatory conditions, Adams et al. [46] demonstrated that collagen degradation products contribute to an osteoarthritis (OA) signature in blood with identification possible between healthy controls, early OA, and end-stage OA. However, the authors did note that these differences may be secondary to age-related chondrocyte changes.

### Neurology

In neurological disorders, the use of metabolomics is rapidly increasing [47]. In the field of Alzheimer's disease, identified metabolites are consistently associated with cognition, dementia, and particular lifestyle factors, which suggests that there may be novel targets for the prevention of cognitive decline and dementia [48]. In Parkinson's disease, metabolomics has deepened the knowledge about alterations in biochemical pathways involved in Parkinson's disease pathogenesis [49]. Metabolomics enables neurodegenerative disease stratification [50] as well as a diagnosis of disease severity in multiple sclerosis [51]. Another promising area is the use of metabolomics in the early diagnosis of traumatic brain injury (TBI) as well as the patient specific prognosis. Recently, a metabolic signature was identified in the serum of patients after TBI that may be indicative of a disrupted blood-brain barrier, offering a new avenue towards more precise TBI patient stratification in clinical practice [52].

### Respirology

NMR-based metabolomic assays of urine and plasma have demonstrated 3 particular urinary metabolites which are correlated with lung function, in cohorts of patients with and without chronic obstructive pulmonary disease [53]. Of the 3 metabolites, trigonelline was noted to have the strongest correlation with baseline pulmonary function. Increased hippurate and formate were associated with better lung function. It has also been determined that children with asthma and allergic rhinitis may be differentiated from age-matched controls based upon the alkane and aldehyde content of their breath condensate [54].

### Gastroenterology

Dawiskiba et al. [55] undertook a profiling study with inflammatory bowel disease (IBD) and concluded that there is no clear separation between ulcerative colitis and Crohn's disease; however, they did note evidence of characteristic signatures between active IBD and IBD in remission. These results were contrasted by Williams et al. [56] who stated that it is possible to distinguish Crohn's disease from ulcerative colitis in a similarly sized age- and gender-matched cohort, based upon choline, lipoprotein, and N-acetylated glycoprotein levels. Between IBD and healthy individuals, there is evidence of increased phenylalanine in the serum samples of patients with IBD, whereas higher glycine and lower acetoacetate levels were observed in urinary samples [57]. With respect to viral hepatitis, urinary metabolomics have shown 11 discriminant metabolites distinguishing patients with hepatitis B from those with cirrhosis or liver cancer [58].

### Cardiovascular Disease

Metabolic profiling of atherosclerosis, the predominant underlying process of cardiovascular disease, has been a longstanding population health priority [59]. Wurtz et al. [60] demonstrated that serum docosahexaenoic acid, glutamine, and tyrosine may be potential predictors for atherosclerosis development. Brindle et al. [61] published a seminal NMR-based study which noted promising separation between blood samples between patients who have coronary heart disease against healthy participants. Sabatine et al. [62] demonstrated that there is an increase in lactic acid and metabolites involved in AMP-mediated skeletal muscle catabolism in both healthy patients as well as those with inducible ischaemia. However, the rise in the citric acid cycle products was seen only in patients with inducible ischaemia. Further

metabolic separation was noted by Vallejo et al. [63] in patients who were diagnosed with acute coronary syndrome against healthy controls both acutely and after 6 months. With respect to heart failure, Dunn et al. [64] demonstrated that pseudouridine, 2-oxoglutarate, 2-hydroxy 2-methylpropanoic acid, erythritol, and 2,4,6-trihydroxypyrimidine were noted as potential biomarkers when discriminating between patients with heart failure and healthy individuals.

*Paediatrics*

Given the relative simplicity, safety, and non-invasiveness of the approach, the application of metabolomic analyses is particularly suitable to the field of paediatrics. There is particular use in the diagnosis inborn errors of metabolism [65]. Metabolic profiling has shown promise in detecting putative gain-of-function of prevalent mutations in genes encoding metabolic enzymes such as isocitrate dehydrogenase (IDH). Metabolomics approaches have shown that mutated IDH1 and IDH2 proteins catalyse an additional reaction resulting in the formation of 2-hydroxyglutarate (2-HG), which is a metabolite absent when IDH is not mutated. This "new" metabolite inhibits α-ketoglutarate-dependent dioxygenases that play a key role in regulating the epigenetic state of cells. On the basis of these findings, phase I trials with selective inhibitors of mutated IDHs have been performed in patients with advanced haematologic malignancies, which have demonstrated an objective response rate ranging from 31 to 40% with durable responses (>1 year). Furthermore, IDH inhibitors have demonstrated early signs of activity in solid tumours with IDH mutations, including cholangiocarcinomas and low-grade gliomas [66].

**Translational Challenges**

Successful translation requires widespread understanding of the underpinning science, acceptance of standard operating procedures, training in the interpretation of results, and presence of reliable metabolomic data libraries. As such, the transition from a research tool to a viable clinical tool requires further co-ordination across a variety of disciplines on a local level. Currently, issues are related to equipment design, experimental validation, standardization of methods, and interpretability in a reliable and reproducible fashion.

Translation hinges upon several hurdles. The first is the standardization of experimental procedures and equipment, particularly through the adoption of standard operating procedures [22]. Data quality and reproducibility requires standardization. In keeping with this, the NIHR-MRC National Phenome Centre has prioritized the standardization of sample collection and analysis [67]. Targeted metabolomic profiling relies on commercial kits which offer highly standardized measurements on a restricted set of known metabolites. While coverage of the metabolic space is limited, the reduced cost means large cohorts can be characterized. This can lead to valuable insights into disease aetiology. Moreover, owing to their affordability and their highly standardized protocol, these kit-based approaches have gained popularity.

Two such suppliers, Biocrates and Metabolon, offer quantification of a few hundred metabolites in a fast and semi-quantitative manner. In so-called "ring trials," Siskos et al. [68] have demonstrated a median inter-laboratory coefficient of variation of 7.6% when using different instrumentation but a common protocol. Inter-laboratory reproducibility of targeted platforms [68], which underlines the potential of the Biocrates platform to monitor human serum in plasma-targeted metabolite profiling. The study also provided a better understanding of the platform error model. These are particularly useful for data scientists carrying out meta-analysis and working towards integrating datasets. Several groups have elected to apply this molecular phenotyping technique to large cohorts. A pertinent example is the KORA cohort, which has been extensively used by Jourdan et al. [69] where 1,614 subjects from KORA S4 (aged 54–75 years at the time of examination) and 3,061 subjects in KORA F4 (aged 31–82 years) have had their serum metabolite profile evaluated and correlated to the body fat-free mass index (FFMI), with a particular focus on branched-chain amino acids. Lending weight to previous findings, recently, nearly 10,000 patients have been profiled to test for a link between diabetes type-2 onset and high levels of circulating branched-chain amino acids [70]. Dutch researchers identified genetic contributors to serum metabolite variation based on Biocrates profiling of nearly 7,500 patients [71]. These studies are testament of how metabolite profiling can be used efficiently to support the understanding of disease and how the technique can inform clinical decision in routine practice.

The second major hurdle is in data analysis and interpretation. As noted, the 2 major analytical approaches, NMR and MS, have significantly different requirements. Firstly, for the purpose of data collection, there are sample size and time issues. Data processing is also troublesome as 1 must account for correction of experimental arte-

facts, equipment differences, batch correction, and normalization. This leads to major differences in the standardization of data for analysis. NMR is intrinsically more reproducible given standard conditions, partly because the sample does not come into physical contact with the detector. MS, despite higher sensitivity, suffers from poorer reproducibility due to changes in instrument response over time and across operators or laboratories [72, 73]. As such, significant efforts are required in mass spectrometry to obtain reproducible data necessary for translation to the clinic.

The third step is the interpretation of data, which requires advanced statistical approaches, such as machine learning. A key stumbling block is the expertise required to fit and interpret these models robustly on a local level if this technology should permeate into routine clinical work. There have been major strides in recent years in both understanding the requirements for data interpretation in this area (eliminating noise, normalization, and non-linear approaches to data analysis) and in the development of new algorithms for metabolite identification [74] and prediction, predominantly through machine learning techniques [75, 76]. In some cases, such as the aforementioned iKnife, this process has been reduced to a binary decision: tumour or not tumour. However, other applications (e.g., analysis of biofluids, such as urine) will probably require standardized application of computational pipelines that result in reproducible and interpretable outputs. This is one of the major aims of the EU-2020 program, PhenoMeNal [77], which is a European e-infrastructure to provide interoperable tools and workflows for modern and large-scale clinical metabolomics. Aside to the incorporation of machine learning, achieving biologically consistent annotation in untargeted metabolomics remains a major challenge within the field metabolomics. The current gold standard for metabolite identification is centred around matching detected features with an authenticated standard, which has been attained using the same methodology and the same equipment. This rigidity understandably leads to several practical challenges when attempting to apply this approach to large datasets. Widely used approaches to overcome this is to use spectral libraries, which are often incomplete, or alternative computational methods, which can match multiple identities with a single feature [78]. This is perceived to be another major hurdle in the widespread dissemination of the technology currently.

The fourth hurdle is centred around the informatic and e-infrastructural support that is required. Many of the human metabolomics data are, as of yet, incomplete and are housed among fractured sets of databases. Initiatives such as PhenoMeNal aim to provide clinical researchers with well-tested and reproducible workflows to perform standard analysis of clinical samples. This includes the computational clustering of metabolomics data for patient stratification as well as workflows for biochemical pathway enrichments. In addition, PhenoMeNal also provides easily accessible computational infrastructures that can sustain large-scale analysis with sufficient compute and storage resources.

The translation of new metabolomics knowledge into health-related practices will require the development of an ethical framework that ensures compliance and enforcement of patient rights. Initiatives such as Tryggve (https://neic.no/tryggve), PhenoMeNal (http://phenomenal-h2020.eu/home/), and Elixir (https://www.elixir-europe.org/) are paving the way towards state-of-the art-scalable cross-border e-infrastructures for efficient, safe, and e-compliant storage of sensitive personal data metabolic phenotyping data. These initiatives are also strongly contributing to establish educational competences to ensure Findable, Accessible, Interoperable, Reusable (FAIR) Data Principles accomplishment in the analysis, sharing, and reuse of sensitive personal data. If governed appropriately, the ultimate direction of these frameworks will allow next-generation big data analysis and artificial intelligence capabilities to combine and coordinate the improved analysis of these metabolomics datasets.

## Conclusions

Metabolomic profiling offers a deeper understanding of metabolic and physiological function. Although it is argued by some that metabolomics has been a vital part of medicine for some time now as well-established laboratory biochemical techniques are essentially targeted quantitative metabolomics. However, it is the novel approach to untargeted analysis using emerging quantitative methods that is unique. Moreover, metabolomics offers a paradigm shift in our perception of disease states as we move away from looking for single molecule disease biomarkers and replace this with a search for more complex and dynamic patterns of metabolite concentrations. There is great potential for metabolomics platforms and technology to be harnessed into routine clinical practice to support clinical decisions and to empower patients through a clearer understanding of underlying

disease dynamics. While the routine utilization of metabolic phenotyping at a patient and population level requires formalized evidence of effectiveness and safety through large-scale randomized or pragmatic clinical trials, the current wave of evidence is advocating the massive benefits of this approach. Ultimately, metabolomics can offer a tangible route through which to translate the depth of innovation in addressing human health and disease through the aid of big data analytics and the enhanced translation of precision mathematics to precision medicine. By doing so, such a multifaceted and wide-ranging technology assessing metabolism can fulfil the 200-year-old legacy of René Laennec's original diagnostic device to render metabolomics the stethoscope of the 21st century.

## Conflict of Interest Statement

The authors declare no conflict of interest.

## References

1 Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0—the human metabolome database in 2013. Nucleic Acids Res. 2012 Nov;41(D1):D801–807.

2 Holmes E, Wijeyesekera A, Taylor-Robinson SD, Nicholson JK. The promise of metabolic phenotyping in gastroenterology and hepatology. Nat Publ Gr. 2015;12:458–71.

3 Puchades-Carrasco L, Pineda-Lucena A. Metabolomics applications in precision edicine: an oncological perspective. Curr Top Med Chem. 2017;17:2740–51.

4 Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, et al. Metabolomics enables precision medicine: 'a white paper, community perspective'. Metabolomics. 2016;12:149.

5 Shah SH, Newgard CB. Integrated metabolomics and genomics: systems approaches to biomarkers and mechanisms of cardiovascular disease. Circ Cardiovasc Genet. 2015 Apr; 8(2):410–9.

6 Sun L, Zou LX, Chen MJ. Make precision medicine work for chronic kidney disease. Med Princ Pract. 2017 Mar;26(2):101–7.

7 Hood L, Friend SH. Predictive, personalized, preventive, participatory (P4) cancer medicine. Nat Rev Clin Oncol. 2011;8(3):184–7.

8 Ashrafian H, Athanasiou T, Nicholson JK, Darzi AW . Unmet medical needs. In: Holmes E, Nicholson JK, Darzi AW, Lindon JC, editors. Metabolic phenotyping in personalized and public healthcare. San Diego: Academic Press; 2016. p. 1–13.

9 Ashrafian H. Mathematics in medicine: the 300-year legacy of iatromathematics. Lancet. 2013 Nov;382(9907):1780.

10 Green ED, Watson JD, Collins FS. Human genome project: twenty-five years of big biology. Nature. 2015 Oct;526(7571):29–31.

11 Patti GJ, Yanes O, Siuzdak G. Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012 Apr;13:263–9.

12 Chetwynd AJ, Dunn WB, Rodriguez-Blanco G. Collection and preparation of clinical samples for metabolomics. Adv Exp Med Biol. 2017;965:19–44.

13 Iwasaki Y, Sawada T, Hatayama K, Ohyagi A, Tsukuda Y, Namekawa K, et al. Separation technique for the determination of highly polar metabolites in biological samples. Metabolites. 2012;2(3):496–515.

14 Amberg A, Riefke B, Schlotterbeck G, Ross A, Senn H, Dieterle F, et al. NMR and MS methods for metabolomics. New York, NY: Humana Press; 2017. p. 229–58.

15 Bingol K. Recent advances in targeted and untargeted metabolomics by NMR and MS/NMR methods. High-throughput. 2018 Apr; 7(2):9.

16 Emwas AH, Al-Talla ZA, Yang Y, Kharbatia NM. Gas chromatography-mass spectrometry of biofluids and extracts. Methods Mol Biol. 2015;1277:91.

17 Gika HG, Theodoridis GA, Plumb RS, Wilson ID. Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. J Pharm Biomed Anal. 2014 Jan;87:12–25.

18 Maier TV, Schmitt-Kopplin P. Capillary electrophoresis in metabolomics. Methods Mol Biol. 2016;1483:437–70.

19 Kirwan JA, Brennan L, Broadhurst D, Fiehn O, Cascante M, Dunn WB, et al. Preanalytical processing and biobanking procedures of biological samples for metabolomics research: a white paper, community perspective (for "precision medicine and pharmacometabolomics task group"-the metabolomics society initiative). Clin Chem. 2018 Aug;64(8):1158–82.

20 Fan TW-M. Considerations of sample preparation for metabolomics investigation. Totowa, NJ: Humana Press; 2012. Vol. 7–27.

21 Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, et al. The future of NMR-based metabolomics. Curr Opin Biotechnol. 2017 Feb;43:34–40.

22 Emwas AH, Luchinat C, Turano P, Tenori L, Roy R, Salek RM, et al. Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review. Metabolomics. 2015 Aug;11(4):872–94.

23 Gowda GA, Djukovic D. Overview of mass spectrometry-based metabolomics: opportunities and challenges. Methods Mol Biol. 2014;1198:3–12.

24 Alonso A, Marsal S, Julià A. Analytical methods in untargeted metabolomics: state of the art in 2015. Front Bioeng Biotechnol. 2015;3: 23.

25 Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. J Proteome Res. 2007 Feb;6(2):469–79.

26 Putri SP, Yamamoto S, Tsugawa H, Fukusaki E. Current metabolomics: technological advances. J Biosci Bioeng. 2013 Jul;116(1):9–16.

27 Worley B, Powers R. Multivariate analysis in metabolomics. Curr Metabolomics. 2012 Nov;1(1):92–107.

28 Cheung PK, Ma MH, Tse HF, Yeung KF, Tsang HF, Chu MKM, et al. The applications of metabolomics in the molecular diagnostics of cancer. Expert Rev Mol Diagn. 2019 Sep; 19(9):785–93.

29 Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. Clin Cancer Res. 2009;15(2):431.

30 Bathen TF, Geurts B, Sitter B, Fjøsne HE, Lundgren S, Buydens LM, et al. Feasibility of MR metabolomics for immediate analysis of resection margins during breast cancer surgery. PLoS One. 2013 Apr;8(4): e61578.

31 Turkoglu O, Zeb A, Graham S, Szyperski T, Szender JB, Odunsi K, et al. Metabolomics of biomarker discovery in ovarian cancer: a systematic review of the current literature. Metabolomics. 2016 Apr;12(4):60.

32 Yu L, Li K, Zhang X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. Oncotarget. 2017 Dec;8(70):115774–86.

33 Troisi J, Sarno L, Landolfi A, Scala G, Martinelli P, Venturella R, et al. Metabolomic signature of endometrial cancer. J Proteome Res. 2018 Feb;17(2):804–12.

34 Inglese P, McKenzie JS, Mroz A, Kinross J, Veselkov K, Holmes E, et al. Deep learning and 3D-DESI imaging reveal the hidden metabolic heterogeneity of cancer. Chem Sci. 2017 May;8(5):3500–11.

35 Balog J, Szaniszlo T, Schaefer KC, Denes J, Lopata A, Godorhazy L, et al. Identification of biological tissues by rapid evaporative ionization mass spectrometry. Anal Chem. 2010 Sep;82(17):7343–50.

36 Alexander J, Gildea L, Balog J, Speller A, McKenzie J, Muirhead L, et al. A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. Surg Endosc. 2017;31(3):1361–70.

37 Balog J, Sasi-Szabó L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Transl Med. 2013 Jul;5(194):194ra93.

38 Glunde K, Penet MF, Jiang L, Jacobs MA, Bhujwalla ZM. Choline metabolism-based molecular diagnosis of cancer: an update. Expert Rev Mol Diagn. 2015 Jun;15(6):735–47.

39 Oakman C, Tenori L, Claudino WM, Cappadona S, Nepi S, Battaglia A, et al. Identification of a serum-detectable metabolomic fingerprint potentially correlated with the presence of micrometastatic disease in early breast cancer patients at varying risks of disease relapse by traditional prognostic methods. Ann Oncol. 2011 Jun;22(6):1295–301.

40 Tenori L, Oakman C, Claudino WM, Bernini P, Cappadona S, Nepi S, et al. Exploration of serum metabolomic profiles and outcomes in women with metastatic breast cancer: a pilot study. Mol Oncol. 2012 Aug;6(4):437–44.

41 Mohan V, Radha V. Precision diabetes is slowly becoming a reality. Med Princ Pract. 2019 Apr;28(1):1–9.

42 Wang TJ, Ngo D, Psychogios N, Dejam A, Larson MG, Vasan RS, et al. 2-Aminoadipic acid is a biomarker for diabetes risk. J Clin Invest. 2013 Oct;123(10):4309–17.

43 Gupta L, Ahmed S, Jain A, Misra R. Emerging role of metabolomics in rheumatology. Int J Rheum Dis. 2018 Aug;21(8):1468–77.

44 Ouyang X, Dai Y, Wen JL, Wang LX. ¹H NMR-based metabolomic study of metabolic profiling for systemic lupus erythematosus. Lupus. 2011 Nov;20(13):1411–20.

45 Kim S, Hwang J, Xuan J, Jung YH, Cha HS, Kim KH. Global metabolite profiling of synovial fluid for the specific diagnosis of rheumatoid arthritis from other inflammatory arthritis. PLoS One. 2014 Jun;9(6):e97501.

46 Adams SB, Setton LA, Kensicki E, Bolognesi MP, Toth AP, Nettles DL. Global metabolic profiling of human osteoarthritic synovium. Osteoarthr Cartil. 2012 Jan;20(1):64–7.

47 Bonomo R, Cavaletti G, Skene DJ. Metabolomics markers in neurology: current knowledge and future perspectives for therapeutic targeting. Expert Rev Neurother. 2020 Jul; 20(7):725–38.

48 van der Lee SJ, Teunissen CE, Pool R, Shipley MJ, Teumer A, Chouraki V, et al. Circulating metabolites and general cognitive ability and dementia: evidence from 11 cohort studies. Alzheimers Dement. 2018 Jun;14(6):707–22.

49 Chang KH, Cheng ML, Tang HY, Huang CY, Wu YR, Chen CM. Alternations of metabolic profile and kynurenine metabolism in the plasma of parkinson's disease. Mol Neurobiol. 2018 Aug;55(8):6319–28.

50 Han W, Sapkota S, Camicioli R, Dixon RA, Li L. Profiling novel metabolic biomarkers for Parkinson's disease using in-depth metabolomic analysis. Mov Disord. 2017 Dec;32:1720–8.

51 Villoslada P, Alonso C, Agirrezabal I, Kotelnikova E, Zubizarreta I, Pulido-Valdeolivas I, et al. Metabolomic signatures associated with disease severity in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm. 2017 Mar; 4(2):e321.

52 Orešič M, Posti JP, Kamstrup-Nielsen MH, Takala RSK, Lingsma HF, Mattila I, et al. Human serum metabolites associate with severity and patient outcomes in traumatic brain injury. EBioMedicine. 2016 Oct;12:118–26.

53 McClay JL, Adkins DE, Isern NG, O'Connell TM, Wooten JB, Zedler BK, et al. (1)H nuclear magnetic resonance metabolomics analysis identifies novel urinary biomarkers for lung function. J Proteome Res. 2010 Jun;9(6): 3083–90.

54 Licari A, Castagnoli R, Brambilla I, Marseglia A, Tosca MA, Marseglia GL, et al. Asthma endotyping and biomarkers in childhood asthma. Pediatr Allergy Immunol Pulmonol. 2018 Jun;31(2):44–55.

55 Dawiskiba T, Deja S, Mulak A, Ząbek A, Jawień E, Pawełka D, et al. Serum and urine metabolomic fingerprinting in diagnostics of inflammatory bowel diseases. World J Gastroenterol. 2014 Jan;20(1):163–74.

56 Williams HR, Willsmore JD, Cox IJ, Walker DG, Cobbold JF, Taylor-Robinson SD, et al. Serum metabolic profiling in inflammatory bowel disease. Dig Dis Sci. 2012 Aug;57(8): 2157–65.

57 Ahmed I, Roy BC, Khan SA, Septer S, Umar S. Microbiome, metabolome and inflammatory bowel disease. Microorganisms. 2016 Jun;4(2):20.

58 Zhang A, Sun H, Han Y, Yan G, Wang X. Urinary metabolic biomarker and pathway study of hepatitis B virus infected patients based on UPLC-MS system. PLoS One. 2013 May;8(5): e64381.

59 Iida M, Harada S, Takebayashi T. Application of metabolomics to epidemiological studies of atherosclerosis and cardiovascular disease. J Atheroscler Thromb. 2019;26(9):747–57.

60 Wurtz P, Raiko JR, Magnussen CG, Soininen P, Kangas AJ, Tynkkynen T, et al. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. Eur Heart J. 2012 Sep;33: 2307–16.

61 Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, et al. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. Nat Med. 2002 Dec;8(12):1439–44.

62 Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, Wiegand R, et al. Metabolomic identification of novel biomarkers of myocardial ischemia. Circulation. 2005 Dec;112(25): 3868–75.

63 Vallejo M, García A, Tuñón J, García-Martínez D, Angulo S, Martin-Ventura JL, et al. Plasma fingerprinting with GC-MS in acute coronary syndrome. Anal Bioanal Chem. 2009 Jul;394(6):1517–24.

64 Dunn WB, Broadhurst DI, Deepak SM, Buch MH, Mcdowell G, Spasic I, et al. Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. 2007. Vol. 3. p. 413–26.

65 Ismail IT, Showalter MR, Fiehn O. Inborn errors of metabolism in the era of untargeted metabolomics and lipidomics. Metabolites. 2019 Oct;9(10):242.

66 Fujii T, Khawaja MR, DiNardo CD, Atkins JT, Janku F. Targeting isocitrate dehydrogenase (IDH) in cancer. Discov Med. 2016;21(117): 373–80.

67 National Phenome Centre [cited 2020 Nov 15]. Available from: https://phenomecentre. org.

68 Siskos AP, Jain P, Römisch-Margl W, Bennett M, Achaintre D, Asad Y, et al. Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. Anal Chem. 2017 Jan;89(1):656–65.

69 Jourdan C, Petersen AK, Gieger C, Döring A, Illig T, Wang-Sattler R, et al. Body fat free mass is associated with the serum metabolite profile in a population-based study. PLoS One. 2012 Jun;7(6):e40009.

70 Lotta LA, Scott RA, Sharp SJ, Burgess S, Luan J, Tillin T, et al. Genetic predisposition to an impaired metabolism of the branched-chain amino acids and risk of type 2 diabetes: a mendelian randomisation analysis. PLoS Med. 2016 Nov;13(11):e1002179.

71 Draisma HHM, Pool R, Kobl M, Jansen R, Petersen AK, Vaarhorst AAM, et al. Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. Nat Commun. 2015 Dec;6: 7208.

72 Katajamaa M, Oresic M. Data processing for mass spectrometry-based metabolomics. J Chromatogr A. 2007 Jul;1158(1–2):318–28.

73 Wu Y, Li L. Sample normalization methods in quantitative metabolomics. J Chromatogr A. 2016 Jan;1430:80–95.

74 Takis PG, Schäfer H, Spraul M, Luchinat C. Deconvoluting interrelationships between concentrations and chemical shifts in urine provides a powerful analysis tool. Nat Commun. 2017 Dec;8(1):1662.

75 Ren S, Hinzman AA, Kang EL, Szczesniak RD, Lu LJ. Computational and statistical analysis of metabolomics data. Metabolomics. 2015 Dec;11(6):1492–513.

76 Alakwaa FM, Chaudhary K, Garmire LX. Deep learning accurately predicts estrogen receptor status in breast cancer metabolomics data. J Proteome Res. 2018 Jan;17(1):337–47.

77 About us – PhenoMeNal [Internet]. [cited 2018 Aug 10]. Available from: http://phenomenal-h2020.eu/home/about/.

78 Alden N, Krishnan S, Porokhin V, Raju R, McElearney K, Gilbert A, et al. Biologically consistent annotation of metabolomics data. Anal Chem. 2017 Dec;89(24):13097–104.