Conference on ‘Over- and undernutrition: challenges and approaches’

Symposium 2: Modern approaches to nutritional research challenges
Targeted and non-targeted approaches for metabolite profiling in nutritional research

John K. Lodge
Cranfield Health, Vincent Building, Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

The present report discusses targeted and non-targeted approaches to monitor single nutrients and global metabolite profiles in nutritional research. Non-targeted approaches such as metabolomics allow for the global description of metabolites in a biological sample and combine an analytical platform with multivariate data analysis to visualise patterns between sample groups. In nutritional research metabolomics has generated much interest as it has the potential to identify changes to metabolic pathways induced by diet or single nutrients, to explore relationships between diet and disease and to discover biomarkers of diet and disease. Although still in its infancy, a number of studies applying this technology have been performed; for example, the first study in 2003 investigated isoflavone metabolism in females, while the most recent study has demonstrated changes to various metabolic pathways during a glucose tolerance test. As a relatively new technology metabolomics is faced with a number of limitations and challenges including the standardisation of study design and methodology and the need for careful consideration of data analysis, interpretation and identification. Targeted approaches are used to monitor single or multiple nutrient and/or metabolite status to obtain information on concentration, absorption, distribution, metabolism and elimination. Such applications are currently widespread in nutritional research and one example, using stable isotopes to monitor nutrient status, is discussed in more detail. These applications represent innovative approaches in nutritional research to investigate the role of both single nutrients and diet in health and disease.

Metabolomics: Metabolite profiling: Stable isotopes: Nutrient physiology: Diet

Nutritional change is accepted as a cornerstone of public health strategy in reducing the risk of many of the chronic diseases that face Western society such as obesity, heart disease and cancer. Nutritional status modulates virtually all body processes, including various ongoing metabolic processes. Currently, however, there is little understanding of the individual response to nutritional intervention and the contribution that specific metabolites and metabolic pathways make in physiological processes related to health and disease. In trying to elucidate the aetiology of multifactorial disease it is well accepted that genetic, environmental and lifestyle factors are important modulators. Nutrition also plays an important role as it supplies the metabolome (all metabolites within a biological sample). During disease development early events lead to a change to nutrient physiology and/or changes to metabolic pathways, which will influence the whole metabolome and may eventually lead to a metabolic disturbance and the origins of a disease phenotype. Using traditional biomarkers and conventional research tools it is not possible to identify changes to nutrients and the metabolome at these early stages when nutritional intervention could potentially reduce disease risk back to a healthy phenotype. However, there are now a number of approaches available to monitor both single-nutrient physiology and global metabolite profiles, and these approaches have shown considerable potential in the study of how diet and the metabolome interact. Such approaches can be defined as...
as either targeted or non-targeted depending on the overall experimental design. Targeted approaches are usually hypothesis testing, as it presumes the target of the investigation is known (nutrient or metabolite) and the aim of the study is to measure its status. A non-targeted approach is usually hypothesis generating, as typically it is designed to provide a more global description of all metabolites present in a biological fluid; thus, the target is actually the whole metabolome and not a specific species. The aim then is to identify species of interest that can be targeted for further investigation and as such both approaches are linked and complementary to each other.

Non-targeted approaches: metabolomics

Metabolomics is a relatively novel technology in nutritional research and offers considerable potential as a result of the wealth of information that potentially can be obtained. There is increasing interest in the use of metabolomics as a tool in understanding the effects of nutritional intervention on the individual (1,3–10). A major challenge is to understand the relationship between nutrition and chronic disease, and a metabolomics-based approach could be used analytically to predict susceptibility to metabolic diseases and elucidate molecular mechanisms that can explain the beneficial effects of dietary interventions. Furthermore, metabolomics is thought to be a promising tool for biomarker discovery (2), with the potential capacity to identify biomarkers of early-stage disease and biomarkers of dietary patterns or individual nutrients. Metabolomics is also seen as providing critical information in the quest to transform current population-based dietary guidelines into future personalised dietary recommendations to help individuals maintain a healthier status and prevent chronic diseases (11).

The technology can be summarised as a global analysis or characterisation of all metabolites measured in a biological sample (metabolome). The metabolome is a complex mixture of thousands of compounds that arise from intracellular and extracellular metabolic pathways and external sources such as food. As a result of the complexity of biological fluids, with the large number of metabolites present, data from biological mixtures are highly complex and generally are processed using multivariate statistical analysis including pattern recognition. These chemometric measurements generate metabolic profiles or patterns of metabolites from which it is then possible to detect changes to the normal phenotype induced by, for example in relation to nutritional studies, diet and metabolic disease. Although this technology has been successfully used in nutritional studies, there are still a number of considerations for its use.

Historically, NMR has been the platform of choice and still represents the most common method for sample analysis in metabolomics. More recently, MS technology has become popular, with both liquid chromatography (LC) and GC being successfully used. NMR requires no sample pretreatment and so the spectrum contains information on both small-molecular-weight compounds and proteins; however, sensitivity is relatively low. MS is highly sensitive; however, it requires sample pretreatment in the removal of proteins and hydrophobic compounds, thus limiting potential information. Furthermore, the method of protein precipitation or sample treatment also has its own analytical considerations and influences the metabolite profiles obtained. As pretreatment and a separation step are needed, MS is not as reproducible as NMR. A general consideration for metabolomic investigations is that no single platform will provide information on all metabolites present and a combination of NMR, LC–MS and GC–MS would be required.

Experimental and analytical variation

Analytical design and variation in relation to metabolomics has been extensively investigated (12–20). As each platform has its own analytical requirements, variation arises from the methodology used in the study and from the instrument itself. Sample collection provides its own issues. Although blood concentrations of metabolites are generally under some sort of metabolic regulation, they are dynamic and responsive to acute changes to the diet. Metabolite profiles have been shown to be influenced by the blood collection tube (21), clotting times and clotting procedure (15). Urine represents an ‘end point’ of metabolism, and as such is not under any metabolic regulatory control as with plasma or serum, and is often the sample of choice in metabolomics investigations (it also requires limited pretreatment compared with blood). However, urinary metabolite profiles are heavily influenced by diurnal variation (12) and the use of sample preservatives (22), especially if 24 h aliquots are being obtained. The extent to which spot urines or first or second void urines are representative has not yet been established. The use of faeces (either lyophilised or faecal water) is relatively new but can be useful for gut microbiota investigations. Sample storage is important; no differences have been found in urinary or plasma metabolite profiles when samples are stored for 0 or 24 h at 4°C (23), but for longer storage times sub-zero conditions (<−20°C) should be used (14,22). Freeze–thaw cycles have not been shown to influence metabolite profiles using LC–MS (14).

While NMR requires no sample pretreatment, MS requires removal of proteins and hydrophobic lipids. Various methods for protein precipitation have been compared in terms of the number of features obtained (13,24,25), with similar observations finding methanol treatment to be the most optimal for plasma samples; however, the variation in profiles obtained from each pretreatment method is extensive. Attempts have been made to further develop and optimise the methodology for LC–MS-based metabolomic studies (13,16,17,20,29). It is clear that the analytical conditions associated with MS need to be further optimised and standardised as sample pretreatment, run time, type of
column, temperature and LC gradient all influence the profiles obtained\(^{(13,16,17,20,26)}\).

**Biological variation**

There is a complex relationship between genotype, metabolic phenotype, diet, lifestyle and environment and the nutritional needs of an individual\(^{(11)}\). Human tissues and cells respond to the delivery of nutrients by adapting their metabolic processes through the regulation of gene transcription, protein levels and enzyme activity. The end point of these processes is that the levels of metabolites change; however, as a result of large inter-individual variations in genotype and phenotype human subjects differ widely in their response to nutrients and diet overall\(^{(11)}\) and thus there is large inter-individual variability in metabolite profiles.

The influence of biological variation has been the subject of some study\(^{(18,19,27,28)}\). Considerable biological variation in metabolite profiles exists as a result of gender, age and BMI\(^{(27,28)}\) and it has been possible to classify samples based on these characteristics. Increased plasma choline, LDL, HDL and unsaturated fatty acids have been found in women and increased VLDL, creatine and isoleucine in men as discriminators of gender\(^{(27)}\). Comparing age-groups, in women plasma valine, isoleucine, alanine and tyrosine are higher in the older population (>46 years v. 18–29 years) while in men plasma lipoproteins and unsaturated lipids are higher in the older population\(^{(27)}\). Discriminators based on BMI include choline and citrate, which are higher in a lean population (BMI <21 kg/m\(^2\)), while tyrosine, isoleucine and glycoproteins are higher in an obese population (BMI >25 kg/m\(^2\))\(^{(27)}\). Thus, in terms of study design there needs to be a careful consideration of the population used, with limited ranges of age and BMI and also taking into account blood pressure\(^{(29)}\), standard clinical chemistry and prescreening questionnaires\(^{(10)}\).

Dietary habits heavily influence metabolite profiles, as foods contain thousands of compounds that supply the metabolome. Populations can be classified according to various dietary factors\(^{(29–31)}\) and certain metabolites have been shown to be related to specific dietary patterns\(^{(30,32,33)}\). For example, trimethylamine N-oxide is related to acute intake of fish\(^{(30,31)}\) while taurine and creatine are increased during high meat intakes\(^{(30,32)}\). Careful dietary standardisation is therefore required before sample collection\(^{(10,34)}\), and indeed this approach has been shown to decrease inter-individual variation in urinary metabolite profiles\(^{(34)}\).

There is a complex interaction between gut microflora and host metabolism\(^{(35,36)}\). As diet modulates microbial physiology, products of microbial metabolism can enter the host metabolome causing substantial variability in profiles. Although urinary metabolite profiles have often been used as a natural point to observe such interactions\(^{(37)}\), recent evidence suggests that there is a large contribution of microbial by-products to the plasma metabolome\(^{(38)}\). A comparison of conventional (with microflora) and germ-free (without microflora) mice has found numerous signals exclusive to the presence of microflora, with approximately a 10% difference in the metabolite profiles between normal and germ-free animals\(^{(36)}\). As well as products of microbial metabolism of amino acids and other compounds, the study has also found products of host phase II metabolism raised in the presence of microflora, demonstrating a direct effect on host drug metabolism capacity\(^{(36)}\). Faeces are an obvious sample to study microflora interactions, as profiles contain a plethora of compounds related to microflora metabolism such as SCFA, organic acids and amino acid derivatives\(^{(38)}\). As the complement of microflora differs between individuals there is considerable inter- and intra-individual variation in profiles\(^{(38)}\), but interestingly this variation has been found to be a result of differences in concentration rather than composition\(^{(38)}\). Nevertheless, as microflora seem to contribute to both blood and urinary profiles they are an important contributor to biological variation.

**Data analysis**

The data produced by metabolomics experiments are highly complex; for example, >2000 features differing in intensity by orders of magnitude are typically represented in LC–MS analysis of plasma or urine. As the technology aims to characterise changes to the metabolome in discreet sets of samples powerful multivariate statistical analysis techniques are required to process the data for sample classification and clustering or discrimination. These techniques essentially are data compression tools; by reducing the dimensionality of the data it becomes easier to visualise graphically so that patterns between groups of data (samples) can be revealed. Both unsupervised and supervised techniques are commonly used. Principal component analysis is an unsupervised method for sample classification. The ‘principal components’ represent variance between samples and so samples that are similar in their metabolite profile will cluster together in a scores plot. An important aspect of this method is that it does not require information about sample classification and as such is a good starting point to visualise the data. However, principal component analysis is often unable to classify similar groups, which is typically found in nutritional studies, and so other techniques are often used. Supervised methods, such as partial least squares discriminant analysis, include information on sample classification, making it more likely that patterns of change between groups are visualised. However, the downside is that it can lead to overanalysis, e.g. partial least squares discriminant analysis will separate groups from random data\(^{(39)}\), and as such these methods require appropriate model validation steps and adequate training and predictive sample sets, usually with independent data.

Clustering methods, as described earlier, only visualise discriminated sample groups demonstrating whether the metabolite profiles are different between sample groups. This feature is important to establish, but of greater interest is identifying what metabolites are responsible for the variation between sample groups, as this information can be used further downstream in pathway analysis and biomarker identification. Loadings plots are an extension to score plots in which those features responsible for group variation are visualised. Although this approach provides
important information on the actual species (e.g. mass and retention time in LC–MS experiments), further investigative work using metabolite databases (e.g. the human metabolome database (40), METLIN (41)) and confirmatory analysis using standards is required to be fully confident of the assignment.

It should be noted that data analysis steps require careful consideration and interpretation in order to be confident of obtaining meaningful data.

### Metabolomics in nutritional research

A number of studies applying metabolomics to nutritional research have been performed (Table 1). A quick comparison of these studies shows that various analytical platforms and biological samples have been used. Generally, NMR has been the most common platform but MS has been gaining popularity over recent years, perhaps because of greater availability and improvement in analytical technology and methodology. Urine is still the commonest biological sample used, especially when NMR is used as the platform, but a number of studies do profile both urine and blood. It is interesting to note that the first ‘real’ nutritional metabolomics study was performed in 2003 (42), which highlights the novelty of the technology. These early studies have shown changes to both the plasma (42) and later the urinary (43) metabolome with standardised meals, which perhaps contributed to identifying subtle changes to the metabolome in a relatively small subject population (five females). A number of studies of differing design have been performed on tea metabolism (44–46). While both black tea and chamomile tea...
confirm hippuric acid as a major species responsible for variation in urinary metabolites after consumption of certain foods (44, 45). Green tea produces additional effects on gut flora fermentation induced by a wine extract (38). More recently, metabolomics studies have investigated carbohydrate intake either during an oral glucose tolerance test (50, 51) or following consumption of energy drinks (52). The former studies have identified various pathways stimulated during the insulin response to glucose, some of which have not previously been reported (50, 51).

Metabolomics has been used to investigate the influence of vitamin E supplementation (400 mg/d, 4 weeks) on plasma metabolite profiles of twelve healthy men using LC–MS (JK Lodge and MC Wong, unpublished results). While principal component analysis has failed to discriminate discrimination of samples, a validated partial least squares discriminant analysis model has been able to discriminate samples pre- and post supplementation. Further investigation of loadings plots has revealed several features for further investigation that have been confirmed to be significantly different pre- and post supplementation (P < 0.05). Preliminary identification of these features has revealed a number of lyso-phospholipids that are significantly increased following vitamin E supplementation (P < 0.05). Although the confirmation of these species is still ongoing, this finding highlights the potential of the technology in being able to identify changes to the metabolome and the identification of novel areas of research that may go unnoticed using conventional techniques.

**Future perspectives for nutritional metabolomics**

Although the potential of metabolomics is without question, the technology is still faced with a number of limitations and challenges. There is no single platform that can analyse all metabolites in a biological sample and so the limitations of each technology have to be taken into account in experimental design. Obviously, a combination of NMR, LC–MS and GC–MS would be ideal to provide more holistic measure of the metabolome, but this approach itself brings important cost implications and an increase in data analysis and complexity. There is currently no standardisation in terms of the methodology associated with sample collection, sample pretreatment and analytical conditions. However, there have been recommendations published to standardise metabolomic investigations and their reporting (53, 54), which will hopefully influence future investigations. A major challenge is the influence of both analytical and biological variability. Analytical variability should improve with the advancement of technology and standardisation of methodology, while the influence of biological variability requires careful understanding of the influence of a variety of factors such as gender, BMI, age, recent dietary intake, gut microflora and general health, as these factors have all been shown to influence the metabolome. Furthermore, the influence of lifestyle, genetic and environmental factors, which are important in the aetiology of chronic disease, need to be further explored. Such considerations have implications for study design to create a relatively homogenous study population to reduce the influence of biological variation.

Another major challenge faced is the identification of metabolites. Although a number of databases are available for most platforms (and should be used in conjunction with each other), the whole metabolome is not yet fully characterised there will always be a problem with the identification of ‘unknowns’. Although useful structural information can be gained from further MS experiments (e.g. fragmentation), identification ultimately requires the use of standards, and if not commercially available the assistance of synthetic chemists.

The future potential of this technology for nutritional research can be envisaged at various levels. It will be possible to integrate metabolomics with other ‘omic’ technologies to provide a more complete picture of the metabolic phenotype. The identification of those metabolites influenced by certain nutrients and/or dietary patterns can be useful for predictive analysis with the possibility of using fluxomics (dynamic changes of metabolites within a cell over time) to investigate the flow of metabolites through metabolic pathways, especially now, as human metabolic networks have been constructed (55). A well-established role for nutritional metabolomics has been identified in the development of the nutritional phenotype and for future personalised nutritional assessment (1, 3, 4, 5). It is thought that metabolomics can be used to predict the response of the metabolome to nutrient or dietary intervention and to predict dietary change required to modify the metabolome to a ‘healthy’ phenotype. Although personalised nutrition is a long way off, metabolomics does offer the possibility of characterising the complex interactions between diet and the metabolome.

**Targeted approaches**

Targeted approaches are more hypothesis testing and include those in which the nutrient(s) or metabolite(s) of interest are known and information on concentration,
bioavailability, turnover or metabolism are required. These approaches can also be seen as a natural progression from non-targeted approaches that are used in the first instance to identify novel targets for such further studies. Targeted analysis can cover either single nutrients or metabolites or an analysis of metabolome subsets including lipids, amino acids and central metabolism. Indeed, the profiling of lipids has developed into its own field of metabolomics and recent studies in this field, which include the analysis of eicosanoid metabolism and signalling, identifying lipids linked to inflammation following fish oil intervention, phospholipid metabolism, monitoring of lipid fluxes using stable isotopes and evaluating differential effects of statins, demonstrate how this particular field is rapidly progressing.

Simply measuring concentration alone provides only limited information; therefore, in order to gain further understanding of the physiology of a nutrient or metabolite other approaches are required. A commonly-used method is to employ a stable-isotope analogue of the target, which as it should behave identically to the parent compound allows the monitoring of that specific form in the body following ingestion, providing information on absorption, bioavailability, distribution, turnover, metabolism and elimination. The use of stable isotopes in nutritional research is extensive and a review of potential applications is beyond the scope of the present report, but there are several reviews that provide a starting point.

As an example of the benefits of using stable isotopes, this approach has been employed to monitor vitamin E status in various populations using \(^{2}\text{H}-\text{labelled } \alpha\)-tocopheryl acetate. Vitamin E transport and delivery within the body is governed by the kinetics of plasma lipoprotein metabolism and much of what is now known about vitamin E transport and distribution has been elucidated using stable-isotope approaches. Intestinal absorption requires the presence of adequate fat, and significant differences have been demonstrated in the plasma uptake of labelled \(\alpha\)-tocopherol between meals containing 2-7 g fat and 17-5 g fat \((P<0.001)\). It has been shown that during the absorptive phase there is no discrimination between vitamin E forms, but a number of biokinetic studies have demonstrated the preferential secretion of \(\alpha\)-tocopherol in VLDL over other vitamin E forms, so that there is a retention of \(\alpha\)-tocopherol in the body. The monitoring of labelled \(\alpha\)-tocopherol in lipoproteins demonstrates that \(\alpha\)-tocopherol transport is dependent on lipoprotein kinetics, with peaks in \(\alpha\)-tocopherol concentration coincident with lipoprotein dynamics. Overall, following plasma uptake of newly-absorbed \(\alpha\)-tocopherol there is a simultaneous decrease in pre-existing or endogenous \(\alpha\)-tocopherol concentration. This rapid turnover of \(\alpha\)-tocopherol in the plasma such that the ‘new’ constantly replaces the ‘old’ is an established phenomenon identified by stable-isotope techniques and also occurs within individual lipoproteins. A similar phenomenon has been found in blood components, with different patterns of uptake in erythrocytes, platelets and lymphocytes representative of simple diffusion or a more-regulated receptor-mediated uptake. Following an understanding of nutrient physiology in a healthy population a comparison can be made with high-risk populations to provide clues as to how nutrient physiology is influenced. Using stable isotopes differences have been found in \(\alpha\)-tocopherol biokinetics between individuals who are normolipidaemic and hyperlipidaemic, smokers and non-smokers and between individuals with the apoE3/3 and apoE3/4 alleles. Overall, these studies highlight that vitamin E homeostasis is ultimately dependent on a combination of factors that control its absorption, transport, distribution and metabolism and that, translated to the field of nutritional research, approaches using stable-isotopes are essential to monitor nutrient status in vivo and their role in the aetiology of disease.

**Conclusions**

With the combined array of targeted and non-targeted approaches it is now possible to not only identify metabolites and nutrients that are influenced by dietary change but also gain a greater understanding of nutrient and metabolite physiology. Both approaches have proved successful in a number of applications and used either alone or in conjunction with one another they represent powerful research methods with the potential of opening novel areas of nutritional research and understanding the role of diet in health and disease.

**Acknowledgements**

The author wishes to thank the British Heart Foundation, BASF AG and the University of Surrey Pump Priming for funding parts of this work. The author declares no conflict of interest.

**References**

Metabolite profiling in nutritional research


70. Traber MG, Burton GW, Ingold KU et al. (1990) RRR- and SRR-α-tocopherols are secreted without discrimination in human chylomicrons, but RRR-α-tocopherol is preferentially secreted in very low density lipoproteins. J Lipid Res 31, 675–685.


