Title: Non-targeted metabolomics in sport and exercise science

Running title: Non-targeted metabolomics in sport and exercise

Authors: Liam M Heaney[1], Kevin Deighton[2] & Toru Suzuki[1]

Affiliations:

[1] Department of Cardiovascular Sciences and NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, LE3 9QP, United Kingdom

[2] Institute for Sport, Physical Activity & Leisure, Leeds Beckett University, LS6 3QS, United Kingdom

Address for correspondence:

Dr Liam Heaney, Department of Cardiovascular Sciences and NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, LE3 9QP, United Kingdom. Email: l.m.heaney@le.ac.uk. Tel: (+44) 0116 258 3041

and

Prof Toru Suzuki, Department of Cardiovascular Sciences and NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, LE3 9QP, United Kingdom. Email: ts263@le.ac.uk. Tel: (+44) 0116 204 4741

Key words: exercise; mass spectrometry; metabolism; metabolomics; NMR

Acknowledgments:

The authors would like to acknowledge the ongoing support from the Practical Research Project for Life-Style related Diseases including Cardiovascular Diseases and Diabetes Mellitus from Japan Agency for Medical Research and Development (AMED), the University of Tokyo, the John and Lucille van Geest Foundation and the National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit.

Disclosure statement:

All authors declare no disclosures nor conflicts of interest.

Ethical statement:

Institutional ethics approval for this review manuscript was confirmed as not being required by the Chair of the University of Leicester Medicine and Biological Sciences Research Ethics Sub-Committee.

Abstract

Metabolomics incorporates the study of metabolites that are produced and released through physiological processes at both the systemic and cellular level. Biological compounds at the metabolite level are of paramount interest in the sport and exercise sciences, although research in this field has rarely been referred to with the global *'omics* terminology. Commonly studied metabolites in exercise science are notably within cellular pathways for ATP production such as glycolysis (e.g. pyruvate and lactate), β-oxidation of free fatty acids (e.g. palmitate) and ketone bodies (e.g. β-hydroxybutyrate). Non-targeted metabolomic technologies are able to simultaneously analyse the large numbers of metabolites present in human biological samples such as plasma, urine and saliva. These analytical technologies predominately employ nuclear magnetic resonance spectroscopy and chromatography coupled to mass spectrometry. Performing experiments based on non-targeted methods allows for systemic metabolite changes to be analysed and compared to a particular physiological state (e.g. pre/post-exercise) and provides an opportunity to prospect for metabolite signatures that offer beneficial information for translation into an exercise science context, for both elite performance and public health monitoring. This narrative review provides an introduction to non-targeted metabolomic technologies and discusses current and potential applications in sport and exercise science.

1. Introduction

The study of human metabolism in response to acute bouts of exercise and chronic exercise training traditionally involves the measurement of selected metabolites, transcription factors and proteins (Egan, Hawley, & Zierath, 2016). This has provided mechanistic insights into the phenotypic changes observed with exercise training and has enabled the development of targeted training and nutritional strategies to maximise adaptations for health and performance, for example, the timing of carbohydrate intake to maximise activation of key cell signalling proteins (Impey et al., 2016) and the application of short-term exercise training to increase insulin-mediated glucose disposal in obese people with type 2 diabetes (O'Gorman et al., 2006). Advanced investigatory techniques such as metabolomics offer the potential to develop the existing knowledge of metabolites and discover novel markers that can provide important information to exercise scientists.

This narrative review is intended to provide an overview of what metabolomics offers as an experimental method and give a brief explanation of the current analytical and bioinformatic techniques employed for non-targeted metabolomics. Examples of where these techniques have been initially employed into exercise-based experiments are illustrated, and potential future directions that metabolomics methodologies may take in providing additive information into the sport and exercise sciences are discussed.

2. What is metabolomics?

Metabolites are defined as *"*low molecular weight organic and inorganic chemicals which are the reactants, intermediates or products of enzyme-mediated biochemical reactions*"* (Dunn, Broadhurst, Atherton, Goodacre, & Griffin, 2011, p. 387), with the term 'metabolome' used to describe the complete array of these metabolites found secreted by a living cell/organism (Nicholson, & Wilson, 2003). The study of the metabolome, metabonomics, was

first described as *"*the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification*"* (Nicholson, Lindon, & Holmes, 1999, p. 1181). In addition to metabonomics where measurements are made of metabolites that are present from metabolic responses in multicellular systems, for example metabolites released by commensal and symbiotic organisms, metabolomics involves measurements of metabolites directly connected to genetic, metabolic and protein-driven processes (Nicholson, & Wilson, 2003). These metabolites may be present in tissues, bio fluids (e.g. plasma/serum and urine) and/or in volatilised form in exhaled breath gases. The premise of metabolomics is that small molecule metabolites measured in a biological medium report on the physiological state, or changes in response to an intervention, occurring within that organism.

Like other methodologies carrying the *'omics* tag, for example genomics and proteomics, metabolomics can employ both targeted and non-targeted strategies. Targeted analyses can be applied to identify metabolites related to specific biological processes/pathways that modulate a biological function of interest (Griffiths et al., 2010). Targeted assays have been developed for commercial distribution, for example the AbsoluteIDQ test kit (Biocrates Life Sciences, Innsbruck, Austria) for the measurement of amino acids, glycerophospholipids and acylcarnitine molecules known to be involved in cell energy metabolism. In contrast, non-targeted metabolomics (also known as unbiased, global or discovery metabolomics) employs a wide-scope analytical collection (Fuhrer, & Zamboni, 2015) and measurement technique whereby all detectable metabolites are (semi-)quantitated, collated and prospected for (bio)markers that are indicative of predefined conditions (e.g. metabolites that discriminate between diagnosis of diseased and healthy individuals). These assigned biomarkers may reflect causative and/or associative relationships with a physiological

state of interest, allowing them to aid in the examination of current health or be applied for diagnostic, prognostic and therapeutic purposes.

Through recent technological and methodological advances in the field of metabolomics, it has been possible to characterise, quantitate and identify an increasing number of analytes. Development and publication of open-source searchable metabolite databases have been possible through the increased availability of metabolite properties (e.g. molecular mass and analytical assay characteristics – discussed in more detail later). Perhaps the most relevant to human investigations is the Human Metabolome Database (HMDB) which was first published in 2007 (Wishart et al., 2007) and has since been updated, with the latest edition released in 2013 (Wishart et al., 2013). HMDB describes over 40,000 metabolites and each is represented with a 'metabocard' that details associated chemical, biochemical, clinical and enzymatic data, with links to website resources of additional information (*The Human Metabolome Database*, 2016). Databases such as HMDB have greatly assisted in the identification and understanding of physiological relevance for metabolites that have been discovered through non-targeted strategies.

3. Commonly employed analytical techniques

Nuclear magnetic resonance spectroscopy (NMR) is an analytical technique often used in organic chemistry to confirm the structure of a synthesised compound. The information provided by NMR can also be applied to biological samples and allows for structural identification of the molecules present.

NMR uses a high-powered magnet to induce a magnetic field that causes some atomic nuclei to spin. For ^{1}H NMR, commonly employed in metabolomics studies (Want et al., 2010), the induced magnetic field causes the protons to align in an orientation corresponding to low or high energy; these are known as α and β states, respectively. The sample is then subjected

to applied radio waves which cause those nuclei in the α state to shift to the β state. Once the applied energy is removed, the nuclei return to their original energy state and an alteration in magnetic field, known as resonance, can be measured and interpreted as peaks on NMR spectra. The magnetic field expressed on the nuclei is influenced by both the externally applied field and the magnetic effect of localised nuclei/electrons, causing changes in the resonance frequency in comparison to that seen from a singular atom. These changes in resonance are compared to a standard, which is defined as zero, and the difference observed is known as the chemical shift. These chemical shifts are characteristic for certain molecular structures and therefore can be used to identify molecules, or parts of molecular structure, present in the sample which aids in identification of the measured metabolites. As not all isotopes exhibit a magnetic spin (for example, ^{12}C is not magnetic, where ^{13}C is), it means that not all molecules can be studied using NMR and therefore the technique is limited in its application to metabolite measurement on a global scale.

Mass spectrometry (MS) is commonly employed for the analysis of metabolites in a global, non-targeted way (Dunn et al., 2011). MS offers highly sensitive, specific, accurate, rapid and robust analytical assays that are perfectly adapted to measure multiple metabolites using singular preparation and analysis methods (Dunn et al., 2011). As modern mass spectrometers are able to utilise large analytical mass scan windows (often in excess of 1000 Da), there is an innate ability to measure all detectable small molecule metabolites present within a sample. Rapid and wide-scan capabilities allow for large databases of study-specific metabolites to be produced and subsequently implemented in multivariate statistical models to discover those which are highly associated with the target state.

One of the most difficult aspects of non-targeted metabolomics is the confident identification of a measured metabolite of interest (Dunn, Broadhurst, Atherton et al., 2011). The application of modern high-resolution mass analysers, such as time-of-flight and the

orbitrap, has improved analyte identification with the ability to measure mass-to-charge ratios (*m/z*) with an error of less than 10 parts per million (<0.001%). Highly accurate measured masses can then be compared to open-source databases and consequently given a tentative identification for a molecule's name/structure. High-resolution measurements provide information that greatly reduces the complexity and, therefore, increases the success of positive identification over the use of nominal mass (i.e. to the nearest Da) systems such as the quadrupole and ion-trap mass analysers.

In addition to the *m/z* measured of the intact molecule, a collision energy can be applied causing the metabolite to fragment. The fragments formed and their ratios to the precursor molecule are reproducible providing the collision energy value is maintained across analytical runs. Known as tandem MS or MS/MS, this process of fragmentation allows for additional validation of metabolite identification through comparison of mass spectra with and without a collision energy applied. MS/MS is particularly useful where two or more isomeric molecules, i.e. possess the same empirical formula and therefore identical *m/z* value, display alternative fragmentation properties and therefore offers the capability of identifying a particular metabolite isomer.

Metabolite identification can be further enhanced through the coupling of chromatographic techniques such as liquid and gas chromatography (LC, GC) to MS (Creek et al., 2011). LC is applied to non-volatile metabolites and is often used for blood and urine based analyses, whilst GC requires the metabolite to enter the analytical system in the gas phase and therefore is most suitable for volatile metabolites, liquid headspace and breath gas analyses.

Chromatography uses the affinity of molecules to a stationary phase for deconvolution of complicated matrices that contain many hundreds of metabolites (e.g. plasma/serum). As the metabolites pass along the chromatographic column their varying affinities to the stationary phase cause them to exhibit different times between entry and exit of the analytical column.

These properties cause metabolites to be separated and introduced into the mass spectrometer at intervals, thereby reducing the complexity of each MS scan. Decreased analytical complexity, through separation of molecules, improves metabolite identification through the reproducibility of metabolite retention times when chromatographic conditions are maintained across analytical runs. GC also offers the use of Kovát's retention index (Kováts, 1958), where a series of homologous alkanes provide comparative retention time data across different chromatographic conditions. These retention indices can be compared to published values (e.g. the NIST mass spectral library) for more confident identification of analytes.

The combination of chromatography with MS allows for the analysis of known standard reference compounds and comparison of retention time, mass spectra and MS/MS spectra for definitive identification of metabolites. Although GC-MS and LC-MS are not currently employed for exercise and sport-based situations in training or performance contexts, they are commonly used in anti-doping strategies for testing of athlete samples for banned substances (Thevis, Kuuranne, Walpurgis, Geyer, & Schanzer, 2016).

A basic workflow for a non-targeted metabolomic experiment can be seen in Figure 1, and the advantages and limitations of using NMR and MS as technologies for non-targeted metabolomics studies are detailed in Table 1.

4. Commonly used bioinformatics for metabolomics

Non-targeted experiments produce large numbers of measured variables and therefore the way in which data are processed is an important factor for isolating meaningful associations with a variable/state of interest. For investigating changes in metabolites between predefined states, principal components analysis (PCA) and [orthogonal] partial least squares-discriminant analysis ([O]PLS-DA) are commonly used statistical techniques. PCA and OPLS-DA allow visualisation of multi-dimensional relationships of measured variables (i.e. metabolites) to

predefined states (e.g. healthy and diseased) (Worley & Powers, 2013). PCA is an unsupervised method which projects data points onto a plot to visualise their distribution dependent on metabolite correlations that show the largest deviations across the dataset. PCA can be utilised to examine trends within a dataset without the force fitting for differences between pre-defined groups. This is particularly important for the analysis of quality control (QC) samples injected at regular intervals throughout experimental periods, with a tight cluster of these samples present within the PCA plot indicative of good analytical reproducibility and, therefore, low study bias (Figure 2). Supervised, or discriminant, analyses are methods that isolate the metabolites with the largest variation between predefined groups, allowing individual metabolites to be isolated by sensitivity and selectivity of group prediction. Methods such as OPLS-DA allow a refocus of analysis to understand differences related specifically to the experimental question, e.g. pre- versus post-exercise state, reducing the impact of systemic variation that may influence PCA models (Wiklund et al., 2008). In addition, OPLS-DA models produce a corresponding S-plot which visualises the covariance and correlation between metabolites and the computed model (Wiklund et al., 2008), indicating potential biomarkers to be isolated to be taken forward for more targeted statistical testing (Figure 2). As these methods use large numbers of variables to compute models, pre-treatment of the data can prove influential on corresponding results. Techniques such as normalisation, missing value imputation, transformation and scaling must be used with care to ensure the validity of results obtained (Di Guida et al., 2016). Metabolite relationships with continuous variables, for example VO_{2max}, are predominantly investigated by correlation analyses for single metabolites, or multiple linear regression for combining several metabolites.

5. Current investigations in sport and exercise science

In the context of sport and exercise science, NMR and MS-based techniques have allowed researchers to explore human metabolism in response to acute and chronic exercise. This includes the use of NMR to measure substrate use (Gonzalez et al., 2015) and metabolic perturbations during exercise (Jones, Wilkerson, & Fulford, 2008). MS has also been extensively used in combination with stable isotope tracers to assess substrate use during exercise (e.g. O'Hara et al., 2012) and to assess the total energy expenditure of elite athletes (Fudge et al., 2006). Alternatively, non-targeted metabolomics experiments are a contemporary adaptation to research methodologies in sport and exercise science. These unbiased 'discovery' methods have shown utility in medical research with prospective novel biomarkers identified in lung cancer (Mathé et al., 2014) and cardiovascular disease (Wang et al., 2011), amongst other conditions. The use of non-targeted metabolomics in sport and exercise science represents an exciting prospect as a method for identifying novel biomarkers relevant to the health and performance effects of sport and exercise interventions. To date, experiments performed using non-targeted metabolomics in this field have predominantly observed fluctuations in metabolites related to energy production pathways, and have been measured in blood, urine and saliva samples. These experiments are reviewed below, with potential future directions also identified.

The principal understanding sought of the exercise metabolome, and the most prevalent in the current literature, is how the metabolite abundances change from rest after an exercise intervention. Understanding these changes and identifying novel biomarkers may provide further insight into the metabolic regulation of adaptation for the future refinement of exercise programs for sporting performance. Exercise programs may also be refined for optimal health benefits based on this rationale and it also allows for further investigation of the interaction

between nutritional interventions, metabolic perturbations and chronic adaptations to exercise training.

The first investigation into multivariate modelling of exercise metabolome profiling was performed by Pohjanen and colleagues (2007) where 402 serum metabolites were used for prediction of the pre- or post-exercise state. Discriminant analyses were used to highlight 34 metabolites with significant changes between groups (paired *t* test $p < 8.2x10^{-5}$) and showed major contributions from increased glycerol and decreased asparagine after a multiple-bout, sub-maximal exercise session. Although this proof of concept study did not investigate the mechanistic effects of these changes, the study highlighted the potential for non-targeted metabolomics to identify changes in metabolites in response to an exercise bout that are beyond the traditional measures used within the sport and exercise sciences.

Further investigations have noted changes in metabolites related to lipid and glucose metabolism at a range of exercise durations and intensities. An intensified training period in endurance athletes which involved a repeated running protocol of 2.5 hrs/day (for 3 days) at 70% VO_{2max}, induced significant elevations in metabolites related to lipid metabolism (Nieman, Shanely, Gillitt, Pappan, & Lila, 2013). Furthermore, the sustained elevation of these metabolites 14 hours after the final exercise bout suggests a prolonged elevation in lipid metabolism during the recovery from an intensified endurance training period (Nieman et al., 2013). Similarly, an acute but prolonged bout of cycling (75 km) caused a 3.1-fold and 1.7-fold increase in linoleic acid oxidation products 13-HODE and 9-HODE, respectively, and showed associations with the onset of oxidative stress as measured by F_2-isoprostane levels ($r = 0.75$, $p < 0.001$, Nieman et al., 2014). This depth of understanding generated from non-targeted metabolomic analysis supports the subsequent use of 13-HODE and 9-HODE as future markers of oxidative stress in response to exercise. The detailed understanding of fatty acid

metabolism during the exercise bout may also help to inform future nutritional strategies that target changes in substrate use and/or oxidative stress responses to exercise.

In addition to products of fatty acid oxidation, induced changes of fatty acid transporter molecules have been measured through increases in medium to long-chain acylcarnitines which have been demonstrated to increase palmitate oxidation in isolated rat muscle (Krug et al., 2012; Lehmann et al., 2010; Nieman, Shanely, et al., 2013). Such measurement of circulating oxidation products and their transporter molecules could show utility in a medical setting, with cardiovascular disease conditions known to display alterations in cardiac tissue metabolism through dysregulated substrate utilisation (e.g. Aubert et al., 2016; Bedi et al., 2016); however, these current investigations involve exercise interventions that would be too strenuous for critically ill patients to perform.

The sensitivity of metabolomics to detect changes in fuel use during exercise has been demonstrated through observed decreases in serum concentrations of branched-chain amino acids in response to repeated 80 m running sprints (Pechlivanis et al., 2013). This is further supported by the identification of increased branched-chain amino acid degradation products in urine samples after exercise. Furthermore, although changes in the metabolome were detected in response to sprint training, there were no observed differences between groups that received either a 10 s or 1 min recovery period between sprint intervals (Pechlivanis et al., 2013). The comprehensive overview of metabolites provided by metabolomic analysis allows for greater certainty that the manipulation of recovery periods did not induce metabolic differences compared with the measurement of a limited number of variables using a traditional approach. More recently, increases in serum lactate, pyruvate, succinate and multiple butyrates, along with a reduction in amino acids, has been recorded after a single bout of resistance exercise (Berton et al., 2016). Urinary increases in lactate, pyruvate and succinate have also been identified as pre- to post-exercise discriminators 30 min after a single 30 s cycle ergometer

sprint (Enea et al., 2010). The use of metabolomics to monitor changes during such high intensity exercise may be particularly beneficial as the large anaerobic contribution to energy provision prevents accurate interpretation of fuel use from gas exchange measurements (Frayn, 1983).

The majority of studies to date have identified biomarkers with known metabolic contributions. However, the benefit of a non-targeted metabolomics approach was demonstrated by Malkar et al. (2013) through the identification of changes in a salivary metabolite of unknown origin in response to exercise. The metabolite was subsequently identified as δ-valerolactam, and although this molecule does not have any known physiological interactions, the isolation of this compound facilitates further investigation into its physiological role as well as the consequences and meaningfulness of the observed changes in response to exercise. The previous identification of stress markers such as cortisol has benefitted the clinical and scientific interpretation of a variety of stressors including exercise (Hough, Corney, Kouris, & Gleeson, 2013). Therefore, although the initial relationship between exercise-induced stress and novel biomarkers is merely correlational, subsequent identification of the role for such biomarkers may make a significant contribution to future research.

Salivary biomarkers of stress and immune function have been extensively measured within sports science settings to minimise the risk of overtraining and upper respiratory tract infections in athletes (Meeusen et al., 2013). Although this represents a common approach in many sporting environments, this relationship is also largely based on correlational evidence and the relationship between markers such as salivary IgA and subsequent infection risk is not perfect, with coefficients of determination typically below 30 % (Gleeson et al., 1999; Neville, Gleeson, & Folland, 2008). Subsequently, it remains feasible that novel markers of suppressed immunity or overtraining may exist which could further assist in the adjustment of training

loads to minimise risk for the athlete. Current evidence supporting the use of non-targeted metabolomics to identify novel markers of stress in salivary samples include observed elevations in metabolic by-products such as 3-methylhistidine (1.5 fold), glucose phosphate isomers (2.5-4.8 fold) and several amino acids (1.2-2.1 fold) in soccer players who expressed signs of fatigue after a 3-day program of matches (Ra, Maeda, Higashino, Imai, & Miyakawa, 2014). A panel of saliva metabolites relevant to cellular energy metabolism (e.g. creatine, glucose, lactate, glutamate, acetate) has also been demonstrated to cluster yo-yo test performance in football players, suggesting that such measurements may be able to predict changes in performance (e.g. performance impairments due to overreaching) which represents an avenue for future research (Santone et al., 2014).

Aside from investigating changes in the exercise metabolome, researchers have also made initial steps into using metabolomics to pinpoint metabolites that may provide insight into an individual's physical capacity, without the need to complete strenuous exercise tests. Routine blood tests that could provide predicted values of physical fitness would be of great benefit to both the medical and sporting communities, providing physiological monitoring and management where a person cannot perform exercise (e.g. critical illness, injury etc.), or where avoiding additional physical exertion would be preferred (e.g. an athlete during busy competition periods). This approach to using the metabolome as an indicator for health and fitness outcomes also aligns with the increasing focus on "precision medicine" (McCarthy, 2015). In a recent investigation, Lustgarten et al. (2013) reported that pyroxidate, 2-hydroxybutyrate and 4-vinylphenol sulphate showed significant associations with VO_{2max} in both males and females. Furthermore, when these were combined with additional metabolites and blood chemistry analytes (e.g. SGOT, blood urea nitrogen), it was possible to explain 58 and 80% of VO_{2max} scores in males and females, respectively. In another study where participants were categorised into high (55 ± 8 mLO_2/kg/min) or low (31 ± 7 mLO_2/kg/min)

fitness groups, a total of 15 amino acids were reported to be different between the groups in both urine and plasma metabolomes (Morris et al., 2013). Further, the authors showed an association between leucine and markers linked to metabolic syndrome and insulin resistance, suggesting a potential link to changes in amino acid utilisation in unfit individuals that could be causative in the development of cardiometabolic disorders. Differences between trained and untrained populations for resting levels of amino acid profiles and metabolites related to energy production and oxidative stress have also been observed by Yan et al. (2009), thereby providing further support for the view that the resting metabolome may reflect physical capacity levels.

A further study comparing distinct categorised groups of low and high fitness demonstrated that the low fitness group had lower levels of phosphatidylcholine and increased free choline (approximately 1.5-fold, $p = 0.017$, Bye et al., 2012). Circulating free choline has recently been implicated in cardiovascular disease (Wang et al., 2014) and is known to be metabolised by the gut microbiome to form an intermediary in the production of trimethylamine N-oxide (Wang et al., 2011), a small molecule metabolite associated with reduced survival in conditions such as heart failure and myocardial infarction (Suzuki, Heaney, Bhandari, Jones, & Ng, 2016; Suzuki, Heaney, Jones, & Ng, 2016; Tang et al., 2014). Bye and colleagues (2012) have provided a basis to further explore exercise training to reduce levels of metabolites such as free choline through improving cardiovascular fitness, with the intention of reducing the risk of later-life development of cardiometabolic disorders.

An additional use of non-targeted strategies in exercise-based investigations has been to understand how nutritional interventions interact with the exercise metabolome. Lee and colleagues (2010) performed a case study experiment on a participant completing an exhaustive submaximal exercise test (75% VO_{2peak} for 45 min followed by 90% VO_{2peak} until fatigue) and red blood cell lysates were analysed with a global metabolomics approach. The test was completed with and without a high-dose oral intake of *N*-acetyl-L-cysteine. The results reported

reduced levels of carnitine, acetyl-L-carnitine, creatine and 3-methylhistidine in the supplemented trial, with exercise-induced changes in reduced and oxidised glutathione blunted. The authors attributed the nutritional supplement as a method to suppress acute exacerbations of oxidative stress, although with only one participant the results may not reflect the general effect across larger sample sizes. Similarly, studies have shown nutritional interventions to alter the circulating levels of molecules related to energy metabolism (Chorell, Moritz, Branth, Antti, & Svensson, 2009; Miccheli et al., 2009; Nieman, Gillitt, et al., 2013), with one showing that individuals supplemented with a low carbohydrate protein drink and classified into a lower fitness group displayed a post-exercise metabolomic profile similar to that of the high fitness group who consumed only water (Chorell et al., 2009). Further research is required to fully appreciate the influence of nutritional interventions on the metabolome at rest and in combination with exercise. The measurement of appropriate outcome variables in response to nutritional manipulations that are designed to influence the metabolome may also help to substantiate the correlations observed between novel metabolites and parameters of health and exercise performance.

The implementation of non-targeted metabolomics into sport and exercise science investigations is in its early stages and is likely to be increasingly utilised in the coming years. The outlined studies present initial exploration and provide a base for future research in the field, with many areas of interest yet to be probed.

6. Future implications and potential hurdles

Non-hypothesis-driven research into the metabolic changes that occur during sport and exercise has recently been termed “Sportomics”. This concept incorporates a top-down study model with the analysis of large datasets of metabolic variables collected in response to sports training and competition (Bassini & Cameron, 2014). The primary distinction between

Sportomics and the use of metabolomics proposed in the present review is that the Sportomics approach typically provides a non-hypothesis driven analysis of a broad range of traditional metabolic variables, whereas non-targeted metabolomics monitors a wider range of metabolites that may be used to identify novel biomarkers of adaptations to exercise for health and performance. We believe that the metabolomics approach discussed in the present review may complement and extend the concept of Sportomics.

In order for novel metabolic biomarkers to be identified for use in exercise and sport, non-targeted strategies must be employed. Research questions must be carefully designed so that correct collection and analysis of samples is performed, allowing for reliable statistical interpretation of the data. Global measurement of metabolites could be applied to help understand the athlete's current state (e.g. fatigue, physical capacity etc.) or for use in aiding prediction of future events such as talent identification, onset of illness, susceptibility to injury or impaired physical performance.

Once investigations have been successful in isolating single or multiple metabolites that offer beneficial measurements, methods to analyse these must be streamlined in order to allow future application with increased throughput (i.e. reduced analysis times) and reduced complexity. Methods to achieve these goals come through adaptations to the non-targeted analytical workflows that can allow for reduced analysis time and increased sensitivity through analyte filtering. Once these targeted strategies have been developed, validation of the usefulness of the metabolite screening must be performed and compared to any alternative measurements/techniques that are currently available.

The major stumbling block for these types of analyses in exercise and sport science is the high cost of purchase and maintenance of NMR and MS systems, with trained personnel required for the everyday functioning of the instrumentation. However, recent advances in technologies have allowed for complex software packages to be simplified and adopt a more

'plug and play' style interface, thus increasing the ease of training for non-specialist users. A further issue is the space and provision required to operate these forms of instrumentation. NMR systems are bulky and extremely heavy due to the housed magnet, and MS systems are required to be under a vacuum with a constant supply of electricity and inert gases during operation. Efforts are being made to overcome these issues and a recent development in portable, compact mass spectrometers (e.g. Heaney et al., 2016) offer a reduced cost and footprint that may provide the important steps for the translation from the laboratory to field-based investigations in exercise and sport.

7. Conclusions

Although in its infancy, there is promise for the development of non-targeted metabolomic analyses applied to exercise and sport-based scenarios. Non-targeted strategies have been employed in the search of new biomarkers for personalised and stratified medicine, and present a new direction for the discovery of metabolite indicators for sport and exercise science. The coupling of chromatography to MS offers attractive methods for the analysis of many hundreds of metabolites in a single analytical run, and can be further developed to employ high-throughput, targeted methods for identified metabolites of interest. Moreover, improved information of fatigue, physical capacity and performance characteristics through metabolomic analyses may be beneficial to sportspeople to achieve their goals with maximum success and efficiency, as well as benefiting exercise testing for health and disease investigation.

References

Aubert, G., Martin, O. J., Horton, J. L., Lai, L., Vega, R. B., Leone, T. C., . . . Kelly, D. P. (2016). The failing heart relies on ketone bodies as a fuel. *Circulation, 133*(8), 698-705. doi:10.1161/circulationaha.115.017355

Bassini, A., & Cameron, L. C. (2014). Sportomics: building a new concept in metabolic studies and exercise science. *Biochemical and Biophysical Research Communications*, *445*(4), 708-716. doi:10.1016/j.bbrc.2013.12.137

Bedi, K. C., Jr., Snyder, N. W., Brandimarto, J., Aziz, M., Mesaros, C., Worth, A. J., . . . Rame, J. E. (2016). Evidence for intramyocardial disruption of lipid metabolism and increased myocardial ketone utilization in advanced human heart failure. *Circulation, 133*(8), 706-716. doi:10.1161/circulationaha.115.017545

Berton, R., Conceição, M. S., Libardi, C. A., Canevarolo, R. R., Gáspari, A. F., Chacon-Mikahil, M. P. T., . . . Cavaglieri, C. R. (2016). Metabolic time-course response after resistance exercise: A metabolomics approach. *Journal of Sports Sciences*. Advance online publication. doi:10.1080/02640414.2016.1218035

Bye, A., Vettukattil, R., Aspenes, S. T., Giskeodegard, G. F., Gribbestad, I. S., Wisloff, U., & Bathen, T. F. (2012). Serum levels of choline-containing compounds are associated with aerobic fitness level: the HUNT-study. *PLoS One, 7*(7), e42330. doi:10.1371/journal.pone.0042330

Chorell, E., Moritz, T., Branth, S., Antti, H., & Svensson, M. B. (2009). Predictive metabolomics evaluation of nutrition-modulated metabolic stress responses in human blood serum during the early recovery phase of strenuous physical exercise. *Journal of Proteome Research, 8*(6), 2966-2977. doi:10.1021/pr900081q

Creek, D. J., Jankevics, A., Breitling R., Watson D. G., Barrett, M. P., & Burgess, K. E. (2011). Toward global metabolomics analysis with hydrophilic interaction liquid chromatography-mass spectrometry: improved metabolite identification by retention time prediction. *Analytical Chemistry*, *83*(22), 8703-8710. doi: 10.1021/ac2021823

Dunn, W. B., Broadhurst, I. B., Atherton, H. J., Goodacre, R. J., & Griffin, J. L. (2011). Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews*, *40*(1), 387-426. doi:10.1039/B906712B

Dunn, W.B., Broadhurst, I. B., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., . . . Goodacre, R. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*, *6*(7), 1060-1083. doi: 10.1038/nprot.2011.335

Di Guida, R., Engel, J., Allwood, J. W., Weber, R. J. M., Jones, M. R., Sommer, U., . . . Dunn, W. B. (2016). Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics, 12*(5), 1-14. doi:10.1007/s11306-016-1030-9

Egan, B., Hawley, J. A., & Zierath, J. R. (2016). Snapshot: exercise metabolism. *Cell Metabolism*, *24*(2), 342. doi:10.1016/j.cmet.2016.07.013

Enea, C., Seguin, F., Petitpas-Mulliez, J., Boildieu, N., Boisseau, N., Delpech, N., . . . Dugue, B. (2010). (1)H NMR-based metabolomics approach for exploring urinary metabolome modifications after acute and chronic physical exercise. *Analytical and Bioanalytical Chemistry, 396*(3), 1167-1176. doi:10.1007/s00216-009-3289-4

Frayn, K. N. (1983). Calculation of substrate oxidation rates in vivo from gaseous exchange. *Journal of Applied Physiology*, *55*(2), 628-634.

Fudge, B. W., Westerterp, K. R., Kiplamai, F. K., Onywera, V. O., Boit, M. K., Kayser, B., & Pitsiladis, Y. P. (2006). Evidence of negative energy balance using doubly labelled water in elite Kenyan endurance runners prior to competition. *British Journal of Nutrition*, *95*(1), 59-66. doi:10.1079/BJN20051608

Fuhrer, T., Zamboni, N. (2015). High-throughput discovery metabolomics. *Current Opinion in Biotechnology*, *31*, 73-78. doi: 10.1016/j.copbio.2014.08.006

Gleeson, M., McDonald, W. A., Pyne, D. B., Cripps, A. W., Francis, J. L., Fricker, P. A., & Clancy, R. L. (1999). Salivary IgA levels and infection risk in elite swimmers. *Medicine and Science in Sports and Exercise*, *31*(1), 67-73. doi:10.1097/00005768-199901000-00012

Gonzalez, J. T., Fuchs, C. J., Smith, F. E., Thelwall, P. E., Taylor, R., Stevenson, E.J., . . . van Loon, L. J. (2015). Ingestion of glucose or sucrose prevents liver but not muscle glycogen depletion during prolonged endurance-type exercise in trained cyclists. *American Journal of Physiology, Endocrinology & Metabolism*, *309*(12), E1032-E1039. doi:10.1152/ajpendo.00376.2015

Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., & Deigner, H. -P. (2010). Targeted metabolomics for biomarker discovery. *Angewandte Chemie International Edition*, *49*(32), 5426-5445. doi: 10.1002/anie.200905579

Heaney, L. M., Ruszkiewicz, D. M., Arthur, K. L., Hadjithekli, A., Aldcroft, C., Lindley, M. R., . . . Reynolds, J. C. (2016). Real-time monitoring of exhaled volatiles using atmospheric pressure chemical ionization on a compact mass spectrometer. *Bioanalysis, 8*(13), 1325-1336. doi:10.4155/bio-2016-0045

Hough, J., Corney, R., Kouris, A., & Gleeson, M. (2013). Salivary cortisol and testosterone responses to high-intensity cycling before and after an 11-day intensified training period. *Journal of Sports Sciences*, *31*(14), 1614-1623. doi:10.1080/02640414.2013.792952

The Human Metabolome Database. (2016, 12 August). Retrieved from http://www.hmdb.ca

Impey, S. G., Hammond, K. M., Shepherd, S. O., Sharples, A. P., Stewart, C., Limb, M., . . . Morton, J. P. (2016). Fuel for the work required: a practical approach to amalgamating train-low paradigms for endurance athletes. *Physiological Reports*, *10*, e12803. doi:10.14814/phy2.12803

Jones, A. M., Wilkerson, D. P., & Fulford, J. (2008). Muscle [phosphocreatine] dynamics following the onset of exercise in humans: the influence of baseline work-rate. *Journal of Physiology*, *586*(3), 889-898. doi:10.1113/jphysiol.2007.142026

Kováts, E. (1958). Gas-chromatographische Charakterisierung organischer Verbindungen. Teil 1: Retentionsindices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone. *Helvetica Chimica Acta, 41*(7), 1915-1932. doi:10.1002/hlca.19580410703

Krug, S., Kastenmuller, G., Stuckler, F., Rist, M. J., Skurk, T., Sailer, M., . . . Daniel, H. (2012). The dynamic range of the human metabolome revealed by challenges. *FASEB Journal, 26*(6), 2607-2619. doi:10.1096/fj.11-198093

Lee, R., West, D., Phillips, S. M., & Britz-McKibbin, P. (2010). Differential metabolomics for quantitative assessment of oxidative stress with strenuous exercise and nutritional intervention: thiol-specific regulation of cellular metabolism with N-acetyl-L-cysteine pretreatment. *Analytical Chemistry, 82*(7), 2959-2968. doi:10.1021/ac9029746

Lehmann, R., Zhao, X., Weigert, C., Simon, P., Fehrenbach, E., Fritsche, J., . . . Niess, A. M. (2010). Medium chain acylcarnitines dominate the metabolite pattern in humans under moderate intensity exercise and support lipid oxidation. *PLoS One, 5*(7), e11519. doi:10.1371/journal.pone.0011519

Lustgarten, M. S., Price, L. L., Logvinenko, T., Hatzis, C., Padukone, N., Reo, N. V., . . . Fielding, R. A. (2013). Identification of serum analytes and metabolites associated with aerobic capacity. *European Journal of Applied Physiology, 113*(5), 1311-1320. doi:10.1007/s00421-012-2555-x

Malkar, A., Devenport, N. A., Martin, H. J., Patel, P., Turner, M. A., Watson, P., . . . Creaser, C. S. (2013). Metabolic profiling of human saliva before and after induced physiological stress by ultra-high performance liquid chromatography–ion mobility–mass spectrometry. *Metabolomics, 9*(6), 1192-1201. doi:10.1007/s11306-013-0541-x

Mathé, E. A., Patterson, A. D., Haznadar, M., Manna, S. K., Krausz, K. W., Bowman, E. D., . . . Harris, C. C. (2014). Noninvasive urinary metabolomic profiling identifies diagnostic and prognostic markers in lung cancer. *Cancer Research, 74*(12), 3259-3270. doi:10.1158/0008-5472.can-14-0109

McCarthy, M. (2015). Obama seeks $213m to fund "precision medicine". *BMJ*, *350*, h578. doi:10.1136/bmj.h587

Meeusen, R., Duclos, M., Foster, C., Fry, A., Gleeson, M., Nieman, D., . . . Urhausen, A. (2013). Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Medicine and Science in Sports and Exercise*, *45*(1), 186-205. doi:10.1249/MSS.0b013e318279a10a

Miccheli, A., Marini, F., Capuani, G., Miccheli, A. T., Delfini, M., Di Cocco, M. E., . . . Spataro, A. (2009). The influence of a sports drink on the postexercise metabolism of elite athletes as investigated by NMR-based metabolomics. *Journal of the American College of Nutrition, 28*(5), 553-564. doi:10.1080/07315724.2009.10719787

Morris, C., Grada, C. O., Ryan, M., Roche, H. M., De Vito, G., Gibney, M. J., . . . Brennan, L. (2013). The relationship between aerobic fitness level and metabolic profiles in healthy adults. *Molecluar Nutrition & Food Research, 57*(7), 1246-1254. doi:10.1002/mnfr.201200629

Neville, V., Gleeson, M., & Folland, J. P. (2008). Salivary IgA as a risk factor for upper respiratory infections in elite professional athletes. *Medicine and Science in Sports and Exercise*, *40*(7), 1228-1236. doi:10.1249/MSS.0b013e31816be9c3

Nicholson, J. K., Lindon, J. C., Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, *29*(11), 1181-1189. doi:10.1080/004982599238047

Nicholson, J. K., & Wilson, I. D. (2003). Understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nature Reviews Drug Discovery*, *2*(8), 668-676. doi:10.1038/nrd1157

Nieman, D. C., Gillitt, N. D., Knab, A. M., Shanely, R. A., Pappan, K. L., Jin, F., & Lila, M. A. (2013). Influence of a polyphenol-enriched protein powder on exercise-induced inflammation and oxidative stress in athletes: a randomized trial using a metabolomics approach. *PLoS One, 8*(8), e72215. doi:10.1371/journal.pone.0072215

Nieman, D. C., Shanely, R. A., Gillitt, N. D., Pappan, K. L., & Lila, M. A. (2013). Serum metabolic signatures induced by a three-day intensified exercise period persist after 14 h of recovery in runners. *Journal of Proteome Research, 12*(10), 4577-4584. doi:10.1021/pr400717j

Nieman, D. C., Shanely, R. A., Luo, B., Meaney, M. P., Dew, D. A., & Pappan, K. L. (2014). Metabolomics approach to assessing plasma 13- and 9-hydroxy-octadecadienoic acid and linoleic acid metabolite responses to 75-km cycling. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology, 307*(1), R68-74. doi:10.1152/ajpregu.00092.2014

O'Gorman, D. J., Karlsson, H. K., McQuaid, S., Yousif, O. Rahman, Y., Gasparro, D., . . . Nolan, J. J. (2006). Exercise training increases insulin-stimulated glucose disposal and GLUT4 (SLC2A4) protein content in patients with type 2 diabetes. *Diabetologia*, *49*(12), 2983-2992. doi:10.1007/s00125-006-0457-3

O'Hara, J. P., Carroll, S., Cooke, C. B., Morrison, D. J., Preston, T., & King, R. F. (2012). Preexercise galactose and glucose ingestion on fuel use during exercise. *Medicine and Science in Sports and Exercise*, *44*(10), 1958-1967. doi:10.1249/MSS.0b013e318258bf85

Pechlivanis, A., Kostidis, S., Saraslanidis, P., Petridou, A., Tsalis, G., Veselkov, K., . . . Theodoridis, G. A. (2013). 1H NMR study on the short- and long-term impact of two training programs of sprint running on the metabolic fingerprint of human serum. *Journal of Proteome Research, 12*(1), 470-480. doi:10.1021/pr300846x

Pohjanen, E., Thysell, E., Jonsson, P., Eklund, C., Silfver, A., Carlsson, I. B., . . . Antti, H. (2007). A multivariate screening strategy for investigating metabolic effects of strenuous physical exercise in human serum. *Journal of Proteome Research, 6*(6), 2113-2120. doi:10.1021/pr070007g

Ra, S. G., Maeda, S., Higashino, R., Imai, T., & Miyakawa, S. (2014). Metabolomics of salivary fatigue markers in soccer players after consecutive games. *Applied Physiology, Nutrition, and Metabolism, 39*(10), 1120-1126. doi:10.1139/apnm-2013-0546

Santone, C., Dinallo, V., Paci, M., D'Ottavio, S., Barbato, G., & Bernardini, S. (2014). Saliva metabolomics by NMR for the evaluation of sport performance. *Journal of Pharmaceutical and Biomedical Analysis, 88*, 441-446. doi:10.1016/j.jpba.2013.09.021

Suzuki, T., Heaney, L. M., Bhandari, S. S., Jones, D. J. L., & Ng, L. L. (2016). Trimethylamine N-oxide and prognosis in acute heart failure. *Heart, 102*(11), 841-848. doi:10.1136/heartjnl-2015-308826

Suzuki, T., Heaney, L. M., Jones, D.J. L., & Ng, L.L. (2016). Trimethylamine N-ozide and risk stratification after acute myocardial infarction. *Clinical Chemistry*, *63*(1), 420-428. doi:10.1373/clinchem.2016.264853

Tang, W. H., Wang, Z., Fan, Y., Levison, B., Hazen, J. E., Donahue, L. M., . . . Hazen, S. L. (2014). Prognostic value of elevated levels of intestinal microbe-generated metabolite trimethylamine-N-oxide in patients with heart failure: refining the gut hypothesis. *Journal of the American College of Cardiology, 64*(18), 1908-1914. doi:10.1016/j.jacc.2014.02.617

Thevis, M., Kuuranne, T., Walpurgis, K., Geyer, H., & Schanzer, W. (2016). Annual banned-substance review: analytical approaches in human sports drug testing. *Drug Testing and Analysis, 8*(1), 7-29. doi:10.1002/dta.1928

Wang, Z., Klipfell, E., Bennett, B. J., Koeth, R., Levison, B. S., Dugar, B., . . . Hazen, S. L. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature, 472*(7341), 57-63. doi:10.1038/nature09922

Wang, Z., Tang, W. H., Buffa, J. A., Fu, X., Britt, E. B., Koeth, R. A., . . . Hazen, S. L. (2014). Prognostic value of choline and betaine depends on intestinal microbiota-generated metabolite trimethylamine-N-oxide. *European Heart Journal, 35*(14), 904-910. doi:10.1093/eurheartj/ehu002

Want, E. J., Wilson, I. D., Gika, H., Theodoridis, G., Plumb, R. S., Shockcor, J., . . . Nicholson, J. K. (2010). Global metabolic profiling procedures for urine using UPLC–MS. *Nature Protocols*, *5*(6), 1005-1018. doi: 10.1038/nprot.2010.50

Wiklund, S., Johansson, E., Sjöström, L., Mallerowicz, E. J., Edlund, U., Shockcor, J. P., . . . Trygg, J. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry*, *80*(1), 115-122. doi: 10.1021/ac0713510

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., . . . Scalbert, A. (2013). HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Research, 41*(Database issue), D801-807. doi:10.1093/nar/gks1065

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., . . . Querengesser, L. (2007). HMDB: the Human Metabolome Database. *Nucleic Acids Research, 35*(Database issue), D521-526. doi:10.1093/nar/gkl923

Worley, B., & Powers, R. (2013). Multivariate analysis in metabolomics. *Current Metabolomics*, *1*(1), 92-107. doi: 10.2174/2213235X11301010092

Yan, B., A, J., Wang, G., Lu, H., Huang, X., Liu, Y., . . . Sun, J. (2009). Metabolomic investigation into variation of endogenous metabolites in professional athletes subject to strength-endurance training. *Journal of Applied Physiology, 106*(2), 531-538. doi:10.1152/japplphysiol.90816.2008

Table Captions

Table 1. A table to show the advantages and limitations of using nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS) for non-targeted metabolomics experiments.

Nuclear magnetic resonance spectroscopy	
Advantages	**Limitations**
Non-destructive - prepared sample can be stored and reanalysed	Lower sensitivity levels when compared to MS
High levels of reproducibility	Complex samples may exhibit overlapping spectral peaks that cannot be deconvoluted
No requirement for vacuum or gas supply	Expensive to purchase and specialised training required to operate
Absolute quantification possible	Large housed magnet can be troublesome for placement of instrument due to weight and interference with surrounding equipment
Structural information available for measured metabolites	

Mass spectrometry	
Advantages	**Limitations**
High levels of sensitivity	Destructive - injected portion of prepared sample cannot be reanalysed
Ability to measure both intact and fragmented ions for improved identification	Subject to instrument fluctuations across experimental periods
Overlapping chromatographic peaks can be deconvoluted by extracting *m/z* values	Instruments must remain in vacuum and most require access to constant supply of gases (e.g. N_2, He)
Highly accurate levels of measured mass achievable (e.g. with use of time-of flight and orbitrap mass analysers)	Multiple ionisation states (positive/negative) and chromatographic techniques (e.g. reverse phase/HILIC) required for full metabolite capture
Increased potential for number of measureable analytes from single analysis over NMR	Constant temperature required to maintain mass accuracy (e.g. time-of-flight mass analyser)
	Expensive to purchase and specialised training required to operate
	Metabolite must ionise to be detected and potential for competition of ionisation in co-eluting metabolites
	Quantitative data requires presence of an internal standard and additional calibration curves performed on known standard after metabolite identification

Figure Captions

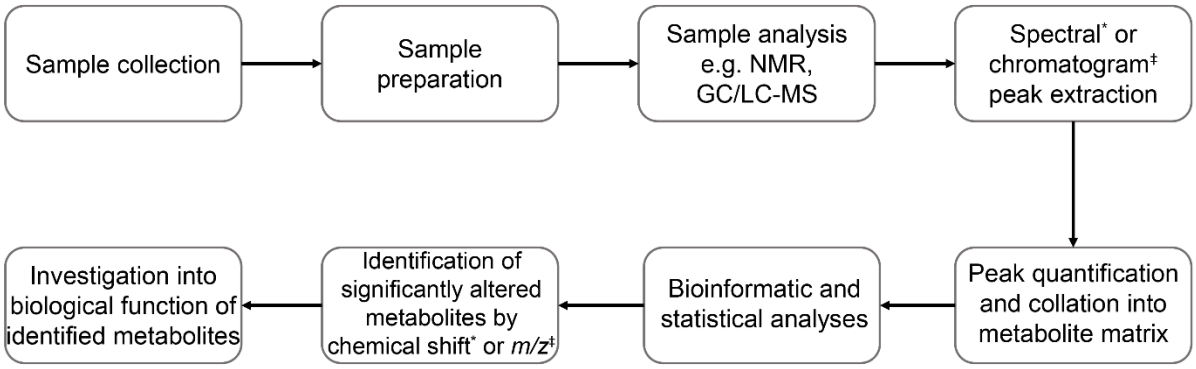

Figure 1. A generalised workflow for metabolite identification using nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS) based technologies for non-targeted metabolomics experiments.

*denotes relation to experiments using NMR; ‡ denotes relation to experiments using GC/LC-MS

GC = gas chromatography; LC = liquid chromatography; *m/z* = mass to charge ratio

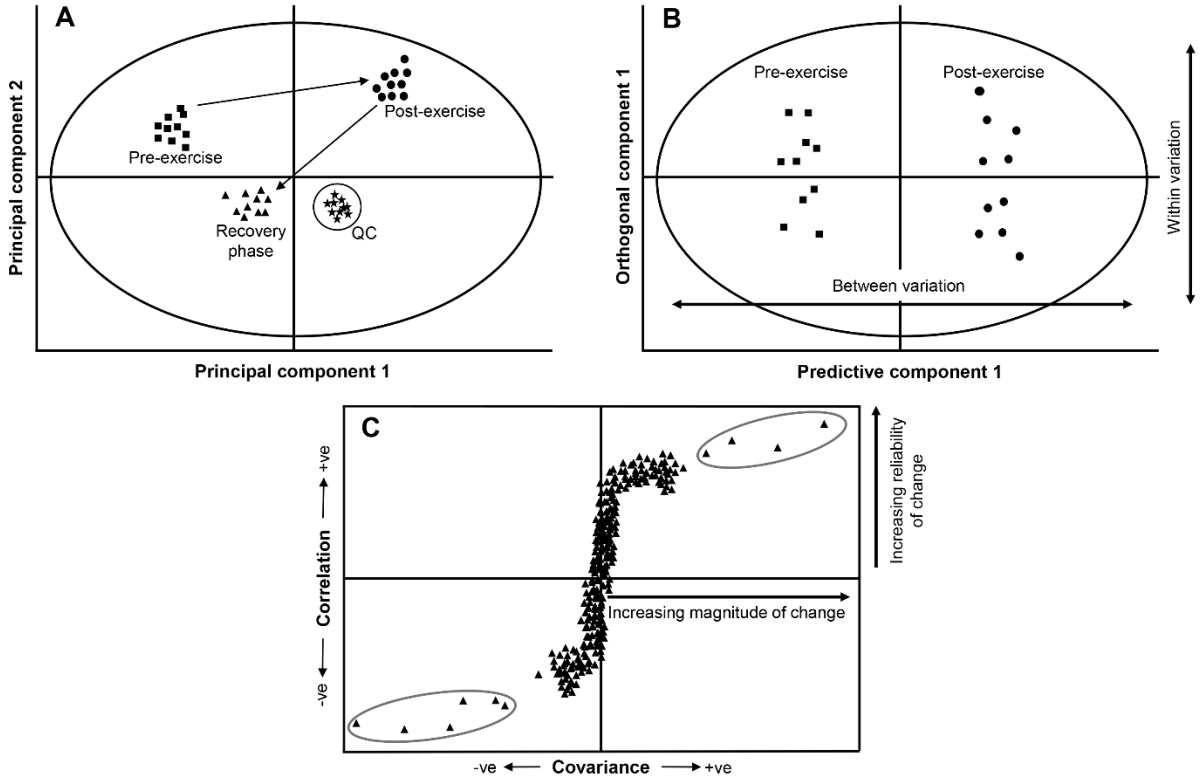

Figure 2. Graphical visualisations of plots for a hypothetical non-targeted metabolomics experiment in sport and exercise science utilising principle components analysis (PCA; **A**) and orthogonal partial least squares-discriminant analysis (OPLS-DA; **B**) with its corresponding S-plot (**C**).

A Example PCA plot to show a shift in metabolic profile of participants from pre- (squares) to post-exercise (circles), with a tendency to return to pre-exercise characteristics after a period of recovery (triangles). Circled stars represent multiple analyses of a quality control (QC) sample at regular intervals throughout the analytical period. Close clustering of these indicates that low levels of instrumental variation (i.e. good reproducibility) are present across the study.

B Example OPLS-DA plot to show a supervised multivariate analysis focussed on identifying metabolites that exhibit significant differences between pre-defined groups. This plot shows differences in metabolite values between pre- (square) and post-exercise (circles) samples.

C Example corresponding S-plot to visualise the most contributory metabolites to the statistical model for separation of pre-defined groups observed from OPLS-DA (Fig **1B**). The triangles in grey ovals signify the metabolites with the greatest magnitude and reliability of change between groups and would commonly be selected for further statistical testing using more targeted approaches (e.g. paired t-test, Wilcoxon test, etc.).