Nutrigenomic approaches for benefit-risk analysis of foods and food components: defining markers of health

Ruan Elliott1*, Catalina Pico2, Yvonne Dommels3, Iwona Wybranska4, John Hesketh5 and Jaap Keijer3

1Institute of Food Research, Norwich, UK
2University of the Balearic Islands, Palma de Mallorca, Spain
3RIKILT-Institute of Food Safety, Wageningen, The Netherlands
4Collegium Medicum, Jagiellonian University, Krakow, Poland
5Institute of Cell and Molecular Biosciences, University of Newcastle, UK

(Received 1 November 2006 – Revised 7 March 2007 – Accepted 29 May 2007)

To be able to perform a comprehensive and rigorous benefit-risk analysis of individual food components, and of foods, a number of fundamental questions need to be addressed first. These include whether it is feasible to detect all relevant biological effects of foods and individual food components, how such effects can confidently be categorised into benefits and risks in relation to health and, for that matter, how health can be quantified. This article examines the last of these issues, focusing upon concepts for the development of new biomarkers of health. Clearly, there is scope for refinement of classical biomarkers so that they may be used to detect even earlier signs of disease, but this approach defines health solely as the absence of detectable disease or disease risk. We suggest that the health of a biological system may better be reflected by its ability to withstand and manage relevant physiological challenges so that homeostasis is maintained. We discuss the potential for expanding the range of current challenge tests for use in conjunction with functional genomic technologies to develop new types of biomarkers of health.

Risk-benefit: Nutrition: Biomarkers: Nutritional genomics

We all routinely perform our own personal version of benefit-risk analysis in the process of making day-to-day decisions, including our dietary choices. Of course, in relation to the food we select, this process will normally be anything but systematic with heavy bias arising from many factors, such as personal likes and dislikes, mood, time pressures, upbringing, seasons, education and awareness of nutritional and health issues.

Even the best informed, most health-conscious and motivated consumer cannot be getting the analysis exactly right because not all the information relevant to them is available. And while the concept of rigorous benefit-risk analysis should underlie the development of dietary recommendations, expert panels developing such recommendations have to work within the constraints of the current limited and fragmented scientific knowledge base. Unexpected biological effects and interactions (be they beneficial or detrimental) and inter-individual differences in nutritional requirements pose a particular problem in this respect. Development of a system or framework for performing comprehensive and rigorous benefit-risk analysis for individual food components, whole foods or diets, and defining a common scale of measurement for comparing the risks and the benefits would be enormously advantageous1. Amongst the challenges involved, several key questions need to be answered. These include:

- How can health be quantified?
- Is it possible to detect all biological effects of a food component, food or diet?
- Can these effects confidently be categorised into benefits and risks?

Within this article we will discuss some limitations of current approaches and then: (i) present some ways of thinking about how to define health; (ii) consider the opportunities for quantifying risks and benefits of dietary (and other environmental) factors that genomic technologies present; (iii) propose possible routes forward, including potentially useful experimental systems and criteria for selecting food components that will allow validation and further concept development.

Developing biomarkers of health

To date, most biomarkers have been developed for the purpose of detecting disease or deviation from the ‘norm’ that may signal disease development. From a clinical standpoint, the most useful biomarkers provide a definitive link to a very specific disease risk or condition. There is much effort going into the identification of markers that provide the earliest possible indication of disease. There is undoubtedly scope to refine

Abbreviations: GSEA, gene set enrichment analysis.
* Corresponding author: Dr Ruan Elliott, fax +44 (0)1603 507723, email ruan.elliott@bbsrc.ac.uk
some current biomarkers so that they can be used to provide indications of more subtle shifts away from a healthy condition. This might be achieved by focusing upon the functional significance of the variations within the range of data that would generally be considered ‘normal’; the aim being to identify successively earlier indicators of any deterioration.

In assessing how nutrition impinges on health, these markers are useful but they can only indicate ‘health’ as the absence of disease. Ideally, we should have markers of optimal health. There are both practical and theoretical considerations that need to be addressed for the development of such markers. Optimal biomarkers must be analytically robust, sensitive, quantitative and practical (e.g. non-invasive). They also should be mechanism-linked so that implications of changes can be understood. In general, the most useful markers would be relevant for the whole population while also being capable of accounting for the effects of parameters such as age, gender, ethnicity and genotype.

Even with such refinements, individual markers used in isolation will not be able to measure health reliably. Instead, integrated multi-component biomarkers are required. Ideally, these would examine a far broader concept of health than simply defining acceptable values for each parameter individually. Since biological processes impinge upon each other, analysis of patterns of parameters will be far more informative. The ‘omic’ technologies that measure large numbers of parameters in parallel offer significant opportunities in this respect. However, even with these new approaches, it is a major challenge to capture the functional status of a biological system with measurements at a single static point. Dynamic measures, taken under varying conditions, may provide a starting point.

To employ ‘omic’ technologies for defining biological parameters that can be used as markers of health, one first needs to identify healthy subjects to work with, but the most obvious way to characterise people as healthy is to use validated markers of health. Overcoming this apparent paradox requires an alternative way of identifying genuinely healthy individuals. We propose that an answer may lie in analysing the robustness of homeostatic control mechanisms, since a key feature of health is the ability to cope with, and respond adequately to, stress (Fig. 1). This approach is already used; for example, cardiac health is assessed by the response to an exercise challenge and oral glucose tolerance tests are used to assess glycaemic control. The concept of using measures of the robustness of the system to cope with a suitable range of moderate challenges (a so-called ‘stress test’) represents a potentially powerful approach for assessing health status, which is the ability to cope with, and respond adequately to, stress (Fig. 1). This approach is already used; for example, cardiac health is assessed by the response to an exercise challenge and oral glucose tolerance tests are used to assess glycaemic control. The concept of using measures of the robustness of the system to cope with a suitable range of moderate challenges (a so-called ‘stress test’) represents a potentially powerful approach for assessing health status, which is the ability to cope with, and respond adequately to, stress (Fig. 1).

Robustness can be defined as ‘the ability to maintain performance in the face of perturbations or uncertainty’ and biological systems employ a range of strategies, such as redundancy, modularity and feedback control, to achieve this. Homeostatic robustness can be evaluated by determining the scale and duration of perturbations elicited in response to a standard challenge. Inappropriately large and/or prolonged deviations from baseline, following an applied stress, may be used as biomarkers of reduction in robustness with associated adverse health implications (e.g. impaired glucose tolerance). On the other hand, it may not be safe simply to assume the reverse: that the smallest observed deviation from baseline following a standard stress test represents the healthiest condition.

An alternative approach would be to test the robustness of a homeostatic system against a series of challenges of increasing magnitude. This would provide an indication of the system’s maximum capacity and a high capacity might be a suitable indicator of health. However, some caution in the interpretation is still required since there are common trade-offs between enhancing the robustness of a system to deal with specific perturbations and increasing its fragility to other perturbations.

The great advantage of applying ‘omic’ approaches in this situation is that they can be used to identify and categorise entire biological responses to different challenges applied during the proposed stress tests and for determining the rate and completeness of the return to the original pre-challenge status for multiple parameters. This information, used in conjunction with available biological knowledge, can be used to develop increasingly detailed and accurate models that describe homeostatic regulatory processes. These kinds of models not only have the potential to provide new insights into how biological systems function, but also may provide a means of identifying the earliest possible indications of long-term disturbances and compensatory processes. In this regard,
current intensive work into the processes underlying chronic diseases is likely to be extremely valuable. This will identify responses characteristic of early pathological changes and thereby help to discriminate these from normal homeostatic processes.

Given the complexity of the human system, it is possible that optimal health cannot be categorised as a single state. Each individual may have the potential to exist in one or more stable healthy states and these may differ from person to person and shift with life-stage and lifestyle and be different for men and women. There may be more appropriate ways to measure health than simply by seeking to define a single averaged normal standard. Individuality can readily be built into the challenge model by selecting stressors most appropriate to an individual’s specific life requirements. In the same way, this approach also provides the scope to account for the effects of genotype.

**Challenge/stress tests for quantification of health status**

A number of standardised biological stress tests already exist. For the determination of health, and benefit-risk analysis of dietary components, it is important to develop a panel of nutrition-relevant stress tests. These tests should encompass the major aspects of healthy biological function and, critically for assessment of diet-related benefit-risk, should involve challenges that are known to impinge upon cell functions that are clearly sensitive to dietary components. We propose the following as examples:

- Infection challenge
- Inflammatory challenge
- Cognitive challenge
- Carbohydrate challenge
- Lipid challenge
- Xenobiotic challenge
- Oxidative challenge
- DNA damage challenge
- Tissue damage challenge
- Exercise challenge.

This list is by no means intended to be exhaustive: substantial expert input would be required to develop and validate a comprehensive panel of tests. In some cases the challenges and/or most appropriate readouts to test the function of these systems are immediately obvious. In others cases, further technical developments are required. Some of these functions may be tested comparatively readily in human subjects (e.g. carbohydrate and lipid metabolism and response to exercise) whereas others would be more difficult to apply in human subjects for obvious ethical reasons. In such cases, the functions may be tested in tissue culture and animal models, and alternative approaches would be needed to take these tests forward for use in human subjects.

**Potential for application of genomic tools**

The ‘omic’ technologies (i.e. transcriptomics, proteomics and metabolomics) are having a substantial impact across the spectrum of life science research. These technologies have potential for the development and validation of novel biomarkers of health because they provide simultaneous measurements covering a wide range of biological processes. The data can be used in pattern-based analysis for defining healthy biomarker profiles and responses. From another perspective, the data generated allow identification of both predicted and unexpected effects; an essential component of any rigorous benefit-risk analysis.

In most cases, moderate nutritional modifications are likely to elicit comparatively subtle effects that, even so, may have profound biological effects in the long term. Detecting such changes is made more difficult by the degree of variation between human volunteers. However, several lines of evidence suggest that the degree of intra-individual variation in multiple parameters determined using ‘omic’ methods is substantially smaller than the inter-individual variation. Thus, study designs that enable individuals to act as their own control are most appropriate. Transcriptomics, proteomics and metabolomics have been applied in a comparatively small number of human nutritional studies to date. Nevertheless, these studies demonstrate that ‘omic’ methods are sufficiently sensitive to detect both acute and chronic effects of nutrition.

Furthermore, integrated analysis of ‘omic’ data, by taking small but consistent changes over several components of a pathway into account, increases the power and the sensitivity of assessment. For example, de Boer and co-workers were only able to identify fatty acid catabolism as an effect of quercetin exposure when a pathway analysis programme was used. Another example is provided by Mootha and colleagues, who described a method, termed gene set enrichment analysis (GSEA), designed to identify sets of functionally related genes that differ in expression between experimental groups. Using this approach, they were able to identify a specific set of genes involved in oxidative phosphorylation (the OXPHOS set), which were expressed at lower levels in muscle of type 2 diabetics than in control subjects. They then confirmed a similar pattern of down regulation of the OXPHOS genes in the muscle of individuals with impaired glucose tolerance. This finding contrasts with gene by gene analysis of the same dataset, employing standard statistical methods with appropriate multiple test correction, which failed to identify any significant differences. The difference in expression of members of the OXPHOS gene set between the diabetic and control groups was typically very modest (approximately 20%) but strikingly consistent across the set (94 of 106 genes). Such small differences would be difficult to detect reliably on a gene by gene basis even by using methods such as real-time PCR, which are widely held to be more sensitive and precise than arrays. Thus, currently, it is possibly only in the context of global gene expression analysis that such small differences can be detected and interpreted with confidence.

In addition to offering improved sensitivity, methods such as GSEA provide at least two other important benefits. First, they have proven to be far more effective than single-gene analyses at revealing common biological effects between independent studies. Second, by presenting observed differences in the form of defined biological processes/pathways, they make interpretation of study results in relation to available knowledge more straightforward than considering the possible biological implications of individual genes in the gene lists typically generated using the more traditional...
methods. For example, Mootha and colleagues were able to take forward the observation described earlier by identifying a subset of the OXPHOS genes that are coordinately expressed in multiple tissues and then providing direct evidence that this co-regulation is mediated by PPAR-\(\gamma\) coactivator 1e\(^{30}\).

Since this original description, GSEA and variations on this type of analysis have repeatedly proven their power by providing new insights into the mechanisms underlying a range of pathological conditions, including prostatic intraepithelial neoplasia\(^2\), myositis and dermomyositis\(^4\), uterine fibroids\(^5\), exposure to high dose statins\(^6\), leukaemia and lung cancer\(^7\). Application and improvement of pathway analyses and GSEA in combination with other statistical methods that take interactions into account, such as Random forests\(^{37}\), will further strengthen sensitivity and accuracy.

Beyond application of these methods individually, integration of all the ‘omic’ technologies represents the most powerful way forward. In the longer term, such a systems biology approach is necessary to cope with the full complexity of human individuals, their diets and the interactions between the two. However, integrating transcriptomic, proteomic and metabolomic datasets is challenging, not least because of the differences in the timescales of responses of RNA, protein and metabolites\(^{38}\). This is a particularly important issue in relation to the challenge test concept, where measuring and interpreting the dynamics of the response is critical. In human subjects, changes in blood concentrations of many metabolites can be detected within minutes of ingesting test meals and these may persist for a number of hours. Alterations in gene transcription profile are unlikely to occur quite so quickly but still can be readily detected within 2h\(^2\). Acute changes in protein expression following food ingestion in human subjects have not yet been investigated using a proteomic approach. The complex dynamics of RNA, protein and metabolite interactions is an issue that is only just beginning to be dealt with in simple model systems\(^{39}\). Nevertheless, while the methods necessary to realise full integration are still in development, examples of studies employing more than one ‘omic’ technology demonstrate that the descriptive power of the individual methods can be further enhanced, even using straightforward integration strategies\(^{40,41}\).

While maximising data capture is the obvious way to start, this is expensive and labour intensive. In the long run, the aim should be to define subsets of key parameters and profiles that will enable the complexity of biomarkers of health to be reduced without compromising their power. It is possible that the dynamics of a response could be inferred from a single sample provided the interrelationships between the key parameters analysed have been defined clearly. The ‘omics’ technologies can provide us with this insight. Low cost, high throughput analytical platforms can then be developed to provide accurate quantification of the specific markers identified\(^45\).

It makes sense to test the new concepts outlined earlier by exploiting established nutritional models. For example, the oral glucose tolerance test provides a reference on which to build a systems biology approach. This could be expanded by systematic analysis of other standardised dietary challenges (e.g. lipid loading, vitamin supplementation). Similarly, the best available models of a healthy phenotype could be used to evaluate challenge tests. For example, long-term energy restriction, which is well known to slow ageing and increase maximum lifespan in Drosophila and C. elegans, rats and mice\(^44\), also improves a host of classical biomarkers of CVD risk and inflammation in human subjects\(^6,45,46\). However, this model may not be appropriate to use in all contexts and advocacy of energy restriction for people must still be considered to be ill advised, based on current knowledge\(^47\). For example, energy restriction can lead to reduction in bone mineral density, muscle size and strength and aerobic capacity\(^48,49\). Exercise overcomes these effects and, therefore, a combination of energy restriction with exercise may be more widely applicable. What is now needed is to use well-characterised model challenge tests to develop the appropriate experimental design and data analysis required to incorporate the ‘omic’ technologies and validate their use for benefit-risk analysis.

Foods and food components

It is essential to start with food components that are well-characterised in terms of their metabolism, the cellular processes they influence and links to specific health outcomes. The availability of comprehensive dose–response data, both acute and chronic, from human subjects and different model organisms would be extremely valuable. Ultimately, the complexity of food has to be taken into account. For example, a more complex food-based test has been developed for analysing response to a combined glucose and lipid challenge\(^50\).

Determining the effects of chronic low exposures on health poses a particular technical and logistical challenge, but one that it is vital to address in the context of nutrition. One approach to overcome this would be to test model organisms with comparatively short lifespans. Alternatively, with other models experimental durations could be used that are just sufficiently long for a new state of equilibrium to be reached. Conveniently, the timescale to reach such a new equilibrium could be determined using genomic technologies. In addition, diet at one stage of life can impact on health at a later stage. In particular, early life nutrition is thought to exert profound effects on metabolic ‘set points’\(^51,52\). Functional genomic technologies may be able to identify early biomarkers of change, for example, in gene methylation\(^54,55\).

Looking beyond the current genomic approach, there are a wide array of technologies and model systems in development that may aid the establishment of biomarkers of health that can then be translated into easy applicable tests. These include imaging and nano sampling, allowing access to tissues that thus far are not accessible in living individuals\(^57–59\). Applied to human subjects, this will diminish uncertainties arising from translation from one species to the other. Furthermore, the availability of these tools, especially when they can be applied to human subjects, will reduce animal testing.

Conclusions

Development of complex markers, based on the integration of challenge tests and genomic technologies represents an exciting, if technically challenging, approach for application as biomarkers of health and in nutritional benefit-risk analysis. To make routine use practical, such biomarkers of health ideally should be based around measurements performed on
acceptable substrates (e.g. biofluids, such as plasma and urine, blood cells and DNA). However, it will be essential that it is possible to interpret these biomarkers in the context of function in multiple tissues. Proof of principle will be facilitated by pragmatic selection of appropriate challenge tests and dietary constituents with well described bioactivities.

To achieve this, research is needed to: (1) identify challenge tests relevant to diet-related benefit-risk that can be suitably adapted to incorporate the ‘omic’ technologies; (2) then to develop the use of ‘omics’ in such new tests, modifying the challenge test if necessary, so as to identify novel complex markers that respond appropriately; (3) finally, to validate the new complex markers with food components that have at least some well-characterised physiological effects. Only then will such tests be ready for the complexities of foods.

Acknowledgements

The discussions that led to the development of concepts put forward in this paper were supported through funding for a series of meetings and workshops and for preparation of this manuscript by The European Nutrigenomics Organisation (NuGO): linking genomics, nutrition and health research (NuGO, CT-2004-505944). NuGO is a Network of Excellence funded by the European Commission’s Research Directorate General under Priority Thematic Area 5 Food Quality and Safety Priority of the Sixth Framework Programme for Research and Technological Development. All authors are members of the risk-benefit assessment work package of NuGO.

References


