Assessing potential biomarkers of micronutrient status by using a systematic review methodology: methods

Lee Hooper, Kate Ashton, Linda J Harvey, Tamás Decsi, and Susan J Fairweather-Tait

ABSTRACT

Background: To explore the relation between micronutrient status and health, it is important to understand which markers of micronutrient status can be relied on and under what circumstances.

Objective: The objective of this article was to develop a common systematic review methodology for use in the assessment of micronutrient status for selenium, iodine, copper, zinc, riboflavin, vitamin B-12, vitamin D, and omega-3 (n-3) long-chain polyunsaturated fatty acids.

Design: We developed a methodology on the basis of defining studies that clearly altered micronutrient status and then pooled data on the effects of this intervention on each specific biomarker to assess objectively the response of various status markers to changes in intake.

Results: The generic methodology included defining, and systematically searching for, studies that resulted in a change in micronutrient status. Study inclusion, data extraction, and assessment of validity were conducted with a minimum of 10% independent duplication. For each study and each potential biomarker, the highest dose and longest duration intervention data were selected to assess the statistical significance of any change in intake on the status biomarker. The consistency of biomarker response was explored by subgrouping the studies according to baseline micronutrient status, sex, population group, supplementation type, dose, duration, and analytic method.

Conclusion: This methodology allows systematic assessment of the usefulness of a number of biomarkers for a selection of micronutrients. Am J Clin Nutr 2009;89(suppl):1953S–9S.

INTRODUCTION

The EURopean micronutrient RECommendations Aligned (EURRECA) Network of Excellence is contributing to food and nutrition policy across Europe by developing the knowledge base and approaches required for the alignment of nutrient recommendations (1). To provide updated micronutrient recommendations across Europe, it is important to be able to assess the micronutrient status of individuals and populations, and to do this we need reliable biomarkers. The National Institutes of Health defines a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” (2). Use of a biomarker that reflects changes in micronutrient status can facilitate understanding of the relationships between dietary micronutrient intake and status and those between micronutrient status and health. This makes the task of determining the effects of micronutrient status on a range of health outcomes more realistic.

Currently, if the results of an epidemiologic study show no substantial relation between micronutrient status and a specific health outcome, it is not clear if the lack of relation is due to the biomarker being an inadequate measure of status or the absence of any relation between status and the health outcome. If a potential micronutrient status biomarker does have an effect on a health outcome, is this because the micronutrient has a specific direct effect on health or because the biomarker is a surrogate variable, namely, that it responds to another factor, eg, obesity, that has a direct effect on health? Until we understand which biomarkers truly reflect nutrient status and under what circumstances (ie, deficiency, adequacy, and excess) and other factors that may be driving biomarker changes, it will be difficult to begin to map real effects of micronutrient status on health. Ideally, biomarkers (or measures of status, as they also are called in this series of articles) are body fluids or tissue components whose composition is directly related to dietary intake over the short, medium, or long term (3).

The aim of the EURRECA RA1.2 Work Package (Biomarkers of Micronutrient Status) was to assess which biomarkers reflect the status of specific micronutrients and under what circumstances. The goal was to develop a tool to enable researchers to

1 From the School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, United Kingdom (LH, KA, LJH, SF-T); the PenTAG, Peninsula Medical School, Universities of Plymouth and Exeter, Exeter, United Kingdom (KA); the Institute of Food Research, Norwich Research Park, Colney, Norwich, United Kingdom (LJH); and the Department of Pediatrics, University of Pécs, Pécs, Hungary (TD).
2 This manuscript does not necessarily reflect the views of the Commission of the European Communities and in no way anticipates the future policy in this area.
3 Presented at the EURRECA workshop “Biomarkers of Micronutrient Status,” held in Sveti Stefan, Montenegro, 9 June 2008.
4 Supported in part by the Commission of the European Communities, specific RTD Programme, “Quality of Life and Management of Living Resources,” within the 6th Framework Programme (contract no. FP6-036196-2 EURRECA: EURopean micronutrient RECommendations Aligned).
5 Address correspondence to L. Hooper, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ, United Kingdom. E-mail: l.hooper@uea.ac.uk.
First published online April 29, 2009; doi: 10.3945/ajcn.2009.27230A.
choose the best and most relevant biomarkers for their work in specific populations.

Conventional methodology for assessing the usefulness of diagnostic tests, in this case biomarkers of micronutrient status, involves evaluating the test against a gold standard. However, there are no generally accepted reference standards for most micronutrients. In the absence of a clear gold standard or accepted health biomarker that reflects micronutrient status in a specific and sensitive way, the only available course of action is to validate the degree to which the marker measures what it is supposed to be measuring, namely micronutrient status. As Rutjes et al state, “Construct validity is not determined by a single statistic, but by a body of research that demonstrates the relationship between the test and the target condition that it is intended to identify or characterize. Validating a test can then be understood as a gradual process whereby we determine the degree of confidence we can place on inferences about the target condition in tested patients based on their index test results” (4).

In this collection of reviews, we have used construct validity as the basis for assessment of whether biomarkers reliably reflect changes in nutrient status. Construct validity relies on accumulated research evidence indicating that if micronutrient status is altered, the biomarker also changes. In the absence of any strong evidence on status (because it is not clear at baseline whether the biomarkers are useful), we explored how well the biomarker reflected changes in status. For this reason, we assessed the effect of micronutrient supplementation or depletion of a type, dose, and duration in which, to the best of our knowledge, change in status would be expected. Given that we included only studies that created circumstances where a change in nutrient status had occurred, the questions were as follows: Did the marker reflect this change in the group overall and in individuals within the group? Were there populations, circumstances, or methodologies where a biomarker that did appear to reflect changes in micronutrient status worked less effectively? At this stage, we were not validating according to health outcomes but simply looking at the accumulated evidence that the biomarker altered with changed status (as defined by a substantial increase or decrease in absorbable micronutrient over a sufficiently long period for status to alter).

The ideal approaches to validate biomarkers are dose-response, repletion-depletion studies that are carried out under carefully controlled conditions over a long enough period of time to allow for adaptation. However, for a variety of reasons, metabolic studies of this kind rarely are undertaken. Therefore, we selected the next best option to establish important principles underlying the relations between micronutrient status and health—conducting systematic searches for studies that would allow us to assess this relation and systematic analyses of the relations so as to include the best available evidence on the effects of changes in micronutrient supply on marker status.

The micronutrients selected for study were chosen on the basis of their importance for European public health and included zinc, copper, selenium, iodine, riboflavin, vitamin B-12, vitamin D (plus calcium), and omega-3 fatty acids (1). Groups with specialist interest in each of these nutrients were assigned 1 or 2 micronutrients and followed a rigorous systematic review methodology to identify and assess the usefulness and responsiveness (under various circumstances) of a variety of biomarkers that have been used to assess the status of the micronutrients.

To enable us to produce a set of systematic reviews, joint methodology was developed (by SJF-T, LH, KA, and TD) and a protocol distributed to all of the participating groups. This written methodology was supported by excellent communication between groups (facilitated by KA), searches run centrally by an expert (KA) to ensure quality searches and to allow access to the core databases by all of the groups, development of a generic Microsoft Access database for data extraction for use by all of the review groups (KA), central development of the tables of characteristics of included studies and validity from the Microsoft Access databases, and 2 workshops (in Gran Canaria, Spain, November 2007, and Norwich, United Kingdom, February 2008) in which practical skills were taught and discussed. Detailed instructions were developed to allow participants to transfer their data from Microsoft Access to Microsoft Excel (Microsoft Corp, Seattle, WA), safeguard their databases, and use the Cochrane software, Review Manager version 4.2 (www.cochrane.org, Cochrane Collaboration; The Nordic Cochrane Centre, Copenhagen, Denmark; from inception to end of 2007), to collate data and run appropriate meta-analyses. Workshops focused on sharing progress on reviews; teaching skills, for instance, structured bibliographic database searching and meta-analysis; expert input on specific micronutrients; discussion of methodologic issues; and ensuring that all of the participants understood the scale of the task undertaken, the importance of careful (and shared) methodology, and networking activities. The results of the systematic reviews were presented at a final workshop in Montenegro (9 June 2008).

Our aim was to develop a generic methodology for systematically reviewing the usefulness of potential biomarkers of micronutrient status. The generic aim of each systematic review was to assess the usefulness of biomarkers of status of the relevant micronutrient in humans.

The primary question to be answered by each review was, Which measures (biomarkers) of status of the micronutrient appropriately reflect change in status of that micronutrient over relevant (stated) time scales?

Secondary questions that were addressed, where the evidence allowed, included the following:

1) What methods (established and novel) are currently used for measuring status of this micronutrient in humans?
2) For each status measure:
   a) In what population subgroups (ie, age, ethnicity, and micronutrient status at baseline) is there evidence that the measure accurately reflects change in status in populations?
   b) Over what time scale does the biomarker respond appreciably?
   c) What other factors affect the status measure (eg, age, body weight, genotype, illness, or sex)?
   d) What methodologies exist for analyzing this biomarker, and which are preferable in terms of accuracy, reproducibility, and cost?
   e) What evidence is there of circumstances in which the measure is not helpful?

METHODS

Methods reflected those used by the Cochrane Collaboration, as expressed in the Cochrane Handbook (5), where relevant, although the Collaboration is concerned primarily with systematic reviews of RCTs of health care interventions.
Inclusion criteria

Specific inclusion criteria were determined for each review, depending on the quantity and quality of literature available; where data were plentiful, studies were restricted to randomized controlled trials (RCTs), but where data were scarce, other types of studies, including controlled clinical trials (ie, nonrandomized studies with a control group) and before-after studies were included. Reviewers also were asked to identify which types of supplemental material were best absorbed and metabolized and to include only studies of these micronutrient supplements so that we could be sure that changes in the status of the micronutrient had actually occurred. Similarly, reviewers were asked to identify the minimum time period over which substantial changes in status may have been expected to occur and to use this as their minimum supplementation or depletion period. Our aim was to ensure that the status of the micronutrient really did change in the included studies, so that a lack of change in the status marker could truly be interpreted as a lack of response of the marker (at least in the specific circumstances) not as an absence of change in status.

Types of study

To be included, a study needed to fulfill all of the following characteristics:

1) Be an intervention study in humans, including a supplementation and/or depletion study (could be narrowed to RCTs where numbers of titles and abstracts identified by the initial search were >2000, suggesting that RCTs would be plentiful).

2) Report the micronutrient status in humans after 2 intervention periods, one in a control group and one in a supplemented or depleted group or report status at baseline and after supplementation or depletion.

3) Have a continued supplementation or depletion over a defined minimum time period; in some cases, the duration could be biomarker specific (an exception was made for instantaneous measures of status, eg, loading tests).

4) Use one of the defined forms of the micronutrient supplements.

Additional criteria could be added as necessary; for example, excluding participants with an illness or at a specific developmental state or excluding those expected to be replete at baseline.

Search strategy

Electronic searches

For each micronutrient review, Ovid MEDLINE (www.ovid.com; from inception to end of 2007), EMBASE (Ovid, www.ovid.com; from inception to end of 2007), and the Cochrane Library central database (www.thecochranelibrary.com) were searched. The search was for intervention studies of the relevant forms of that micronutrient using text terms with appropriate truncation and relevant indexing terms. Each search was in the form [micronutrient terms] and [intervention study terms] and [human studies]. The searches for all of the reviews were developed and run by experts in search methodology (KA with support and advice from LH). Development of the searches was iterative, using relevant indexing and text terms, exploding indexing terms where subterms were useful, and rarely focusing as a way to limit search size. As different search variants were tried, results were assessed to ensure that useful articles were not being lost and that new useful articles were being added. Centers were welcome to run additional searches themselves if they felt it was appropriate. The full search from the selenium review is shown as an example in Figure 1. The full Ovid MEDLINE search strategy for each review can be found at www.ajcn.org, and the strategies for the other databases were based on the relevant Ovid MEDLINE strategy.

Reference search

An additional Ovid MEDLINE search was conducted for each review, searching for reviews of methods of assessing the status of each micronutrient (KA). A maximum of 10 reviews was collected in full text and the reference lists checked. Studies that appeared to be intervention studies but that had not already been assessed for inclusion were collected. Reviewers also decided whether or not to check the reference lists of their included studies.

Experts

Each review group was encouraged to contact experts in their micronutrient with their list of included studies and protocol, explain the aims of the review, and ask if the experts knew of additional intervention studies that may have been useful within the review. Suggested studies were collected in full text and assessed for inclusion. This contact with experts was facilitated by experts who were invited to attend the February 2008 EURRECA workshop on Biomarkers of Status (6). This meeting was aimed at increasing interest in the reviews, improving dissemination among the scientific community, and helping to identify new approaches.

Data collection

Titles and abstracts were screened for inclusion by a single reviewer with independent duplicate assessment of ≥10% of the studies by a second reviewer. Only when it could be ascertained that titles/abstracts did not meet the inclusion criteria were they excluded. When a title/abstract could not be rejected with certainty, the full text of the article was obtained and evaluated further. The full text of all of the articles collected was screened for inclusion using an inclusion/exclusion form by a single reviewer with independent duplicate assessment of a random sample of ≥10% by a second reviewer. Where the first reviewer found some studies especially complicated to assess for inclusion (or data extraction), they could suggest some of the 10% to be duplicated but not the whole sample. Where the 2 reviewers disagreed, the study was discussed and a consensus decision reached where possible. If this was not possible, then a third reviewer was asked to arbitrate.

Although we encouraged reviewers to aim for 100% duplication of assessment of inclusion and data extraction for most review groups, with only one member of staff employed on the EURRECA project, this was unrealistic because both tasks are time consuming. However, the 10% duplication was important in ensuring that high-quality decisions were made and justified, raising issues about those decisions, challenging assumptions of the main reviewers, and ensuring that other team members better understood the review process.

For each included study, data were extracted onto the Microsoft Access database file; a separate database was kept for each review, modified slightly from the generic tool produced centrally.

For each included study, data were extracted onto the Microsoft Access database file; a separate database was kept for each review, modified slightly from the generic tool produced centrally.
for all of the reviews [see Figure 2 for a screenshot of the database as used for the selenium review (7)]. Data were extracted by a single reviewer with independent duplicate assessment of a random sample of ≥10% by a second reviewer. It was recommended that the form was piloted by using it to data extract several articles by each of the reviewers before the main body of data extraction began, discussed within the review team and/or with KA, then amended as appropriate before beginning full data extraction. Data extracted included bibliographic data, study design, location of study, study aim, intervention (including dose, form, duration, and whether oral, enteral, or parenteral) and control, status methods used, description of method, characteristics of population group (ie, age, sex, baseline status details, subpopulation details, disease status, etc.), sample size, mean status measure and variability for the intervention and control arms at specific time points, related health biomarkers, and/or status.

There was no formal policy on contacting authors for further or missing data, but most reviewers did contact some study authors, especially to collect data from articles that appeared large or robust but in which important details were missing. The ability to contact all of the authors of included studies was limited by reviewer time constraints.

Quality assessment

Quality assessment was completed as part of the data extraction (questions on validity were built into the Access database). The issues addressed for study quality included randomization, reasons for dropouts, methods for checking compliance and levels of compliance, similarity between intervention and control arms at baseline, and verification of supplementation dose used.

Data synthesis

Standard flow charts for keeping track of the number of studies assessed and included in the review were provided (see Figure 3, and for active examples see each of the reviews), and standard tables [tables of characteristics of included studies and validity of included studies, eg, Tables 1 and 2 in the copper review (8) or selenium review (7)] were generated centrally for each review team from their Microsoft Access database (or by the reviewers themselves where they preferred).

We extracted the number of participants included (and assessed) in each arm of each study and the following: 1) means and SDs of the baseline and change in status measure (from baseline to each assessed time point) for each measure in the intervention and control groups, 2) means and SDs of the baseline and final readings in the treatment and control arms, or 3) where there was no control arm, the baseline and later readings in the intervention group at each time point and for each micronutrient dose.

For each status measure, we first attempted to assess if there was an overall response to change in intake. To do this we carried out a meta-analysis of the change in status measure in the intervention group compared with the change in status measure in the control group for all of the included studies that assessed the measure. For each study we chose the time point with the longest duration of supplementation or depletion and in which the micronutrient was provided in an appropriate form and at the
highest dose (or depletion) concentration available (in studies where both supplementation and depletion occurred, the supplementation data were preferred as being most likely to alter status). Studies were subgrouped by their type (RCT, controlled clinical trial, or before-after study) and meta-analysis was carried out with Cochrane software, Review Manager version 4.2 (Cochrane Collaboration; www.cochrane.org), with random-effects meta-analysis. A statistically significant result indicated that the marker was indeed responding to supplementation and/or depletion. This was the answer to the primary review question. Levels of heterogeneity were noted (heterogeneity was considered significant where $P < 0.1$ on the chi-square test or $I^2 > 50\%$).

The secondary questions were answered as follows:

1) What methods (established and novel) are currently used for measuring status of this micronutrient in humans? The range of potential markers identified within the review answered this question; results were discussed narratively.

2) For each status measure:
   a) In what population subgroups (age, sex, genotype, and micronutrient status at baseline) is there evidence that the measure accurately reflects change in status in populations? Studies were subgrouped (where either the whole population was of one group or where data were provided by a subgroup) as follows: i) population—infants, children and adolescents, pregnant and lactating women, adults, postmenopausal women, the elderly, low income, and immigrants; ii) population—male or female; iii) population baseline status—low, moderate, or high (define for each status measure); iv) population—genotype (as appropriate); v) intervention—type of micronutrient; and vi) intervention—dose.

Mixed studies may have been represented in several subgroups; for example, a study with pre- and postmenopausal women included would have had data for the whole study represented in the “adults” group. If it also presented data for the postmenopausal women separately, then these data also could be represented in the “postmenopausal women” group.

Our ability to answer this question was limited by the scope of available data. For some reviews, eg, the omega-3 review (9), there were plenty of studies available to use for subgrouping for at least some outcomes, but for other reviews, eg, the copper review (8), studies were scarce and subgrouping only occasionally was helpful.

b) Over what time scale does the biomarker respond substantially? To assess the time scale of the biomarker, we plotted the time course for each study with data for at least 2 time points after baseline onto a single graph by using a moderate supplemental dose and plotting the biomarker measure (y axis) against the supplementation or depletion duration (x axis in weeks). Alternatively, we ordered studies in meta-analyses by intervention duration and assessed whether there appeared to be a trend over time. This process was restricted by our choice to limit studies to a minimum duration, so that for many biomarkers there was little evidence of change in status over time, because the changes would have occurred before outcome measures were assessed in our studies.

c) What other factors may affect the status measure (eg, body weight, genotype, illness, sex, etc)? Any data reported on other variables that affected the biomarker readings (including participant characteristics, eg, weight, sex, or other dietary intake or status) were collated and reported narratively. Where appropriate, formal subgrouping in
meta-analysis by the relevant factors (e.g., healthy compared with unhealthy or underweight compared with normal weight compared with obese) also was undertaken.

Overall, this question was not well addressed within the reviews because the studies collected did not record and assess this type of data. In the future, a different type of study may need to be collected to truly address this question.

d) What methodologies exist for analyzing this biomarker, and which are preferable in terms of accuracy, reproducibility, and cost? What techniques were used in assessing the biomarker? Relevant data were discussed narratively. Again, where relevant, subgrouping by different techniques in meta-analysis was carried out.

e) What evidence is there of circumstances in which the measure is not helpful? Any evidence of circumstances or populations in which the biomarker did not provide a good marker of nutrient status was discussed narratively.

Because there was a danger of categorizing some biomarkers as ineffective when there actually was a shortage of data, such that one would not expect a statistically significant effect size on pooling, the reviewers agreed that we would declare a biomarker effective (statistically significant pooled effect size; \( P < 0.05 \)) or ineffective (statistically insignificant pooled effect size; \( P > 0.05 \)) only where the pooling included \( \geq 3 \) studies and \( \geq 50 \) participants overall. When reporting results of the reviews in the abstracts, we asked that reviewers reported effectiveness this way and also commented on the study designs included (and other validity criteria where relevant). Where there were \(< 3 \) studies or \(< 50 \) participants but a statistically significant (\( P < 0.05 \)) effect was seen nonetheless, the biomarkers were likely to be effective, but if the effect was statistically insignificant (\( P \geq 0.05 \)), then it was stated that there were insufficient data to make a decision. Reviewers also were asked to reflect on the methodologic quality of their included studies as part of their assessment and description of the data.

RESULTS

The reviews discussed here are presented as individual articles in the remainder of this supplement (7–14). Each includes a short methodology section, which primarily reports the areas in which the methodology used in the review differs from that given here and sets out the specific criteria used for each review. Results are tabulated and also discussed as forest plots where appropriate.

DISCUSSION

The problem of identifying biomarkers for characterizing micronutrient status has gained increasing importance in the past few years. For example, the results of a clinical study designed to search for an optimal biomarker for fatty acid intake were reported in 2007 (15), but it can be assumed, with good reason, that there will be a limited number of clinical studies with the sole aim of evaluating the potential use of biomarkers of micronutrient status. Given these constraints, a more practical approach is to systematically review data published in studies originally designed to answer various clinical questions. Within the field of micronutrients, this meta-analytic approach has been used so far primarily for the investigation of the relation between intake and biomarker (16) or between biomarker and some clinical outcome parameter (17). Meta-analyses rarely have been undertaken to compare the efficacy of different biomarkers in the characterization of micronutrient status (18). Hence, our attempts to systematically review data by using a standard operating procedure is novel and will contribute to the selection of most appropriate biomarkers for micronutrients in future studies.

The EURRECA project has supported and trained scientists in a number of European countries to carry out high-quality systematic reviews to assess which biomarkers of micronutrient status are useful and under what circumstances. The time and support needed for groups who have not previously been involved in systematic reviewing to develop the skills needed has been substantial, but the evidence gained on micronutrient biomarkers will lead to the publication of a set of best practice guidelines on biomarkers to be used by researchers in future. This will allow increased understanding of previously conducted studies that link micronutrient status and health and will improve design of studies in future.
One generic issue for all of the reviews included here was that there were far fewer studies available for biomarker assessment than initially predicted, and the risk of bias of included studies was greater than expected. The latter varied between micronutrients, but for some there were very few high-quality studies, and a primary finding from this exercise was that further research is needed to assess the usefulness of many potential biomarkers. Even where there was good evidence that a biomarker is useful, there often were large gaps in our understanding of how well it works in particular population groups or in people of different baseline micronutrient status.

The selenium review (7) explicitly examined the problematic methodologic issues in the area. Data often were not presented in a way that made them easily accessible (eg, only in graphic form, so that graphs had to be expanded and measured to extract data, increasing the risk of inaccuracy); the numbers of participants involved at each stage and reasons for dropouts often were not mentioned or not clear; data often were presented without any information on variance (no SD, SE, or 95% CI); timing of sample collections was unclear; there were large differences between intervention and control arms at baseline (very common because many small studies were included) or only pooled baseline data were presented; data for crossover studies were reported as means and SDs of the 2 groups at the end of each arm but not in a way that made statistical use of the crossover element of the studies; and analytic methods often were ill described. These kinds of issues arose repeatedly in many of the reviews.

The novelty of using systematic review methodology to assess the usefulness of biomarkers of nutritional status does not lie in any extensive new methodology because the methodology is firmly grounded in high-quality international standards for systematic reviewing, eg, those from QUOROM (Quality of Reporting of Meta-analyses) (19), the Cochrane Collaboration (6), and the Agency for Healthcare Research and Quality, although the latter were published after we completed these reviews (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credibility to our methodology.

The importance of this methodology in developing a solid base on which to establish which biomarkers truly appear to respond to intervention (20). However, the method of assessing the validity of a diagnostic test where there is no gold standard is much less explored, and our approach in extending this to assessing the usefulness of biomarkers explores new territory (4). The methodology was partially validated in the vitamin D review (10), from which one could argue that a gold standard biomarker already exists in the form of 25-hydroxyvitamin D (although it has not been rigorously tested). In this review, 25-hydroxyvitamin D was clearly a good marker of vitamin D status, both confirming the marker and lending credi...