British Journal of Nutrition (2009), **102**, S1–S2

Foreword

The community of epidemiologists and nutritionists interested in diet and health will be grateful to the team that has provided the papers in the present volume reviewing the validity of dietary assessment methods for use in epidemiologic studies. This has been a huge task because the number of validation studies has become enormous (over 1000 are found in the NIH online registry: www.appliedresearch.cancer.gov/cgi-bin/dacv/index.pl); and summarising the complex and heterogeneous data in these reports is a daunting challenge. The papers in the present review have been grouped by population type and class of micronutrients, which will be helpful for those planning a study. This analysis was conducted as part of the European Commission's EURopean micronutrient RECommendations Aligned Network of Excellence, (EURRECA) whose aim is to address the harmonisation of micronutrient recommendations between European countries, as well as to understand how nutritional information including requirements and intake adequacy is processed among different population groups. The present review focuses on the fifteen priority micronutrients and *n*-3 fatty acids targeted by EURRECA.

As a summary of validity, the authors have used the correlation coefficient comparing the method being evaluated with a presumably superior method or a biomarker. This coefficient is appropriate as it takes into account both the error in measurement and the between-person variation in the dietary factor. In calculating average correlation coefficients across studies, they have weighted the results from each study according to a new quality score. The elements of this score generally make sense, in particular the size of the study is very important and this might be given even more weight. One component, giving greater weight to studies with more heterogeneous populations, should be used cautiously because heterogeneity in variables such as sex or age can magnify between-person variation in diet and thus also the correlation coefficients. Because these variables would routinely be controlled in an epidemiologic analysis, the results of the validation study could be misleading if correlation coefficients were also not controlled for these variables. In principle, it is most important that the population of a validation study be representative of the population in which the dietary assessment would be used, and the variables that will be adjusted for in the ultimate epidemiologic analyses should also be adjusted for in the validation study. Interestingly, the weighting of results from different validation studies did not materially affect the average correlation coefficients, probably mainly because the quality scores did not vary greatly. Nevertheless, this was a useful process as it documents the robustness of the findings, and the quality score can be useful to those planning or analysing a validation study. The authors have also provided ranges of correlation coefficients that can be considered poor, acceptable, good and very good; these can serve as useful adjectives in future reports. If the focus is on dietary intake, these should probably not be applied to correlations with biomarkers because they are influenced by absorption and metabolism in addition to intake.

Broadly, the conclusions of this massive review support the value of FFQ in epidemiologic studies as the large majority of correlations were in the acceptable to good range (r 0·30–0·70), especially when the comparison was with measures of long-term intake. This general finding is consistent with the results from applications of FFQ, which have documented reproducible findings for many diet and disease associations. These findings have played an important role in modifying dietary recommendations (e.g. from a focus on total fat to fat quality), food manufacturing (e.g. elimination of *trans* fatty acids), formulation of supplements (e.g. reduction in retinol) and medical practice (e.g. eliminating the avoidance of dietary Ca to prevent nephrolithiasis).

Although the documentation of the validity of dietary assessment methods has been an essential step forward, we cannot be satisfied or rest at this point. First, diets are moving targets because they change with the evolution of personal preferences and shifts in the food supply; thus, a questionnaire with good validity 20 years ago may no longer perform as well. More importantly, all of the dietary assessment methods are imperfect; in this review, only rarely were correlations in the very good range ($r > 0{\cdot}70$). The overall evidence suggests that there may be a ceiling to the validity of FFQ, although the ceiling is in part due to the imperfect nature of the comparison methods. The present review did document that questionnaires with more items produced modestly higher correlations, but the differences were not large and gains from further lengthening of questionnaires will be limited. One approach to increasing validity for the assessment of long-term intake is to use the average of repeated administrations of a dietary questionnaire; for example, when we used the average of three questionnaires over an 8-year period, we did find correlations in the very good range. Repeated dietary assessments over time, whatever the method used, will be important in studies of chronic disease, especially cancer, because of changing diets and long and variable latencies. The low cost and burden on the subjects with FFQ make these repeated measures feasible. In addition to repeated measures, some have suggested that the validity of dietary assessments might be increased by combining the results of short- and long-term assessment methods, but this has not been formally evaluated.

Although the number of validation studies is now large and the general validity of dietary assessment methods has been established, the need for these studies has not decreased. In addition to providing evidence on overall validity, these studies are essential for quantitative corrections for measurement error in epidemiologic applications of dietary

questionnaires in specific populations. Also, validation studies can be used to improve dietary questionnaires, particularly when analysed at the level of individual foods rather than nutrients. In these validation studies, diet records presently remain the best gold standard if this is feasible, although it is possible that web-based 24-h recalls might be found to play an important role. The integration of biomarkers in validation studies is highly desirable whenever possible. This can provide a measure of relative validity of the questionnaire and gold standard methods and, using the method of triads, can provide an estimate of the correlation of each method with true intake. In such studies, it is important to include repeated samples of the biomarker to account for within-person variation, which can create the false appearance of correlated errors between diet assessment methods. Furthermore, the timing of collection of biomarkers is important; if they are collected close in time to a short-term method, such as a 1-week record, this can exaggerate their correlation.

The papers in the present volume provide a valuable milestone on our journey to understand the impacts of diet on health and disease. These effects have already been found to be far greater and more diverse than anyone anticipated several decades ago, and we have come to understand that they operate throughout the life cycle. As we move forward, our progress will continue to depend heavily on dietary assessment methods, and underpinning this is the need to further refine and document the validity of our methods. Again, those of us engaged in research using dietary assessments will be highly appreciative of the great efforts of our colleagues who have produced this volume documenting our progress to date.

Walter Willett

Boston
September 2009

doi:10.1017/S0007114509993102