CHAPTER 7:CONCLUSIONS AND RECOMMENDATIONS

"*Ever since records of diet were first kept, we had an unerring but misguided confidence in our ability to measure our own food consumption - until, that is, nutritional epidemiology revealed to us the error of our ways*."

These words of Nelson [154] summarize the essence of this study and confirm that the contribution of a dietary validation study primarily lies in describing the nature and magnitude of the error structure of the measurement of diet. In this regard the present study additionally indicated how the error structure of a FFQ type dietary fat screener might differ for subgroups and different food categories, how it might be affected by the choice of reference method and statistics used and, finally, how the aim (outcome measure) could determine the conclusions. Thus, in the present study the measurement of grade six learners with the dietary fat screener

- was internally consistent
- was reproducible in the girls, but random error characterized the boys' assessment
- did not agree with screening performed by parents, with systematically lower overall reporting by the parents
- had a different error structure for the different food categories in the tool and for the various quantification elements (portion size and frequency of intake and the resultant scores)
- showed selective agreement with measures of high fat intake (PFE, PSFE, and cholesterol intake) from a three-day food record. The agreement depended on the statistical analysis and the outcome measure used for comparison
- was highly sensitive, but could not achieve good specificity simultaneously.

Nutritional exposure can be defined on three different levels of measurement: as food, nutrients and biomarkers.[162] The test method in this study measured on the food level, the primary level of exposure, but it was intended to relate to PFE, PSFE and cholesterol intakes, all on the nutrient level. Thus the research design of this study, where the test-retest reproducibility and screener completed by the parents (reference method 2) addressed the food-level agreement, and the three-day food record (reference method 1) focused on nutrient-level agreement, was considered an appropriate and strong approach, also incorporating the triangulation principle. The inclusion of anthropometric and design quality control specifically with respect to the food record further strengthened the internal validity of the study. Complementing the design with a biomarker, with a completely independent error structure, could have 'perfected' the comparative validation.[317]

The validation process is sometimes considered to relate to the measurement and not to the method from which the measurement is derived,[154] meaning that validation considers the context within which dietary assessment methods are used. Consequently all conclusions derived from this study primarily relate to the given context: A public, urban, middle-class primary school in South Africa, accommodating mainly white, Afrikaans speaking, children. Nevertheless, the application of scientific design principles and quality control measures within the quantitative domain of investigation do allow some generalizations.

The test method is in essence a FFQ. Drewnowski [98] has argued that FFQ estimates do not appear to be based on memory for actual events, but that food frequencies are inferred, as opposed to remembered, and are based on some subjective image of a usual or typical diet. As such, he argues, FFQ's cannot be 'validated' since they measure primarily predispositions and attitudes. They can thus not be compared to instruments that capture actual behaviour in the short term. Even though this is not a commonly held perspective in the nutrition literature, the reasoning does provide some additional explanation to the limited agreement between the test method as a FFQ and the three-day food record in the current study.

A similar note is struck when the question arises whether a 'usual / habitual / typical' diet exists. Is it only a construct in the minds of dietitians and nutritionists or is it an objective entity? Whilst it has been operationally defined (for example the average in a long series of food records [157]) ever since the mentioned study of Huenemann and Turner (see review of literature)[238] the existence of a 'usual' diet has from time to time been debated. Thus, again, should a 'usual diet' not exist, validation is either not possible or, at best, construct validity (in contrast to criterion validity) would be an option.

If, however, a 'usual diet' does exist, then the three-day food record could rightly be criticized as being an inappropriate reference method for validating the test method. Furthermore, given the fact that the three-day food record is an imperfect measure of dietary intake, it would have been ideal had the test-retest reproducibility thereof also been determined. Only then, when the variability (random error structure) of the reference method is also known, can more definitive conclusions about potential relationships be drawn. This, of course, also applies to the screener completed by the parents.

The above has implications for the selection of the sample. The lack of due consideration of sample size for a validation study has been addressed (for example references [185, 166, 318]) but

Keller et al,[184] have suggested that, in general, a sample size of 100 should be adequate. Representativeness and composition of the sample also demands attention: The high prevalence of high fat intakes observed in the present study may be a true reflection of reality in the study group, but it may have reduced variability and may have affected some statistics for example some of the statistical indicators of agreement such as the kappa statistic, are affected by prevalence.[305] By trying to obtain a population with trait prevalence near 50% this could be addressed.[305] However, 'manipulation' of the sample affects the relevance of other indicators, such as positive and negative predictive value.

There are many ways of analyzing and expressing reliability and validation studies [319] and the most appropriate statistical analysis has not been established.[28, 42, 122, 154, 304] This also appears to be true for methods specifically aiming to measure fat intake: Simon et al [320] compared standard methods based on a null hypothesis of no agreement between instruments (FFQ, 24h recall and three-day food record) and an alternative method of analysis based on a null hypothesis that the instruments should be in agreement. They conclude that the latter is more appropriate. Jones [185] reviewed and critically appraised the scientific merit and methodologies used for nutritional screening and assessment tools and concluded, "no one tool is judged to have been published with sufficient care given to its application, development and evaluation." This was confirmed by Dennis et al [147] specifically in respect of the design and reporting of FFQ. Their scoring method and the Consensus Document on the development of FFQ [114] should provide more design and analysis guidance for the future. Close collaboration with bio-statisticians seems to be indicated. In this study a variety of well-established and novel statistical analyses were reported in order to provide a comprehensive picture and enable comparisons to other studies.

In conclusion, the dietary fat screener should not yet be used in grade six learners as a sole assessment method within the South African primary health care context, given the country's present, overall nutrition profile [321] and available health care resources. Screening is an inexact science. For that reason ethical and legal responsibility should rest on those administering it to inform the public of a particular tool's discriminatory properties.[313] The data obtained in this study suggest that if intervention or monitoring of dietary intake trends are to be based on only the dietary fat screener, further developments and / or modifications to increase its validity are needed.

Possibilities for structural changes to the tool include re-scrutinizing the item list (for example critically evaluating the role of eggs as food category or adjusting the relative weight of the

individual food items), ideally based on a food consumption survey in a (nationally) representative sample including the target group. Alternatively, statistical modeling of expert judgment matrices could be used to obtain an indication of the relative importance of the individual food categories. Critical investigation of alternative scoring principles is another avenue to investigate. It is furthermore possible that including more covariates (for example BMI or gender) into the logistic regression could result in improved discrimination abilities of the screener. Checking the assumptions of the nutrient data for South Africa could also be useful. Finally, deeper insight into the cognitive processes of dietary assessment of children in the target group could be helpful, but without sacrificing the inherent strengths of a screening tool.

In spite of the identified limitations, given the high prevalence of high fat intakes in the target group (and thus the risk for developing CNCD), the tool may in the interim be very valuable for creating awareness of high fat intakes. The food-based nature of the screener should be a practical starting point for providing needs-driven nutrition education and anticipatory guidance (similar to the approach used with the REAP and WAVE tools),[34] within population-wide promotion of the dietary guidelines.

Once measurement with the dietary fat screener has been shown to be reproducible and valid in this target group, expansion of the target group and context in line with the outline in Figure 1.1 is recommended. In doing so a greater segment of the South African society may eventually benefit from the research. This should be followed up with randomized controlled trails of screening using the designs suggested by Barrett et al [309] in order to ascertain cost-effectiveness of the process in the South African context.