European Journal of Clinical Nutrition (2004) 58, 1297–1305

0954-3007/04 $30.00
www.nature.com/ejcn

ORIGINAL COMMUNICATION

Validity and reproducibility of a computerised tool for assessing the iron, calcium and vitamin C intake of Belgian women

C Matthys[1]*, I Pynaert[1], M Roe[2], SJ Fairweather-Tait[2], A-LM Heath[3] and S De Henauw[1]

[1]*Department of Public Health, Ghent University, Ghent, Belgium;* [2]*Institute of Food Research, Colney, Norwich, UK; and* [3]*Department of Human Nutrition, University of Otago, Dunedin, New Zealand*

Objective: To determine the relative validity of a newly developed iron intake assessment tool, designed specifically to assess iron, calcium and vitamin C intake.
Design: Estimates of iron, calcium and vitamin C intake from a computerised iron intake assessment tool compared with those from 11-day estimated dietary records.
Setting: Region of Ghent ($N = \pm 225\,000$), a city in Flanders, the Dutch-speaking part of Belgium.
Subjects: In all, 50 women aged 18–39 y, participating in a large-scale epidemiological study on iron intake and iron status.
Main results: Mean dietary iron intake from the 11-day food record, the unadjusted dietary iron intake assessment tool and the adjusted tool was, respectively, 10.5 ± 2.7, 10.4 ± 4.3 and 9.6 ± 2.9 mg. For the different nutrients, the correlation coefficients vary from 0.45 to 0.60 for adjusted intake. The mean difference of iron intake by the two methods (0.8 ± 2.9 mg) did not differ significantly from zero. The new method correctly classified 38% (iron), 38% (calcium) and 58% (vitamin C) of the subjects to the correct tertile. The correlation coefficients ranged from 0.48 for adjusted vitamin C intake to 0.73 for adjusted calcium intake between two administrations.
Conclusion: The newly developed instrument can be used to assess mean group intakes of iron, calcium and vitamin C in women consuming a Western diet. However, since the ranking capability of the new tool is rather weak, further refinement of the tool is required to produce a robust method for assessing iron, calcium and vitamin C intakes of individuals.
Sponsorship: This research was supported by the Belgian National Fund for Scientific Research/Flemish Division (Fund No G.0152.01).
European Journal of Clinical Nutrition (2004) **58,** 1297–1305. doi:10.1038/sj.ejcn.1601965
Published online 31 March 2004

Keywords: iron; dietary assessment method; dietary record

*Correspondence: C Matthys, Department of Public Health, Ghent University, UZ – 2 Blok A, De Pintelaan 185, Ghent B-9000, Belgium.
E-mail: christophe.matthys@UGent.be
Guarantor: C Matthys.
Contributors: CM designed the study, analysed and interpreted the data and wrote the initial draft of the manuscript. IP assisted in collecting data and assisted in the preparation of the manuscript. MR assisted in the design of the computerised questionnaire and assisted in the preparation of the manuscript. SJF-T assisted in interpreting the data and in preparing the manuscript. A-LMH assisted in the design of the computerised questionnaire and in the preparation of the manuscript. SDH designed the study, interpreted the data and assisted in the preparation of the manuscript.
Received 25 September 2003; revised 23 January 2004; accepted 3 February 2004; published online 31 March 2004

Introduction

Iron intake and iron status are important nutritional topics, not only in the developing world but also in developed countries. Both iron deficiency and iron overload are considered to be important public health issues, affecting different subgroups of the population (Spanjersberg & Jansen, 2000).

In a recent dietary survey in a small area in Flanders, the Dutch-speaking part of Belgium, it has been observed that the dietary intake of iron in adolescent girls was low and, at a population level, below recommended intakes (Matthys *et al*, 2003). Analogous observations have been reported in other European countries (Michaud *et al*, 1989; Belton *et al*, 1997; Roma-Giannikou *et al*, 1997; Cruz, 2000; Rolland-Cachera *et al*, 2000; Samuelson, 2000). It is however, not clear to what extent this observation of low iron intake is also translated into low iron status in these young girls, and to what extent this problem would persist during the reproductive period of life of adult women and how it possibly affects their pregnancies and offspring. This issue is currently the subject of an ongoing study on iron intake and iron status in adult (pregnant) women, carried out by Ghent University.

For the purpose of this large-scale epidemiological study, a new dietary intake questionnaire has been developed and validated. This new dietary assessment tool was intended to measure the intake of total iron with a high precision, but also to allow for corrections in terms of bioavailability of iron as influenced by the presence of dietary enhancers (vitamin C) (Lynch & Cook, 1980) and inhibitors (calcium) (Hallberg *et al*, 1991) of iron absorption in the gastrointestinal tract. The present paper focused on two modifiers, namely vitamin C and calcium. However, one must be aware of the presence of other dietary modifiers in the iron absorption mechanism (eg tea (Disler *et al*, 1975)).

The questionnaire was developed in collaboration with the Institute of Food Research in Norwich (UK) and was to a large extent inspired by an existing, validated food frequency questionnaire (FFQ) developed at the University of Otago, Dunedin, New Zealand (Heath *et al*, 2000). The aim of the present study was to validate the newly developed Belgian version of the iron intake assessment tool (IIAT), designed specifically to assess iron, calcium and vitamin C intake, in women aged 18–39 y.

Materials and methods

Participants

Participants in the validation substudy were recruited from the pool of nonpregnant participants in a large epidemiological study of iron intake and status. This project included in total over 800 women aged 18–39 y, randomly selected from the population register of the region of Ghent, a medium-sized city in Flanders, with a population of 225 000. Subjects were excluded if they were not familiar with the Dutch language. The overall routine set of investigations included a food questionnaire (IIAT), a 2-day food diary, a general sociodemographic questionnaire and a fasting blood sample.

Women were invited to participate in the study by mail. On receipt of written informed consent, they were invited for a computer-assisted dietary assessment session at the Department of Public Health of Ghent University. These sessions were organised for groups of, on average, 12 subjects, guided and supervised by dietitians with extensive experience in conducting dietary interviews and quantifying and coding foods. At the end of each session, participants were asked to complete an estimated 2-day food record as a complementary part of the study and were invited to participate in a validation study by completing an 11-day estimated food record for the purpose of the validation substudy. A total of 69 women volunteered to take part in the validation study. At 1 month after the completion of the 11-day dietary records all participants were invited to complete the computerised IIAT for a second time.

From the original 69 volunteers, 16 did not complete the full 11 days and one was pregnant. Another two women were excluded from the validation sample due to computer technical problems during their completion of the questionnaire. Therefore, the final validation data set included 50 women. Of these women, 47 also participated in the reproducibility test. Another four subjects—originally not volunteering for the 11-day food record—completed the computerised IIAT twice. The extra four volunteers were selected because the authors wished to have at least 50 subjects for the reproducibility analysis. Finally, 51 subjects were included in the reproducibility test.

IIAT

The IIAT is a computerised questionnaire based on the general concept of the diet history, adapted to a self-administered setting. A pretesting phase among people without special computer skills or specific nutritional knowledge was performed to optimise user-friendliness. The test showed that the participants were comfortable using the program, under supervision, once they had been taught how to 'point and click' a computer mouse on food items.

At the beginning of the session, all participants received a standardised audiovisual explanation on how to use the computerised IIAT.

The architecture of the assessment tool and the underlying software contains three main parts: an estimate of the overall meal frequency, a meal-based diet history and a checklist of specific food items. The first part was an inventory of the number of times per week people ate breakfast, lunch, dinner, and morning, afternoon and evening snacks (overall meal frequency). A week in this context means a normal week during the previous month.

The second part (the meal-based diet history) allowed respondents to report their individual 'usual' pattern of food intake in an interactive way on the computer screen. For this purpose, all meals from an average week appear separately on the screen and invite respondents to choose from 16 food groups, containing a total of 209 food items. In this way respondents could describe each meal and snack eaten. For each chosen food item, a serving size was automatically suggested by the computer. The participants could multiply or divide the proposed serving size in order to match it to their own usual serving size. The proposed serving sizes were based on the Belgian standard guide on household weights and measures (Health Council Belgium, 1997). The issue of portion size was included in the audiovisual introduction. Once each individual meal was completed, the subjects were asked to report the exact frequency of consumption for that specific meal (individual meal frequency) (Tylavsky & Sharp, 1995). The third part of the tool is a checklist of 77 food items that appears on the screen when the subjects have entered all meals. This list contains food items available in

Belgium that do not necessarily contribute substantially to iron intake on a population level, but could, however, due to a high content of iron, vitamin C or calcium, substantially influence the reported iron intake on an individual level. These food items could be added to any meal. Finally, the participants were shown the overall meal frequency as originally reported (Part one) and could adjust it when necessary.

An adjustment factor was introduced in the calculation of the iron intake. The aim of the factor is to investigate whether the respondents were able to estimate the relative frequency of consumption of specific foods better than the absolute frequency of their consumption. The adjustment factor was equal to the overall meal frequency divided by the sum of the individual meal frequencies, and was calculated for each meal and snack category. The 'adjusted' nutrient intakes were calculated by multiplying the individual meal frequencies reported for each meal by the corresponding adjustment factor (Heath *et al*, 2000).

Food groups and individual food items compiled in the meal-based history were all food items identified as contributing substantially to the overall iron intake, or containing a dietary component that affects iron absorption, in the Flemish meal pattern. In order to determine the food sources that, on a population level, contributed 95% of the intake of the dietary components of interest, two recent epidemiological surveys were used. Both studies, one in adolescents (1997) (Matthys *et al*, 2003) and the other in pregnant women (1996) (De Vriese *et al*, 2001), used the same dietary methodology, namely a consecutive 7-day estimated diary. Foods with a very high iron content that are part of the Belgian dietary pattern were also included in the food list.

The subject-specific average total intake of each dietary component was computed by the sum of the products of the nutrient content of the food items in each meal and the individual meal frequencies, and subsequently divided by seven. The food composition data for total iron, calcium and vitamin C were based on the following tables: Dutch food composition tables (NEVO, 1996, 2001), the Belgian food composition tables from 1995 and 1999 (NUBEL, 1995, 1999) and the McCance and Widdowson food composition table (Holland *et al*, 1991). The Dutch food composition table was the main source of nutrient content data, the Belgian version was used for typical Belgian food items, and the English table was used when data were missing in the Dutch or Belgian version.

The completion time of the whole procedure (explanation of the IIAT (20 min), completing the IIAT (60 min), explanation of the 2-day food diary, measuring height and weight) was between 90 and 120 min.

Reference method

The estimated food record was chosen as the reference method and a semistructured diary was used. Special attention was given to the issue of the estimation of portion sizes of food items and this was demonstrated with a number of standardised examples.

In all, 11 days of estimated dietary record were collected. The number of days is based on the formula of Beaton *et al*. (1979), and the within-person coefficient of variation of iron intake for women based on Willett's data (1998). The subjects started to record a food diary the day after the first completion of the computerised questionnaire. The 11 record days were not consecutive because of the high burden for the respondents and to minimise recording fatigue. The recording days were grouped in blocks of two or three consecutive days, each separated by 1 week. In this way the recording days were spread over a period of 1 month and included all days of the week.

In the diaries, days were truncated into six eating occasions, namely breakfast, lunch, dinner and snacks (divided into morning, afternoon and evening snacks). Information on the type (including brand names) and amount of food consumed was collected through an open entry format. After completion, the diaries were processed into food quantities and codes by experienced dietitians on the basis of a standard protocol, including a standard manual on food portions and household measures (Health Council Belgium, 1997). The same food composition tables as in the IIAT were used. Calculation of nutrients was done by means of a nutritional software package developed by Unilever in the Netherlands (Unilever, 1992). The average energy intake and nutrient intakes were calculated as the mean of the 11-day intake period.

Members of the research unit measured height and weight of all subjects when they completed the IIAT. The measurements were carried out according to the standardised method as described in WHO, Technical Report Series 854 (World Health Organization, 1995).

Statistical analysis

Statistical analysis was performed with the SPSS software (SPSS, 1999). A *P*-value of 0.05 was used as the threshold for significance. Tests for normality were performed using a Kolmogorov – Smirnov test. The intake of some nutrients was normally distributed. The difference between means was tested using the paired *t*-test and Wilcoxon matched-pairs signed rank test. Associations between nutrient intakes by each dietary method were described using Spearman rank correlation coefficients, because some nutrients were not normally distributed and the ranking of the individual in the current study was of particular importance. The Bland – Altman method (1986) was used to assess the agreement between the methods across the range of intakes. As the aim of quick methods of dietary assessment is to permit ranking so that subjects at the extremes of the distribution are correctly classified, both the IIAT and dietary record results were divided into tertiles in order to examine whether subjects were classified in the same or different categories by the two methods (Sempos, 1992). The results permit an assessment of the proportion of subjects who were classified correctly (Cade *et al*, 2002). The results can be reported as an exact agreement and extreme misclassification. Agreement has also been assessed using the weighed κ statistic. Values of κ over 0.80 indicate very good agreement, between 0.61 and 0.80 good agreement, 0.41–0.60 moderate agreement, 0.21–0.40 fair agreement and <0.20 poor agreement (Altman, 1991). Actual values for surrogate categories, as described by Willett (1998), were calculated by grouping subjects in tertiles on the basis of the surrogate method, in this case the IIAT. The 'true mean value' was calculated for each tertile using intake determined by the 11-day estimated dietary record. This gives an indication of the 'true' intakes that are indicated by the IIAT tertiles. These categories were compared using one-way ANOVA. The reproducibility was assessed using correlation coefficients (Spearman); and paired *t*-test or Wilcoxon matched-pairs signed rank test to test whether there was a significant difference between the nutrient intakes reported at the first and second administration. The repeatability was also tested by calculating the mean and standard deviation of the differences between two administrations of the IIAT (Bland & Altman, 1986).

There were sufficient participants to be able to detect a significant difference between mean iron intakes of 1.55 mg with a power of 80% and a significance level of 0.05.

The study was approved by the Ethical Committee of the Faculty of Medicine and Health Sciences of Ghent University.

Results

Subjects included in the final validation data set ($n=50$) had a mean age of 31 ± 6 y (range 19–40 y) and a mean BMI of $23.9\pm4.2\,kg/m^2$ (range $17.6–36.7\,kg/m^2$).

Mean dietary iron intake was 10.5 ± 2.7 mg from the food record, 10.4 ± 4.3 mg from the unadjusted dietary iron intake assessment tool and 9.6 ± 2.9 mg from the adjusted tool. There were no significant differences between the mean intakes of vitamin C or iron according to the different instruments, although the adjusted tool agreed less well with the dietary record than the unadjusted tool. The mean intake of calcium estimated by the adjusted tool was significantly lower than the dietary record (see Table 1).

The correlation coefficients between the 11-day dietary record and the IIAT (unadjusted and adjusted) are shown in Table 2. The Spearman correlation coefficient varies from 0.45 to 0.60 for adjusted intake. The adjusted iron and calcium dietary intakes have stronger correlations with the food records than the unadjusted intakes.

The mean difference between the 11-day dietary record and the adjusted IIAT (see Table 2) was 0.8 ± 2.9 mg. Hence, 95% of the individual iron intakes as assessed by the adjusted IIAT varied between 6.7 mg above and 5.1 mg below the estimated dietary record value. This is graphically shown in a Bland and Altman plot. A visual inspection of the data suggests that the difference between the two methods remains stable over the whole range of mean intake. The plot is given for adjusted and unadjusted iron intake in Figure 1.

Table 3 shows the percentages of subjects correctly classified and classified in extreme tertiles by the IIAT into estimated dietary record tertiles. The adjusted IIAT classified 38% of the people into the correct tertile for iron intake, while 6% are grossly misclassified. For vitamin C better results were obtained, the adjusted tool classified 58% of the participants correctly and only two subjects (4%) were grossly misclassified. Crossclassification of calcium was similar to iron. The weighed κ statistic varied from 0.20 for adjusted calcium intake to 0.48 for adjusted vitamin C intake.

Table 4 shows the actual value for surrogate categories comparing the IIAT tertiles with the estimated dietary records tertiles. The actual values show a progressive increase over the surrogate categories. Significant differences were observed between the extreme tertiles for all dietary components. Significant differences between extreme tertiles suggest that the IIAT can distinguish groups at extreme levels of intake.

The IIAT was completed twice by 51 subjects to assess the instrument's repeatability. No significant differences between the mean dietary component intakes assessed by the two administrations were established. The correlation coefficients ranged from 0.48 for adjusted vitamin C intake to 0.73 for adjusted calcium intake (see Table 5). The Spearman correlation coefficient of adjusted iron intake was 0.66. The mean difference between the adjusted iron intake reported in the two administrations of the IIAT was 0.2 ± 2.5 mg. This

Table 1 Mean (s.d.) intakes of iron, calcium and vitamin C estimated by the dietary record and the iron intake assessment tool ($n=50$)

	Estimated dietary record	*Unadjusted iron intake assessment tool*	*Adjusted iron intake assessment tool*	P-*value dietary record vs unadjusted IIAT*[a]	P-*value dietary record vs adjusted IIAT*[b]
Iron (mg)	10.5 (2.7)	10.4 (4.3)	9.6 (2.9)	0.833*	0.051*
Calcium (mg)	866 (334.8)	809 (388.2)	738 (268.8)	0.296*	0.003*
Vitamin C (mg)	106 (51.4)	110 (67.9)	106 (71.8)	0.670*	0.972*

[a]Difference between estimated dietary record and unadjusted iron intake assessment tool (IIAT).
[b]Difference between estimated dietary record and adjusted iron intake assessment tool (IIAT).
*Paired *t*-test to test difference between means.

Table 2 Correlation coefficients and mean differences (s.d.) between the dietary record and the iron intake assessment tool ($n = 50$)

	Spearman's correlation between dietary record and		*Difference (mean (s.d.)) between dietary record and*	
	Unadjusted IIAT[a]	*Adjusted IIAT*[a]	*Unadjusted IIAT*[a]	*Adjusted IIAT*[a]
Iron (mg)	0.40	0.45	0.1 (4.3)	0.8 (2.9)
Calcium (mg)	0.48	0.52	57.4 (383.8)	128.1 (289.3)
Vitamin C (mg)	0.63	0.60	−3.4 (55.1)	−0.3 (63.0)

[a]IIAT = Iron Intake Assessment Tool (IIAT).

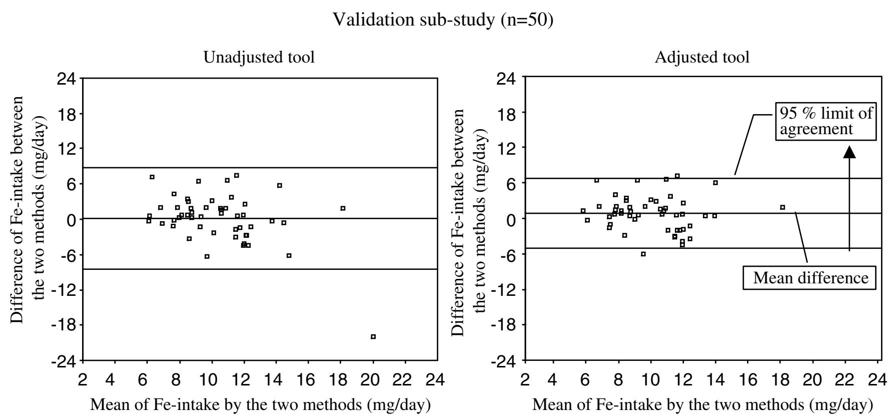

Figure 1 Differences between the mean iron intake for 11-day dietary record and IIAT (mg/day) in the validation subanalyses.

Table 3 Crossclassification of the dietary record and the iron intake assessment tool tertiles. ($n = 50$) (percent (95% confidence interval for proportion))

	Percentage classified in		
	Same tertile	*Opposite tertile*[a]	*Weighed κ-coefficient*
Chance	*33*	*22*	
Iron unadjusted IIAT[b]	36 (23–49)	8 (0–16)	0.17 (−0.02–0.37)
Iron adjusted IIAT[b]	38 (25–51)	6 (0–13)	0.22 (0.02–0.42)
Calcium unadjusted IIAT[b]	42 (28–56)	8 (0–16)	0.25 (0.05–0.44)
Calcium adjusted IIAT[b]	38 (25–51)	8 (0–16)	0.20 (0.01–0.40)
Vitamin C unadjusted IIAT[b]	60 (46–74)	6 (0–13)	0.48 (0.28–0.67)
Vitamin C adjusted IIAT[b]	58 (44–72)	4 (0–9)	0.48 (0.28–0.67)

[a]Opposite tertile of the first tertile is the third and *vice versa*.
[b]IIAT = Iron Intake Assessment Tool.

means that 95% of repeated IIAT adjusted iron intakes fell between 4.8 mg below and 5.2 mg above the first iron mean intake. A Bland and Altman plot is given for adjusted and unadjusted iron intake in Figure 2.

Discussion

The development and the validation of the new dietary instrument was in the context of an ongoing large-scale epidemiological study investigating the relationship between iron intake and iron status in young adult women. The new tool was designed to measure the intake of total iron, calcium (an iron absorption inhibitor) and vitamin C (an iron absorption enhancer). Other enhancers and inhibitors (eg tea) could be estimated by the current version of the assessment tool. The consumption of tea was not included because tea is rarely consumed in this study population. Validation studies were carried out to measure the extent to which a method actually measured the aspect of the diet for which it was designed. The issue of how to assess the validity of a new dietary assessment method is frequently debated (de Groot *et al*, 1998). Different techniques are suggested by different authors. Nelson (1996) recommended the use of six techniques, while Willett (1998) recommended seven approaches to evaluate dietary questionnaires. In this study, six methods to assess the relationship and agreement between the newly developed IIAT and estimated dietary records were used, namely paired comparisons of means, correlation analysis, Bland–Altman analysis, crossclassification, κ-statistic and actual values for surrogate categories. Based on the comparison of means, the IIAT can estimate mean intakes for iron, vitamin C and unadjusted calcium intake, although there is a tendency for the newly developed tool to underestimate the intake measured by dietary record. It is not possible to determine whether this is due to under-reporting because the new dietary instrument is not

Table 4 Use of actual values for surrogate categories to compare the iron intake assessment tool with the estimated dietary records ($n=50$)

		Mean dietary record intake			Statistical Test		
	Tertiles defined by	*T1*	*T2*	*T3*	*T1/T3*	*T1/T2*	*T2/T3*
Iron (mg)	Dietary record	7.8	10.3	13.4			
	Unadjusted IIAT[a]	9.4	10.8	11.3	0.013	0.075	0.772
	Adjusted IIAT[a]	9.2	10.9	11.4	0.003	0.050	0.654
Calcium (mg)	Dietary record	572	815	1261			
	Unadjusted IIAT[a]	722	835	1052	0.003	0.326	0.090
	Adjusted IIAT[a]	721	844	1025	0.006	0.249	0.174
Vitamin C (mg)	Dietary record	57	97	168			
	Unadjusted IIAT[a]	77	94	149	0.000	0.067	0.002
	Adjusted IIAT[a]	72	99	149	0.000	0.023	0.005

[a]IIAT = Iron Intake Assessment Tool.

Table 5 Comparison of two iron intake assessment tools administered 2 months apart ($n=51$)

		Intake of dietary components (mean (s.d.))				
		First administration	*Second administration*	*P-value first and second administration*	*Difference (mean (s.d.))*	*Spearman correlation*
Iron (mg)	Unadjusted IIAT[a]	10.5 (4.3)	10.5 (4.3)	0.913[b]	0.1 (4.4)	0.60
	Adjusted IIAT[a]	9.7 (2.8)	9.6 (3.4)	0.644[b]	0.2 (2.5)	0.66
Calcium (mg)	Unadjusted IIAT[a]	806 (375.9)	811 (343.8)	0.924[b]	−5.5 (410.1)	0.64
	Adjusted IIAT[a]	731 (256.3)	735 (308.6)	0.884[b]	−4.6 (221.9)	0.73
Vitamin C (mg)	Unadjusted IIAT[a]	102 (65.9)	92 (53.8)	0.353[c]	9.9 (61.9)	0.58
	Adjusted IIAT[a]	100 (69.9)	86 (51.6)	0.183[c]	13.3 (59.2)	0.48

[a]IIAT = Iron Intake Assessment Tool.
[b]Paired *t*-test.
[c]Wilcoxon matched-pairs signed rank test.

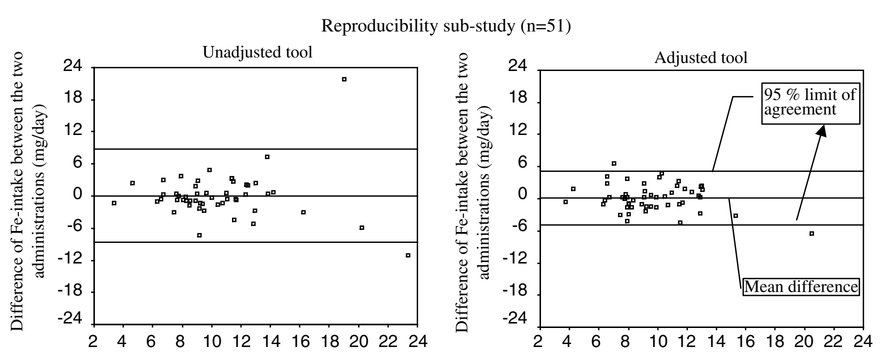

Figure 2 Differences between the mean of iron intake for repeated administrations of IIAT (mg/day) in the reproducibility subanalyses.

designed to estimate energy intake (this would require a more extensive food list including fats).

The correlation coefficients range from 0.45 to 0.60 for adjusted intake. When the correlation is below 0.3 or 0.4, the attenuation is so severe that it is difficult to detect true associations (Cade *et al*, 2002), but all the correlations for the adjusted tool are higher than the 0.4 threshold proposed by Cade *et al* (2002). Masson *et al* (2003) reported that Spearman correlation coefficients are useful for assessing the relative validity of estimates of nutrient intake by FFQs. Spearman correlation coefficients above 0.5 are recommended. Brunner *et al* (2001) have suggested that correlations between

FFQs and weighed records of 'about 0.5 for most nutrients' are 'good evidence that the FFQ has the ability to rank individuals', while Willett (1994) suggests that when FFQs are compared with multiple records of diet, correlation coefficients may reach 0.6–0.7. In view of these methodological considerations, the authors conclude that the correlation coefficients achieved in the present study are rather low for iron, while the calcium and vitamin C correlations for the adjusted tool are acceptable. The low correlation coefficient for iron indicates a relatively weak relationship between the new method and the reference method and renders the ranking capability of the instrument questionable.

The Bland and Altman test confirms the ability of the new instrument to estimate group means for iron and vitamin C. The large standard deviation of the mean differences between the two methods suggests that the two methods yield unacceptably different results in terms of individual iron intake for a substantial number of participants. However, it is not necessary for the assessment tool to accurately estimate absolute intakes of individuals in order to be useful in an epidemiological setting where extremes of intake are more relevant to health outcomes.

Crossclassification was included in the analysis to test the ability of the assessment tool to assign individuals to broad categories of intake. In the present study, crossclassification according to tertiles of intake showed reasonable agreement between the two methods. Masson *et al* (2003) suggested that in dietary validation studies more than 50% of subjects being correctly classified and less than 10% of subjects having grossly misclassified is acceptable. More than 50% of subjects were classified into the same third of the nutrient intake for vitamin C, but correct classification was below 50% for iron and calcium. On the other hand, for each nutrient the percentage of being grossly misclassified was below 10%. Crossclassification according to quartiles was also performed. The percentages of correctly classified and misclassified did not differ from tertile analyses (data not shown).

The authors introduced the weighed κ-statistic, although its use is controversial (Maclure & Willett, 1987). The weighed κ can be valuable in that it gives a single value to represent agreement. Vitamin C intake showed a moderate agreement, whereas the other dietary components had a fair agreement. Masson *et al* (2003) suggested that the weighed κ statistic is meaningful to present in association with the percentages of crossclassification. It is desirable that the weighed κ-value is above 0.4 and that 50% of subjects are correctly classified and less than 10% of subjects are grossly misclassified. The adjusted iron intake had a κ-value of 0.22 and the percentage of correctly classified was 38% and grossly misclassified was 6%. One could conclude that the IIAT showed only a fair agreement with the reference method, but since subjects with widely differing intakes may be grouped into one category while subjects with very similar intakes close to the cutoff point may be grouped into different categories, agreement between the two approaches should not be expected. In studies with small numbers of subjects, misclassification of a few subjects can make a large difference to the κ value (Masson *et al*, 2003). The final technique used to assess the validity of the IIAT was 'actual values for surrogate categories'. The significant increase of all dietary components between the first and third tertile suggests that the IIAT reliably distinguishes extremes of intake for all three nutrients.

Acceptable repeatability was established for the IIAT. No significant differences could be distinguished between the first and the second administration. The Spearman correlation coefficients fell within the common range 0.5–0.7 (Willett, 1998). The comment made by Willett that 'a high degree of reproducibility does not ensure validity because high correlation can be simply the result of correlated error' applies to the present study. The mean differences between the two measurements were limited and showed the ability of the new instrument to estimate mean dietary component intake.

In general, the validation and repeatability analyses gave better results for the adjusted version of the tool as compared to the unadjusted version. Therefore, it is concluded that the former version is preferable and will be used for future analyses of the data.

At the start of this study, 69 women volunteered to fulfil the whole set of investigations. Of these, 16 did not complete the 11-day dietary record. The authors compared general characteristics of both groups of volunteers, those who completed the study and those who did not. There were no differences in the mean age and BMI of the groups. Moreover, the mean intake of iron, calcium and vitamin C assessed by the IIAT did not differ between the two groups of volunteers (data not shown). The educational level was similar for both groups of volunteers.

The observed differences, on an individual level, between the IIAT and the dietary records are sometimes quite large and raise questions that need to be addressed. In general, questionnaires based on frequency data have a tendency to stimulate over-reporting, especially for socially acceptable food items (Nelson & Bingham, 1996). On the other hand, the checklist with high iron content food items might for a number of subjects still be incomplete and cause substantial underestimations. The large differences between the two methods for a number of individuals may also reflect the constraints with respect to the ability to complete questionnaires properly as this technique relies heavily on the ability to conceptualise cognitively such aspects of the diet like 'frequencies' and 'portion sizes' (Nelson *et al*, 1994).

More general methodological aspects of the diet record method that was used as a 'gold standard' should also be taken into consideration. Firstly, there is the aspect of different interpretation of portion size in the two methods. The IIAT used standard portion sizes, while the diet record used more detailed descriptions or expressions in grams of portion sizes. The different kind of food portions could introduce food quantification errors. In the current version of the IIAT, the authors did not use visual or three-dimensional models to estimate portion size, although this is recommended in the literature (Nelson *et al*, 1996; Robson & Livingstone, 2000). To reduce the completion time of the new instrument, the authors preferred not to use photographs. Secondly, differences in food items were also found. For example, some individuals reported the consumption of an iron-fortified breakfast cereal ('Special K') in one tool and not in the other. As this is a food item containing more than 20 mg iron/100 g, its presence or absence in the diet makes a very significant difference. *A posteriori*, it is impossible to differentiate between 'erroneous reporting' and 'natural variation in food consumption'. Thirdly, the diet record method is considered an important reference method for questionnaire validation. The 'estimated' technique was chosen in preference to the 'weighed' technique because of the high respondent burden and time-consuming characteristic of the latter. Moreover, Nettleton *et al* (1980) found differences of only 2–5% between estimated and weighed records depending on the type of nutrient and population studied. For groups, the error may be small and of little importance, but for individuals, it may be large. Although estimated records are less accurate than weighed records of individuals' diets, they have the same order of accuracy when ranking subjects into thirds or fifths (Bingham *et al*, 1988). Fourthly, there is the possibility of an influence of the number of days that were assessed by each method. The IIAT asked for information about a normal week (7 days) of last month, while the 11-day dietary record was spread out over a month after the completion of the questionnaire. A change in the diet could occur during this period.

There are no existing iron-specific questionnaires in the literature with which to compare the validity of the newly developed dietary instrument, except for the meal-based FFQ developed in New Zealand by Heath *et al* (2000), on which the current instrument was based. On the whole, the results of the current analyses are highly comparable to the New Zealand data. In both studies an adjustment factor was introduced, which appeared to effect a considerable improvement compared to unadjusted data. Comparing the newly developed instrument with the New Zealand version, the correlation coefficients in the present study are lower for iron but higher for calcium and vitamin C. The percentage of subjects misclassified to extreme tertiles in the present study is similar to the percentage of subjects classified to extreme quartiles in the New Zealand study. The correlation coefficients between two administrations are in the present study similar for iron, but lower for calcium and vitamin C.

A number of general questionnaires have attempted to estimate iron intake, and a small review of general questionnaires that measured iron intake in women gave a range of Spearman correlation coefficients (relation between new questionnaires and reference method) from 0.07 (Munger *et al*, 1992) to 0.54 (Masson *et al*, 2003). The findings of the present study are, therefore, comparable with the results of earlier studies. The newly developed instrument represents a substantially lower respondent burden than 11-day dietary records and the elimination of possible researcher coding and entry errors because the subjects enter their own food items in the computer. The analysis of the current study suggests that the IIAT is appropriate for assessing group mean intakes for iron, calcium and vitamin C, and allows for statistical testing for differences in intake of these nutrients between selected subgroups of the population. It is advised that the adjusted version would be used for these purposes. However, for the purpose of studying associations between iron intake and specific outcome variables (like eg iron status), further refinements of the tool would be advantageous.

Acknowledgements

We thank the dietitians Mrs M Bellemans and Mrs M De Maeyer for their important contribution to the fieldwork and the data input, Mr R Preece (Institute of Food Research, UK) and Mr B Sterken for their contribution in the software development, and Professor Dr D De Bacquer for his statistical advice. The assessment tool is available on request. This work was financially supported by the Belgian National Fund for Scientific Research/Flemish Division (Fund No G.0152.01). Iron research at the Institute of Food Research is funded by the Biotechnology and Biological Sciences Research Council and the UK Food Standards Agency.

References

Altman DG (1991): *Practical Statistics for Medical Research*. London: Chapman & Hall.

Beaton GH, Milner J, Corey P, McGuire V, Cousins M, Stewart E, de Ramos M, Hewitt D, Grambsch PV, Kassim N & Little JA (1979): Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am. J. Clin. Nutr.* **32**, 2546–2549.

Belton NR, Macvean ADL, Richards ND, Elton RA, Moffat WMU & Beattie TF (1997): Nutrient intake in Scottish adolescents. *Proc. Nutr. Soc. B* **56**, 303A.

Bingham SA, Nelson M, Paul AA, Haraldsdottir J, Björge Löke E & van Staveren WA (1988): Methods for data collection at an individual level, In Manual on Methodology for Food Consumption Studies, eds ME Cameron & WA van Staveren, pp 53–106. Oxford: Oxford University Press.

Bland JM & Altman DG (1986): Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.

Brunner E, Stallone D, Juneja M, Bingham S & Marmot M (2001): Dietary assessment in Whitehall II: comparison of 7 d diet diary and food-frequency questionnaire and validity against biomarkers. *Br. J. Nutr.* **86**, 405–414.

Cade J, Thompson R, Burley V & Warm D (2002): Development, validation and utilisation of food-frequency questionnaires—a review. *Public Health Nutr.* **5**, 567–587.

Cruz JA (2000): Dietary habits and nutritional status in adolescents over Europe—Southern Europe. *Eur. J. Clin. Nutr.* **54**(Suppl 1), S29–S35.

de Groot C, van Staveren WA & Riboli E (1998): Abstracts of the Third International Conference on Dietary Assessment Methods. *Eur. J. Clin. Nutr.* **52**, 51–89.

De Vriese SR, De Henauw S, De Backer G, Dhont M & Christophe AB (2001): Estimation of dietary fat intake of Belgian pregnant women. Comparison of two methods. *Ann. Nutr. Metab.* **45**, 273–278.
Disler PB, Lynch SR, Charlton RW, Torrance JD, Bothwell TH, Walker RB & Mayet F (1975): The effect of tea on iron absorption. *Gut* **16**, 193–200.
Hallberg L, Brune M, Erlandsson M, Sandberg AS & Rossander-Hulten L (1991): Calcium: effect of different amounts on nonheme- and heme-iron absorption in humans. *Am. J. Clin. Nutr.* **53**, 112–119.
Health Council Belgium (1997): *Household Weights and Measures; A Manual for a Standardised Quantification of Food Items in Belgium*. Brussels: Ministry of Social Affaires, Public Health and Environment, Health Council Belgium.
Heath AL, Skeaff CM & Gibson RS (2000): The relative validity of a computerized food frequency questionnaire for estimating intake of dietary iron and its absorption modifiers. *Eur. J. Clin. Nutr.* **54**, 592–599.
Holland B, Welch A, Unwin I, Buss D, Paul A & Southgate D (1991): *McCance and Widdowson's, The Composition of Foods, Fifth Revised and Extended Edition*. Cambridge: Royal Society of Chemistry.
Lynch SR & Cook JD (1980): Interaction of vitamin C and iron. *Ann. NY. Acad. Sci.* **355**, 32–44.
Maclure M & Willett WC (1987): Misinterpretation and misuse of the kappa statistic. *Am. J. Epidemiol.* **126**, 161–169.
Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, Grubb DA & Bolton-Smith C (2003): Statistical approaches for assessing the relative validity of a food- frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutr.* **6**, 313–321.
Matthys C, De Henauw S, Devos C & De Backer G (2003): Estimated energy intake, macronutrient intake and meal pattern of Flemish adolescents. *Eur. J. Clin. Nutr.* **57**, 366–375.
Michaud C, Musse N, Kahn JP, Grebert M, Burlet C & Mejean L (1989): Nutrition behavior in adolescent students (15–19 years of age) in the Nancy metropolitan area. A comparison with the recommended nutritional intake of the French population. *Rev. Epidemiol. Sante Publique* **37**, 149–159.
Munger RG, Folsom AR, Kushi LH, Kaye SA & Sellers TA (1992): Dietary assessment of older Iowa women with a food frequency questionnaire: nutrient intake, reproducibility, and comparison with 24-hour dietary recall interviews. *Am. J. Epidemiol.* **136**, 192–200.
Nelson M (1996): The validation of dietary assessment, In *Design Concepts in Nutritional Epidemiology*, eds BM Margetts & M Nelson, pp 241–272. Oxford: Oxford University Press.
Nelson M, Atkinson M & Darbyshire S (1994): Food photography. I: The perception of food portion size from photographs. *Br. J. Nutr.* **72**, 649–663.
Nelson M, Atkinson M & Darbyshire S (1996): Food photography. II: Use of food photographs for estimating portion size and the nutrient content of meals. *Br. J Nutr.* **76**, 31–49.
Nelson M & Bingham S (1996): Assessment of food consumption and nutrient intake, In *Design Concepts in Nutritional Epidemiology*, eds BM Margetts & M Nelson, pp 123–169. Oxford: Oxford University Press.
Nettleton PA, Day KC & Nelson M (1980): Dietary survey methods. 2. A comparison of nutrient intakes within families assessed by household measures and the semi-weighed method. *J. Hum. Nutr.* **34**, 349–354.
NEVO (1996): *NEVO-Table, Dutch Food Composition Table 1996*. Zeist: NEVO Foundation (in Dutch).
NEVO (2001): *NEVO-Table, Dutch Food Composition Table 2001*. Zeist: NEVO Foundation (in Dutch).
NUBEL (1995): *Belgian Food Composition Table, 2nd Edition*. Brussels: Ministry of Public Health (in Dutch).
NUBEL (1999): *Belgian Food Composition Table, 3rd Edition*. Brussels: Ministry of Public Health (in Dutch).
Robson PJ & Livingstone MB (2000): An evaluation of food photographs as a tool for quantifying food and nutrient intakes. *Public Health Nutr.* **3**, 183–192.
Rolland-Cachera MF, Bellisle F & Deheeger M (2000): Nutritional status and food intake in adolescents living in Western Europe. *Eur. J. Clin. Nutr.* **54**(Suppl 1), S41–S46.
Roma-Giannikou E, Adamidis D, Gianniou M, Nikolara R & Matsaniotis N (1997): Nutritional survey in Greek children: nutrient intake. *Eur. J. Clin. Nutr.* **51**, 273–285.
Samuelson G (2000): Dietary habits and nutritional status in adolescents over Europe. An overview of current studies in the Nordic countries. *Eur. J. Clin. Nutr.* **54**(Suppl 1), S21–S28.
Sempos CT (1992): Some limitations of semiquantitative food frequency questionnaires. *Am. J. Epidemiol.* **135**, 1127–1132, [invited commentary].
Spanjersberg MQI & Jansen EHJM (2000): *Iron deficiency and Overload in Relation to Nutrition*. Bilthoven, The Netherlands: Rijksinstituut voor Volksgezondheid en Milieu.
SPSS (1999): *Statistical Package for the Social Sciences*, 10th Edition Chicago: SPSS.
Tylavsky FA & Sharp GB (1995): Misclassification of nutrient and energy intake from use of closed-ended questions in epidemiologic research. *Am. J. Epidemiol.* **142**, 342–352.
Unilever (1992): *Becel Nutrient Calculation Program*. Rotterdam: Nederlandse Unilever Bedrijven B.V.
Willett WC (1994): Future directions in the development of food-frequency questionnaires. *Am. J. Clin. Nutr.* **59**, 171S–174S.
Willett WC (1998): *Nutritional Epidemiology*. New York: Oxford University Press.
World Health Organization (1995): *Physical Status: The Use and Interpretation of Anthropometry*, Technical Report Series 854 Geneva: WHO.