

## Reproducibility and relative validity of dietary glycaemic index and glycaemic load assessed by the food-frequency questionnaire used in the Dutch cohorts of the European Prospective Investigation into Cancer and Nutrition

Huaidong Du<sup>1,2\*</sup>, Daphne L. van der A<sup>1</sup>, Marit M. E. van Bakel<sup>3</sup>, Lisa D. M. Verberne<sup>1,4</sup>, Marga Ocké<sup>1</sup> and Edith J. M. Feskens<sup>4</sup>

<sup>1</sup>National Institute for Public Health and the Environment (RIVM), PO Box 1, 3710BA Bilthoven, The Netherlands

<sup>2</sup>Department of Human Biology, Nutrition and Toxicology Research Institute of Maastricht (NUTRIM), Maastricht, The Netherlands

<sup>3</sup>Nutritional and Database Resource Team, International Agency for Research on Cancer (IARC), Lyon, France

<sup>4</sup>Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands

(Received 20 August 2008 – Revised 1 December 2008 – Accepted 12 December 2008 – First published online 23 March 2009)

Limited information is available on the reproducibility and validity of dietary glycaemic index (GI) and glycaemic load (GL) estimated by habitual diet assessment methods such as FFQ, including the FFQ used in the Dutch cohorts of the European Prospective Investigation into Cancer and Nutrition study. To examine the reproducibility and relative validity of GI and GL, we used data from 121 Dutch men and women aged 23–72 years. They completed the FFQ three times at intervals of 6 months and twelve 24-h dietary recalls (24HDR) monthly during 1991–2. GI and GL were calculated using published values. Intra-class correlation coefficients of the three repeated FFQ were 0.78 for GI and 0.74 for GL. Pearson correlation coefficients between the first FFQ and the weighted average of the 24HDR were 0.63 for both GI and GL. Weighted  $\kappa$  values between the first FFQ and the average of the 24HDR (in quintiles) were 0.40 for GI and 0.41 for GL. Bland–Altman plots showed a proportional bias in GI ( $\beta = 0.46$ ), but not in GL ( $\beta = 0.06$ ). In conclusion, this FFQ can be used in epidemiological studies to investigate the relationship of GI and GL with disease risks, but the proportional bias should be taken into account when using this FFQ to assess the absolute GI values.

**Glycaemic index: Glycaemic load: Reproducibility: Validity: FFQ**

The FFQ is the most frequently used method to assess habitual diets in large-scale epidemiological studies, given that it is relatively inexpensive and less labour demanding when compared with, for example, food diaries or 24-h dietary recalls (24HDR). Dietary glycaemic index (GI) and glycaemic load (GL) have recently been considered as potential risk factors of chronic diseases such as type 2 diabetes, heart diseases and certain types of cancer<sup>(1)</sup>. Several epidemiological studies have been conducted investigating the potential role of dietary GI and GL on disease risks<sup>(2–5)</sup>. However, none of the FFQ used in these studies are specifically designed for measuring dietary GI and GL and the reproducibility and validity of GI and GL assessed by these FFQ are mostly unknown.

The FFQ used in the Dutch cohorts of the European Prospective Investigation into Cancer and Nutrition (EPIC) study was designed to capture the habitual intake of nutrients and major food items of the Dutch adult population. It is a self-administered semi-quantitative FFQ, contains 178 food items and has also been used by other Dutch studies apart

from EPIC<sup>(3)</sup>. The reproducibility and validity of energy, food and nutrient intake have been investigated before<sup>(6,7)</sup>. The present study was conducted, using previously collected data, to examine the reproducibility and relative validity of GI and GL estimated by this FFQ.

### Methods

#### *Subjects, study design and dietary assessment*

A detailed description of the subject recruitment, FFQ, reference method, data collection and data processing has been published previously<sup>(6)</sup>. In brief, the subjects were recruited from the study populations of the Dutch part of the EPIC study from four towns in The Netherlands. Of the 960 people invited by mail, 240 (25%) responded positively, 288 (30%) refused to participate and 432 (45%) did not respond. Out of the 240 who responded positively, we selected 134 subjects equally distributed across the four towns, in 20-year age

groups and of both sexes. In total, 121 subjects, 63 men and 58 women, completed the study. They were between 23 and 72 years old and had a mean BMI of 25.2 kg/m<sup>2</sup>. During 1991–2, the FFQ was administered three times with intervals of 6 months by mailing to the participants at home. 24HDR (one for each weekend day and two for each weekday) were administered monthly throughout a year, either face to face or by telephone interview. Nutrient intake was calculated according to the 1993 version of the Dutch food composition table (Nederlands Voedingsstoffenbestand table)<sup>(8)</sup>. The weighted average of 24HDR was calculated with a weight of one for weekdays and two for weekend days.

#### Glycaemic index and glycaemic load calculation

In order to calculate dietary GI and GL, GI values were assigned to every carbohydrate-containing food item included in the FFQ and 24HDR. The GI database developed in the EPIC study was used as the main source of information. Methodologies used to compile this GI database have been briefly explained before<sup>(2,3)</sup> and a full descriptive paper is going to be published elsewhere<sup>(9)</sup>. In brief, this GI table was compiled from different published sources<sup>(10–12)</sup>, including 415 values from the international table of GI and GL values from Foster-Powell *et al.*<sup>(10)</sup>, 31 published values for British foods<sup>(11)</sup>, 9 foods from the official website of the GI and GL database from the Sydney University<sup>(12)</sup> and 20 communicated values for food (groups) with no equivalence in the sources mentioned previously (Wolever T & Brand-Miller J, personal communication). For the Foster-Powell table, a few adaptations were made according to a prioritised selection list of criteria developed by the FAO/WHO Expert Consultation. For example, the items that were not measured under standard conditions were excluded if the items analysed under standard conditions were already available in the table (standard conditions include the following: 50 g or, in the case of low carbohydrate-containing foods; 25 g of available carbohydrate, measurement time for healthy and diabetic subjects of 2 and 3 h, respectively; at least six participants)<sup>(9)</sup>. All food items in the FFQ could be linked to a GI value in this GI database. For about 30% (258 out of 764) of the items in the 24HDR, no direct link was possible. Therefore, a GI value from other sources had to be assigned. For this, the same sources of information and criteria as used in the compilation of the EPIC GI database were applied. Dietary GI and GL were calculated as the weighted mean of GI and GL from all foods consumed daily using the commonly used formulas<sup>(3)</sup>.

#### Statistical method

To assess the reproducibility, intra-class correlation coefficients were calculated.

To assess relative validity, Pearson correlation coefficients and weighted  $\kappa$  statistics were calculated for the first FFQ and the average of the 24HDR. Absolute agreement between two measurements was determined using the Bland–Altman method, in which the difference was plotted against the mean of the two dietary assessment methods. Linear regression analysis was performed to investigate whether the bias was proportional to the levels of GI and GL. Residual method was used to adjust for energy intake<sup>(13)</sup>.

Statistical analyses were performed using Statistical Analysis Systems version 9.1 (SAS Institute, Cary, NC, USA) and SPSS 15.0 for Windows (SPSS, Inc., Chicago, IL, USA).

#### Results

Means and standard deviations of GI and GL as assessed by the FFQ and 24HDR are shown in Table 1. Both GI and GL were roughly normally distributed (data not shown).

No sex differences were observed; therefore, analyses were performed in the combined population of men and women. Crude intra-class correlation coefficients between three repeated FFQ were 0.82 for GI and 0.86 for GL (Table 2). Adjusting for energy intake slightly attenuated the correlation coefficients to 0.78 for GI and 0.74 for GL.

Pearson correlation coefficients between the first FFQ measurement and the average of 24HDR were 0.69 for GI and 0.79 for GL before adjusting for energy intake, and 0.63 for both GI and GL after the adjustment for total energy intake. Weighted  $\kappa$  values between the first FFQ and the average of 24HDR (in quintiles) were 0.42 for GI and 0.53 for GL. After the adjustment for energy intake, these values changed to 0.40 and 0.41, respectively.

There was a systematic under-reporting in the FFQ for both GI (mean difference 2.2) and GL (mean difference 3.0) when compared with the average of 24HDR (Table 1). A positive association was found between the difference and mean of GI ( $\beta = 0.46$ , 95% CI 0.28, 0.63; Fig. 1), suggesting the presence of proportional bias in the FFQ: under-reporting at lower GI level and over-reporting at higher GI level. No such association was found for GL ( $\beta = 0.06$ , 95% CI  $-0.11$ , 0.24) (Fig. 2), although fairly wide limits of agreement were observed ( $-43.7$  to  $37.7$ ). Logarithmic transformation, as proposed by Bland & Altman<sup>(14)</sup>, did not remove the proportional bias in GI (data not shown).

#### Discussion

We found good reproducibility and relative validity of GI and GL assessed by the FFQ used in the Dutch EPIC cohorts.

To our knowledge, only three dietary questionnaires have been evaluated for their validity of GI and GL measurements<sup>(15–17)</sup>, of which the Swedish FFQ<sup>(16)</sup> and the Japanese dietary history questionnaire<sup>(17)</sup> have also been evaluated for their reproducibility. The reproducibility of the present FFQ was higher than those found in these two studies.

**Table 1.** Mean glycaemic index (GI) and glycaemic load (GL) measured by the Dutch European Prospective Investigation into Cancer and Nutrition FFQ and 24-h dietary recalls (24HDR; *n* 121)

(Means values and standard deviations)

	GI		GL	
	Mean	SD	Mean	SD
FFQ1	56.8	4.0	143.7	57.5
FFQ2	57.0	4.1	138.2	49.1
FFQ3	57.1	3.9	132.6	46.3
Weighted average of 24HDR	58.9	2.8	146.7	48.6

FFQ1, the first FFQ; FFQ2, the second FFQ; FFQ3, the third FFQ.

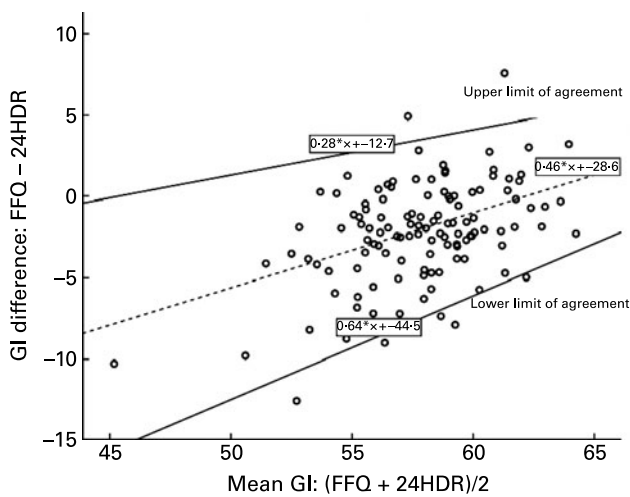
**Table 2.** Reproducibility and relative validity of glycaemic index (GI) and glycaemic load (GL) measured by the FFQ used in the Dutch cohorts of the European Prospective Investigation into Cancer and Nutrition study ( $n$  121)  
(Mean values and 95% confidence intervals)

Reproducibility and validity	GI				GL			
	Crude		Adjusted*		Crude		Adjusted*	
	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI
Intra-class correlation coefficient of repeated FFQ	0.82	0.77, 0.86	0.78	0.71, 0.83	0.86	0.82, 0.90	0.74	0.67, 0.80
Pearson correlation†	0.69	0.58, 0.77	0.63	0.51, 0.73	0.79	0.71, 0.85	0.63	0.51, 0.73
Spearman correlation†	0.67	0.56, 0.76	0.62	0.49, 0.72	0.83	0.76, 0.88	0.60	0.48, 0.71
Weighted $\kappa$ ‡	0.42	0.31, 0.53	0.40	0.29, 0.51	0.53	0.43, 0.63	0.41	0.30, 0.52

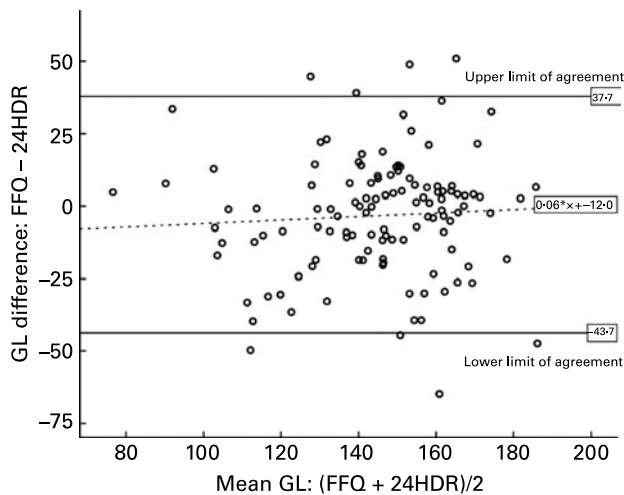
\* Adjusted for energy intake.

† The first FFQ (FFQ1) was compared with the average of 24-h dietary recalls (24HDR).

‡ Agreement between quintiles of FFQ1 and the average of the 24HDR.



**Fig. 1.** Bland-Altman plots of glycaemic index (GI) from the first FFQ and the weighted average of twelve 24-h dietary recalls (24HDR;  $n$  121).



**Fig. 2.** Bland-Altman plots of glycaemic load (GL) from the first FFQ and weighted average of twelve 24-h dietary recalls (24HDR;  $n$  121).

Reproducibility of questionnaires could be influenced by the time interval between administrations. Too long, real change in eating habits may occur, which may underestimate the reproducibility. On the other hand, if the interval is too short, participants may remember the previous questionnaire and therefore respond with the same answers, thus overestimating the reproducibility. Compared with the two previous studies, the time interval of the present study (6 months) is shorter than that from Levitan *et al.*<sup>(16)</sup> (1 year) but longer than the administration time of the Japanese dietary history questionnaire (3 months)<sup>(17)</sup>. Therefore, most likely the time interval is not the reason for the higher reproducibility observed in the present study.

The validity of dietary assessment methods can be tested with different statistical tools depending on the intended use. The Bland-Altman method has been suggested as the best method for assessing the agreement between two measurement methods<sup>(18)</sup>. In our case, the plots indicate a systematic underestimation of GL by the FFQ, whereas for GI there is an underestimation at the lower GI level and an overestimation at the higher GI level. However, given that FFQ are, most of the time, used in association studies to rank individuals according to their intake level rather than for measuring the absolute level of intake, calculating correlation coefficients and weighted  $\kappa$  statistics is more appropriate<sup>(19)</sup>. The validity of GI and GL in the present study, as indicated by the Pearson correlation coefficients, was similar to the results from three previous studies<sup>(15-17)</sup>. Only one of these three studies used weighted  $\kappa$  values to assess the agreement between the modified Willett FFQ and three 4-d weighted food records<sup>(15)</sup>. Values from the present study were higher for GL (0.41 v. 0.30) but lower for GI (0.40 v. 0.53), when compared with their findings. It has been suggested that correlation coefficients of 0.5 or higher indicate that the FFQ has sufficient ability to rank individuals according to their nutrient intake as continuous variable, while for categorical variables, weighted  $\kappa$  values are recommended to be above 0.4 in order to draw valid conclusions<sup>(19)</sup>.

In conclusion, the findings of the present study support the use of this FFQ in epidemiological studies in which associations of GI and GL, either as continuous or categorical variables, with disease risks are under investigation. However, the proportional bias should be taken into account when using this FFQ to assess absolute GI values.

## Acknowledgements

The present study is supported by the DiOGenes project. DiOGenes is the acronym of the project 'Diet, Obesity and Genes' supported by the European Community (Contract No. FOOD-CT-2005-513946; <http://www.diogenes-eu.org>). H. D. conducted the statistical analyses and drafted the manuscript, under the supervision of D. L. v. d. A. and E. J. M. F. M. M. E. v. B. provided the EPIC GI database, L. D. M. V. worked on assigning GI values to the items in 24 HDR data and M. O. developed the FFQ. All co-authors provided comments and suggestions on the manuscript and approved the final version. We thank Vanessa Ginder from the Wageningen University for preparing the dataset ready to be used for the present analyses. None of the authors has a conflict of interest.

## References

1. Barclay AW, Petocz P, McMillan-Price J, *et al.* (2008) Glycemic index, glycemic load, and chronic disease risk – a meta-analysis of observational studies. *Am J Clin Nutr* **87**, 627–637.
2. Cust AE, Slimani N, Kaaks R, *et al.* (2007) Dietary carbohydrates, glycemic index, glycemic load, and endometrial cancer risk within the European Prospective Investigation into Cancer and Nutrition cohort. *Am J Epidemiol* **166**, 912–923.
3. Du H, van der A DL, van Bakel MM, *et al.* (2008) Glycemic index and glycemic load in relation to food and nutrient intake and metabolic risk factors in a Dutch population. *Am J Clin Nutr* **87**, 655–661.
4. Liu S, Manson JE, Stampfer MJ, *et al.* (2001) Dietary glycemic load assessed by food-frequency questionnaire in relation to plasma high-density-lipoprotein cholesterol and fasting plasma triacylglycerols in postmenopausal women. *Am J Clin Nutr* **73**, 560–566.
5. Silvera SA, Jain M, Howe GR, *et al.* (2007) Glycaemic index, glycaemic load and ovarian cancer risk: a prospective cohort study. *Public Health Nutr* **10**, 1–6.
6. Ocke MC, Bueno-de-Mesquita HB, Goddijn HE, *et al.* (1997) The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol* **26**, Suppl. 1, S37–S48.
7. Ocke MC, Bueno-de-Mesquita HB, Pols MA, *et al.* (1997) The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int J Epidemiol* **26**, Suppl. 1, S49–S58.
8. NEVO Foundation (Stichting Nederlands Voedingsstoffenbestand) (1996) *Dutch Food Composition Table (Nederlands voedingsstoffenbestand (Dutch Nutrients File), NEVO Table)*. Den Haag: Dutch Nutrition Center/Information Office for Nutrition (Voedingscentrum/Voorlichtingsbureau voor de Voeding).
9. van Bakel MME, Slimani N, Feskens EJM, *et al.* (2009) Methodological challenges in the application of the glycemic index in epidemiological studies using data from the European Prospective Investigation into Cancer and Nutrition. *J Nutr* **139**, 568–575.
10. Foster-Powell K, Holt SH & Brand-Miller JC (2002) International table of glycemic index and glycemic load values: 2002. *Am J Clin Nutr* **76**, 5–56.
11. Henry CJ, Lightowler HJ, Strik CM, *et al.* (2005) Glycaemic index and glycaemic load values of commercially available products in the UK. *Br J Nutr* **94**, 922–930.
12. The official website of the glycemic index and GI database. Human Nutrition Unit, University of Sydney 2006. [www.glycemicindex.com](http://www.glycemicindex.com) (accessed 30 November 2007).
13. Willett WC (1998) *Nutritional Epidemiology*, 2nd ed. New York: Oxford University Press.
14. Bland JM & Altman DG (1996) Measurement error proportional to the mean. *Br Med J* **313**, 106.
15. Barclay AW, Flood VM, Brand-Miller JC, *et al.* (2008) Validity of carbohydrate, glycaemic index and glycaemic load data obtained using a semi-quantitative food-frequency questionnaire. *Public Health Nutr* **11**, 573–580.
16. Levitan EB, Westgren CW, Liu S, *et al.* (2007) Reproducibility and validity of dietary glycemic index, dietary glycemic load, and total carbohydrate intake in 141 Swedish men. *Am J Clin Nutr* **85**, 548–553.
17. Murakami K, Sasaki S, Takahashi Y, *et al.* (2008) Reproducibility and relative validity of dietary glycaemic index and load assessed with a self-administered diet-history questionnaire in Japanese adults. *Br J Nutr* **99**, 639–648.
18. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
19. Masson LF, McNeill G, Tomany JO, *et al.* (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.