Methodological Challenges in the Application of the Glycemic Index in Epidemiological Studies Using Data from the European Prospective Investigation into Cancer and Nutrition 1–3


1. Department of Community Medicine, University of Tromsø, N 9037 Tromsø, Norway; 2. Department of Cardiology and Department of Clinical Epidemiology, Aarhus University Hospital, 8200 Aalborg, Denmark; 3. Institute for Cancer and Nutritional Epidemiology, Maastricht, The Netherlands; 4. National Institute for Public Health and the Environment, 3721 MA Bilthoven, The Netherlands; 5. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 8535GS Utrecht, The Netherlands; 6. Human Nutrition Unit, Department of Public Health, University of Parma, IT-43100 Parma, Italy; 7. Danish Cancer Society, Institute of Cancer Epidemiology, 2100 Copenhagen, Denmark; 8. School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006 Sydney, Australia; 9. Institute of Public Health of Navarra, Pamplona and CIBER en Epidemiología y Salud Pública, E-30008 Pamplona, Spain; 10. Public Health Division of Gipuzkoa, 4-20013 San Sebastian, Spain; 11. Cancer Epidemiology Unit, University of Oxford, OX3 7LF Oxford, UK; 12. Department of Public Health and Primary Care, MRC Centre for Nutritional Epidemiology in Cancer Prevention and Survival, CB1 8RN Cambridge, UK; 13. Division of Cancer Epidemiology, German Cancer Research Center, 69120 Heidelberg, Germany; 14. Department of Epidemiology, German Institute of Human Nutrition, Potsdam-Rehbrücke, Germany; 15. Department of Odontology, Carolinum, Umeå University Hospital, 901 87 Umeå, Sweden; 16. Department of Surgery, Malmo University Hospital, 205 02 Malmo, Sweden; 17. Department of Cardiology and Department of Clinical Epidemiology, Aarhus University Hospital, 8200 Aalborg, Denmark; 18. Institute of Community Medicine, University of Tromsø, N 9037 Tromsø, Norway; 19. Department of Clinical and Experimental Medicine, Federico II University, I-80131 Naples, Italy; 20. Molecular and Nutritional Epidemiology Unit, ISPO Cancer Prevention and Research Institute, 30135 Florence, Italy; 21. Iserm, ERI 20, Université Paris-Sud, EA 4045, IFR 69, Institut Gustave Roussy, 94805 Villejuif, France; and 22. Division of Cancer Epidemiology, DKFZ, 69120 Heidelberg, Germany

Abstract

Associations between the glycemic index (GI) or glycemic load (GL) and diseases are heterogeneous in epidemiological studies. Differences in assigning GI values to food items may contribute to this inconsistency. Our objective was to address methodological issues related to the use of current GI and GL values in epidemiological studies. We performed ecological comparison and correlation studies by calculating dietary GI and GL from country-specific dietary questionnaires (DQ) from 422,837 participants from 9 countries participating in the European Prospective Investigation into Cancer and Nutrition study and single standardized 24-h dietary recalls (24-HDR) obtained from a representative sample (n = 33,404) using mainly Foster Powell’s international table as a reference source. Further, 2 inter-rater and 1 inter-method comparison were conducted, comparing DQ GI values assigned by independent groups with values linked by us. The ecological correlation between DQ and 24-HDR was good for GL (overall r = 0.76; P < 0.005) and moderate for GI (r = 0.57; P < 0.05). Mean GI/GL differences between DQ and 24-HDR were significant for most centers. GL but not GI from DQ was highly correlated with total carbohydrate (r = 0.98 and 0.15, respectively; P < 0.0001) and this was higher for starch (r = 0.72; P < 0.0001) than for sugars (r = 0.36; P < 0.0001). The inter-rater and inter-method variations were considerable for GI (weighted x coefficients of 0.49 and 0.65 for inter-rater and 0.25 for inter-method variation, respectively) but only mild for GL (weighted x coefficients < 0.80). A more consistent methodology to attribute GI values to foods and validated DQ is needed to derive meaningful GI/GL estimates for nutritional epidemiology.


Introduction

In 1981, Jenkins et al. (1) introduced the concept of the glycemic index (GI)28 to provide a classification of carbohydrate-containing foods based on their ability to raise blood glucose concentrations. Later, the term glycemic load (GL = GI × amount of carbohydrate) of the diet was introduced to take both the quantity and quality of carbohydrates into account. Low overall dietary GI and GL were shown to improve blood and urinary parameters related to the diabetic syndrome (2–5). Initial studies of GI and

1 Supported by grants from the U.S. National Cancer Institute (5RO1CA102460-02), the DIOGenes project (FOOD-CT-2005-513946), and a National Health and Medical Research Council Public Health Fellowship (I520018).

0022-3166/08 $8.00 © 2009 American Society for Nutrition.


First published online January 21, 2009; doi:10.3945/jn.108.097121.
GL were conducted in highly controlled settings and included only a few food items with well-known GI values.

The GI concept, as a possible predictor of metabolic disorders, has encouraged researchers in diverse fields to apply existing GI tables to dietary data that were already available for other purposes and study contexts. However, several methodological issues make it difficult to use the physiological GI data in nutritional epidemiological studies and may explain the inconsistent findings for the association between GI/GL and disease outcomes (6, 7). First, the number of food items reported in epidemiological studies largely surpasses the number of items for which a GI value has been determined. Because most of the information comes from the Foster-Powell table (8), which is limited to ~750 foods, many assumptions have to be made when linking them to dietary data. This is particularly problematic for European studies, because this table contains very few European data. Second, in contrast to nutrients, GI values are not concentration values of the food item but the physiological response to its consumption. Although the GI concept has existed for 25 y, values have not been added to any national nutrient databases, leading to a subjective judgment of assigning GI values from general GI tables to local food items. Lastly, most dietary questionnaires (DQ) used in epidemiological studies estimate usual long-term individual intakes based on a restricted number of foods not specifically selected and validated for dietary GI or GL. In addition, they do not measure mixed dishes and individual recipes accurately.

Our main objectives in this study were to address methodological issues related to the use of GI/GL in large epidemiological studies on diet and disease not designed for this purpose by means of 1) ecological comparison between DQ and 24-h dietary recall (24-HDR); 2) inter-rater; and 3) inter-method studies in the assignment of GI values to DQ food items. Furthermore, the aims were to help establish a consistent methodology for the assignment of GI values to different dietary tools and European foods and to provide recommendations for improvements.

To address these methodological objectives, we calculated dietary GI and GL using DQ and 24-HDR data from the European Prospective Investigation into Cancer and Nutrition study (EPIC), as well as GI and GL from participants of a study in Parma, Italy (for the inter-method study).

Materials and Methods

The EPIC and Parma study. The EPIC study is a prospective cohort study investigating the relation between diet, lifestyle, metabolic and genetic factors, and the risk of developing cancer and other chronic diseases (9, 10). This cohort study involved a nonrepresentative sample of 519,978 participants residing in 23 centers of 10 countries (Denmark, France, Germany, Greece, Italy, The Netherlands, Norway, Spain, Sweden, and the UK). Centers recruited both men and women, except in France, Norway, Utrecht (The Netherlands), and Naples (Italy), where only women were involved. Participants from Greece were excluded from the current analyses, because not all relevant variables for this analysis were available in the central database.

The Parma study is a validation study for GI and GL involving 354 adult participants from the Parma area (F. Brighenti, F. Scazzina, N. Pellegrini, D. Del Rio, D. Ardigo, S. Valtenuea, V. Pala, S. Sieri, V. Krogh, and I. Zavaroni, unpublished data).

**Dietary data.** The EPIC dietary data were assessed at baseline by means of quantitative DQ with individual portion sizes (in France, Spain, The Netherlands, Germany and Italy, except Naples), semiquantitative FFQ [in Denmark, Norway, Naples, Umeå (Sweden), and the UK], or combined dietary methods (semiquantitative FFQ and a 14-d record) in Malmo (Sweden) that were developed and validated locally (11, 12). The number of food items varied from 88 in Norway to 2443 in Malmo. Detailed descriptions on the study populations and the usual dietary intake have been described elsewhere (13, 14).

Furthermore, single 24-HDR were collected between 1995 and 2000 from a representative sample of ~8% of each EPIC center (n = 37,000). The 24-HDR were highly standardized across countries using a computerized program (EPIC-SOFT) and contained detailed information on foods and recipes reported by the study subjects (15, 16). As only 1 24-HDR was available per subject, comparisons could be made at the population mean level only.

For the purpose of the inter-method study, a validated FFQ specifically designed to capture foods that contribute most to dietary GI and GL in northern Italy (the so-called Parma questionnaire), using analyzed GI values for locally consumed foods (17), was used to be compared with the EPIC north Italian questionnaire. Briefly, the questionnaire included questions on 42 main food groups and 131 specific food items representing over 95% of the carbohydrate intake in northern Italy.

**Matching of food items from DQ and 24-HDR to table GI values.** Details of the stepwise approach to assign GI values to dietary items of both the DQ and the 24-HDR are summarized (Fig. 1; Supplemental Material).

In total, GI values were assigned to 79% of all food items (range 48–94%) of each of the EPIC DQ and 75% of all items from the 24-HDR (range 68–80%) covering >99% of total carbohydrate intake.

**Participants in the ecological, correlation, and dietary comparison studies.** For the EPIC study, participants at the top and bottom 1% of the energy intake:energy requirement ratio reported in the DQ were excluded (n = 9,928) as a routine procedure to remove outliers. Furthermore, for the current analyses, participants with outlier values in the DQ (2–3 times the P99 value, based on common sense) for carbohydrate, sugar, starch, fruit, bread, pasta/rice, crisp bread, potato, cakes, and dry biscuits were excluded (n = 193). The participants were then grouped into 14 centers and 9 countries according to the center or country-specific DQ and whether the participants were health conscious or from the general population (for the UK). For the comparison between DQ and 24-HDR, only participants having both dietary measurements were considered and thus involved 11,786 men and 21,618 women (3.4% exclusions). For the ecological comparison and correlation studies, all participants having DQ data were involved, but diabetic participants (DQ, n = 13,840; 24-HDR, n = 1095) were excluded because they may have changed their diet in favor of low-GI foods. Thus, these analyses involved 117,149 men and 305,688 women (14.0% exclusions). Each participant provided informed consent and the original project was approved by the International Agency for Research on Cancer (IARC) ethical review committee.

**Methods and participants for the inter-rater comparison studies.** The Dutch and Danish groups had, independently from IARC, already assigned GI values to the foods in their national EPIC DQ. These data were used to determine the inter-rater agreement (i.e. agreement between the people assigning GI values to the same questionnaire). Because the Dutch and Danish groups both used the Foster-Powell table as the only

---

* To whom correspondence should be addressed. E-mail: slimani@iarc.fr.

27 Details of the stepwise approach to assign glycemic index values to dietary items of both dietary questionnaire and 24-h dietary recall as well as Supplemental Tables 1 and 2 are available with the online posting of this paper at jn.nutrition.org.

28 Present address: Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Australia.

29 Present address: Data Collection and Exposure Unit, European Food Safety Authority, Parma, Italy.

30 Abbreviations used: DQ, dietary questionnaire; EPIC, European Prospective Investigation into Cancer and Nutrition; GI, glycemic index; GL, glycemic load; 24-HDR, 24-hour dietary recall; IARC, International Agency for Research on Cancer.

4 To whom correspondence should be addressed. E-mail: slimani@iarc.fr.
source to assign GI values to their center-specific data, to make fair comparisons possible, IARC assigned for this exercise GI values to the EPIC Dutch and Danish questionnaires as described in Fig. 1 but using only the Foster-Powell table as reference. Similar to IARC, the Danish group based the linking of GI values to foods on the criteria developed by the FAO/WHO expert consultation. However, in contrast to IARC, they calculated the GI of both simple and complex foods from their individual ingredients using the foods available in the Danish food tables. Also, the Danish group assigned the GI value of cooked foods to raw foods for foods not present in the Foster-Powell table. The Dutch group did not have any particular criterion for inclusion or assignment established (18). The individual’s average dietary GI and GL were calculated as indicated above. We then compared the dietary GI and GL results from Denmark (n = 55,852 from 2 centers) and 1 center from The Netherlands (Utrecht, n = 16,885) with those for the same participants from IARC.

**Method and participants in the inter-method comparison study**. For the purpose of the inter-method study (i.e. agreement between the people assigning GI values to different questionnaires while also using different GI databases), IARC assigned GI values to those foods of the Parma questionnaire that matched the EPIC north Italian questionnaire. For 57 of the 131 items, a corresponding item was available in the EPIC DQ and 24-HDR. The Dutch group did not have any particular criterion for inclusion or assignment established (18). The individual’s average dietary GI and GL were calculated as indicated above. We then compared the dietary GI and GL results from Denmark (n = 55,852 from 2 centers) and 1 center from The Netherlands (Utrecht, n = 16,885) with those for the same participants from IARC.

**Calculation of dietary GL and GI**. An individual’s average daily or dietary GI is an indicator of the average quality of the carbohydrates consumed in terms of glycemic response and dietary GI integrates both the quantity and quality of carbohydrates consumed, thus representing the total glycemic burden of the diet. The computation included digestible carbohydrate (total carbohydrate minus fiber) only. Average dietary GI and GL were calculated as follows:

\[
GI = \frac{\sum_{i=1}^{n}[CHO_i \cdot QF_i \cdot GL_i]}{\text{total digestible carbohydrate intake per day}}
\]

\[
GL = \frac{\sum_{i=1}^{n}[CHO_i \cdot QF_i \cdot GL_i]}{\text{total digestible carbohydrate intake per day}}
\]

where CHO is the digestible carbohydrate content of food item \(i\) (g/g), QF is the quantity of food item \(i\) per day (portion \(\times\) frequency), GI is the GI for food \(i\), and \(n\) is the number of foods eaten per day. For the DQ, the quantity of food item \(i\) represents the amount of food eaten daily during the past year. GI and GL were expressed as a percentage of the glycemic response elicited, using 50 or 25 g glucose as a reference food.

**Statistical analyses**. Center-specific data of crude and adjusted daily GI, GL, and total carbohydrate for the 14 EPIC centers having both DQ and 24-HDR data are presented as means. Adjustments were made for energy, age, weight, and height. In addition, GI and GL estimates were controlled for season and day of the week (the latter for 24-HDR only) using internal weighting factors.

For ecological comparison of GI and GL, Pearson correlation coefficients between adjusted DQ and 24-HDR means were calculated.

To evaluate the linear association between GI, GL, and the daily nutrient intakes of carbohydrate, sugar, starch, fiber, and energy, we computed crude and adjusted Pearson partial correlation coefficients using log-transformed variables. In addition to the adjustments mentioned above, values were adjusted for center. Adjustment using the residuals of GI and GL on each of the nutrients was also tested, but results were similar. All these statistical analyses were stratified by gender.

Inter-rater and inter-method agreement was assessed using the weighted \(k\) statistic calculated with a linear set of weights (19) for quartiles of GI and GL. Pearson correlation coefficients were determined at the level of food GI and GL values are presented as means ± SD. Stepwise linear regression was used to evaluate the predictive power of foods explaining the variability of dietary GI and GL between the assigning centers. For Denmark and Parma, this analysis was adjusted for gender (Utrecht had only women). An \(\alpha\) of 0.05 was considered significant. The statistical analyses were performed using SAS version 9.1.

**Results**

**Ecological correlations between mean GI and GL from DQ and 24-HDR**. Pearson ecological correlations between the 24-HDR and the DQ was moderate for GI (Fig. 2A) in both men (\(r = 0.59; P = 0.08\)) and women (\(r = 0.52; P < 0.05\)) and good for GL (Fig. 2B) in men (\(r = 0.81; P < 0.01\)) and women (\(r = 0.69; P < 0.01\)).
Comparisons of gender-specific population GI and GL means across centers between DQ and 24-HDR data. Crude GI means from the DQ were significantly different from those of the 24-HDR, except for men in Heidelberg and Potsdam (Table 1) and for women in Ragusa, Naples, The Netherlands, and Potsdam (Table 2). Energy adjustment affected neither these results nor the ranking of the centers. There was no systematic trend of over- or underestimation among the different DQ compared with the 24-HDR, but the differences between the 2 dietary methods were greater ($P < 0.05$) for the semiquantitative questionnaires compared with the quantitative questionnaires.

For GL, crude DQ means were significantly different from those of the 24-HDR, except for Heidelberg and Denmark in men and for Heidelberg and the UK health-conscious population in women. Energy adjustment changed these results for several centers and also changed the ranking of most of the centers. Only for women, the GL was more often overestimated with the DQ as compared with the 24-HDR, independent of the type of questionnaire.

For the 24-HDR, the center rankings for crude GI in both genders and adjusted GL in men were similar to the one for carbohydrate but not the adjusted GI in women. For the DQ, the center ranking for crude GI was similar to the one for carbohydrate in women, whereas the center rankings for crude GI in men and adjusted GL in both genders were different from the one for carbohydrate.

### Correlations between GI/GL and individual nutrients from DQ

Crude dietary GI values were significantly inversely correlated with sugar ($r = -0.24$ and $-0.26$ in men and women, respectively) and significantly and positively related to starch ($r = 0.15$ and $0.42$ for men and women, respectively) and increased upon adjustment for energy, whereas they were not correlated with total energy, percent energy from total carbohydrate, carbohydrate, and fiber (Supplemental Tables 1 and 2).

### Table 1: Dietary GI, GL, and total carbohydrate values for men from DQ and 24-HDR

<table>
<thead>
<tr>
<th>Country and center</th>
<th>n</th>
<th>GI (Crude)</th>
<th>GI (Adjusted)</th>
<th>GL (Crude)</th>
<th>GL (Adjusted)</th>
<th>Total carbohydrate, g/d (Crude)</th>
<th>Total carbohydrate, g/d (Adjusted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spain</td>
<td>1743</td>
<td>56.9 ± 0.1a</td>
<td>56.7 ± 0.1b</td>
<td>55.3 ± 0.0a</td>
<td>55.0 ± 0.1b</td>
<td>146.2 ± 1.1a</td>
<td>133.9 ± 0.7a</td>
</tr>
<tr>
<td>Italy</td>
<td>166</td>
<td>59.2 ± 0.3a</td>
<td>58.9 ± 0.3b</td>
<td>58.1 ± 0.4</td>
<td>57.7 ± 0.4</td>
<td>211.6 ± 3.6a</td>
<td>187.6 ± 2.0a</td>
</tr>
<tr>
<td>North Italy</td>
<td>1251</td>
<td>55.0 ± 0.1a</td>
<td>54.9 ± 0.1b</td>
<td>56.5 ± 0.2</td>
<td>56.3 ± 0.2</td>
<td>158.4 ± 1.4a</td>
<td>148.1 ± 0.8a</td>
</tr>
<tr>
<td>Germany</td>
<td>1012</td>
<td>54.5 ± 0.1a</td>
<td>54.6 ± 0.1b</td>
<td>54.7 ± 0.2</td>
<td>54.8 ± 0.2</td>
<td>132.6 ± 1.4</td>
<td>147.5 ± 0.8</td>
</tr>
<tr>
<td>Potsdam</td>
<td>1215</td>
<td>54.6 ± 0.1a</td>
<td>54.7 ± 0.1b</td>
<td>55.0 ± 0.2</td>
<td>55.0 ± 0.2</td>
<td>141.3 ± 1.3</td>
<td>146.4 ± 0.7</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>1251</td>
<td>59.1 ± 0.1a</td>
<td>59.0 ± 0.1b</td>
<td>59.8 ± 0.2</td>
<td>59.5 ± 0.2</td>
<td>164.8 ± 1.3</td>
<td>159.4 ± 0.8</td>
</tr>
<tr>
<td>UK</td>
<td>398</td>
<td>55.9 ± 0.2a</td>
<td>56.0 ± 0.2b</td>
<td>58.0 ± 0.3</td>
<td>58.1 ± 0.3</td>
<td>138.5 ± 2.3</td>
<td>153.4 ± 1.3</td>
</tr>
<tr>
<td>Health conscious population</td>
<td>109</td>
<td>56.4 ± 0.3a</td>
<td>56.5 ± 0.3b</td>
<td>59.0 ± 0.5</td>
<td>59.2 ± 0.5</td>
<td>138.8 ± 4.3</td>
<td>165.5 ± 2.4</td>
</tr>
<tr>
<td>Denmark</td>
<td>1898</td>
<td>61.2 ± 0.1a</td>
<td>61.3 ± 0.1b</td>
<td>58.6 ± 0.1</td>
<td>58.7 ± 0.1</td>
<td>148.0 ± 1.0</td>
<td>148.1 ± 0.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>1394</td>
<td>57.2 ± 0.1a</td>
<td>57.3 ± 0.1b</td>
<td>56.4 ± 0.1</td>
<td>56.8 ± 0.2</td>
<td>155.5 ± 1.2</td>
<td>151.5 ± 0.7</td>
</tr>
<tr>
<td>Malmo</td>
<td>1349</td>
<td>59.3 ± 0.1a</td>
<td>59.5 ± 0.1b</td>
<td>56.4 ± 0.1</td>
<td>56.4 ± 0.1</td>
<td>150.4 ± 1.2</td>
<td>173.4 ± 0.7</td>
</tr>
<tr>
<td>Umea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Values are means ± SE. Letters indicate different from 24-HDR: *$P < 0.0001$; **$P < 0.05$; ***$P < 0.001$; ****$P < 0.01$.

2 Based on the glucose = 100 scale.

3 Standardized for season.

4 Adjusted for energy, age, height, and weight and standardized for season.

5 Adjusted for energy, age, height, and weight and standardized for day of the week and season.

6 Adjusted for energy, age, height, and weight and standardized for day of the week and season.
TABLE 2 | Dietary GI, GL, and total carbohydrate values for women from DQ and 24-HDR\(^1,2\)

<table>
<thead>
<tr>
<th>Country and center</th>
<th>DQ</th>
<th>24-HDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spain</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>英国</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Values are means ± SE. Letters indicate different from 24-HDR: *P < 0.0001; **P < 0.001; ***P < 0.05; ****P < 0.1.
2. Based on the glucose = 100 scale.
3. Standardized for season.
4. Adjusted for energy, age, height, and weight and standardized for season.
5. Standardized for days of the week and season.
6. Adjusted for energy, age, height, and weight and standardized for day of the week and season.

All carbohydrate-related nutrients were positively correlated with crude dietary GI. Total carbohydrate correlated highly with GI (r > 0.97 for both sexes) and was similar across centers, whereas the correlation with the percentage of energy from total carbohydrate was more moderate and varied more across centers. GI was more strongly correlated with starchy than with sugar. Northern European centers generally showed higher correlations of GI with sugar and lower correlations with starch compared with the southern European centers (results not shown).

Inter-rater comparisons between IARC and Dutch or Danish GI linking methods. From the 244 items in the Dutch questionnaire, 133 items were allocated a GI value at our institute (IARC), whereas the Dutch center allocated a value to only 90 items. For Denmark, the questionnaire included 218 items, 133 of which were given a GI value at IARC and 83 in Denmark. The Danish and Dutch groups did not always allocate values to items for which similar foods were present in the table. At IARC, the approach was to try to find the best possible link with existing foods for these items. To better evaluate the inter-rater variation, only those items to which both IARC and the 2 local groups had assigned GI values were taken into account.

Pearson correlations between food GI values assigned to the Dutch and Danish DQ items at IARC and locally were good, with r = 0.76 for the Dutch DQ and r = 0.80 for the Danish DQ (P < 0.0001). The overall impact on daily GI and GL when the centers assigned a different GI for a particular food depended on the carbohydrate content of the food, as well as the frequency and quantity of the food consumed. Mean GI and GL differed (P < 0.0001) between IARC and Utrecht but did not differ between IARC and Denmark (Table 3). For the quartile ranking of participants according to their GI values, the IARC method of linking GI values to questionnaire items agreed well with the Danish method, but moderately with the Utrecht method (weighted κ coefficients of agreement were 0.65 and 0.49, respectively). For the ranking of participants according to GL values, the methods of IARC and each of the local centers agreed very well, with weighted κ coefficients of agreement of 0.88 between IARC and Utrecht and 0.91 between IARC and Denmark.

TABLE 3 | GI and GL values obtained at IARC and locally at the individual EPIC centers from Utrecht (The Netherlands) and Denmark, and Parma (Italy) for the respective DQ\(^1,2\)

<table>
<thead>
<tr>
<th>Working group</th>
<th>Foods with assigned GI value, n</th>
<th>Gl</th>
<th>Gl</th>
<th>Gl</th>
</tr>
</thead>
<tbody>
<tr>
<td>IARC-Utrecht()</td>
<td>90</td>
<td>16,885</td>
<td>55.1 ± 3.5</td>
<td>110.2 ± 29.1</td>
</tr>
<tr>
<td>Utrecht</td>
<td>90</td>
<td>16,885</td>
<td>51.4 ± 3.3()</td>
<td>102.9 ± 28.0()</td>
</tr>
<tr>
<td>IARC-Denmark()</td>
<td>83</td>
<td>55,852</td>
<td>50.5 ± 4.4</td>
<td>114.4 ± 36.5</td>
</tr>
<tr>
<td>Denmark</td>
<td>83</td>
<td>55,852</td>
<td>50.5 ± 4.4</td>
<td>114.5 ± 36.9</td>
</tr>
<tr>
<td>IARC57()</td>
<td>57</td>
<td>354</td>
<td>53.7 ± 2.9</td>
<td>161.1 ± 55.0</td>
</tr>
<tr>
<td>Parma57()</td>
<td>57</td>
<td>354</td>
<td>56.2 ± 3.5()</td>
<td>168.3 ± 56.9()</td>
</tr>
<tr>
<td>Parma131()</td>
<td>131</td>
<td>354</td>
<td>56.1 ± 3.4()</td>
<td>177.0 ± 58.1()</td>
</tr>
</tbody>
</table>

1. Values are means ± SD. *Different from the center-specific values from IARC, P < 0.0001; **Different from the values from IARC57, P < 0.0001.
2. Based on the glucose = 100 scale.
3. IARC-Utrecht: GI and GL values assigned and calculated by IARC for the Utrecht DQ.
4. IARC-Denmark: GI and GL values assigned and calculated by IARC for the Danish DQ.
5. IARC57: GI and GL values calculated and assigned to 57 Parma foods using the IARC GI-database.
6. Parma57: GI and GL values using the Parma GI-database including only 57 foods for which IARC has a match.
7. Parma131: GI and GL using the Parma GI-database including 131 Parma foods.
Finally, we looked specifically at the food subgroups that together explained >90% of the variation in GI and GL between the assigning groups (Table 4). For Utrecht, the milk products (milk, milk beverages, and fermented milk) and potatoes appeared to have the greatest impact on GI and GL differences, followed by bread. For both GI and GL, the results were essentially the same. In Denmark, bread explained the majority of the differences in both GI and GL.

**Inter-method comparison between IARC and the Parma GI/GL validation research group.** GI and GL were lower (P < 0.0001) at IARC than in Parma, both when including 57 and 123 foods (Table 3). In addition, the method of Parma revealed broader ranges of GI and GL values than our method (GI, 46.2 – 68.7 vs. 45.2 – 62.4, respectively; GL, 52.7 – 440.4 vs. 50.1 – 403.9, respectively) and was even broader when all 131 foods were taken into account (Parma GI, 46.2 – 68.5; GL, 54.3 – 481.4).

The agreement between the 2 methods was good for GL (weighted κ coefficients of 0.82 and 0.77 for 57 and 131 foods, respectively) but only fair for GI (both weighted κ coefficients = 0.25 for 57 and 131 foods, respectively).

For both GI and GL, bread accounted for most of the differences between the assignment methods of the 2 centers (Table 4).

**Discussion**

To date, only a few studies have described the methodology of assigning GI values to items obtained from different dietary methods. Three publications have described the methods to assign GI values to the items of semiquantitative DQ (20–22), whereas only 1 study has done this using 24-HDR items (23). Only 3 studies reported on validity and/or reproducibility of GI/GL from FFQ data (24–26). In the present article, we revealed several methodological problems when assigning GI values to food items using existing GI tables and DQ that were not designed specifically for this purpose. Although we could not perform a validity study for GI and GL at the individual level, at the population level, differences in GI, but not GL, tended to be greater for semiquantitative DQ than for quantitative DQ. Thus, the availability of individual portion sizes in the quantitative DQ (France, northern Italy, Ragusa, Spain, The Netherlands, and Germany) did not help to better estimate GI values but indirectly influenced GI (GI/total carbohydrates). However, independent of the type of DQ, the relative differences between DQ and 24-HDR were higher for GI than for GL and tended especially in women to be overestimated for the DQ. It may be that DQ (standard) portion sizes fit men better. However, other possible explanations more related to the dietary methods are that women, compared with men, may provide more “socially desirable” answers such as a higher intake of carbohydrates, when filling out a DQ. Or it may be that women have greater intra-individual variation in their diets and therefore the mean of 1 24-HDR may be less representative of the true mean intake for women than for men.

The ranking of participants according to GI was not consistent with moderate ecological correlations. The lower correlations for GI compared with GL may partly be explained by the narrow range in GI, making a shift in the ranking number more likely to occur (Tables 1 and 2). For GI, absolute values differed more between the 2 dietary methods, especially in women. However, the ranking was similar, resulting in better ecological correlations, particularly in men. On the other hand, ecological correlations of nutrients are mostly in the same order of magnitude as we found for GI and GL. In the specific context of the EPIC study, future analyses of GI/GL should use DQ data calibrated with the 24-HDR data to correct measurement errors (27).

The questionnaire data of this study revealed a very high correlation between GL and digestible carbohydrate at the individual level, which was in agreement with the findings of Schulz et al. (22). This suggests that with the current instruments, the GL estimates are potentially “surrogates” for carbohydrate intake. Another possibility may be that in affluent societies where fewer unprocessed grains and legumes with considerable lower GI values are eaten, the additive value of GL compared with total carbohydrate intake may be limited. Despite this high correlation, Beulens et al. (18) demonstrated that GL but not total carbohydrates was associated with cardiovascular disease risk.
which argues against GL being a simple surrogate of digestible carbohydrate intake. However, the most striking difference with our results was that upon adjustment for energy, the correlation of GL with digestible carbohydrate remained very high in our study, whereas it became weaker in the study of Schulz et al. (22). One possible explanation for this discrepancy may be the differences in dietary habits between populations. Their population consists of Americans, among which 29% were African Americans, whereas our population consisted of almost exclusively Caucasian white Europeans (28,29). Compared with a recent study (22), we found weaker negative correlations for energy adjusted GI values with fiber intakes, confirming the observation that low-GI foods do not necessarily reflect high-fiber foods (30).

By studying the inter-rater agreement, we were able to test whether the application of existing GI tables is prone to subjective use. The Dutch DQ has previously been validated for nutrients and food groups (31,32) and the Danish DQ for nutrients and energy (33). For The Netherlands, the relative validity of carbohydrates was 0.76, and for potatoes, bread, fruit, sugar and sweets, biscuits and pastry, it was 0.70, 0.78, 0.56, 0.69, and 0.45, respectively. For Denmark, the relative validities of carbohydrates and sugar were 0.40 and 0.50, respectively, for men and 0.47 and 0.41, respectively, for women. Thus, we expected the GI and GL values obtained for Denmark to be less accurate than those for The Netherlands. Consistent with this, the discrepancy between mean DQ and 24-HDR values was larger in Denmark than in The Netherlands. From the weighted \( \kappa \) coefficients for GI and GL, we concluded that small discrepancies in the assignment of GI values affects the ranking of participants for GI more importantly than for GL. This may again be due to the narrow ranges for GI.

To improve concordance between data, we propose to apply a more standardized methodology for the linking of GI values (Supplemental Material; Fig. 1), in particular to: 1) exclude items that were not measured under standard conditions (34); and 2) compute mean GI values weighted according to the relative frequency of consumption of their contributing foods reported in the 24-HDR as is practiced for nutrients. In addition, for foods that are not present in the Foster-Powell table, we suggest using caution with the breakdown of these foods to single ingredients. Although this is currently done for the GI calculation of a mixed meal (35,36) and for some composite foods this may be the best way to proceed, in specific cases, the interference, of e.g. fats or acidic ingredients, will result in a lower GI then when obtained by simple calculation (37,38).

For the inter-method study, more than one-half of the questionnaire items to which the Parma research group attributed a GI value had no corresponding item in the IARC food item list, resulting in more narrow ranges of GI and GL values at IARC. This may be an indication that the EPIC Italian DQ has low specificity and may be inadequate for measuring dietary GI and GL. Our simple comparison of the DQ and 24-HDR values from northern Italy confirmed the underestimation of the GI and GL with the DQ. Also, the fact that the agreement for GI was inferior for the inter-method study compared with the inter-rater studies (0.25 vs. 0.49 and 0.65), although not performed for the same centers, may indicate that country-specific food GI data are lacking. In fact, Henry et al. (39) found notable differences between some of the British foods and similar items in the Foster-Powell table and highlighted the need for testing local foods for use in European studies. However, from a recent study, it seems that, despite these limitations, the EPIC Italian DQ is still able to discriminate among different levels of intake (40).

In conclusion, when assigning existing GI table values to foods, especially those consumed in Europe, several considerations have to be made. Limitations to assigning appropriate GI values include the restricted number of items in the Foster-Powell table (which to date is the main source of GI values), the inclusion of mainly American or Australian food items, lack of values for mixed dishes, and the lack of information on differences in variety (e.g. potato, rice), degree of ripeness (e.g. banana), composition (e.g. more or less fat), cooking methods, and product formulations of the same brand (41). It is of great importance for future research on GI and GL to generate more European-specific GI values, in particular for items such as bread, potatoes, cakes, and breakfast cereals, which contribute more to differences in dietary GI and GL within and between subjects and populations. More work is also required to design and validate DQ for measuring dietary GI and GL. Despite these limitations, the ranking of participants to their GI value seems acceptable with the existing measurements of dietary intake, but the ranking of participants according to GI should be used with caution.

Acknowledgments

We thank Mathieu Mazuir, Carine Biessy, and Corinne Casagrande at the IARC, Lyon, France, for their technical and statistical support. We also thank Thomas Wolever from the Department of Nutritional Sciences, University of Toronto, and Glycaemic Index Testing, Inc. Toronto, Ontario, Canada for his advice on the assignment of GI table values to food items.

Literature Cited