The glycaemic index (GI) concept was originally introduced to classify different sources of carbohydrate (CHO)-rich foods, usually having an energy content of >80% from CHO, to their effect on post-meal glycaemia. It was assumed to apply to foods that primarily deliver CHO, causing hyperglycaemia. Low-GI foods were classified as being digested and absorbed slowly and high-GI foods as being rapidly digested and absorbed, resulting in different glycaemic responses. Low-GI foods were found to induce benefits on certain risk factors for CVD and diabetes. Accordingly it has been proposed that GI classification of foods and drinks could be useful to help consumers make ‘healthy food choices’ within specific food groups. Classification of foods according to their impact on blood glucose responses requires a standardised way of measuring such responses. The present review discusses the most relevant methodological considerations and highlights specific recommendations regarding number of subjects, sex, subject status, inclusion and exclusion criteria, pre-test conditions, CHO test dose, blood sampling procedures, sampling times, test randomisation and calculation of glycaemic response area under the curve. All together, these technical recommendations will help to implement or reinforce measurement of GI in laboratories and help to ensure quality of results. Since there is current international interest in alternative ways of expressing glycaemic responses to foods, some of these methods are discussed.

**Glycaemic index: Carbohydrates: Blood glucose response: Classification of foods:**

**Glycaemic response: Glycaemic load**

**Introduction**

The glycaemic index (GI) concept was originally introduced as a means of classifying different sources of carbohydrate (CHO) and CHO-rich foods in the diet, according to their effect on postprandial glycaemia (Jenkins et al. 1981). It was assumed to apply to foods that primarily deliver available CHO such as potatoes, rice, cereals, etc., usually having an energy content of >80% from CHO. The usual 50 g CHO test load has traditionally referred to available CHO providing sugars for absorption from the small intestine at a certain rate. As such, low-GI CHO were classified as those that are digested and absorbed slowly and lead to a low glycaemic response, whereas high-GI CHO are rapidly digested and absorbed and show a high glycaemic response. This relationship between the rate of digestion and absorption and glycaemic response is also shown using various in vitro digestion models that mimic the in vivo situation. A very high correlation exists between the rate of in vitro glucose release from starchy foods, using pancreatic and brush-border enzymes, and the glycaemic response in vivo (Granfeldt et al. 2005; Englyst et al. 2003). Recently various food factors that may influence in vivo absorption and to a certain degree affect the outcome of GI values were summarised by Arvidsson-Lenner et al. 2004 (see Table 1).

The rate of glucose entry into blood and the duration of the elevated blood glucose is known to induce many hormonal and metabolic changes that may affect health and...
Table 1. Food factors affecting the glycaemic response of foods and meals (adapted from Arvidsson-Lenner et al. 2004)

<table>
<thead>
<tr>
<th>Food factor</th>
<th>Examples of influencing factors</th>
<th>Effect on glycaemic response and glycaemic index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross matrix structure</td>
<td>Grinding</td>
<td>Higher when homogenised</td>
</tr>
<tr>
<td>Cell-wall and starch structure</td>
<td>Degree of ripening</td>
<td>Higher with ripening</td>
</tr>
<tr>
<td>Granular starch structure</td>
<td>Heat treatment</td>
<td>Higher when gelatinised</td>
</tr>
<tr>
<td>Amylose and amylopectin content</td>
<td>Amylopectin is branched and more rapidly digestible than amylose</td>
<td>Lower with higher amylose content</td>
</tr>
<tr>
<td>Gelling dietary fibre content</td>
<td>Added gelling fibres</td>
<td>Higher with increased amylopectin content</td>
</tr>
<tr>
<td>Organic acids, e.g. acetic acid</td>
<td>Added acids</td>
<td>Reduced</td>
</tr>
<tr>
<td>Amylase inhibitor</td>
<td>Added</td>
<td>Reduced</td>
</tr>
<tr>
<td>Monosaccharide composition</td>
<td>Type of added sugars, e.g. glucose:fructose ratio</td>
<td>Reduced with increased fructose content</td>
</tr>
<tr>
<td>Molecular composition of carbohydrate</td>
<td>Type of raw material</td>
<td>Reduced with increased number of bonds other than α 1–4 and α 1–6</td>
</tr>
<tr>
<td>Resistant starch content</td>
<td>Heating—cooling cycles</td>
<td>Indifferent when testing equal amounts of available carbohydrate</td>
</tr>
</tbody>
</table>

This has led to review papers (Wolever, 1990a; Wolever et al. 1991), which discussed the influence of methodological variation and provided several recommendations for GI measurement. More recently the methodology of measuring GI was discussed by an expert panel, as part of a global discussion on the role of dietary CHO in nutrition (Food and Agriculture Organization, 1998). This panel agreed on a reference methodology and provided guidelines for measurement in future GI testing. In line with this development, there is an increasing number of foods which have been characterised regarding glycaemic response as attested by the last version of the international table of GI (Foster-Powell et al. 2002). Increasing information is also at hand regarding the food mechanisms responsible for differences in GI between foods and disease parameters. In this respect, low-GI foods were often found to induce benefits on risk factors for certain chronic diseases. Because of these observations it was proposed that GI data for foods could be used to make priorities for food selection within food groups.

Meanwhile, many studies have examined the short-term biological and health effects of foods, meals and diets of varying GI (Jenkins et al. 1987; Brand-Miller, 1994; Wolever & Bolognesi, 1996a; Järvi et al. 1999; Kaplan et al. 2000; Foster-Powell et al. 2002; Benton et al. 2003; Wolever & Mehling, 2003). More recently, intervention studies were developed (Brand et al. 1991; Frost et al. 1994, 1998; Giacco et al. 2000; Gilbertson et al. 2001; Wolever & Mehling, 2002; Bouché et al. 2002; Rizkalla et al. 2004) and some epidemiological studies based on prospective cohorts have provided new conclusions about the possible implications of GI on health; for example, diabetes (Salmeron et al. 1997a,b; Meyer et al. 2000), CVD (Frost et al. 1999; Liu et al. 2000, 2001; van Dam et al. 2000) and cancer (Slattery et al. 1997; Augustin et al. 2001, 2003; Franceschi et al. 2001; Jenkins & Franceschi, 2001). GI may also have relevance for sports performance (Thomas & Franceschi, 2003), whereas its role for obesity has recently been debated (Pawlak et al. 2002; Raben, 2002).

Recently Livesey (2002) addressed the effects of low- and high-glycaemic meals and diets on health and disease-related parameters. Based on recent observations it is expected that reductions in daily glycaemic load (GL) may lead to a reduced risk for developing diabetes and CVD. For example, Salmeron (1997a,b) showed that the GL of the daily diet correlates with the risk of developing diabetes in women but not in men. Brand-Miller et al. (2003a) observed a clinically significant decrease of protein glycation with a reduction of GL of the diet. Such observations encompass the potential to reduce the glycaemic response to foods by such means as the substitution of available CHO with indigestible or non-available CHO or with protein and/or fat. In contrast, in the Framingham Offspring Cohort it was recently demonstrated that whole-grain intake is inversely associated with homeostasis model assessment of relative insulin resistance (HOMA-IR), and a lower prevalence of the dysmetabolic syndrome, whereas dietary GI, but not GL, is positively associated with HOMA-IR and the prevalence of the dysmetabolic syndrome (McKeown et al. 2004). According to Brand-Miller (2004), the quality of CHO (i.e. GI) more often shows a significant association with disease risk (diabetes, CVD, cancer) than does the CHO content or GL of the diet. Very recently, several other papers and reviews addressed the impact of GI and GL on health aspects (Tavani et al. 2003; Frost et al. 2004; Kelly et al. 2004; Opperman et al. 2004). No benefits occurred in the Frost study, but according to the authors it cannot be excluded that potential effects may have been concealed due to drug therapy. In the Tavani study, a higher GI slightly increased risk for acute myocardial infarction but only in elderly individuals (more than 60 years) in association with overweight. On the other hand, the meta-analysis by Opperman et al. (2004) supports the value of low-GI foods to lower total cholesterol and improve metabolic control of diabetes. The benefits in the Kelly study were modest, and appeared mainly on total cholesterol and glycated haemoglobin. Based on the publications listed above there is accumulating evidence (Brand-Miller, 2004) that diets containing a preponderance of foods that elicit low glycaemic responses (‘low-GI foods or diets’), as originally defined by Jenkins et al. (1981) induce modest to clinically important benefits in the intermediate term as shown by intervention studies, and from epidemiological studies of health benefits in the longer (6–10 year) term.
modulation of GI of foods (Björck et al. 2000; Augustin et al. 2002; Björck & Elmstahl, 2003).

There is also a growing interest in GI from research, public health and industrial bodies. Recently, the FAO and WHO recommended that the bulk of CHO should be of low GI and rich in NSP (Food and Agriculture Organization, 1998) In several countries (Australia, France, Sweden, Canada and South Africa), the use of the GI concept has been integrated in dietary guidelines given by health professionals, and an increasing number of food companies market low-GI products. In line with these developments there is an increasing interest in the measurement of glycaemic response and GI of foods by a large number of academic and commercial laboratories, for both research and commercial application purposes.

The glycaemic index concept, as any concept, has ‘pros’ and ‘cons’

Recently there has been debate on the accuracy of some aspects of the method for the measurement of GI. This was summarised by Pi-Sunyer (2002) and Monro (2003). In addition, there are other ways of describing glycaemic responses of foods or drinks of mixed macronutrient composition, such as the glycaemic glucose equivalent (GGE; Monro, 2002, 2003), which ultimately is an index that arguably more closely represents food portion sizes, and GI (GL refers to the product of the amount of available CHO in a certain amount of food and its GI, divided by 100; Liu et al. 2003). These enable an index of glycaemic effects to be assigned to all types of food.

Another issue relates to CHO types (Asp, 1995) and analytical methods. According to the current physiological definition of dietary fibre (Champ et al. 2003), all indigestible CHO are considered to be dietary fibre; thus also resistant starch (RS), non-digestible oligosaccharides and sugar alcohols (polyols). Yet the classical method of measuring dietary fibre did not measure these CHO appropriately, leading to an underestimation of the true content of unavailable CHO in foods and diets. The question ‘has this influenced the presently published GI values of foods’ (Foster-Powell et al. 2002) is valid in this respect. It should be noted, though, that the majority of commercial foods included in the international GI tables contain low levels of these sources of indigestible CHO.

Thus, it is acknowledged in this respect that it is impossible to answer all questions, as the required data are not always available. It is also acknowledged that alternative concepts of communicating the glycaemic responses to foods and mixed meals might be more appropriate to inform healthy consumers as well as patients, including diabetics, about the health impacts of this aspect of their diets. In all cases, however, long-term studies (and none less than 8 weeks) are needed to establish the extent, nature and circumstances of health benefits, this including studies on GI. The current intensified interest in glycaemic response will undoubtedly lead to further progress of refining or extending the current methods as well as of defining the best measure for clinical and epidemiological research.

Since there are differing opinions on various aspects of measuring glycaemic response to foods and drinks and the possible use of a descriptive value to support well-informed food and drink choices, there is a need for consensus on the methodology of measuring glycaemic responses and converting the results to comparative values (Arvidsson-Lenner et al. 2004; Laville, 2004). Therefore, in the present paper we aim at discussing frequently asked questions on how to measure a glycaemic response and calculate a GI value according to the classical GI method appropriately.

We provide information about the scientific backgrounds, the methodological steps to be preferred, and comment on how deviations can be accepted without detrimental effect to the results. We describe key criteria related to choice of an appropriate control, number of subjects, the importance of comparing equal amounts of available CHO and how to deal with determination and description of these. We also give recommendations for future research and further developments related to measuring the glycaemic response to foods and drinks in our daily diet and its possible impact on health and disease.

Questions on methodological aspects of glycaemic index

Subject number

How many subjects should be enrolled?

Scientific background. As in all studies, the number of subjects enrolled determines, in part, the width of the CI for the estimates obtained, and the power of the study to detect differences in GI. Using more subjects provides better power and more precise results, but at a higher cost. The CI and experimental power of a study depend upon the number of subjects studied and the variability of the endpoint measured. A recent interlaboratory study (Wolever et al. 2003), in which GI values of five foods were determined in forty-seven subjects, provided a reasonable estimate of the inherent variability of results from GI determinations. The SD of GI values were linearly related to their means (Fig. 1). Thus, the width of the CI and power analysis for GI values depends not only on the number of subjects but also on the mean GI. Fig. 2 shows half-widths (or margins of error) of 95 % CI by number of subjects and mean GI in healthy individuals for GI values based on glucose (i.e. the GI

Fig. 1. Relationship between mean glycaemic index (GI) and SD of GI value according to the classical GI method appropriately.
of glucose = 100). Fig. 3 shows the difference in GI that can be detected with 80% power at a level of $P < 0.05$ (two-tailed) by number of subjects and mean GI. It can be seen that use of ten subjects provides useful results; large improvements in power and precision would require two to three times more subjects.

To further examine this question, a simulation analysis was carried out to examine the effect of sample size on the variability of the precision of estimates of the mean GI. The simulation used data from the thirty-seven subjects from whom capillary blood was taken and who performed exactly three repeats of the reference food glucose (Wolever et al. 2003). For each of the five foods (instant potato, bread, rice, spaghetti and barley), 1000 bootstrap samples of sizes eight, ten, twelve, twenty and thirty-seven (twenty-seven for spaghetti and barley), were taken from the individual GI values and the corresponding mean and 95% CI for the margin of error were calculated for each bootstrap sample (Fig. 4). As can be seen from Fig. 4, the width of 95% CI, which is an estimate of the precision of the estimated margin of error, decreases with sample size and can vary widely in small sample sizes.

**Recommendation.** The inclusion of ten subjects provides a reasonable degree of power and precision for most purposes of measuring GI. The number of subjects can be increased if the aim of the study is to detect small differences in GI or when greater precision is required.

**Test number**

How many times should the test be replicated for the standard and test foods?

**Scientific background.** In determining the GI of a series of test foods, the area under the curve (AUC) of the reference food is used as the denominator of every other test food. Precision will be improved if the measurements on the test food and the reference food are repeated among individuals. However, repeating all measurements adds undesirable expense.

Variation in the response to the reference food has greater effect on the results than the corresponding variation in the test foods, because the former is used to calculate the GI value of every test food in the series. Thus, the focus of this discussion will be how many times the reference food trials should be repeated. It is desirable to obtain as representative a value for the AUC for the reference food for each subject as possible. This can be achieved in practice by using the mean value of several trials of the reference food. The average of three trials of the reference food has been shown to reduce the variation of mean GI values (Wolever et al. 1991, 2002). The suggestion to use three repeated trials, as opposed to some other number, was based on a relatively small number of clinical observations. Although not enough real data exist to determine exactly how many repeats of the reference food should be done, both a simulation study and a theoretical argument presented later indicate that two or three measurements are appropriate.

From a theoretical perspective, we are interested in examining how the precision of a ratio of two measurements can be improved by replacing the denominator by the mean of several measurements. Using a linear approximation from a Taylor Series’ expansion, an approximate expression for the variance of $Y/X$ is:

$$\frac{\mu_Y^2}{\mu_X^2} \left( \frac{\sigma_X^2}{\mu_X^2} + \frac{\sigma_Y^2}{\mu_Y^2} - 2 \rho \frac{\sigma_X \sigma_Y}{\mu_X \mu_Y} \right),$$

where $\mu_X$ and $\mu_Y$ are the means of $X$ and $Y$, $\sigma_X$ and $\sigma_Y$ are their standard deviations, and $\rho$ is the correlation between $X$ and $Y$. We will assume that the CV ($\sigma/\mu$) are constant and the same for both $X$ and $Y$. We will denote this CV by $\nu$. Replacing $X$ by the mean of $k$ measurements reduces its sd to $\sigma_X/\sqrt{k}$ and as a result the sd of the ratio $Y/X$ is approximately:

$$\sqrt{\frac{\mu_Y^2}{\mu_X^2} \left( 1 + \frac{1}{k} - \frac{2\rho \sigma_X \sigma_Y}{\mu_X \mu_Y} \right)} \left( 1 + \frac{1}{k} \right),$$

where $\rho$ is the correlation between $Y$ and the mean of the $k$ values of $X$. It should be noted that as $k$ increases, so will the correlation $\rho$. Our goal is to minimise the sd of the ratio. For the interlaboratory study (Wolever et al. 2003), correlations between test food and reference food AUC measurements were typically 0.6 to 0.7. As an approximation, if we fix $\rho$ at 0.6, the ratio sd is minimised for $k = 2.78$ and for $\rho = 0.7$, which is the typical correlation of two measurements in a clinical setting. Therefore, three repeats of the reference food are appropriate.
the SD is minimised for $k = 2.06$, i.e., a minimum of two reference food measurements should be taken and three measurements is appropriate when the correlation in AUC measurements between the reference and test foods is more moderate.

In order to get a further indication of the effect of the number of reference food measurements on the precision of the resulting estimates of GI, a simulation study was carried out based on the data of Wolever et al. (2003). The CV of glucose AUC values was assumed to be constant over all subjects, and taken to be the pooled value of CV for all thirty-seven subjects for whom there existed exactly three repeated trials of oral glucose (reference food) and from whom capillary blood was taken. This pooled CV was 27.9%. For each subject, a number of AUC values were simulated from a normal distribution whose mean was the mean AUC elicited by oral glucose (reference food) for that subject and whose CV was the pooled estimate. The observed AUC values for five foods (instant potato, bread, rice, spaghetti and barley) were divided by the mean of the simulated responses to the reference food to give the GI. Based on the resulting GI, the margin of error (the half-width) for a 95% CI was calculated. The analysis was run using one or the average of two, three, four, five, ten and 100 simulated repeated values of the reference food. The results are shown in Fig. 5. The margin of error of the estimate of mean GI decreased substantially from one to two reference measurements. There was virtually no benefit to be gained from taking more than three to four measurements of the reference food AUC. The decrease in the margin of error from one to two reference foods is statistically significant; there are no statistically significant differences among the margins of error for any other number of reference food measurements.

**Recommendation.** We recommend repeating the trials of the reference food at least once, to obtain at least two values, in each subject.

**Subject status**

What should be the physiological or pathophysiological status of the subject?

**Scientific background.** As it is known that subject characteristics such as the insulin sensitivity and glucose tolerance status influence the glycaemic response to a food, it is of interest whether the physiological or pathophysiological status of the subjects needs to be considered for reasons of precision and comparability of GI values, as well.

Theoretically, the highest precision of GI values is expected when the within-subject variation of the incremental AUC of the blood glucose response is minimised. The variability of the glycaemic response to the same food tested on several occasions by the same subject can be expressed as the CV (CV = SD as a percentage of the mean). This question was addressed in one study that compared the average CV of glycaemic responses for repeated tests of white bread of subjects without diabetes.

![Fig. 4. Bootstrap estimates (---) and 95% CI (---) for the margins of error of 95% CI for mean glycaemic index (GI) for five foods (potato (a), bread (b), rice (c), spaghetti (d), barley (e)) for different numbers of subjects (GI of glucose = 100).](image1)

![Fig. 5. Margin of error (half-width of 95% CI) of simulated glycaemic index (GI) values calculated in thirty-seven subjects in which the area under the curve for the reference food was the mean of 1 to 100 simulated trials (GI of glucose = 100). (---), Mean; (---), potato; (---), bread; (---), rice; (---), spaghetti; (---), barley.](image2)
with those of subjects with type 2 diabetes and subjects with type 1 diabetes. Subjects with type 2 diabetes displayed significantly less within-subject variation (CV 15 %) than subjects with type 1 diabetes (CV 29 %; \( P = 0.008 \)) with normal subjects being intermediate (CV 22 %) (Wolever et al. 1985; Wolever, 1990a).

The question, whether GI values as relative glycaemic responses (RGR) are still affected by subject characteristics in general and whether GI values obtained in one kind of subjects are comparable or valid in different subjects, was addressed in several studies. Subject characteristics that have been examined specifically and were found to have no significant effect on mean GI include: normal vs. diabetic subjects (Jenkins et al. 1983); individuals with type 2 diabetes vs. individuals with type 1 diabetes (Wolever et al. 1987; Livesey, 2002); type 2 diabetic subjects on oral agents vs. on insulin (Jenkins et al. 1986); type 2 diabetic subjects in good vs. poor metabolic control (Wolever et al. 1986b); children with type 1 diabetes vs. adults with type 1 diabetes (Wolever et al. 1988a; Livesey, 2002); rural African vs. normal Western subjects (Jenkins et al. 1981; Walker & Walker, 1984); glucose tolerance status and BMI (Wolever, 1990a; Wolever et al. 1998).

**Recommendation.** Subject characteristics do not appear to have a significant effect on mean GI values but the variation of the values may differ in various groups, being highest in individuals suffering from type 1 diabetes. Therefore, for routine testing, healthy human volunteers are recommended.

What should be the reference food?

**Scientific background.** A variety of foodstuffs have been used as reference for GI measurement. In the updated GI database (Foster-Powell et al. 2002), we counted ten different reference foods for almost 1300 measurements: glucose, bread, white bread, wholemeal barley bread, wheat chapatti, potato, rice, wheat and arepa (a Mexican starchy food). However, most of the measurements were done using glucose or white bread as the reference food (> 90 % of studies).

In studies using rice, potato or some other local foods as the reference, these were chosen because of the large part taken by these foods in the local CHO intake. In the case of polyols, sucrose was often used as control as this is usually the CHO replaced. However, detailed information about the reference foods are sometimes not provided in reports, which limits the applicability of such data in those cases.

Furthermore, the correspondence between the glucose-raising effect of such or many local reference foods and the glucose-raising effect of the frequently studied white bread and glucose reference foods has been poorly studied. This greatly limits comparison of data from different laboratories and experiments, and it also limits dissemination of the results at an international level. By contrast, some laboratories have studied systematically the correspondence between GI on either white bread or glucose reference basis (Sugiyama et al. 2003). Of the reference foods used, bread or glucose have been discussed most. Since bread is a common food, it has been argued that choosing this reference allows the determination of GI in a more physiological manner. However, the composition of white bread may vary from one experiment to another and this may make comparison of results from various experiments difficult. Nevertheless, in a recent interlaboratory study it was concluded that GI values for locally obtained bread are no more variable than those for centrally provided (comparable) foods (Wolever & Mehling, 2002). In contrast, white bread from a certain local area may differ in GI compared with a common white bread from another area, as happened for the French white bread baguette (Foster-Powell et al. 2002). The GI values obtained when white bread is used are typically about 1.4 times those obtained if glucose is used as standard (Food and Agriculture Organization, 1998).

Glucose is sometimes preferred since the CHO source is more standardised. However, occasionally subjects can get nauseous after taking a concentrated glucose drink in the morning after an overnight fast.

**Recommendation.** We recommend the index is expressed relative to glucose = 100. However, for practical purposes, we feel it is acceptable to use reference foods other than glucose (such as white bread) during the measurement of GI. This can be done as long as they have been calibrated against glucose and the condition of preparation of this food is standardised. Dissemination of results should include the calibration data.

**Meal volume, composition and consumption time**

(a) Does drink CHO concentration and osmolarity play a role?
(b) Within how much time do the subjects have to consume the test food?
(c) Should the test-meal volume be adjusted?

**Scientific background.** The time required for consumption of the test food should allow convenience for the test subjects while minimising effects on gastric emptying rate and feelings of discomfort. The osmolarity of a test drink seems to play no important role but the CHO content does. It was shown that there is a linear relationship between CHO content of the drink and gastric emptying time. Because of this, the rate of delivery of CHO from the stomach to the small intestine is fairly constant for soluble available CHO. For example Brouns et al. (1995) observed a constant delivery of about 1.2 g CHO/min for various solutions containing a range of 45–90 g CHO/l and an osmolarity of 243–374 mOsm/kg. Accordingly, the CHO concentration and osmolarity of the test drinks, at least in the range tested, are not expected to have any impact on glycaemic response.

In the case of foods with low to moderate CHO density, in order to prevent a too large meal size, the available CHO content of the test portion is lowered to 25 g (see question on CHO load; pp. 155–156). Moreover, as the sessions take place in the morning, after a 10–14 h overnight fast (see question on timing of the test; p. 156), the subjects receive a drink during the test, for their rehydration and comfort. However, even if the amount of drink taken by the subjects is controlled, it differs from one study, or laboratory, to another, being a second source of variation of the volume.
The determination of the GI of a food usually involves the test of a portion of this food containing 50 g available CHO (see later; p. 156). As all foods do not have the same CHO content, the volume and weight of the test portions varies. Both volume and meal weight will have an impact on the time required for consumption. The question arises as to whether this will influence glycaemic response. It is known for both solid and liquid test meals, in normal and diabetic subjects, that prolonging the time of ingestion from a few minutes to several hours markedly flattens the glycaemic response (Jenkins et al. 1990, 1992). Consuming 75 g glucose solution over 10 min as compared with 1 min tended to increase blood glucose 60 to 120 min later, possibly because of a significantly smaller insulin response during the first 30 min (Heine et al. 1983). The impact of the volume, CHO content and osmolarity on gastric emptying has been well studied (Brouns et al. 1995; Brouns, 1998). However, the effect of volume of CHO foods per se has been poorly studied. Fluids have a strong effect of initial volume on the gastric emptying rate. Accordingly, ingestion time should best be standardised (Brouns, 1998). Unfortunately, the results of studies undertaken to date to evaluate the effect of volume on glycaemic response are contradictory. Some authors found no significant effect of water volume (Gregersen et al. 1990), whereas others observed or indicate a relationship between the volume of water or food and postprandial responses, and sometimes incremental areas (Torsdottir & Anderson, 1989; Young & Wolever, 1998; Sievenpiper et al. 2001). These contradictory results could be explained by differences in physiology between diabetic and normal subjects and also by large variations between the protocols used in these studies. Based on current information, it is recommended to avoid potential variability in results due to the amount of drink given to the subjects during the GI testing. Thus, what can be proposed is to give the subjects a standardised amount of drink, such as 250 ml (one glass), without taking into account water included in the food, and the reference used for the GI testing should be 50 g glucose diluted into 250 ml of water, or a portion of white bread containing 50 g available CHO given with 250 ml of a drink.

For solid meals the effect of varying time of ingestion over periods of less than 30 min has not been investigated. Based on practical experience it seems that consumption within 10–15 min depending on the type of food is appropriate. For example, very sweet foods with a substantial volume may not be consumable for every individual within 10 min.

The first bite or sip in the mouth is set as time 0 and the first blood sample is taken conventionally at exactly 15 min afterwards.

Other factors. In GI testing, the drink given to the subjects is not always water; coffee or tea have also been used. Young & Wolever (1998) showed that neither tea nor coffee significantly affect the incremental area under the glycaemic response curve; however, as volume, they influence the pattern of blood glucose response. Moreover, caffeine is known to acutely decrease insulin sensitivity in human subjects (Graham et al. 2000; Keijzers et al. 2002). In the study of Graham et al. (2000), ingestion of 375 mg caffeine resulted in a significant increase in both C-peptide and insulin as well as an increase in the glucose AUC of 24 %. Thus, the best drink to administer seems to be water. As a second step, non-energetic or stimulant beverages (coffee, tea) may be supplied as long as an identical quantity and type of beverage is administered during the different sessions of each subject.

Recommendations. We recommend supplying a standard amount of 250 ml water to the subjects with the test portions, and with the white bread portion if it is used as referent. If the referent is glucose, we recommend using a solution of 50 g glucose diluted into 250 ml water. Fluid ingestion is advised to take place within 5–10 min. Solids and semi-solids should be ingested within 10–20 min, depending on the type and taste of the food. The first blood sample should be taken exactly 15 min after the first bite of the food or first sip of the drink.

Carbohydrate basis for glycaemic index determination

Should the test serving be based on available (glycaemic) CHO?

Scientific background. In the classical GI concept only CHO sources that are assumed to be fully digestible, absorbable and glycaemic are included in the calculation of the 50 g CHO portion size. However, this approach has been criticised for several reasons (Monro, 2003). One is the suggestion that ‘available’ CHO is difficult to measure accurately. Accordingly, the usual way of determining ‘available CHO’, by subtracting ‘dietary fibre’ from total CHO, is not really a true reflection of in vivo available CHO content and may lead to overestimation in the case of products containing indigestible CHO, which are not recovered as dietary fibre. Thus, it could be argued that ‘available’ CHO is only a concept rather than a definable quantity. The truly available CHO fraction is difficult to determine by in vitro models, since these do not account for any influences within the gastrointestinal tract, for example the fact that some in vitro digestible CHO may escape small-intestinal digestion and absorption, thus ending unabsorbed in the colon. The analytical aspect is particularly important in the case of starchy foods. Until recently it was difficult to measure the amount of type 3 RS in a food. The former AOAC analytical dietary fibre method did not measure type 3 RS appropriately, leading to some underestimation of the total RS content of many foods in the past. However, in 2002 AOAC recognised an additional new method (McCleary & Monaghan, 2002) that measures type 3-retrograded RS specifically and accurately.

This analytical progress seems of relevance since an unknown and significant fraction of native starch, once it has been prepared for consumption, may be resistant to digestion. This fraction is usually referred to as intrinsic RS. The amount of RS in a food is further influenced by the food processing and preparation conditions applied as well as by adding RS as a food ingredient. Accordingly, starch may be resistant due to the native (raw and indigestible) botanical structure or starch crystallinity being formed after cooking, during cooling down, by a process called...
retrogradation (formation of indigestible crystalline structures). Alternatively, RS can be added as an ingredient during food processing. The RS content may also change depending on the degree of ripening of a fruit or vegetable. For example, a green banana has a very high content of RS, which is lowered during ripening to almost zero in a fully ripe banana. As a consequence, the amount of RS in a starch-based food may range from a few percent to as much as 40% depending on the native source and/or the way of meal preparation (Champ et al. 2003). Analysis of total starch less RS obtained by approved methods makes possible estimation of the available starch fraction. There are also several enzymic procedures available simulating the digestion in the gastrointestinal tract and making it possible to analyse the available and RS content of foods (Englyst et al. 2005). This will be discussed later (p. 153) under the heading of second-meal effects.

In order to address the issue of whether a GI value should be based explicitly on available CHO, we need to ask also to what extent the low glycaemic responses elicited by low-GI foods (based on available CHO) are due to CHO malabsorption. If low glycaemic responses occur due to CHO malabsorption, then there would be no value in retaining the concept that the GI should be based on available CHO. However, this seems not to be the case. Although low-GI foods tend to be associated with a decreased amount of CHO absorbed (i.e. the amount of available CHO they contain has been overestimated), this only accounts for maximally 15% of the blood glucose-lowering effect. In addition, the effects of a low-GI food, lentils, on the second-meal glycaemic response is the opposite to that seen with a reduced amount of CHO absorption, but the same as that seen by slowing CHO absorption (Jenkins et al. 1990) (see also second-meal effects later; p. 153).

Amount of carbohydrate absorbed. Wolever et al. (1986a) compared the amount of CHO malabsorbed from whole-wheat bread and lentils using a breath hydrogen method and comparing the obtained data with those of directly measured contents in ileal effluent from individuals with an ileostomy. Both methods indicated that about 10% of the ‘available’ CHO in bread and about 18% in lentils was malabsorbed. Thus, portions of bread and lentils containing 50 g ‘available’ CHO (defined as total CHO minus dietary fibre) actually contained about 45 g and 41 g, respectively, of the truly measured absorbed CHO. This difference in absorbed CHO would account for about a 9% difference in glycaemic response (100 x (45 – 41)/45 = 8.9%). In fact, however, the glycaemic response elicited by lentils was 75% less than that of bread (Wolever et al. 1986a). Thus, malabsorption only accounted for about 12% of the reduction in glycaemic response (100 x 9/75 = 12%).

Fig. 6 shows the relationship between GI and percentage CHO malabsorbed for twenty foods as directly measured in a subject with an ileostomy. The dotted line indicates the magnitude of reduction in glycaemic response which would be expected based on the reduced amount of CHO absorbed. Even at high levels of CHO malabsorption seen with legumes, the reduction in CHO malabsorption accounts for < 20% of the reduced glycaemic response. It is therefore believed that the major determinant of the low GI values of starchy foods in this study is a reduced rate of CHO absorption. Of course other mechanisms are possible, but it is true that slowing the rate of absorption of CHO, for example by sipping glucose slowly over 3 h, reduces glucose and insulin responses (Jenkins et al. 1990).

If, as suggested by Fig. 6, not all of the ‘available’ CHO in foods is actually absorbed, the question is raised as to how much GI values are underestimated because the portion sizes used contained less than 50 g available CHO. This question has been addressed by determining the glycaemic responses of starchy foods containing high amounts of intrinsic RS (15 or 35% on total starch basis) using two different portion sizes.
Glycaemic index as a function of the digestibility of available carbohydrate. In several comparative studies, the glycaemic impact of CHO has been related to the rate of digestion of available CHO. Englyst et al. (2003) correlated twenty-three cereal-based starchy foods for their GI in vitro and their rate of CHO digestion in vitro (as assessed by the content in rapidly available glucose). Foods were ingested on a 50 ± 2 g available CHO basis. According to the results of analytical measurement of NSP and RS contents the twenty-three foods had very similar content in non-available CHO. However, the GI was highly variable between the different foods studied but was significantly related to the in vitro rate of digestion (rapidly available glucose content). Accordingly, GI was more explained by the rate of digestion of the available CHO fraction than by the content in non-available CHO (dietary fibre and RS).

Second-meal effects. To see if the acute metabolic effects of a low-GI food, lentils, was due to a reduced rate or a reduced amount of available CHO absorption, Jenkins et al. (1982) studied the effects of different breakfast test meal on the glycaemic response after a standard lunch. Subjects consumed four different breakfast meals, followed 4 h later by a standard lunch. The two basic breakfast meals were matched for available CHO (127–128 g), dietary fibre (26–29 g), protein (57 g) and fat (13–18 g) and consisted of lentils, butter and tomato, or wholemeal bread, cheese and tomatoes. The breakfasts were consumed over a period of 15–20 min. Two further breakfasts were taken; one, representing reduced CHO intake, contained only one-quarter of the wholemeal bread but was identical in all other respects. The other, representing slow CHO absorption, contained the full amount of bread, cheese and tomato but eaten in small divided portions over 4 h. The glycaemic response elicited by lentils and the continuous and the quarter bread meals were significantly less than that of the control bread breakfast (Fig. 7). However, the pattern of blood glucose response after lentils (Fig. 8) was followed much more closely by the continuous meal than by the quarter bread meal (Fig. 9). The second-meal response refers to the glycaemic response elicited by the standard lunch meal consumed 4 h after breakfast. The second-meal response after lentils was 38 % less than that after the control bread breakfast (Fig. 7). The second-meal response tended to be less than control after the continuous bread breakfast but was significantly increased after the quarter bread breakfast (Fig. 7). These results suggest that the acute metabolic effects of slowing CHO absorption differ from those of reducing the amount of CHO absorbed, and that the metabolic effect of lentils is not due to CHO malabsorption. Similar conclusions were drawn from second-meal experiments done by Liljeberg et al. (1999).

These data suggest that the concept of GI based on 50 g available CHO differs from the concept of the GL or GGE based on glycaemic response as a food effect and it is therefore essential to distinguish between these measures.

Other proposed methods of classifying the glycaemic impact of foods. One method is an index based on total CHO. A relative glycaemic effect (RGE) could be determined by...
measuring the glycaemic responses elicited by amounts of foods containing 50 g total CHO, expressed relative to 50 g glucose; that is, using a total CHO basis.

There seem to us several reasons why it is necessary to distinguish between GI, based on available CHO, and RGE based on total CHO. First, unavailable CHO may or may not influence the glycaemic impact of digestible CHO consumed at the same time with them. Not controlling for the amount of available CHO would mean that valid conclusions as to whether reduced glycaemic responses were due to a reduced amount of available CHO or to some other effect cannot be drawn. Second, as described elsewhere in the present paper, the metabolic effects of consuming a certain amount of CHO while slowing its absorption are different from those of reducing the amount of CHO absorbed. Regarding the latter, most of the studies in this area compare the effects of exchanging sucrose or dextrins with fructo-oligosaccharides or inulin (Williams & Jackson, 2002). Effects that are seen in these cases may, along with a reduction of the amount of glycaemic CHO ingested, be due to the production of short-chain fatty acids during colonic fermentation (Brighenti et al. 1999).

Another method is based upon GL or GGE. The term GL was first proposed by Salmeron et al. (1997a) and was defined as the quantity GI \times g, where GI is the food GI and g is the weight of CHO the food portion contains. Thus, a particular GL may be achieved either from low-GI foods, or from high-GI foods at a lower available CHO intake. The ability of GL to predict glycaemic responses has been tested (Brand-Miller et al. 2003b), and GL values for a large number of foods have been published (Foster-Powell et al. 2002). It has been assumed that if more servings of food are consumed, GL values are additive without energy adjustment (Brand-Miller et al. 2003b); however, whether this is so is not clear since the original rationale for the use of GL was based on epidemiological studies in which GL was adjusted for energy intake.

Monro introduced the term GGE (Monro, 1990a, b, 2002; Monro & Williams, 2000) which is defined as: ‘the weight of glucose equivalent in its effect to a given weight of food’. GGE can be calculated using existing GI tables as follows:

\[
\text{GGE} = \text{food weight (g)} \times \% \text{CHOAVL} \times (\text{GI}_{\text{food}}/\text{GI}_{\text{glucose}}),
\]

where %CHOAVL is the amount of available CHO contained in 100 g food, GI_{food} is the GI value of the food and GI_{glucose} is the GI of glucose (\(= 100\)). Thus, GGE is equivalent to GL.

It has been suggested that GGE is an accurate predictor of glycaemic responses (Liu et al. 2003), but the validity of this conclusion has been questioned (Wolever, 2004a).

In general the utility of the above-mentioned methods of classifying glycaemic responses, i.e. RGE, GL and GGE, remains to be better established. Issues which need to be addressed include: what is the reproducibility of the values; are the values the same in different subjects; are the values affected by glucose intolerance or insulin resistance; do they hold in mixed meals; does altering dietary intake based on them result in long-term improvement in glycaemic control, blood lipids, body weight, or other health-related indicators?

In general the validity of the above estimates of GI of available CHO and glycaemic impact remains to be better established. In this respect it should also be noted that a value of GI might not be appropriate when the intake of fat is high (see Flint et al. 2004).

**Recommendation.** It is recommended that the measurement of GI of test foods is based on comparing equivalent amounts of available CHO (for further explanation on the definition of available CHO, see later; p. 156). For clarity and precision in terminology, it is recommended that the term ‘GI’ be restricted to tests of foods based on available CHO, and that different terms, such as RGE, GL or GGE, be used to classify the glycaemic impact of foods based on total CHO or serving size.

**What should be the carbohydrate dose?**

**Scientific background.** Generally, most of the tested foods contain significant amount of available CHO (as g eaten in a
serving, or as g/100 g food). A 50 g available CHO load has been initially proposed as a reference amount for GI testing. However, in the case of foods with low to moderate CHO density, this may lead to large volumes having to be ingested. Increasing the intake quantity of CHO increases incremental AUC glucose response in a dose–response manner in healthy subjects (Wolever & Bolognesi, 1996a; Lee & Wolever, 1998). However, the dose–response curve is curvilinear and shows a trend to plateau at intakes greater than 50 g CHO. Small intakes of high-GI foods also elicit a glucose response, whereas glucose only slightly increases after the consumption of a 25 g CHO load of low-GI foods. In addition, after an intake of less than 10 g CHO, a significant change in blood glucose may not be detected. Therefore, a range of 25 to 50 g available CHO appears to be most appropriate for GI testing experiments, with 50 g as the preferred target dose. Less than 25 g available CHO (but higher than 10 g) may be considered for foods with very low available CHO content, but can at present not be recommended due to lack of data. Accordingly, this requires further research. Finally, when the reference food is glucose, the nature of the glucose source has to be taken into account. In the case of monohydrated glucose, a correction factor, taking into account the water linked to the glucose, has to be used to calculate the quantity of CHO provided from glucose. Indeed, 1·1 g monohydrated glucose provides 1·0 g glucose.

**Recommendation.** We recommend testing 50 g available CHO loads. In the case of foods with low to moderate CHO density, it is justified to lower the CHO load to 25 g to avoid an unrealistically large meal size and to adjust to 25 g CHO the referent food portion.

**How should carbohydrates be analysed and calculated?**

**Scientific background.** According to the FAO definition currently used in many countries, ‘total carbohydrate less dietary fibre’ is defined as available CHO. At present we know that this may often be incorrect, since this definition does not account for indigestible CHO that from a physiological point are dietary fibres (Champ et al. 2003) but are not measured using the classical dietary fibre analytical method. Some examples of such CHO are the fructans inulin and fructo-oligosaccharides and type 3 RS. Accordingly the available CHO should, as far as possible, be analysed with approved specific analytical methods. Simple and appropriate enzymic procedures are available for analysis of the main potentially available CHO in foods i.e. available starch, lactose, sucrose, glucose and fructose (maltodextrins, maltose). For practical purposes, these methods can be expected to give a reasonably good estimate of the available CHO load in foods with only intrinsic CHO. In the case of starchy foods, the bioavailable starch fraction as well as RS can be analysed with specific analytical procedures (McCleary & Monaghan, 2002; for a review, see Champ et al. 2003). When some CHO components such as RS or sugar alcohols may occur in non-transformed or manufactured foodstuffs, the evaluation of these components, using accurate procedures, is recommended.

Fig. 10 shows possible ways to obtain the analytical contents in CHO components and to calculate the load of a defined amount of available CHO to be tested. When required, a conversion factor is recommended to adjust CHO that are being tested (available starch, disaccharides, monosaccharides) to a comparable monosaccharide level. ‘Indeed 1 g available starch delivers 1·1 glucose units during digestion because of hydration during hydrolysis process, whereas the conversion factor to be used for available disaccharides is 1·05.’

**Recommendation.** It is recommended that the main potentially available CHO in foods, i.e. available starch, lactose, sucrose, glucose and fructose, are analysed with accurate analytical methods. For starch-containing foods, analytical procedures are recommended that allow separate determination of the available starch and the RS.

**When should the test take place?**

**Scientific background.** To be comparable with other foods in the GI tables, the test should be taken at breakfast time in the morning. This fasted condition is most stable concerning possible intra-individual differences due to time of the day and meal influences. There is evidence that the glycaemic response data obtained from a test at lunchtime, after a standard breakfast, differ significantly from those obtained after an overnight fast. For example, RGR of two breakfast cereals were studied after a 10–14 h overnight fast or at lunchtime after a standard breakfast. It was observed that the difference in glycaemic response between the cereals when tested in the morning was significantly greater than that observed at lunchtime (Wolever & Bolognesi, 1996c).

**Recommendation.** The test should preferably take place before 10.00 hours in the morning, after a 10–14 h overnight fast.

**Subject preparation**

What should be the preparation (instructions to subjects) before the measurement?

**Scientific background.** Many factors modulate the response to a test meal in any one subject and differences in these factors between eating occasions would be expected to be a potential cause of inaccuracies in GI, GL, GGE, etc. It has long been known that the CHO content of the diet for the few days before an oral glucose tolerance test markedly affects the result, with poorer glucose tolerance being observed after a low-CHO diet (Hales & Randle, 1963). Acute physical exercise can increase muscle glucose uptake on the following day (Malkova et al. 2000), and improve insulin sensitivity for 48 h (Mikines et al. 1988). This usually results in lower plasma insulin concentrations in response to nutrient ingestion (Tsatsonis & Hardman, 1996; Tsatsonis et al. 1997), although generally there is no effect on systemic fasting plasma glucose concentrations (Tsatsonis & Hardman, 1996; Tsatsonis et al. 1997; Malkova et al. 2000). Folch et al. (2003) recently studied the effect of a large starch meal at rest and after exercise in men and women and observed that the contribution of substrate oxidation to energy expenditure as well as fat and glycogen balance and the effect of previous exercise were similar in men and
women. Post-exercise, however, postprandial plasma glucose levels differed significantly from rest. Alcohol consumption may have profound effects on glucose homeostasis (Shelmet et al. 1988; Siler et al. 1998), especially in the fasting state, although the effects of alcohol consumption on metabolism the following day do not appear to have been tested. Cigarette smoking is another potential confounder since it may cause acute insulin resistance (Attvall et al. 1993; Frati et al. 1996). Intake of dietary RS on the day before a meal challenge may improve glucose tolerance (Robertson et al. 2002). Even the composition of the evening meal the day before a test may affect the result. Meals rich in either CHO or fat on the evening before a meal-challenge test have significant effects the following day: fat tolerance is poorer following a CHO-rich evening meal; glucose tolerance is poorer following a fat-rich (low-CHO) evening meal (Robertson et al. 2002). However, the effect of evening meal macronutrient composition on fat tolerance is much more marked than the effect on glucose tolerance (Robertson et al. 2002). The GI of the evening meal may also exert an effect independent of macronutrient composition. Evening meals of low GI produce better glucose tolerance the following morning compared with evening meals of high GI (Wolfver et al. 1988b; Thorburn et al. 1993). This effect may be due to colonic fermentation, since it is seen with barley but not with spaghetti with an identical GI (Granfeldt et al. 2005).

The exact length of the overnight fast may also be important: the steady state often supposed to exist after overnight fast is, in fact, a time of significant change in terms of falling plasma insulin concentrations and increasing lipolysis (Klein et al. 1993; Samra et al. 1996).

In view of all these factors, it might well be supposed that strict standardisation of diet and activity for 24 h before the measurement of GI would improve the reproducibility of results. In practice, however, such standardisation considerably increases the burden on participants and investigators, and it is important, therefore, to ask whether the supposed benefits are seen in reality. This question was studied by Campbell et al. (2003). They compared tests in which exercise and evening meal on the day before the test, and the length of overnight fast, were strictly controlled, with ‘uncontrolled’ tests. No smoking was allowed on the morning of the test on either occasion. In contrast to expectations, the variability of results was no greater in the ‘uncontrolled’ tests and in fact tended to be lower. There was no effect on mean values of glycaemic response. Participants who were smokers (three out of thirteen) reported ‘craving’ during the controlled tests, but nevertheless returned similar values for glycaemic response and variability under the two conditions. Vigorous physical activity on the day before the ‘uncontrolled’ tests did not produce a significant effect, although there was a tendency to lower blood glucose concentrations at 90 min. Whilst the effects of many of these measures have been shown under other laboratory conditions, it appears that they are not of sufficient significance to warrant stricter controls when applied to GI testing.

**Fig. 10.** Principles of analysis and calculation of available carbohydrate (CHO) in test foods. * This analysis is optional, and depends on the estimated content (step 2). † This analysis is recommended when possible. Results allow to calculate the total sugars content. A correction factor can also be applied to disaccharides (× 1·05) to convert into monosaccharide equivalents, depending on the method used. ‡ Correction factor of 1·1 is to be applied to convert starch content into monosaccharides (glucose) equivalents, and a correction factor of 1·05 is to be applied to convert disaccharides content into monosaccharides (glucose) equivalent (depending on the method used). § This last step of calculation is to be adapted according to the decided available CHO served load.
Re-utilisation of control test results

Can the results from the reference glucose tests from a selected group of subjects from one study be used for subsequent studies?

Scientific background. Using reference values that have been obtained in one study as reference values in another subsequent study would only be acceptable by respecting the following two conditions:

(1) Between two studies, the subjects must not undergo physiological changes (for example, significant weight change, starting medical treatment, changes in physical activity) leading to important changes in their fasting blood glucose or in their glycaemic response to foods;

(2) The position of the glucose tests in the studies is relevant (see question ‘how many times should the test be reproduced for the standard and test foods?’, p. 148).

If a glucose test is always done at the beginning and at the end of the study, the last glucose test of the first study can be used as the first glucose test of the second study. However, since no data are available on the impact of time on changes in response to the reference food, we recommend limiting the time between tests to a maximum of 2–3 months. Applying reference data from one study to data collected in another is generally considered poor scientific practice, although it may yield reasonable results. Although practiced often it is not recommended to do so.

Recommendation. In case that the same subjects are tested in subsequent trials with short intervals, test data of the reference food in a particular group of subjects can be used for different studies in the same subjects, if the test of the reference food occurred within 3 months. However, at least one new test of the reference food should be done in every study, and, in each study, one test of the reference food should be done at least every 3 months, or for every six foods tested by the subject, whichever is more frequent.

How should the randomisation be done?

Scientific background. Randomisation in clinical trials is necessary to eliminate bias in treatment assignment, facilitate blinding of the identity of the treatments from the investigators and to permit the use of probability theory to determine the statistical significance of the results (Schulz & Grimes, 2002a). As important as the treatment assignment is allocation concealment, to prevent the investigators from deciphering assignments, which subverts the purpose of randomisation (Schulz & Grimes, 2002a). These considerations apply to parallel-design clinical trials of drugs, but are difficult or impossible to apply to studies in which the GI of foods is determined. One reason for this is that GI trials are cross-over studies in which all subjects undergo all treatments. In addition, it is usually impossible to blind either the subjects or the investigator from the treatments being undertaken, and efforts to do so may change the nature of the food tested and hence alter the results. Bias in GI trials is possible if either subjects or investigators try to modify the results of a test by changing their behaviour based on knowledge of the upcoming test, or change the order of the treatments based on their behaviour. Creating a list of the order in which tests will be done, and not telling the subject about which food product will be received before it is presented as the test meal may help minimise these problems but is insufficient to prevent bias from occurring. The advantages of using true randomisation to allocate the order of test of the foods, as opposed to non-randomised methods of mixing up the order of treatments, is not clear. In addition to minimising bias, randomisation of the treatments in GI studies is desirable to control for systematic ‘order’ effects, and for non-systematic random events, both of which may influence the results. It is possible to achieve better balance among the possible orderings of test foods in time by using an adaptive randomisation procedure or minimisation. In this approach, orderings that include foods that have already been assigned to subjects disproportionately early in the order will become less likely to be assigned than orderings in which these foods are scheduled later. However, because of the additional complexity this creates in the study design and analysis, and because practical considerations as noted later may be of overriding concern, we do not feel there is a need for adaptive randomisation and opt instead to include order effects as a covariate in the analysis.

Examples of systematic order effects might be related to stress or seasonal effects. Subjects who are new to the testing procedures may experience more anxiety during the first few tests they perform than they do after becoming used to the procedures. Anxiety may influence glycaemic responses via delayed gastric emptying or insulin antagonistic effects of stress hormones. If subjects are studied over a period of several months, seasonal changes may influence the diet and activity patterns of the individual, which, in turn, may affect glycaemic responses. Examples of random effects might be a sudden hot spell, or a breakdown in the air-conditioning system in the laboratory, which affects the performance of the glucose analysers. Factors that influence how treatments are randomised include the nature and number of foods tested, the number of times the reference food is repeated and the number of subjects being studied. Consideration of how the number of subjects influences the methods of randomisation and concealment is critical for parallel studies, but not quite so important for cross-over studies. The main issue is that, with small numbers of subjects, simple randomisation procedures commonly result in unequal group sizes. For example, a sample size of ten,
over 33% of sequences generated with simple randomisation, would yield a ratio imbalance of 3:7 or worse (for a sample size of twenty, the probability is nearly 12%). Various strategies for randomising in blocks can prevent unequal group sizes, though it has been argued that some inequality in group size is desirable to enhance concealment (Schultz & Grimes, 2002b). Complete randomisation of all the test reference foods could result in all the reference food tests being conducted first, or last or all together in the middle. If a large number of foods were being tested, the study may last for 3–4 months, and the grouping together of the reference food tests would make it difficult to detect any systematic order effects. To determine whether systematic order effects are present, the three trials of the reference food could be conducted first, last and in the middle of the sequence of test foods. To minimise bias due to random events, it is most desirable to randomise the order of the foods tested. However, in certain situations it is not possible or desired to completely randomise all the food tests; for example, the desire to obtain results for some foods first, or the testing of foods with a short shelf-life (for example, fresh potatoes). Randomisation in blocks is acceptable, but doing reference food tests after each block test may be considered to control for order effects. True randomisation of the order of testing of the foods would be ideal, but is not commonly practiced, and is not considered necessary due to the absence of data about the magnitude of bias introduced by non-randomised allocation of the order of testing. More sophisticated experimental designs, such as partial latin square, where each subject ingests several but not all foods, may allow more foods to be tested in a single experiment in less time and with less burden on the subjects. Thus far, data using such approaches for GI testing are lacking, thus not allowing for making recommendations.

**Recommendation.** It is recommended that foods for testing be randomised in blocks of up to six foods with a trial of the reference food being done before and after each block. The reference food trial performed after the first block of test foods can be used as the first reference food trial in the next block of test foods. Randomisation is not the only approach to assign tests and subjects.

**Should we include as many males as females?**

No differences have been observed in glycaemic response between males and females (Wolever et al. 2002). Accordingly, there are no grounds to avoid the common practice of inclusion of both sexes in a study.

**How many foods in one batch of tests?**

(a) Can we test as many foods as we want in one test?

(b) Is there a maximal number of test foods that can be compared with the mean of two to three reference foods?

**Scientific background.** Studies that have been done in the past have evaluated as much as up to twelve foods in one GI experiment, along with three times the control test. The reference test has been repeated every 6–8 weeks. The total test duration should not exceed 4 months, to avoid possible seasonal variations in GI. An ideal set-up may be to test three test foods, along with three control-reference trials, in a period of 2 weeks, allowing to test on Monday, Wednesday and Friday.

**Recommendation.** The total test duration should not exceed 4 months. With multiple food testing a reference test should be done at the beginning and a repetition should take place after every 6–8 weeks.

**How should blood be sampled?**

**Scientific background.** Arterial blood is delivered from the heart to the tissues of the body. The concentration of glucose in arterial blood is therefore in principle the measurement of interest because this is the concentration to which the tissues are exposed. However, sampling of arterial blood is an invasive procedure that carries some risk and is unlikely to be applied to the measurement of GI. Alternatives are: (i) capillary blood, which approximates arterial blood in composition – this necessitates puncture of the skin (usually on the finger or earlobe); (ii) venous blood taken from a cannula in a vein, usually on the forearm or in the antecubital fossa; (iii) ‘arterialised venous’ blood, taken from a cannula in a vein draining the hand, the hand then being warmed to open arterio-venous anastomoses (so that substrate concentrations resemble arterial blood). Because the tissues consume glucose, the concentration of glucose in a peripheral vein will be lower than that in an artery. The magnitude of this arterio-venous difference may be considerable, especially after glucose ingestion when insulin stimulates glucose uptake. In one study, the glucose concentration in blood taken from an antecubital vein was 4 mmol/l lower than that in arterised venous blood at 60 min after ingestion of glucose (Frayn et al. 1989). Furthermore, this difference is not consistent; ambient temperature, for instance, affects the rate of blood flow through the forearm and has a marked effect on venous blood glucose concentrations (Frayn et al. 1989). Therefore, venous glucose measurement might be expected to be more variable than capillary. This has been confirmed directly in the interlaboratory study of GI, in which laboratories taking venous blood returned CI of >50%, compared with <30% in all laboratories using capillary blood (Wolever et al. 2002, 2004).

Direct comparisons between capillary and venous blood also bear out this difference. Capillary blood glucose concentrations were consistently greater than venous blood glucose concentrations when assessed during the measurement of GI of various products, so that AUC were 33–40% lower for venous blood glucose (Granfeldt et al. 1995b). However, the measurements of GI were not affected as both reference and test measurements were similarly affected. Sensitivity of the measurement, however, was greater using capillary blood sampling (Granfeldt et al. 1995b). The site at which capillary blood is taken may affect the result. When glucose concentrations are rising or have just risen sharply, fingertip capillary glucose concentrations are greater than those in capillary blood from other sites including forearm, thigh and abdomen (Ellison et al. 2002; Jungheim &
Koschinsky, 2002; van der Valk et al. 2002), with variable differences in repeated tests on the same subject (Ellison et al. 2002); conversely, when blood glucose concentrations fall, fingertip capillary glucose concentrations are lower than those on the forearm (Jungheim & Koschinsky, 2002). The message is that fingertip capillary sampling is preferable to other sites. Earlobe and fingertip capillary blood do not seem to have been compared directly.

Practicalities. Placement of a venous cannula requires some medical or nursing supervision; on the other hand, once in place, blood sampling is painless. Many subjects do not like repeated stabs for capillary blood; on the other hand, again, methods for sampling capillary blood are being refined and with practice the technique can be made almost painless. Placement of a cannula for sampling arterialised venous blood is made in a retrograde direction, which is slightly more difficult than normal cannulation. The hand can then be warmed in a variety of ways, although hot water is not recommended as heat transfer to the body is significant, and a box with relatively static air at 55–60°C works best in practice (Frayn & Macdonald, 1992). Sampling capillary blood results in relatively smaller volumes. If it is desired to measure substances other than glucose (for example, insulin) this may be possible using micro-methods. However, if a number of hormones or metabolites are to be measured, then larger samples will be needed, and arteriaлизed venous blood may be the method of choice (Wolever, 2004b). Measurement in whole blood is also possible. Whole-blood and plasma glucose concentrations are closely related and one can be calculated from the other if the packed cell volume is known (Dillon, 1965). Provided that either plasma or whole blood are being used consistently in any series of measurements, no bias should be caused. However, it would not be appropriate to use plasma on one occasion and whole blood on another.

Recommendation. We recommend the use of either capillary or arterialised venous blood and discourage the use of normal venous blood. Fingertip capillary blood appears to give the greatest sensitivity. We recommend the use of either capillary or arterialised venous blood and discourage the use of normal venous blood in order to improve sensitivity and to remove the potential for variations in measured GI due to fluctuations in factors such as ambient temperature. Fingertip capillary blood appears at present to give the greatest sensitivity, but with improved methods for sampling of capillary blood becoming available, more research could be conducted on alternative sites. Whole blood and plasma concentrations are closely related and one can be calculated from the other if the packed cell volume is known.

Measure both glucose and insulin?

Scientific background. A growing body of research resulting from dietary trials and large epidemiological studies indicates that diets with a high GI or high GL are associated with an increased risk of insulin resistance, dyslipidaemia, and the development of CVD and type 2 diabetic subjects (Slabber et al. 1994; Salmeron et al. 1997a,b; Frost et al. 1999). In addition, hyperinsulinaemia has also been implicated as a risk factor for the development of arteriosclerosis (Leeds, 2002), CVD (Frost et al. 1999, Leeds, 2002), diabetes (Willett et al. 2002), cancer (McKeown-Eyssen, 1994; Giovannucci, 1995) and juvenile myopia (Cordain et al. 2002).

Measurement of insulin response to foods, and the introduction of an insulinemic index concept (Holt et al. 1997) could be of value in investigating health issues. However, more data are necessary concerning potential differences in GI and insulinemic index between foods. Moreover, an insulin index is likely to be influenced by the pathophysiological status of the test subject (Wolever et al. 2004).

To date, most GI studies have not measured the postprandial insulinemic responses accompanying the glycaemic responses of the test foods examined. However, for many CHO-rich foods and pure CHO, there appears to be a close correlation between glycaemic and insulinemic responses (Holt et al. 1997; Englyst et al. 2003). It should be acknowledged though that certain foods produce higher insulin responses than expected from their GI (Ostman et al. 2001). Especially, protein is known to induce a larger insulin secretion, when combined with CHO. Thus, to obtain more data on glucose homeostasis of nutrient combinations, insulin measurement may be desirable. In case of the latter, most often venous blood will be needed to obtain blood samples that are large enough. Measurement of insulin in small capillary samples or using arterialised blood using a hand placed in a heat box is the method of choice but this needs more sophisticated equipment and experience.

Recommendation. For routine use of the GI method, glucose measurement is enough and we do not recommend to systematically measure the insulin level. When the measurement of insulin response in addition to glucose response is of interest for other reasons, or when budget limitations do not play a role, we recommend the measurement of both for completeness.

Blood-sampling time schedule

What should be the optimal blood-sampling schedule?

Scientific background. Usually blood is sampled in the fasting state and at 15, 30, 45, 60, 90 and 120 min after starting to eat the test meal in individuals without diabetes. And in subjects with diabetes, usually blood is obtained in the fasting state (0) and at 30, 60, 90, 120, 150 and 180 min after starting to eat. Taking blood samples less frequently (sampling blood every 30 min instead of every 15 min) or for less than 2 h in normal subjects significantly influences the mean and variation of the resulting AUC (Wolever, 2004b). The mean and variation of the resulting GI values also tend to increase as the frequency and duration of blood sampling decreases. In addition, reduced frequency and duration of blood sampling resulted in the development of statistically significant correlations between the mean AUC of the reference food and the mean GI of the test foods in the different subjects, a correlation which is not significant’ (Wolever, 2004b). The existence of a correlation between
AUC and GI indicates that GI is not adequately controlling for differences in glycaemic response between subjects, and hence that the GI varies in different subjects. The GI is known to be a property of the food independent of subject characteristics. Therefore, methods that result in GI values that do not control for differences in AUC between individuals cannot be considered to be valid.

Gannon & Nuttal (1987) analysed the effect on RGR of altering the length of time over which blood samples were taken from 1 to 5 h. RGR is defined as the net incremental glucose AUC (see Fig. 11) after a test meal expressed as a percentage of the net area after 50 g glucose. The results are not necessarily able to apply to the GI because RGR is based on calculating AUC using a different method than used for determining GI (incremental AUC). For legumes, the RGR markedly increased as the length of time of blood sampling increased. This is presumably because, after consuming foods such as legumes, which are slowly digested and absorbed, blood glucose concentrations tend to remain above the fasting level for several hours, in contrast to the consumption of glucose after which the blood glucose concentrations undershoot the baseline. Thus, as the time over which net AUC is calculated is extended, positive area accumulates for the legumes (i.e. net AUC becomes larger), while negative area accumulates for glucose (i.e. net AUC becomes smaller). After sucrose, milk and fructose, the RGR markedly decreased as the length of time over which blood was sampled increased, presumably because blood glucose concentrations after these test meals undershot more than after glucose. For starchy foods such as bread, oats and potato, there was only a small effect of time of blood sampling on RGR. Although an analysis to see if extending the length of blood sampling beyond 2 h affects the GI value has not been performed, it seems probable that there may be effects, which may vary for different types of foods.

Thus, adhering to a standardised schedule of blood sampling with respect to both frequency and length of time is likely to be important for obtaining valid and consistent GI values. Analysing more than one baseline sample may theoretically reduce the variability (more sampling) or increase variability (baseline may fall with time). In practice, analysing data from nineteen subjects showed this to have only a very small effect overall on the incremental AUC.

Recommendation. We recommend the following blood sampling schedule in subjects without diabetes: fasting (0) and at 15, 30, 45, 60, 90 and 120 min after starting to eat the test meal.

Washout period between two consecutive tests

What should be the duration of the washout period?

The food tested on a first session must not influence the glycaemic response of the subject to the food tested on the next session. Thus, making two test sessions on two successive days, or with a washout period of 1 d, may not be enough. But it seems that a washout period of 2 d is enough. In fact, the impact of the food tested on the metabolism of the subject 2 d later, if any, is negligible. Moreover, the subject eats between the two sessions, and these meals have more effect on the next test than the portion test eaten 2 d before.

Recommendation. There are no data to make any recommendation at this stage. Therefore, we recommend further research addressing this issue.

How should the glycaemic index be calculated?

Which type of calculation of AUC should be applied?

Scientific background. Many methods can be used to calculate AUC (Fig. 11), and since use of the different methods can result in different GI values (Table 2), this should be standardised.

The methods mostly discussed are:
(1) Total AUC;
(2) Incremental area until first return to baseline (incremental AUCcut);
(3) Incremental AUCcut, method 2;
(4) Incremental AUCcut, method 3;
(5) Incremental AUCmin, method 4;
(6) Net incremental AUC, method 5.

![Fig. 11. Different methods to determine area under the curve (AUC). (a), Total AUC, method 1; (b), incremental AUCcut, method 2; (c), incremental AUC, method 3; (d), incremental AUCmin, method 4; (e), net incremental AUC, method 5.](https://doi.org/10.1079/NRR2005100)
(3) The area over the baseline under the curve, ignoring area beneath the curve (incremental AUC);
(4) Incremental area using the lowest blood glucose as the baseline (incremental AUCmin);
(5) The net incremental AUC (apply trapezoid rule for all increments positive and negative) (net incremental AUC).

Method 1, total AUC, includes all the area under the blood glucose response curve down to a blood glucose concentration of 0. Fasting glucose is independent of the food consumed. Thus variations in total AUC due to variation in fasting glucose are not due to the test meal consumed. In addition, the portion of total AUC below fasting glucose is independent of the test meal. The test meal can only affect AUC above the fasting level (unless there is an undershoot). Since the area above the fasting glucose is only a small proportion of the total AUC, total AUC is insensitive for distinguishing between foods with different glucose-raising effects. For example, from Table 2, it can be seen that incremental AUC accounted for only 14 to 38 % of the total AUC.

For methods 2, 4 and 5 there are insufficient data to make any recommendation.

Method 3, incremental AUC, is the method recommended by the Food and Agriculture Organization (1998), and the method used for most calculations of GI up to the present time. Since use of different methods of calculating AUC results in different GI values (Tables 2 and 3), one argument for retaining this method is that adoption of a new method might render all previous GI values inaccurate. However, this is not a satisfactory argument if the method is wrong. This raises the question as to whether there is any objective advantage of method 3 (incremental AUC) over methods 2 (incremental AUCcut), 4 (incremental AUCmin) and 5 (net incremental AUC). Recently the performance of some of these and other methods of calculating incremental AUC was compared using a data-set of five foods tested in forty-seven subjects (Wolever, 2004b) (Table 2).

GI values based on the use of incremental AUCmin (method 4) and net incremental AUC (method 5) did not agree well with those derived from the recommended method, and GI values based on incremental AUCmin were significantly less than those from the recommended method (Table 2). This in itself does not mean that these methods should not be used. However, GI values based on incremental AUCmin were significantly correlated with the subjects’ glycaemic response to the reference food (oral glucose), suggesting that this is not a good method of determining GI, since the GI values obtained depend on the glucose tolerance status of the subject. In addition, the use of incremental AUCmax is not recommended because the SD of the GI values, 24-0, is 18 % greater than that for the recommended method. Incremental AUCcut produced GI values which agreed well with the recommended method, but the SD of the GI values tended to be greater than for the recommended method, and although the difference was small, it was present for each of the five foods, and the difference was statistically significant by paired t test.

The length of the time period during which post-ingestion testing is done, as well as subject characteristics, may play a significant role. In certain subjects, foods with readily available CHO, such as white bread, may induce a sharp and rapid decline in blood glucose within the 120 min period, and a hypoglycaemia in the late postprandial phase. Thus, the GI of pasta products as calculated from the 2 h AUC in healthy elderly men were not significantly different from the white reference bread despite prominent differences in course of glycaemia, with a low but sustained net increment in blood glucose following the pasta. The small glucose

<table>
<thead>
<tr>
<th>Method</th>
<th>Food</th>
<th>Gluc 1</th>
<th>Gluc 2</th>
<th>Gluc 3</th>
<th>Mean glucose</th>
<th>GI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total AUC</td>
<td>654·2</td>
<td>819·4</td>
<td>758·8</td>
<td>785·1</td>
<td>787·8</td>
</tr>
<tr>
<td>2</td>
<td>AUCcut</td>
<td>92·2</td>
<td>309·4</td>
<td>91·0</td>
<td>264·3</td>
<td>221·6</td>
</tr>
<tr>
<td>4</td>
<td>AUCmin</td>
<td>173·0</td>
<td>309·4</td>
<td>147·0</td>
<td>264·3</td>
<td>240·2</td>
</tr>
<tr>
<td>5</td>
<td>Net AUC</td>
<td>72·2</td>
<td>309·4</td>
<td>145·6</td>
<td>264·3</td>
<td>239·8</td>
</tr>
</tbody>
</table>

Gluc. glucose measurement; GI, glycaemic index.

* Method shown on Fig. 12.
AUC following white bread was related to a substantially higher insulin response. More studies are needed to evaluate if the 2 h time period commonly used needs to be modified in certain subjects (Granfeldt et al. 1991).

**Calculation of glycaemic index**

How should the ratio be calculated?

**Scientific background.** At present, a ratio we shall call $f:r$ is calculated for each subject where $f$ is an individual subject’s incremental AUC after consuming the test food and $r$ is the incremental AUC for the same subject after consuming the reference food. The individual values of $f:r$ are then averaged over all subjects to give the GI for the test food. This method will be referred to as the mean of the ratios.

Another method would be to calculate $F:R$ where $F$ is the mean incremental AUC for a group of subjects after consuming the test food and $R$ is the mean incremental AUC for the same group of subjects after consuming the reference food. This method is called the ratio of the means. Using the same set of data, these methods do not yield exactly the same result (Wolever et al. 2003). The advantage of the mean of the ratios method is that one can derive an error term more easily, which allows the GI of different foods to be compared statistically. The disadvantage is that this may require more tests (to repeat the reference food testing in each subject).

**Recommendation.** Whatever the remaining points still to be elucidated, we recommend to calculate AUC using the incremental AUC (method 3), ignoring area under the baseline. We recommend calculating GI as the mean of the individual ratios. These calculations are explained in detail below (please refer to Fig. 12 and Table 3).

**Illustration and formulas.** The general formula for calculating AUC is as follows:

Assuming that at times $t_0$, $t_1$, … $t_n$ (here equalling 0, 15 … 120 min, respectively), the blood glucose concentrations are $G_0$, $G_1$, … $G_n$, respectively:

$$\text{AUC} = \sum_{n=1}^{n} A_n,$$

where $A_n$ is the AUC for the $n$th time interval, and the $n$th time interval is the interval between times $t_{n-1}$ and $t_n$.

**Interlaboratory values and variation**

How comparable are GI values measured by different laboratories?

**Scientific background.** Mean GI values measured in different laboratories vary. Fig. 13 shows the distribution of mean GI values reported for several foods (Foster-Powell et al. 2002). By just looking at these data, it is difficult to draw conclusions about whether the variation represents methodological differences, true differences between foods of similar description, or random variation. It is also difficult to determine the magnitude of true random variation of mean GI values between different laboratories. For some foods, such as apple, lentils and brown bread, the interlaboratory sd (the sd of the mean GI calculated within each laboratory) of the GI values is <10, whereas for others, such as potato, the interlaboratory sd of the mean values is >20. An interlaboratory sd of <10 would likely be acceptable because this would mean that the 95 % CI for any GI value would be <±20. Such a value is larger than ideal, but does allow one to know with confidence that the GI values of many low-GI foods (for example, GI of lentils = 28) will almost always have a lower GI than many high-GI foods (for example, GI of bread = 71) when determined by different laboratories using standardised and correct methods. However, if the interlaboratory sd was much over 10, then the GI formed by combining the mean
GI from each laboratory into an overall mean would become so large that it would be difficult to distinguish between foods, and the GI would lose practical utility.

Recently an interlaboratory study was performed to estimate the between-laboratory variation of mean GI values (Wolever et al. 2003). It included five laboratories, which used capillary blood sampling to determine the GI values (glucose = 100) of four centrally provided foods and locally obtained white bread (Fig. 14). The pooled interlaboratory SD of the GI values for the five foods was approximately 9.

From this the probability of two laboratories obtaining GI values for the same food within limits from 0 to 20 have been calculated as 1 – 2N, where N equals the area under the normal distribution curve between −C and minus infinity, where C = D/(sD x √2), where D is the GI difference and sD is the interlaboratory SD (Fig. 15). With an interlaboratory SD of 9, the chance of the outlying point(s) in relation to the rest of the data, the outlying point(s) should be given along with the position excluded, this fact should be stated, and the results including the results of a statistical analysis (Altman, 1991). Outliers could be due to incorrect subject preparation (for example, non-fasting), analytical error, or errors in data calculation. In the absence of errors, an outlier could be due to an unrepresentative response to the standard or test food or to the subject having a true idiosyncratic response. An unrepresentative response to the standard food is suspected if the subject’s GI values for other test foods are also outliers in the same direction, or if there is large variability for the data obtained for the individual subjects may sometimes reveal values that are implausible by comparison with the rest of the data from other subjects. Outliers are particularly important because they can have considerable influence on the results of a statistical analysis (Altman, 1991). Outliers could be due to incorrect subject preparation (for example, non-fasting), analytical error, or errors in data calculation. In the absence of errors, an outlier could be due to an unrepresentative response to the standard test food or to the subject having a true idiosyncratic response. An unrepresentative response to the standard food is suspected if the subject’s GI values for other test foods are also outliers in the same direction, or if there is large variability for the repeated standard tests (Wolever et al. 1991). If outliers are excluded, this fact should be stated, and the results including the outlying point(s) should be given along with the position of the outlying point(s) in relation to the rest of the data, i.e. number of SD from the mean (including the outlier(s)). This will allow others to assess the impact of removing the outlier(s) on the results and conclusions and to form an opinion of the reasonableness of excluding the outlier(s).

**Conclusion.** When standardised methods are applied to determining the GI (glucose = 100) of identical high-CHO foods, the interlaboratory SD of GI values is approximately 9.

**How to handle outlier values?**

The advantage in dealing with GI values is that by indexing the individual glycaemic response to a standard, the variation that occurs due to several factors, for example, subject and food-related ones, the between-individual variation is minimised (Wolever, 1990a). Checking the data obtained for the individual subjects may sometimes reveal values that are implausible by comparison with the rest of the data from other subjects. Outliers are particularly important because they can have considerable influence on the results of a statistical analysis (Altman, 1991). Outliers could be due to incorrect subject preparation (for example, non-fasting), analytical error, or errors in data calculation. In the absence of errors, an outlier could be due to an unrepresentative response to the standard or test food or to the subject having a true idiosyncratic response. An unrepresentative response to the standard food is suspected if the subject’s GI values for other test foods are also outliers in the same direction, or if there is large variability for the repeated standard tests (Wolever et al. 1991). If outliers are excluded, this fact should be stated, and the results including the outlying point(s) should be given along with the position of the outlying point(s) in relation to the rest of the data, i.e. number of SD from the mean (including the outlier(s)). This will allow others to assess the impact of removing the outlier(s) on the results and conclusions and to form an opinion of the reasonableness of excluding the outlier(s).

**Recommendation.** There is no general advice or rule in statistical analysis to exclude outlying data from the complete data-set. Nevertheless, if an outlying value does not seem to be plausible or if any mistake in determining this data point is evident, this value could be regarded as suspicious and removed from the data-set (Altman, 1991). Therefore, under certain circumstances and for reasons of practicability an outlier could be removed from the data-set without reassessment of that particular GI value.

**Glycaemic index of mixed meals**

What is the impact of measurement in the context of mixed meals? Can we measure the GI of a meal (composed of different foods)?

---

**Table 4. Blood glucose concentrations (mmol/l)**

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>45</th>
<th>60</th>
<th>90</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>4.85</td>
<td>6.78</td>
<td>6.84</td>
<td>6.15</td>
<td>5.76</td>
<td>4.01</td>
<td>4.76</td>
</tr>
<tr>
<td>Gluc 1</td>
<td>4.25</td>
<td>5.21</td>
<td>7.84</td>
<td>7.68</td>
<td>7.32</td>
<td>6.91</td>
<td>6.97</td>
</tr>
<tr>
<td>Gluc 2</td>
<td>5.11</td>
<td>6.91</td>
<td>7.74</td>
<td>6.89</td>
<td>4.76</td>
<td>6.62</td>
<td>6.11</td>
</tr>
<tr>
<td>Gluc 3</td>
<td>4.34</td>
<td>5.45</td>
<td>8.08</td>
<td>7.49</td>
<td>7.56</td>
<td>6.25</td>
<td>5.31</td>
</tr>
</tbody>
</table>

Gluc, glucose measurement.

**Table 5. Illustration of calculation of incremental area under the curve with one set of blood glucose concentrations values**

<table>
<thead>
<tr>
<th>Concentrations</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>45</th>
<th>60</th>
<th>90</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increment</td>
<td>–</td>
<td>2.44</td>
<td>2.39</td>
<td>0.77</td>
<td>−0.50</td>
<td>−0.06</td>
<td>0.33</td>
</tr>
</tbody>
</table>

*See Fig. 12.
Scientific background. In general, foods are not eaten alone but in the context of a meal. This leads to two questions:

1. Does a food have the same GI when eaten alone or during a meal?
2. Can we predict the GI of a meal by the GI of the different foods consumed within the meal, or, in another way, can we predict the glucose response to a meal by the glucose responses to the different components within this meal?

First of all, the GI of a single food is dependent on the composition of this food. Actually, different elements can affect the GI of a food, such as the presence of fat (Collier & O’Dea, 1983), the presence of protein (Granfeldt et al., 1991), the presence of some anti-nutrients (Thorne et al., 1983), and acidity or the presence of acidic compounds (Liljeberg et al., 1995; Liljeberg & Bjorck, 1996, 1998). Moreover, other factors linked to the quality and the quantity of CHO present in the food can also affect its GI (see also Table 1). When a food is included into a meal, the glucose response per se to this meal will not be the same as the glucose response to the single food. Indeed, the other macronutrients (fat, protein, CHO and other nutrients)
provided by the co-ingested foods have an influence on both glucose and insulin responses.

However, it remains the quality and the quantity of CHO that determine the glucose response to the meal (Wolever & Bolognesi, 1996b). A study evaluated the glucose response to six starch-rich foods with varying GI (beans, lentils, rice, spaghetti, potato and bread) either tested alone or in mixed meals (Bornet et al. 1987). All tests contained 50 g CHO, and meals were adjusted to bring the same amount of fat, protein, water and energy, but not the same amount of fibres. The study demonstrated that, in respect to glucose-raising potential, the hierarchy between foods tested alone (GI testing) was respected once the same foods were tested in these meals. This means that the GI concept remains discriminating in the context of a meal (Bornet et al. 1987).

Concerning the glycaemic response to a meal, several authors agree that this can be predicted by the GI of each food composing this meal (Collier et al. 1986; Wolever & Jenkins, 1986; Chew et al. 1988), whereas others think such a prediction is not justified (Coulston et al. 1987), this even though some of their results seem to differ from their conclusions, as commented in Augustin et al. (2002).

Very recently Flint et al. (2004) published data on thirteen breakfast meals and a control meal of differing composition. In their study they could not show that the GI of mixed breakfast meals, as calculated by using GI table values, correlated with the GI as measured in vivo. A stronger relationship was observed for either fat or protein or energy content than for CHO alone. These results are not unexpected because, although the test meals contained 50 g available CHO, there was a large range of fat (3–42 g), protein (5–28 g) and energy (1130–2990 kJ). In the face of such large differences in fat and protein, nutrients, which are known to influence glucose and insulin responses (Nuttall et al. 1985; Ercan et al. 1994), it is expected that GI of the meal components alone would not predict glycaemic responses. Another problem is that Flint et al. (2004) determined meal GI based on food GI values obtained from the international table of GI (Foster-Powell & Brand-Miller, 1995) and it is not clear whether the correct GI values were ascribed to the foods used. For example, Finnish and German breads are ascribed a GI value of 92 (white bread = 100) based on entries 45 to 54 in the table, which are for rye breads with a range of GI values from 58 to 123; of these ten different breads, only one actually was a German bread and it had a GI value of 58 (Schauberger et al. 1978). The value used for oatmeal, entries 116 and 117, are not for oat meal but for raw oat bran (high-fibre extract of oats), etc. Accordingly, the data presented in the paper by Flint et al. (2004) are difficult to interpret and render the conclusions of these authors that in general the GI of mixed meals can not be predicted invalid. Instead, available evidence indicates that GI values of composite high-CHO meals can be predicted well based on the meal components.

Recommendations. We recommend calculating the GI of a meal from the separately measured GI of the individual foods. For most accurate prediction, and particularly in studies testing the utility of the GI in mixed meals, the GI values of the individual foods should be measured rather than estimated from GI tables.

Examples for the calculation of the glycaemic index of mixed meals. Tables 6 and 7 show the contribution to meal GI of breakfast and lunch foods.

**Glucose analysis**

What is the standard method of measurement of glucose concentration in body fluids used in GI determinations?

Glucose concentration can be determined by enzymic methods such as used normally in blood gluco-meters (used for diabetics) or enzymic determination in plasma samples. However, different methods for the measurement of glucose may differ in their precision.

Recommendation. The precision of the method chosen should be evaluated. Ideally, methods with a CV > 3 % should not be used for scientific purpose. Glucose methods should return 98 % of a glucose spike to be acceptable.

The precision and accuracy of the method chosen should be evaluated. Methods for glucose estimation with a CV > 3 % will introduce significant analytical variation into the measurement of GI. Accuracy may be estimated from the recovery of added glucose (a ‘glucose spike’) and

<table>
<thead>
<tr>
<th>Mixed meal</th>
<th>CHO (g)</th>
<th>Total CHO (%)</th>
<th>GI</th>
<th>Contribution to meal GI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breakfast biscuits (50g)</td>
<td>36</td>
<td>40</td>
<td>52</td>
<td>40 % x 52 = 21</td>
</tr>
<tr>
<td>Milk + chocolate powder (250ml)</td>
<td>30</td>
<td>33</td>
<td>38</td>
<td>33 % x 38 = 13</td>
</tr>
<tr>
<td>One banana</td>
<td>25</td>
<td>27</td>
<td>58</td>
<td>27 % x 58 = 16</td>
</tr>
<tr>
<td>Totals</td>
<td>91</td>
<td>100</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

CHO, carbohydrate; GI, glycaemic index.
Glucose or white bread are recommended as reference test foods. Test foods should be randomised in blocks of maximal 6; it is recommended to calculate GI as the mean of the AUC calculation should be based on incremental AUC, Blood sampling times should be at 0 min (baseline sample), followed by 15, 30, 45, 60, 90 and 120 min after starting to eat the test meal; AUC calculation should be based on incremental AUC, ignoring area under the baseline; It is recommended to calculate GI as the mean of the individual ratios.

Future research and development directions

Some important directions for future research concerning the measurement of GI can be given:

- Not enough data exist to determine exactly how many repeats of a reference food test should be done.
- Are mean and variation of GI data comparable in:
  - (a) Insulin sensitive v. insulin resistant non-diabetic subjects?
  - (b) Obese, overweight and lean subjects?
  - (c) Normal children v. normal adults?
  - (d) Various ethnic groups?
  - (e) Different age states – children, middle-aged, elderly?
  - (f) Type 1 or 2 diabetes?

Conclusions

Based on the currently available science described in the present review, the following conclusions and recommendations are made:

1. At least ten test subjects should be tested to obtain a sufficient statistical power;
2. Inclusion of both sexes in a study is acceptable;
3. For GI measurement the CHO portion should be based on available CHO;
4. A test dose of 50 g available CHO is recommended;
5. For foods with a low CHO content it is justified to lower the test dose to 25 g;
6. The reference food should be measured at least twice.
7. The use of healthy human volunteers is recommended;
8. Fluid ingestion, 250 ml, should take place within 5–10 min;
9. Solids and semi-solids ingestion within 10–15 min;
10. Glucose or white bread are recommended as reference foods;
11. The evening before a test each subject should consume a meal of choice and repeat that meal before each subsequent test. Unusual vigorous physical activity should be avoided;
12. Test foods should be randomised in blocks of maximal 6;
13. The total duration of the tests should not exceed 4 months;
14. In the case of testing multiple foods, a reference test should be done at the beginning and a repetition should take place after every 6–8 weeks;
15. Glucose measurement alone is appropriate;
16. For more mechanistic and/or metabolic studies, measurement of both glucose and insulin is recommended;
17. Blood sampling times should be at 0 min (baseline sample), followed by 15, 30, 45, 60, 90 and 120 min after starting to eat the test meal;
18. AUC calculation should be based on incremental AUC, ignoring area under the baseline;
19. It is recommended to calculate GI as the mean of the individual ratios.

Acknowledgements

This work was performed by an ILSI Europe invited working group, who had multiple meetings in the period 2002–2004. The work was supported by an unconditional grant from the ‘Dietary Carbohydrates Task Force’ of the European Branch of the ‘International Life Sciences Institute’ (ILSI Europe). Industry members of this task...
force are CocaCola, Cerestar-Cargill, Danisco, Danone, Kellogg, Masterfoods, Südzucker, Unilever, Royal Cosun. For further information about ILSI Europe, call +32 2 771.00.14 or email info@ilsieurope.be. The opinions expressed herein are those of the authors and do not necessarily represent the views of ILSI or ILSI Europe.

References


McKeown-Eyssen G (1994) Epidemiology of colorectal cancer revisited: are serum triglycerides and/or plasma glucose associated with risk? *Cancer Epidemiology, Biomarkers and Prevention* 3, 687–695.


Wolever TMS (2004b) Effect of blood sampling schedule and method calculating the area under the curve on validity and precision of glycaemic index values. British Journal of Nutrition 91, 295–300.