Glycemic Index: The Analytical Perspective

JON W. DeVRIES
GENERAL MILLS, INC.
Minneapolis, MN

The purpose of science, and thus, the role of the scientist, is generally believed to be to pursue the discovery, development, and application of new knowledge. A second purpose, perhaps of equal or greater importance, is to evaluate knowledge of a particular subject matter in a balanced manner. This evaluation not only consists of assessing the knowledge of the subject, but of assessing the subject in relation to other relevant, existent, scientific knowledge. In many cases, as disappointing as it may be to those immersed in a particular scientific topic, exciting research findings turn out to be inconsequential when realistically assessed. Examples from recent history include cold fusion (16), which created a substantial frenzy in the scientific community until it was shown that the observed phenomenon was the result of using inadequate measurement systems; gravity shields (16), again a phenomenon that proved to be nothing once adequate measurement procedures were incorporated; and “polywater.” (In the 1960s, a Russian scientist reported creating polywater by drawing water into quartz capillary tubes where it “polymerized” into a thick, viscous fluid that was then analyzed as “polymerized water” by infrared spectroscopy. The fluid had a boiling point of >500°C and a freezing point of –40°C, and there was immediate worldwide concern about whether this polymerization might occur in larger bodies of water or, for that matter, in human bodies. It turned out that the viscous fluid was nothing more than a silicon polymer residue from inside the capillary tubes that the water was drawn into, the residue having an infrared spectrum very similar to that conjectured to be polywater by its “discoverers” (1–3). It turned out to be a crock of polywater indeed, similar to the finding in the 1700s that “water converted to earth” when boiled in a sealed glass vial (6).

Much of the scientific research in food nutrition is supported by public funding, i.e., through academic institutions or, if not directly or indirectly by public funds, by funds generated from the sale of foods to the public. All would agree that the goals of ethical nutrition research are to increase human life span and to improve the quality of life over that lifespan through improved overall health, improved overall bodily function, and improved bodily comfort. Often the results of ethical research result in food regulations, including food labeling regulations, focused on improving overall individual and public health.

When food labeling regulations are in place, such regulations should provide for accurate and factual information with regard to the food being labeled to allow consumers to improve their diets by wise food selection and thus enjoy the benefits of the health and nutrition research that they have directly or indirectly helped to fund. Good labeling regulations minimize unfair competition and allow opportunities for the development of enhanced products. Good labeling regulations also require clear definitions of quantitated components and require adequate (accurate and reproducible) analyses to quantitate those components.

Recently, there has been an effort within segments of the scientific community to promote the inclusion of the glycemic index (GI) on food labels as an indicator of the quality of the food in the package. The GI has been variously claimed to assist in reducing body weight, stabilizing blood sugar, decreasing visceral fat, controlling appetite, improving energy level, enhancing memory, balancing mood, promoting regularity (10), and helping to prevent and control diabetes, cardiovascular disease, some cancers, and other chronic conditions (11). It is important that we, as a scientific community, conduct a scientific assessment of relevant research to determine whether the measurement tool being utilized is actually a relevant measurement and whether the research using that tool is supportive of including the GI as part of a packaged food label. The assessment should consider the evidence with regard to the measurement of the GI of foods and with regard to whether or not the relative ranking of foods by GI is related to improved states of health. Addressing these questions is essential before embarking upon a course involving tremendous cost that may result in little, if any, health improvement.

What Is GI?

The GI method was developed by Jenkins et al (12) to assist diabetes patients in controlling postprandial glucose through diet. Perhaps because the researchers were appreciative of the fact that different indi-

All methods require careful scientific assessment for validity.

A multilaboratory study demonstrated that the GI measurement purporting to differentiate one food from another is not reproducible from one eating occasion to the next and therefore does not differentiate foods.

Extensive research on the GI method results attempting to show cause and effect relationships with health states has been conducted.

Research results on GI-characterized foods have generally not shown significant relationships with health outcomes.

The inability of the GI method to differentiate between foods on eating occasions leads to the conclusion that the food itself is a minor contributor to a given GI measurement, and therefore the GI method does not measure a meaningful property of a food.
with test sample areas A–F including only those areas above the baseline.

**Predictability**—An Essential Outcome of Analyses

One of the key criticalities of any measurement procedure is a predictability or prediction function. If a certain parameter is measured a number of times and the method of measurement is a valid measurement technique, then, upon repeating the measurement one more time, one expects to obtain the same result. Say, for example, that one uses a tape measure to measure a piece of lumber five times and obtains the following in inches: 96.0, 96.1, 95.9, 96.05, and 95.95. One can logically predict that the next time it is measured, the lumber’s measurement will be in the range of 95.9–96.1 in. Contrast this to flipping a coin five times with the coin coming up heads each time and then using the coin flipping data to predict that the sixth flip will also come up heads. We know that this prediction is absurd unless the coin is asymmetrical or that, perhaps, its two sides are identical. Even worse, we might have to fight the “gambler’s fallacy” urge to believe that since the first five tosses came up heads, it’s best to bet that the next toss will be tails, even though our logic tells us there is still only a 50:50 chance.

These two extremes (the lumber measurement and the coin toss) are easier to understand than a situation in which the relationship of the measurement to the prediction of the next measurement is not so obvious. For example, suppose we tried to measure the length of our piece of lumber above using an elastic tape measure (the kind that one expects certain fishermen use in the unstretched mode for characterizing their catch) as the measurement tool and obtains the following in inches: 96.0, 10.3, 141.7, 50.8, and 181.2. What would we predict the next measurement to be? Would we predict 96.0 inches, the average measurement obtained in this case? No, we cannot predict what it will be if we only have the elastic tape measure as our measurement tool and do not know the results we obtained with the regular tape measure. Even though the average measurement in each case is 96.0, the tool used in the second series of lumber measurements is clearly not an effective measurement tool since we cannot use any of its previous results to predict the true length of the lumber. One needs to appreciate that this same principle applies to all types and systems of measurement.

Table I lists the typical variability for methods used for nutrition labeling of foods. The method variability (shown as standard deviation) listed for each method is from reports of multilaboratory studies made on the methods involved. $S_r$ represents the standard deviation within a single laboratory. If that laboratory repeats an analysis on the same sample, one would expect the second result to be within a range ($R$) of 2.8 $S_r$. For example, taking the upper end of the protein variability range, if a laboratory analyzes protein and obtains 50% protein on a given sample, one can predict that the same laboratory will obtain a value within the range of 49.0–51.0% if it analyzes the sample again and that another laboratory analyzing the same sample will obtain a value within the range of 48.5–51.5%. Although this variability may be greater than analysts would like in some cases, data acceptable for labeling purposes can be generated, and the result can be predicted before the sample is analyzed an additional time.

**Predictability—How Does the Glycemic Index Fare?**

Let’s consider the predictability of GI in contrast to the nutrients discussed above. For any given combination of tester, food, or laboratory (often referred to as a test center), we cannot accurately or precisely predict the GI of the next eating occasion of a particular sample with any degree of confidence. Recently the results of a multicenter (multilaboratory) study were reported (19). The study involved seven laboratories (research centers), all experienced with GI testing. Using glucose as the control, each center conducted its part of the study using 8–12 normally healthy testers and four common, easily prepared and homogenized foods, namely instant mashed potatoes (prepared with water in-

<table>
<thead>
<tr>
<th>Analyte</th>
<th>$S_b$</th>
<th>$S_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
<td>0.047–0.37</td>
<td>0.08–0.54</td>
</tr>
<tr>
<td>Sugars</td>
<td>0.06–0.29</td>
<td>0.13–0.85</td>
</tr>
<tr>
<td>Moisture</td>
<td>0.09–0.51</td>
<td>0.42–0.72</td>
</tr>
<tr>
<td>Dietary fiber</td>
<td>0.01–1.63</td>
<td>0.04–2.37</td>
</tr>
<tr>
<td>Fat</td>
<td>0.09–1.11</td>
<td>0.15–4.17</td>
</tr>
</tbody>
</table>

$^a$ All analytes are measured in percent.

$^b$ $S_b$ is the standard deviation of the method when performed within a single laboratory.

$^c$ $S_a$ is the standard deviation of the method when performed in multiple laboratories.
Table II. Results of interlaboratory studya,b

<table>
<thead>
<tr>
<th>Test Food</th>
<th>Mean Glycemic Index</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potato</td>
<td>84.5</td>
<td>±32.7</td>
</tr>
<tr>
<td>Rice</td>
<td>71.1</td>
<td>±38.2</td>
</tr>
<tr>
<td>Spaghetti</td>
<td>46.9</td>
<td>±26.7</td>
</tr>
<tr>
<td>Barley</td>
<td>34.7</td>
<td>±24.7</td>
</tr>
<tr>
<td>White bread</td>
<td>72.5</td>
<td>±35.8</td>
</tr>
</tbody>
</table>

a Derived from data of Wolover et al (19), Table I.

b All results are dimensionless, relative to glucose = 100.

Does the GI Indicate a Meaningful Property of a Food?

What does it mean for an individual to consume a food one time and obtain a very small area under the blood glucose response curve but to obtain a very large area under the blood glucose response curve on another eating occasion? As we saw above, a single eating occasion for a particular food can result in a GI ranging from no response to a response significantly larger than that of glucose itself. Considering that this occurs with the same sample of a food, it is hard to perceive (assuming a low GI is better for the consumer than a high GI) that said food is healthier to eat on one eating occasion than on another. Perhaps the fact that the GI measurement is not predictable explains why research attempting to relate the GI of foods and diets to states of health is equivocal, i.e., that which is not predictable can hardly be predictive.

As we review the results of various studies investigating the potential impact of low-GI foods on health states, we realize that the majority of the studies show no significant effect. For those studies concluding that a significant impact on a health state can be related to the GI of a food or diet, there are typically other factors that have not been considered that might provide the same impact. Factors such as the energy density of the foods or diets being compared or the type of dietary fiber present in the low-GI versus the high-GI food or diet must be taken into consideration. Interestingly, this was noted in the original publication by Jenkins et al (12), who said, "Some subjects found the volume of garden peas, soya beans, apples, peanuts, and some of the root vegetables difficult to complete in the allotted time." Energy density should likely be accounted for in any GI-health relationship study. It is further interesting to note that foods currently recommended as healthy diet foods by authoritative bodies on nutrition (namely fruits, vegetables, and whole grains) were distributed across the range of glycemic indices measured, i.e., carrots GI = 92, wholemeal bread GI = 72, kidney beans GI = 29, banana GI = 62. The same is true for foods recommended to be consumed in moderation, i.e., candy bar GI = 68, potato chips GI = 51, sponge cake GI = 46, and sausage GI = 28. It should be further noted that low-GI foods can be produced by increasing the fat content, the fructose content, and in some cases the organic acid content, and it is unclear whether these components of the diets have been adjusted for in the evaluation of GI versus health states.

Franz (8) summarized studies on outcome factors for diabetes patients on low-versus high-GI diets of greater than two week duration (Table III). In addition, Franz reports that three studies show a positive association between low GI (or low glycomic load [GL]) and reduced risk of diabetes and insulin resistance, while five studies show no effect. In some of the five showing no effect, however, a beneficial effect of dietary fiber was observed. The report does not indicate whether energy density, dietary fiber, and dietary fiber type were considered in the positively associated studies. Franz further points out that, of more than 30 weight-loss studies on the effect of the GI of the diet, only five reported greater weight loss for low GI; three reported greater weight loss for high...
GI; and the rest showed no significant difference.

Feskens and Du (7) reported that epidemiological studies on GI, GL, and diabetes do not support the GI hypothesis fully, and they conclude that it is unclear to what extent chronic diseases such as type 2 diabetes, obesity, and cardiovascular disease can really be prevented. Sloth and Astrup (17) reported being unable to find convincing evidence in the existing literature to suggest that a low-GI diet is superior in achieving improvement in cardiovascular health and in reducing body weight in healthy subjects when compared to official dietary advice. They also found that it is impossible for the consumer to predict the GI of individual foods—and even more so of mixed meals. They agree with McMillan-Price and Brand-Miller (13) that, in practice, it is difficult to tease out the separate effects of GI, palatability, volume, fiber, and other factors that influence satiety. Is it possible that the fact that only a minority of the published studies show a positive impact in relation to a low GI is due to the fact that the GI is extremely variable among any series of eating occasions for the same food?

When other relevant variables (known or unknown) are not controlled, it is likely that some percentage of studies purporting to establish a cause-effect relationship will turn up positive. For example, in the coin toss example, even though we do not control for the force of ejection of the coin from the thumb, the angle of ejection of the coin from the thumb, and the vertical distance between the thumb and the surface upon which the coin lands, we still come up with positive results (heads) 50% of the time. On the other hand, if we had control over the three factors in the coin toss, or if we knew the value for each of those factors, we could either control how the coin would land or predict how it would land if we corrected for those factors and considered the orientation of the coin before it left the thumb. With full control, we could flip heads 100% of the time or, with full correction based on knowledge, we could predict the final state of the coin 100% of the time.

That it is important to take into account other factors in the foods or diets of GI studies is best illustrated by examples. For example, Nilsson et al (15) studied a “high”-GI food (white wheat bread) compared to four “low”-GI foods (wheat kernels, barley kernels, spaghetti, and spaghetti with wheat bran) eaten in the evening for their effect on glucose tolerance at breakfast the next morning. Only the barley sample showed improved glucose tolerance. The GI is not a factor in the effect observed, which is more likely due to the β-glucan soluble fiber or other compositional factors such as biologically active phytochemicals of the barley compared to those of the wheat. In a study by Frost et al (9) on insulin sensitivity, comparing a low-GI diet to a high-GI diet, the researchers concluded that a low-GI diet is beneficial. However, if one looks closely at the diet, low-GI foods included pasta, oats, whole-grain products, pulse vegetables such as beans and lentils, and whole fruits. The high-GI diet consisted of avoidance of these foods. Thus, the low-GI diet reflected the recommendations of current dietary guidelines, while the high-GI diet did not necessarily do so. Differences in quantity and types of fiber were likely present to significantly extents; therefore, these factors, which were not accounted for in the conclusions, may have had a significant effect.

It could be argued that individuals are different in their response to ingested carbohydrate (even the tester’s state of physical fitness has an effect [14]), but ostensibly, the purpose of establishing an index is to adjust for that variability and compare foods one to another. Obviously, with the result ranges that have been obtained, the GI is not a tool for achieving that comparison. The attempt of Jenkins et al (7) to move individual response variability by creating an index comparing all results to glucose was ineffective. If it had been effective, the ratio of response of a particular food to glucose would be consistent for the same sample tested repeatedly by one individual on different occasions, and it would be consistent for the same sample tested by various individuals on different occasions.

Fundamental issues must exist with the GI measurement, which explains why predictable results cannot be obtained. To determine available carbohydrate, and thus establish a test portion of the food that provides 50 g of available carbohydrate, solid analytical techniques are used. Measurement of time for the curve of blood glucose vs. time presents no issues. Blood glucose concentration measurement, per se, presents no particular issues, with typical clinical blood glucose analyzers having low coefficients of variation (<3%) and at-home blood glucose analyzers having somewhat higher variability (<8%) (18). Both are considered adequate for assessing a patient’s blood glucose levels with regard to diagnosis and treatment of disease. In a study comparing a handheld glucose meter to a clinical meter, the handheld meter was found to produce values 12.7% higher than the clinical meter; however, the handheld instrument was calibrated for plasma glucose while the clinical meter was calibrated for whole blood glucose. The authors did not correct for the calibration differences when reporting the results of their GI study, although they indicate that they would expect a 10–15% difference between the two.

The coefficient of variation of the clinical meter was determined to be 1.26%, but the coefficient of variation of the handheld meter was not determined, so a direct comparison cannot be made (18). Since a single GI determination involves a minimum of 14 measurements using either analyzer, the differences in variation between the clinical and handheld meters would not result in a significant contribution to the variation found in GI measurements.

### Why the Glycemic Index Cannot Be a Predictive Measure

The fundamental measurements (available carbohydrate, time, blood glucose concentration) used to make the measurements for calculating the glycemic indices are solid measurements. Therefore, one must conclude that index determination itself is deficient and must conduct a deeper search for root causes regarding the reason or reasons that GI results are unpredictable.

In as much as the GI cannot be predicted (as evidenced by the fact that it cannot be reproducibly measured) even when the

---

**Table III. High glycemic index (GI) versus low GI diet studies in diabetes, number of studies showing various results**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Type 1 Diabetes</th>
<th>Type 2 Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low GI SB&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Low GI Not SB</td>
</tr>
<tr>
<td>HbA1c</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Fructosamine</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>FPG</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Insulin requirements or level</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>HDL</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>LDL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Data from Franz (11), table 2, page 61.
<sup>b</sup> FPG = fasting plasma glucose, HDL = high-density lipoproteins, LDL = low-density lipoproteins.
<sup>c</sup> SB = significantly better.
food sample is held constant, alternate factors must be far more relevant than the food itself, with the food itself making a relatively minor contribution to the GI on any given eating occasion. The properties of the carbohydrate (other than digestibility, as measured by the total carbohydrate minus dietary fiber) are a minor component of the GI and are probably not significant in individual GI determinations. Trying to correlate the GI with various physiological effects is futile, inasmuch as there is no factor to hold constant in order to test the variables against.

References
11. Glycemic Index Laboratories Inc. What is the glycaemic index? Toronto, ON, Canada, 2005.

Jonathan W. DeVries received his bachelors degree in chemistry from Augsburg College and his Ph.D. in organic chemistry from the University of Minnesota. He has been active in industrial chemistry and analytical research for nearly 40 years in the areas of nutrition and food safety. For 26 years, DeVries has been working for validation and international standardization of analytical methods through such organizations as AACC International, AOAC International, ASTM, and AOCS. Jon recently completed a term as chairman of the AACC International vitamins committee and has also been active in the oat and dietary fiber committees. Jon served as a member of the Foods Chemicals Codex Committee of the National Academies of Science and chair of the Carbohydrates Technical Committee of ILSI (International Life Sciences Institute) North America. He has authored numerous papers on foods analyses, including approved methods of analysis for dietary fiber, vitamin A, vitamin E, and several water-soluble vitamins. He is currently a senior principal scientist at General Mills Inc., where he serves as the senior technical manager for the Medallion Laboratories, a company division that provides analytical services to the food and other industries. DeVries can be reached at Jon.DeVries@genmills.com.