
Cees van Dijk¹* and Gerben H van Dijk²

¹Food and Bio based Research, Wageningen University and Research centre, The Netherlands
²Department of Cardiology, Rijnstate Hospital, Arnhem, The Netherlands.

*Corresponding Author: Cees van Dijk, Food and Bio based Research, Wageningen University and Research centre, 6700 AA Wageningen, The Netherlands. E-mail: vandijk.cees@gmail.com.

Received: September 18, 2015; Published: November 03, 2015

Abstract

Background: The glycemic index (GI) of common, processed food products can only be determined experimentally. Adequate numerical models capable to predict the GI value of a product based on the macronutrient composition of the raw material are not available.

Objective: To develop predictive Partial Least Squares Regression (PLSR) models based on the relation between the macronutrient composition of a wide range of common, processed food products and their glycemic response, taking the processing method as selection criterion into account.

Method: Published GI data (n = 601) of common, processed food products were used for PLSR based model development. Prerequisites to develop reliable, predictive models are: i) the use of a proper reference product, ii) an adequate numerical description of the macronutrient composition of the processed products, iii) the use of equi-energy (1MJ) servings rather than products containing 50 g carbohydrate.

Results: After transformation of the published data into 1 MJ containing servings, they were clustered into four distinguishable product groups, each group characterized by its specific processing method. Processed legumes were recognized as a separate, fifth group. The glycemic response differs between diabetic and non-diabetic test person (P < 0.05), resulting in ten different predictive PLSR models. In all cases sugar, starch and total dietary (TDF) were as the only regression factors. For the data of these ten models the correlation coefficient between the 1 MJ based and predicted glycemic values was r = 0.88. In addition it was also shown that the higher the percentage energy contained by “protein + fat”, the lower the glycemic response.

Conclusions: Relevant predictive PLSR models of common, processed food products can only than be developed if, i) equi-energetic servings rather than 50 g carbohydrate servings are used, ii) the processed food products are clustered into four product groups, each group based on major characteristics of its processing method.

Keywords: Predictive models; Equi-energetic servings; Glycemic response; Common processed food products; Macronutrient composition; (non-) diabetic

Abbreviations: PLSR: Partial Least Squares Regression; RMSEC/P: Root Mean Square Error of Calibration/Prediction; TDF: Total Dietary Fiber; IAUC: Average Integrated Area’s Under the Curve; GlyS: Glycemic Score (1 MJ servings); GI: Glycemic Index (50g carbohydrate servings)

Introduction

Since the introduction of the glycemic index (GI) concept [1] numerous publications have addressed this issue. An overview of the results concerning the GI value of individual food products [2], or up to date information about the GI values of individual, common food products can be found on internet [3]. Food products with similar GI values can be clustered [4] ranging from low (GI < 55), medium (55 < GI < 70) to high GI products (GI > 70). Modification of a food product, either by adapting its macronutrient composition, or by external addition of another food product (e.g. bread versus bread plus cheese), currently requires the assessment of the GI value of the new product respectively composite, according to the standard method [5]. The current state of affairs is that, given the absence of adequate predictive models, neither the prediction of the GI values of common, individual food products, nor products with an adjusted composition, nor meals is possible. The primary goal of this study was to develop predictive Partial Least Squares Regression (PLSR) models based on the quantitative relation between relevant macronutrients as predictor of the glycemic response of processed, common consumers' food products. PLSR is particularly useful to predict a dependent variable, like the glycemic response, from a (very) large set of independent variables, e.g. the macronutrients of a broad spectrum of food products. These latter variables can serve as predictors for the dependent variable, e.g. the glycemic response. As shown previously [6] using PLSR the perceived satiety of a wide range of common fresh and processed food products can adequately be related to their macronutrient composition, with emphasis on protein, and sugars. The resulting models are not only able to describe the relation between macronutrient composition and perceived satiety, but also characterized by their predictive capability. A prerequisite to develop these models was the use of 1 MJ servings. For the determination of the glycemic response of a food product 50g carbohydrate (starch plus sugars) present in a test product is used for the determination of the glycemic response, the GI (glycemic index) value of the test product [5]. A trigger to initiate the development of predictive models for the glycemic response of common, consumers food products was the observation that 1 MJ servings, in case the assessment of the perceived satiety of food products, resulted in applicable predictive models. Here it was realized that the way of processing, besides the macronutrient composition of the products, had also be taken into consideration.

This was based on the observation that, starting with the same raw material, a different processing method leads to a significant different GI value between the products obtained. Obvious examples are cooked versus instant mashed, baked, or roasted potatoes [1,7,8], cooked versus parboiled rice [9], and bread versus pasta respectively [10]. These examples suggest that the processing method is a relevant factor contributing to the glycemic response of a product. Given the observation that, using equi-energetic servings [6], the satiety response can be modeled, published data of the glycemic response of common, processed food products, together with their macronutrient composition, were transformed from 50g carbohydrate into 1MJ containing servings. As indicated above, the way of processing affects the glycemic response of food products for this reason food products were clustered into product groups based on the processing method applied. The determination of the minimal amount of relevant product groups was performed iteratively. For each product group the data, based on 1MJ servings, were used to construct product-group specific, predictive PLSR models. These PLSR models, using the processing method as selection criterion, are relating relevant macronutrients of the food products analyzed to their glycemic responses. The models developed, were based on the independent, integrated information originating from 58 studies, from various laboratories over a period of more than three decades. These models can be useful to enable the production of food products with desired glycemic properties.

Material and Methods

Data sources and chemical information

Information concerning the GI values of individual, processed common food products was obtained from peer reviewed published manuscripts. Experimental products were beyond the scope of this study and therefore excluded. Only those manuscripts were selected that, for the determination of the GI value of a product, contained a proper reference test food, either 50g glucose, or 50g carbohydrates present in white bread. In addition, an adequate numerical description of the macronutrient composition of each individual product was also required. This compositional information concerned the amount of protein, fat, sugar, starch and total dietary (TDF). Sugar is defined here as the sum of all common mono-, and disaccharides, carbohydrate as the sum of starch and sugar. The data concerning the


Transformation into equivalent energy servings

PLSR model formation was performed on groups of food products that contained 1 MJ energy rather than 50g carbohydrate. For each individual product, its total energy content was calculated based on its protein (17 kJ/g), carbohydrate (17 kJ/g), fat (37 kJ/g) and organic acid (13 kJ/g) content [11]. The energy content of dietary is ignored in view of its different metabolic routing compared to carbohydrate, protein and fat. Resistant starch [12,13] is considered as. For each individual product its total energy content, the individual macronutrients and its GI value were multiplied with a factor required to adjust the total energy content to 1 MJ. This latter amount of energy per sample was chosen since a serving of white bread containing 50g carbohydrates, frequently used as reference test food, comes close 1 MJ. In addition, for reasons of comparison, the modeling of the perceived satiety of food products [6] was also performed using 1 MJ servings [14]. All further calculations were performed using food products with their energy content adjusted to 1MJ. To distinguish from the glycemic index “GI” that is based in 50g carbohydrate servings, the glycemic response of 1MJ servings is referred to as glycemic score (GlyS).

The consequences of using 1 MJ rather than 50g carbohydrate servings

To determination the GI value of a product (ref 10) either glucose (50g) or white bread (50g carbohydrate) is used as reference. The energy content of 50g glucose is 0.85MJ, the energy content of white bread, based on 50g carbohydrate is, in this study, set at 1050 kJ. This value was obtained as the average energy content of the white bread samples studied (n=41; 1050 ± 50 kJ). Based on literature information [2,15] if the GI value of white bread is set at 100, the GI value estimated for glucose is 141. Inversely, if the GI of glucose is set at 100, the estimated GI value of white bread is 100/1.41 = 70.9. These values are used throughout this study. Before starting the data analysis the energy content of all products were set at 1 MJ.

In Scheme 1 the approach is given to come from GI values, based on 50g carbohydrate, to GlyS values, based on 1 MJ servings, for glucose, bread and product “P”.

<table>
<thead>
<tr>
<th>Product</th>
<th>Glucose</th>
<th>White Bread</th>
<th>Product P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbohydrate (g)</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Energy (MJ)</td>
<td>0.85</td>
<td>1.05</td>
<td>A</td>
</tr>
<tr>
<td>GI values per 50g carbohydrate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI (Glu ≡ 100)</td>
<td>≡100</td>
<td>70.9</td>
<td>B/1.41</td>
</tr>
<tr>
<td>GI (WhiBr ≡ 100)</td>
<td>141 (=100/0.709)</td>
<td>≡100</td>
<td>B</td>
</tr>
<tr>
<td>Product description per 1 MJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy (1MJ)</td>
<td>1 MJ</td>
<td>1 MJ</td>
<td>1 MJ</td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>58.82 (=50/0.85)</td>
<td>47.17 (=50/1.05*)</td>
<td>50/A</td>
</tr>
<tr>
<td>GlyS values per 1MJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GlyS per 1MJ (Glu as reference)</td>
<td>117.6 (=100*58.82/50)</td>
<td>67.5 (=70.9/1.05)</td>
<td>B/1.41A</td>
</tr>
<tr>
<td>GlyS (WhiBr ≡ 100)</td>
<td>176 (= 117.6 * 105/70.9)</td>
<td>≡100</td>
<td>1,05B/A (=105B/(1.41*70.9A))</td>
</tr>
</tbody>
</table>

Scheme 1: Transformation scheme to come from a GI to GlyS with either glucose or white bread as reference. WhiBr, white bread; *1.05 is the amount of energy contained in white bread, containing 50g carbohydrate


In case the glycemic response is expressed as GI value the reference material contains 50g glucose as carbohydrate, with a GI value set at 100. As discussed above, white bread contains 1.05 MJ (on average), 50g of carbohydrate and a GI value of 70.9. Product “P” is characterized by an energy content of “A” MJ and a GI value of “B/(1.41A)”. Defining 1 MJ of energy, rather than 50g carbohydrate as benchmark, the amount of glucose increases to (1000/17) 58.8g and consequently, its GlyS value to 117.6. For 1 MJ white bread the amount of carbohydrate decreases to 47.2g and its GlyS value to 67.5. For product P the amount of carbohydrate becomes (50/A)g and its GI value B/(1.41A). Within this study, white bread, having an energy content of 1 MJ is used as reference with a GlyS value set at 100. To obtain the correct GlyS value for all products, the calculated GI values at 1 MJ energy content have to be multiplied with a factor of 1.05 (MJ), the (average) energy content of a piece of white bread containing 50g carbohydrate, and divided by the energy contain of any product also containing 50g of starch plus sugars.

Model development and statistical analysis

Partial Least Squares Regression model development and validation testing were used as described previously [6,16,17]. Outliers were excluded both from model development and validation testing. Assessment of the significance of difference between PLSR models was performed as follows. Model 1 is explained by:

\[
\text{Model 1} = C_1 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n \tag{Eqn. 1a}
\]

and Model 2 by:

\[
\text{Model 2} = C_2 + \beta_{n+1} X_{n+1} + \beta_{n+2} X_{n+2} + \beta_{n+3} X_{n+3} + \cdots + \beta_{2n} X_{2n} \tag{Eqn. 1b}
\]

SPSS20 was used to perform the F-test to assess the consequences of addition of the regression parameters of Model 2 to Model 1. The null-hypothesis here is that \( \beta_{n+1} = \beta_{n+2} = \beta_{n+3} = \cdots = \beta_{2n} = 0 \). The alternative hypothesis is that at least one of the additional variables has an effect on Model 1, with P < 0.05.

Outliers

Outliers are defined as described previously [6] using Dixon’s Q-Test at P < 0.05 [17]. In addition to the TN7 discordance test, dependent on the regression residuals, the TN7 (TN7 = \( (x_n - x_{(n-1)})/(x_n - x_1) \)) to TN13 (TN13 = \( (x_n - x_{(n-2)})/(x_n - x_3) \)) were used [18]. In some cases the sample size was larger than 100 and no critical values were available [19]. In case the sample size \( N > 100 \) the entire sample set was analyzed and the five most outlying samples were selected. The remainder of the samples was randomly selected to form test groups of 95 samples. The five most outlying samples were added to the test groups and these new sample groups were further analyzed for outliers.

Results and Discussion

Different processing methods leads to five distinguishable product groups each characterized by its own regression model

In first instance the identification of distinguishable product groups was performed without making a distinction between non-diabetic and diabetic test persons. Five distinguishable groups (Table 1) were identified, each characterized by its own PLSR model. These five regression models are mutually, statistically significantly different from each other (P < 0.05). Of these five distinguishable groups, four are based on different ways of processing (Table 1). For product group 1 either baking or frying is the major processing step. Baking and frying refer to the integral process at elevated temperatures, either in air, or fat/oil respectively, where starch gelatinization, water evaporation, Maillard reactions and crust formation occur simultaneously. In general, baking is applied to all kinds of cereals (wheat, rye, barley, rice, spelt, com, millets, oats, buck wheat) to produce bread, cake, crackers, cookies etc. Frying is e.g. applied to r potatoes, either to produce fried potatoes, French fries, or potato chips. Product group 2 consists of breakfast cereals and instant mashed potatoes. A major processing step of industrial produced breakfast cereals is the gelatinization of the starch of the raw material by (extrusion-) cooking, followed by product drying to produce “ready to eat” breakfast cereals, that are rehydrated prior to consumption [20]. Similar to breakfast cereals this cooking and drying process, followed by rehydration before consumption is also used to produce instant

mashed potatoes. Product group 3A mainly consists of (steam) cooked products like potatoes, rice and corn. Included in this group is (German) pumpernickel bread. This product is prepared in the presence of 100% humidity, comparable with (steam) cooking. Product group 3B just contains cooked/sterilized pulses. The (physico-) chemical properties of processed pulses apparently, with emphasis on the gelatinized starch, differ in such a way from the products of group 3A that they form a different group. Product group 4 consists of products, obtained by different ways of processing. These products are pasta, bulgur and parboiled grain products. Pasta is made from dough, followed by shaping the dough at about 50°C and drying [21]. Starch gelatinization takes place during cooking the pasta prior to consumption. Bulgur [22] is basically produced using (steam) cooking, drying, dehulling, grinding and sorting. Parboiled products are produced by a process of soaking, pressure steaming and drying prior to milling. Despite obvious differences in their processing history and product properties, pasta, bulgur and parboiled products can be clustered into one group. The overall chemical composition of the five product groups with regard to the carbohydrate and composition is given in Table 2. Included in this table is the average percentage of energy of contained in the form or protein and fat.

<table>
<thead>
<tr>
<th>Product Group</th>
<th>Processing method</th>
<th>Products</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Baking or frying</td>
<td>Bread products, cookies, cakes, muffins, crackers etc., fried potatoes, French fries, chips</td>
<td>1, 7, 8, 9, 23-64.</td>
</tr>
<tr>
<td>Group 2</td>
<td>(Extrusion-) cooking, followed by drying and rehydration</td>
<td>Breakfast cereals, porridge, instant mashed potatoes</td>
<td>1, 7, 8, 9, 25, 26, 28, 30, 33, 34, 36, 40, 41, 46, 51, 55, 59, 65-68.</td>
</tr>
<tr>
<td>Group 3A</td>
<td>(Steam) cooking</td>
<td>Cooked potatoes, -rice, and -corn, pumpernickel bread.</td>
<td>1, 7, 8, 9, 23-25, 27, 30, 32, 33, 36, 38, 41, 42, 46, 50, 52, 54, 60, 66, 69.</td>
</tr>
<tr>
<td>Group 3B</td>
<td>Cooking, or sterilisation</td>
<td>Processed legumes</td>
<td>1, 7, 9, 26, 28, 30, 35, 36, 39, 41, 52, 56, 66, 73-75.</td>
</tr>
<tr>
<td>Group 4</td>
<td>Heating/partial cooking, followed by drying and rehydration</td>
<td>Pasta, bulgur, parboiled grain products</td>
<td>1, 7, 9, 23, 24, 26, 28, 29, 30, 32, 33, 36-38, 41, 46, 48, 49, 50, 54, 58, 59, 60, 70, 76</td>
</tr>
</tbody>
</table>

**Table 1**: Characterization of the processing methods and their products, enabling the prediction of the Glycemic Score of individual processed, common food products.

<table>
<thead>
<tr>
<th>Product Group</th>
<th>Sugar (g)</th>
<th>Starch (g)</th>
<th>TDF (g)</th>
<th>Energy % (Protein + Fat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>6.6 ± 6.5</td>
<td>33.9 ± 10.9</td>
<td>4.2 ± 4.3</td>
<td>31.3 ± 13.9</td>
</tr>
<tr>
<td>Group 2</td>
<td>10.7 ± 7.1</td>
<td>31.3 ± 10.4</td>
<td>5.5 ± 5.5</td>
<td>28.6 ± 10.9</td>
</tr>
<tr>
<td>Group 3A</td>
<td>2.7 ± 4.6</td>
<td>44.5 ± 9.5</td>
<td>3.9 ± 3.8</td>
<td>19.6 ± 14.1</td>
</tr>
<tr>
<td>Group 3B</td>
<td>4.1 ± 5.2</td>
<td>30.1 ± 10.2</td>
<td>11.4 ± 6.0</td>
<td>41.8 ± 16.7</td>
</tr>
<tr>
<td>Group 4</td>
<td>1.4 ± 1.3</td>
<td>43.4 ± 7.0</td>
<td>4.9 ± 4.4</td>
<td>23.8 ± 11.8</td>
</tr>
</tbody>
</table>

**Table 2**: Overall chemical composition of the five distinguishable product groups according to Table 1.

**Difference in glycemic response between diabetic and non-diabetic test persons**

Having defined these five distinguishable groups, the questions were addressed if, per product group, the glycemic response significantly differs between diabetic and non-diabetic test persons. For all groups together, it was observed that the glycemic response of diabetic test persons significantly differs from that of non-diabetic test persons (P < 0.05). For this reason, per product group a PLSR model had to be developed both for non-diabetic and for diabetic test persons. No difference in glycemic response was observed between Type 1 (n = 58) and Type 2 (n = 185) diabetic test persons. In addition, products just containing sugars and no starch, the glycemic response also differ between diabetic and non-diabetic test persons [76].

PLSR models for diabetic and non-diabetic test persons relating processing and macronutrients to the glycemic score

For the ten PLSR models (see Table 3) developed, the same three regression factors, starch, sugar and TDF are required, only their values differ per product group. In all cases, for all product groups starch is the most dominant factor. In all cases the value of TDF as regression factor is negative, with exception of legumes (group 3 B). A negative value for TDF factor implies that the higher the amount of the lower the glycemic response. The reason for a positive value of TDF as regression factor for legumes can only be explained if it is assumed that, during processing, the heat induced breakdown products of the cell walls of legumes [78] are, in contrast to the heat induced cell wall breakdown products of all other processed food products, capable to enhance the glycemic response of its starch.

<table>
<thead>
<tr>
<th>Product groups and test persons</th>
<th>Number of samples</th>
<th>Number of outliers</th>
<th>r</th>
<th>RMSEC</th>
<th>Regression coefficients</th>
<th>Estimated GlyS value of starch</th>
<th>Estimate ΔGS/gram TDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1; Diabetic</td>
<td>91</td>
<td>2</td>
<td>0.91</td>
<td>8.5</td>
<td>0.56 1.13 -0.07</td>
<td>123 ± 18</td>
<td>-0.63</td>
</tr>
<tr>
<td>Group 1; Non-diabetic</td>
<td>146</td>
<td>2</td>
<td>0.84</td>
<td>13.5</td>
<td>0.52 1.14 -0.15</td>
<td>130 ± 16</td>
<td>-0.76</td>
</tr>
<tr>
<td>Group 2; Diabetic</td>
<td>45</td>
<td>1</td>
<td>0.85</td>
<td>10.6</td>
<td>0.99 1.41 -0.26</td>
<td>127 ± 6</td>
<td>-0.85</td>
</tr>
<tr>
<td>Group 2; Non-diabetic</td>
<td>69</td>
<td>1</td>
<td>0.87</td>
<td>16.1</td>
<td>0.79 1.25 -0.11</td>
<td>152 ± 6</td>
<td>-0.71</td>
</tr>
<tr>
<td>Group 3A; Diabetic</td>
<td>43</td>
<td>2</td>
<td>0.88</td>
<td>10.4</td>
<td>0.35 0.92 -0.12</td>
<td>109 ± 13</td>
<td>-0.55</td>
</tr>
<tr>
<td>Group 3A; Non-diabetic</td>
<td>70</td>
<td>2</td>
<td>0.72</td>
<td>16.3</td>
<td>0.39 0.85 -0.07</td>
<td>114 ± 5</td>
<td>-0.69</td>
</tr>
<tr>
<td>Group 3B; Diabetic</td>
<td>24</td>
<td>2</td>
<td>0.90</td>
<td>6.6</td>
<td>0.01 0.38 0.82</td>
<td>34 ± 18</td>
<td>2.05</td>
</tr>
<tr>
<td>Group 3B; Non-diabetic</td>
<td>40</td>
<td>2</td>
<td>0.80</td>
<td>11.0</td>
<td>0.61 0.47 0.20</td>
<td>42 ± 8</td>
<td>0.62</td>
</tr>
<tr>
<td>Group 4; Diabetic</td>
<td>44</td>
<td>2</td>
<td>0.79</td>
<td>9.6</td>
<td>0.03 0.79 -0.29</td>
<td>82 ± 12</td>
<td>-1.11</td>
</tr>
<tr>
<td>Group 4; Non-diabetic</td>
<td>30</td>
<td>3</td>
<td>0.71</td>
<td>10.1</td>
<td>0.07 0.60 -0.66</td>
<td>96 ± 16</td>
<td>-1.96</td>
</tr>
<tr>
<td>Diabetic, total</td>
<td>247</td>
<td>9</td>
<td>0.92</td>
<td>9.6</td>
<td>NR NR NR NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Non-diabetic, total</td>
<td>354</td>
<td>10</td>
<td>0.87</td>
<td>14.6</td>
<td>NR NR NR NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Total</td>
<td>601</td>
<td>19</td>
<td>0.88</td>
<td>13.3</td>
<td>NR NR NR NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

Table 3: Results of the PLSR model formation for diabetic and non-diabetic test persons, of the five identified product groups, the numerical values of the regression coefficients of the PLSR models, the numerical values of the regression correlation of the RMSEC, and of the estimate per product group of the Glycemic Score of 1 MJ starch in the absence of sugar and TDF, and the change in the GS value of 1MJ starch per gram TDF. *Exclusive white bread with GlyS ≡ 100 (n = 22), †Exclusive white bread with GlyS ≡ 100 (n = 26), NR; not relevant.

Noticeable is that for all five groups the values of the Root Mean Square Error of Calibration (RMSEC), the modeling error, of the models for the diabetic test persons are, per product group, lower than the corresponding RMSEC values for the models of the non-diabetic test persons. The drug supported glycemic response in case of diabetic test persons obviously resulted in a decrease in the glycemic response variability. Included in Table 2 is a compilation of the results for the samples tested by either diabetic, non-diabetic and the total amount of samples tested. From all samples 3% was designated as outlier.

An analysis of the regression results

A linear regression between all measured and predicted GlyS values resulted into the following equations for both groups of test persons:

Non-diabetic: \[ \text{GlyS-predicted} = 1.01 \times \text{GlyS-reference} - 0.53 \quad (r=0.88; \text{RMSEC}=14.2) \] (Eqn. 2a)

Diabetic: \[ \text{GlyS-predicted} = 0.99 \times \text{GlyS-reference} + 0.33 \quad (r=0.92; \text{RMSEC}=9.2) \] (Eqn. 2b)

For control purposes for the 10 separate models, the models for the non-diabetic test persons were used to predict the data belonging to the diabetic test persons and vice versa. The following regression equations were obtained:

Non-diabetic models predict diabetic data:

\[ \text{GlyS-predicted} = 0.88 \times \text{GlyS-reference} + 8.6 \quad (r=0.90; \text{RMSEC}=11.4) \] (Eqn. 3a)

Diabetic models predict non-diabetic data:

\[ \text{GlyS-predicted} = 0.71 \times \text{GlyS-reference} + 20.4 \quad (r=0.83; \text{RMSEC}=15.9) \] (Eqn. 3b)

Based on the difference between the slope and intercept values between Equations 2a, b and Equations 3a, b respectively, it is obvious that the non-diabetic models are not suited to predict the diabetic data and the diabetic models are not suited to predict the non-diabetic data. In case of an ideal situation the values for slope and intercept should be 1 and 0 respectively. This situation is almost reached in Equations 2a, b. This results strongly also supports the conclusion that the glycemic response of diabetic persons differs from that of non-diabetic persons.

The regression models developed enable, per group and type of test person, the assessment of the GlyS value of 1 MJ starch (= 58.8g). Group 4 and, especially group 3B, clearly exhibit substantial lower GlyS values for their starches than the first three groups. In the last column of Table 2 the effect of TDF on the change in the glycemic response per g TDF is given per group. This effect is estimated to be linear with the amount of TDF.

From the samples belonging to each product-, and test group the average amount of starch and TDF per product and test group was estimated. From the calculated GlyS value of 1 MJ starch in the absence of TDF (Table 3), the GlyS value of the average amount of starch, present within each group, was obtained (Table 4). In addition, from the estimated average amount of TDF and the estimate of the change in GlyS per gram TDF (ΔGlyS/gram TDF; see Table 3) the absolute and percentage change in GlyS caused by the average amount of TDF was calculated (Table 3). Based on these estimates it is obvious that only for legumes (group 3B), and to a lesser extent, group 4, has a substantial contribution to the change in the glycemic response of its starch.

From the information in Tables 3 and 4, the estimated GlyS values of groups 1, 2 and 3A represent a different GlyS range compared with groups 3B and 4. For this reason the plot of the measured against the predicted GlyS values, including the 95% confidential interval, for the product groups 1, 2 and 3A is presented in Fig. 1A, for groups 3B and 4 in Figure 1B.

Taking the average measured GlyS value ± two times the SD, the majority of the samples of Figure 1A is located between 35 < GlyS < 120 and of Figure 1B between 15 < GlyS < 75. From these figures it is obvious that the GlyS range of the samples of groups 1, 2 and 3A, presented in Figure 1A, is substantially larger than for the samples of groups 3B and 4. Compared to groups 3B and 4 the samples of product groups 1, 2 and 3A also cover, despite the relatively high GlyS values of their starches, the low GlyS range (see Figure 1A). This can only be the case if, within a given product, the amount of glycemic macronutrients, sugar and starch, is low and, consequently, the amount of protein and/or fat is high.

<table>
<thead>
<tr>
<th>Product Group</th>
<th>Test panel</th>
<th>Average amount of starch (g/MJ)</th>
<th>Average amount of TDF (g/MJ)</th>
<th>GlyS of average amount of starch</th>
<th>ΔGlyS/gram TDF</th>
<th>ΔGlyS for the average amount of TDF</th>
<th>% change of GlyS by TDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Diabetic</td>
<td>37</td>
<td>4,3</td>
<td>77</td>
<td>-0,63</td>
<td>-2,7</td>
<td>-3,5</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>32</td>
<td>4,1</td>
<td>71</td>
<td>-0,76</td>
<td>-3,1</td>
<td>-4,4</td>
</tr>
<tr>
<td>Group 2</td>
<td>Diabetic</td>
<td>32</td>
<td>6,0</td>
<td>68</td>
<td>-0,85</td>
<td>-5,1</td>
<td>-7,5</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>31</td>
<td>5,1</td>
<td>80</td>
<td>-0,71</td>
<td>-3,6</td>
<td>-4,5</td>
</tr>
<tr>
<td>Group 3A</td>
<td>Diabetic</td>
<td>45</td>
<td>5,7</td>
<td>84</td>
<td>-0,55</td>
<td>-3,1</td>
<td>-3,7</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>44</td>
<td>2,8</td>
<td>86</td>
<td>-0,69</td>
<td>-2,0</td>
<td>-2,3</td>
</tr>
<tr>
<td>Group 3B</td>
<td>Diabetic</td>
<td>36</td>
<td>10,6</td>
<td>21</td>
<td>2,05</td>
<td>21,7</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>26</td>
<td>11,7</td>
<td>19</td>
<td>0,62</td>
<td>7,3</td>
<td>39</td>
</tr>
<tr>
<td>Group 4</td>
<td>Diabetic</td>
<td>43</td>
<td>5,0</td>
<td>60</td>
<td>-1,11</td>
<td>-5,6</td>
<td>-9,3</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>44</td>
<td>4,8</td>
<td>72</td>
<td>-1,96</td>
<td>-9,4</td>
<td>-13,1</td>
</tr>
</tbody>
</table>

Table 4: Estimate of the GlyS value of the average amount of starch per product group and the change in this GlyS value by the average amount of TDF.

Figure 1A: Measured versus predicted GlyS values of baked or fried products, of cooked products and of breakfast cereals. ▲: Baked or fried products; △: Cooked products; □: Breakfast cereals; *: Outlier; ·····: 95% confidence interval.

The consequences of the presence of non-glycemic macronutrients

Given the prerequisite of equi-energetic servings the lower the amount of starch and sugar in a product, the higher the amount of protein and/or fat and, consequently, the amount of energy contained by these macronutrients. To assess the consequences of the increase of the amount of energy within a product in the form of protein and/or fat, for each individual food product, per group, the sum of the percentage energy of protein and of fat, contained in each individual sample (n = 601) was calculated. Per group these sample specific percentages were related with respectively the measured and the predicted GlyS values of these individual products. The results of linear regression between the percentage energy in the form of protein and fat versus both, the measured and the predicted GlyS values are given in Table 5.

For groups 1, 2 and 3A the correlation between the GlyS value and the percentage energy contained in protein and fat is fair/good for the measured and excellent for the predicted GlyS values. For group 3B and 4 this correlation is (very) poor for the measured GlyS values and poor/fair for the predicted GlyS values. From Table 4 it can furthermore be concluded that per group the values for slope and intercept are similar for the measured and predicted GlyS values. With exception of group 3B the values for the intercept are similar to the values of the estimated GlyS values of 1MJ starch given in Table 2. These latter values only refer to the product group specific GlyS value of starch, in the absence of TDF and sugar. The intercept values of Table 4 include, in contrast to Table 2, besides starch also the effects of TDF and sugar. Obviously in case of Table 4, with exception of group 3B, the GlyS enhancing effect of sugars is compensated by the GlyS decreasing effect of TDF. For group 3B the higher value of the intercept (Table 4) compared with the GlyS value starch alone (Table 2) is probably caused by the substantial change of the glycemic score of processed legumes by TDF (Table 3).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>Diabetic</td>
<td>-1.23</td>
<td>117</td>
<td>-0.89</td>
<td>-1.39</td>
<td>120</td>
<td>-0.98</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>-1.53</td>
<td>120</td>
<td>-0.82</td>
<td>-1.53</td>
<td>120</td>
<td>-0.94</td>
</tr>
<tr>
<td>Group 2</td>
<td>Diabetic</td>
<td>-1.94</td>
<td>132</td>
<td>-0.82</td>
<td>-2.30</td>
<td>144</td>
<td>-0.98</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>-2.42</td>
<td>148</td>
<td>-0.86</td>
<td>-2.26</td>
<td>143</td>
<td>-0.99</td>
</tr>
<tr>
<td>Group 3A</td>
<td>Diabetic</td>
<td>-1.30</td>
<td>105</td>
<td>-0.83</td>
<td>-1.26</td>
<td>108</td>
<td>-0.99</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>-1.24</td>
<td>112</td>
<td>-0.73</td>
<td>-1.27</td>
<td>110</td>
<td>-1.00</td>
</tr>
<tr>
<td>Group 3B</td>
<td>Diabetic</td>
<td>-0.51</td>
<td>56</td>
<td>-0.39</td>
<td>-0.53</td>
<td>55</td>
<td>-0.63</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>-0.65</td>
<td>62</td>
<td>-0.66</td>
<td>-0.66</td>
<td>64</td>
<td>-0.81</td>
</tr>
<tr>
<td>Group 4</td>
<td>Diabetic</td>
<td>-0.81</td>
<td>71</td>
<td>-0.71</td>
<td>-0.78</td>
<td>69</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td>Non-diabetic</td>
<td>-0.47</td>
<td>69</td>
<td>-0.24</td>
<td>-0.61</td>
<td>69</td>
<td>-0.58</td>
</tr>
</tbody>
</table>

Table 5: Linear regression group between the percentage energy contained by “protein + fat” and both, the measured and the predicted GlyS values per product contained in a product group and the slope and intercept of the regression line.

White bread as reference

Besides glucose, white bread is most frequently used as reference in the determination of the glycemic index of food products. As indicated previously [6] it is anticipated that starch retrogradation significantly increases the perceived satiety of white bread. Using the same data to assess the effect of starch retrogradation on the perceived satiety [6] in combination with the regression model for non-diabetic test persons of group 1, the GlyS value of white bread decreases with 12%, 15% and 16% after 1, 3 and 5 days of storage respectively. As indicated previously [76] these data suggest that the use of white bread as reference requires a re-evaluation.

Conclusions

The glycemic behavior of common, processed food products was analyzed to evaluate if predictive models could be developed, relating the macronutrient composition of the processed products with their glycemic response, taking the way of processing into account. The main reason to study common, processed food products is to obtain processed products with quality attributes, including the glycemic response, that are constant in time [16]. This can only be reached if properties of the raw materials and the way of processing, together with their processing conditions, are constant. For this reason products developed for pure research purposes, e.g. pasta varying in its thickness or processing conditions [9] are beyond the scope of this study. As shown previously [6] predictive PLSR models can be developed for the perceived satiety of fresh and processed food products with the constraint of equi-energetic (1MJ) servings. To come to equi-energetic servings, a reversible transformation scheme was developed to come from 50g carbohydrate into 1MJ servings taking the reference product, either glucose, of white bread, into consideration (Scheme 1); an operation comparable in going from degrees Centigrade to degrees Fahrenheit (and vice versa). Applying this constraint of 1MJ servings to a large range of common, processed food products, it is possible to develop predictive PLSR models relating relevant macronutrients (starch, sugar and) of individual food products to their glycemic response. To obtain these models however, besides the constraint of 1MJ servings, a prerequisite is that also the processing method has to be taken into consideration (Table 1) together with the type of assessor (diabetic, or non-diabetic). Taking these constraints into account, five different groups were defined and ten PLSR models. These ten models are mutually significantly different (P < 0.05) from each other: Legumes (group 3B) exhibit a deviant glycemic behavior compared to the four other groups; a glycemic enhancing effect by its possibly caused by the heat degraded cell wall products [78]. Product group 4, consisting of pasta, bulgur and par-boiled products, is less consistent in its processing method compared to the other groups. With respect to the estimated GS value starch it is obvious that this value decreases in the range Group 2 > Group 1 > Group 3A > Group 4 > Group 3B. Part of this difference in GS value of starch can be ascribed to the way of processing.

Clustering the predicted GlyS values from the regression models of the diabetic and of the non-diabetic test persons clearly shows that the glycemic response between these two groups differ (Equations. 2a, b and 3a, b). This clearly shows that the glycemic response of diabetic persons differs from that of non-diabetic persons.

With reference to Table 4 the question was addressed regarding the relation between the energy content of a food product, either present in the form of carbohydrates, or in the form of fat plus protein, and its glycemic response. The results given in this table clearly show that a good to excellent relation exists for all food groups between the predicted GlyS values and the percentage energy contained within protein plus fat; the more protein plus fat a sample contains the lower its GlyS value. Here the difference between the GI and the GlyS approach becomes obvious. The dimensionless GI value of a food product is determined on basis of the glycemic response of 50g carbohydrate present within that product [1,5]. In this case the other macronutrients (protein, fat and) are formally of no relevance and ignored, though they actually do affect the glycemic response of any product. This decrease in GlyS value by either protein and, or fat in our opinion also explains the not well understood observation that the addition of fat or protein to a product decreases its GS value (80.81). In other words, though neither formally protein nor fat are not directly responsible for the glycemic response of a food product their presence in any food product does have a significant effect on the glycemic response of that product.

As such, the GI is a very valid concept with its major focus on the starch plus sugar based glycemic response of a product. Per product, the consequences of the variable amounts of protein, fat and are already integrated in the specific GI value of that product. The GlyS approach recognizes the role of the product specific amounts of protein and fat by using equi-energetic rather than 50 g carbohydrate servings. Besides the amount of sugar and starch, protein, fat and their mutual relations, characteristic for each individual food product also play, though indirectly, a role. The consequence of this is that, for example in case of the determination of the GI of apples almost no energy in the form of protein and fat is ignored, in case of almonds about 24, 5MJ is ignored. This will be worked out in greater detail [77].

The overall observation that the 601 samples analyzed here (see Table 2), originating from 58 manuscripts (Table 1), after transformation from 50g carbohydrate into 1 MJ servings, and after clustering into five product groups, that are based on the way the products have been processed, have a correlation of 0.88 between their measured and their predicted GlyS value strongly suggest the correctness of the approach to use equi-energetic servings for modeling purposes. Using this approach it is also clearly shown that diabetic and non-diabetic persons have significant different glycemic response to the same product. The consequence of this difference in glycemic response is that per food group a predictive PLSR model for diabetic and for non-diabetic persons had to be developed.

In addition, using the above described approach all the data analyzed here can be considered to be analyzed in a greater perspective. Data analysis is not just limited to one of the studies mentioned in the Table 1, but part of all these studies. These studies were performed at different places during a times scale of more than three decades. This all together enabled the development of the PLSR based, predictive models. This study also shows that, starting with the same raw material, the processing method strongly determines the glycemic response of the product obtained. For example wheat based products can be found in groups 1, 2 and 4, and potato based products in group 1, 2 and 3. A wheat or potato product of e.g. group 2 is characterized as outlier in group 1 indicating the relevance of processing on the GlyS values of the products obtained. Furthermore the approach chosen here enables the recognition and definition of sample outliers, inconsistent results, which recognition and definition is virtually impossible just studying the GI value of individual products. The models developed here can be used to develop products with a desired GlyS thus also GI value. Confirmation always requires a conventional GI measurement.

**Conflict of Interest**

Declaration of personal or funding interests: None.

Bibliography


