TITLE:

Modeling carbohydrate counting error in type 1 diabetes management[1]

AUTHORS:

Chiara Roversi[1], Martina Vettoretti[1], Simone Del Favero[1], Andrea Facchinetti[1], and Giovanni Sparacino[1], on behalf of the Hypo-RESOLVE Consortium

[1]Department of Information Engineering, University of Padova, Padova, Italy.

Address correspondence to:

Prof. Giovanni Sparacino

Department of Information Engineering, University of Padova

Via G. Gradenigo, 6, 35131, Padova (PD), Italy

Tel: +39 049 827 7741, Fax: +39 049 827 7699

ORCID iD: https://orcid.org/0000-0002-3248-1393

Email: gianni@dei.unipd.it

WORD COUNT: 5403 words

Keywords: type 1 diabetes, insulin therapy, carbohydrates, carbohydrates counting error, mathematical modeling, simulation

---

[1] Part of this work was presented as poster at the 41st Engineering in Medicine and Biology Conference (Berlin, 23-27 July 2019) and as an oral presentation at the 55th Annual Meeting of the European Association for the Study of Diabetes (Barcelona, 16-20 September 2019).

2

Abstract

Background: The error in estimating meal carbohydrates (CHO) amount is a critical mistake committed by type 1 diabetes (T1D) subjects. The aim of this study is both to investigate which factors, related to meals and subjects, affect the CHO-counting error most and, to develop a mathematical model of CHO-counting error embeddable in T1D patient decision simulators to do in silico clinical trials.

Methods: A published dataset of 50 T1D adults is used, which includes patient's CHO-count of 692 meals, dietitian's estimates of meal composition (used as reference), and several potential explanatory factors. The CHO-counting error is modeled by multiple linear regression with stepwise variable selection starting from 10 candidate predictors, i.e. education level, insulin treatment duration, age, body weight, meal type, CHO, lipid, energy, protein and fiber content. Inclusion of quadratic and interaction terms is also evaluated.

Results: Larger errors correspond to larger meals, most of the large meals are underestimated. The linear model selects CHO ($p<0.00001$), meal type ($p<0.00001$) and body weight ($p=0.047$), while its extended version embeds a quadratic term of CHO ($p<0.00001$) and interaction terms of meal type with CHO ($p=0.0001$) and fiber amount ($p=0.001$). The extended model explains 34.9% of the CHO-counting error variance. Comparison with the CHO-counting error description previously used in the T1D patient decision simulator shows that the proposed models return more credible realizations.

Conclusions: The most important predictors of CHO-counting errors are CHO and meal type. The mathematical models proposed improve the description of patients' behavior in the T1D patient decision simulator.

3

1. Introduction

Therapy for type 1 diabetes (T1D) consists of exogenous insulin administrations, aimed at maintaining blood glucose (BG) concentration within a safe range [1]. Meal insulin doses are tuned according to the amount of carbohydrates (CHO) ingested at meals [2]. In particular, the American Diabetes Association recommends the carbohydrate counting, or CHO counting approach, which consists of estimating the amount of CHO in the meal and administering an insulin dose proportional to this amount [3]. Accurate CHO counting requires specific training and knowledge about the CHO content of different foods and meals. Consequently, CHO counting is a difficult task for T1D subjects, who frequently commit errors. Remarkably, the smartphone applications proposed in the literature as an automatic aid for counting CHO [4-6] are, at present, only used by a minority of patients.

Currently, little quantitative knowledge is available in the literature regarding the accuracy of CHO estimations made by patients and the impact of CHO counting errors. One common practice is to consider an error of ±10 g per meal or snack acceptable, based on the study performed by Smart et al.[7] which demonstrated that, a ±10 g error in the estimate of 60 g of CHO (i.e. a relative error of 17%) did not lead to any difference in postprandial glycemic control. Nevertheless, several studies have suggested that patients, in reality, make much larger errors [8-14].

Other literature has investigated the impact of CHO counting errors on glycemic control, through both in vivo [15-17] and in silico [11,12,18,19] clinical trials and has shown that CHO counting errors can strongly influence post-prandial BG excursions: CHO underestimation can cause post-prandial hyperglycemia, while CHO overestimation can lead to hypoglycemic episodes.

In order to extensively assess the effect of CHO counting errors on the quality of glycemic control via computer simulations and, more in general, to perform more reliable in silico trials, an accurate model of the CHO counting error is needed. One of the most recently presented tools for performing in silico trials is the T1D patient decision simulator [20], which complements the UVA/Padova model of glucose, insulin and glucagon dynamics in T1D patients [21] with a model of T1D patients' behavior when making treatment decisions. The latter includes a very simple CHO counting error description, i.e. a

4

probability density function, which does not take into account any correlation of the CHO counting error with any patient or meal covariates that might influence the accuracy of CHO counting. Nevertheless, apart from this simple model, to the best of our knowledge no other model of the CHO counting error has, as yet, been proposed in the literature. Indeed, a new, more accurate model of the CHO counting error is needed so as to improve the description of patient behavior when making treatment decisions and so be able to simulate T1D treatment scenarios even more realistically.

The main goals of this work are i) to study the factors that can influence the CHO counting error, including both patient and meal characteristics, ii) to develop a model of the CHO counting error which takes these characteristics into account, and iii) to incorporate the new model in the T1D patient decision simulator [20]. To achieve these aims, we used the dataset already published in Brazeau et al. [10], which, uniquely, offers very rich information about meals (e.g. accurate estimates of meal nutrient content) and patients (e.g. level of education, duration of insulin treatment, body weight, etc.), which, to the best of our knowledge, has not been published in any other literature studies. After a first exploratory analysis of the available meals and of CHO counting error data, CHO counting error models are developed using multiple linear regression with stepwise variable selection applied on 10 candidate predictors and their interaction and quadratic terms. The generalizability of the findings is assessed using a leave-one-out cross-validation strategy. The models developed are then incorporated into the T1D patient decision simulator [20].

2. Methods

2.1 Dataset description

The available data come from a published study [10] that involved 50 T1D adults who estimated their CHO quantity in meals for about 3 days, while maintaining their usual physical activities and food habits. Participants wore a continuous glucose monitoring (CGM) sensor throughout the study. Participants were 48% women, 42.7±11.1 years old,26% of whom had attended Secondary School,22% CEGEP,52% had attended university. Participants had a mean diabetes duration of 21.4±12.7 years, HbA1c of

5

7.6±1.2% (60±10 mmol/mol),body weight of 72.7±14.8 kg and body mass index (BMI) of 25.1±3.6 kg/m2.The data are:

- the CHO amount estimated by each subject for all meals, including breakfast, lunch, dinner and snacks;
- the CHO amount of each meal determined by an expert dietitian using a computerized analysis program;
- other covariates concerning meal composition (e.g. proteins, energy content, lipids and fibers), information on subjects and therapies (e.g. age, level of education, duration of T1D, body weight, BMI, HbA1c, etc.) and glucose variability metrics extracted from CGM sensor data.

The food diaries completed by each participant were analyzed by a dietitian using the Food Processor SQL (ESHA Research, Salem, OR) with the 2007 Canadian Nutrient File and, if or when necessary, food label information was added to the database. The analyses were verified by an independent expert [10].

A total of food records of 692 meals (146 breakfasts, 156 lunches, 146 dinners and 244 snacks) are available for, on average, 13.8±3 meals per subject. Both the median and interquartile range of the macronutrient content for different types of meal are reported, for reasons of space, in supplementary Table S1.

2.2 Exploratory analysis of meals and CHO counting errors

We calculated the CHO estimation error, $\widetilde{CHO}$, as:

$$\widetilde{CHO} = \widehat{CHO} - CHO \tag{1}$$

i.e. the difference between the patient's CHO estimate ($\widehat{CHO}$) and the CHO amount determined by the dietitian (CHO), which was considered as the reference CHO count. An exploratory analysis of the meals consumed and the CHO counting errors committed by subjects was performed using both boxplots and scatterplots.

6

2.3 Developing of a CHO counting error model

2.3.1 Linear model

A CHO counting error model was developed using the multiple linear regression approach. For our purpose, the CHO counting error, $\widetilde{CHO}$, is the dependent variable, while the other variables of the dataset (e.g. level of education, duration of insulin treatment, meal CHO content, fiber amount, type and others) are the independent variables. The coefficients of the model are estimated using the linear least squares approach [22,23].

At first, the full model, i.e. the model with all the predictors taken into account, was fitted and statistical tests were performed to investigate which coefficients are significantly different from zero. In particular, an F-test with null hypothesis H_0: *all the slopes of the linear model are equal to 0* and a t-test for each of the model's coefficients with null hypothesis H_0: *the coefficient is equal to 0* were both performed with a 5% level of significance [22,23]. If the F-test detected that at least one of the regressors was related to the response, and the t-tests showed that not all the coefficients were significantly different from 0, the model complexity could be reduced by selecting the most important predictors to explain the response. The reduction of model complexity was performed using stepwise variable selection with bidirectional elimination [22,23]. The stopping rule employed adopted the p-value of an F-statistic, so as to test models both with, and without, a potential variable at each step. The threshold on the p-value for entering was set at 0.05, while on the p-value for removing, was set at 0.1.

2.3.2 Model with interactions and quadratic terms

The linear model considered may be restrictive. A simple way of introducing more flexibility into the model would be to add interaction and polynomial terms [22,23].

In our case, we extended the linear model by introducing interaction and quadratic terms. In our investigation we performed the same statistical tests and methods to reduce model complexity as those used for the linear model case (Section 2.3.1). The coefficient of determination, R^2, and the *adjusted R^2* parameter, i.e. R^2 adjusted for the number of predictors in the model, were used to compare the amount of CHO counting error variance explained by the different models developed.

7

2.4 Selection of the regressors

Given the large number of possible regressors (risk of overfitting) and the fact that they are correlated (making the interpretation of the coefficients difficult), when developing the model we restricted our analysis to a subgroup of regressors, manually selected, to remove the redundant terms and covariates not relevant for the CHO counting error.

Firstly, we removed variables strictly correlated with each other, to reduce collinearity. Thus, from a group of highly correlated variables, we kept only those which best explained the others, so as to preserve as much information as possible. For example, this was done for the variables related to obesity, i.e. body weight, body mass index, waist and lean body mass, which are strongly correlated with each other. Among these, only body weight was kept. Variables relating to macronutrient meal content (i.e. CHO, lipids, energy and fiber) are correlated with each other, but the correlation is not high enough to assume collinearity, except for the energy content, which has a Pearson correlation coefficient of 0.79 with CHO, 0.86 with proteins and 0.89 with lipids. We decided to keep all the variables related to meals in the model and perform a check of collinearity once the CHO counting error model had been developed (Section 3.2.4).

Then, since the regression technique is not able to give information about the direction of the cause-and-effect relationship between the CHO counting error and the available covariates, some variables that seemed to be more the effect, rather than the cause, of the error, i.e. the number of units of rapid insulin and HbA1c level, were dropped. The same was done for glycemic control metrics, also because these metrics were calculated as the mean over 72-h of the study, thus they represent glycemic control over the entire duration of the experiment, not just before meals.

2.5 Assessment of the generalizability of the models through cross-validation

A second independent dataset, including all the covariates used as explanatory factors of the CHO counting error, that could have been used for the purpose of model validation, is not available in the published literature, at least to the best of our knowledge. Therefore, the models developed in this paper are validated by a cross-validation strategy. More precisely, we performed a leave-one-out cross validation "per subject", i.e. at each iteration we used the data of one subject as a test set, while data of all the other 49 subjects were adopted as the training set. In this way, at each iteration, the test set was

8

composed of data on a subject which were not used for training the models. For each iteration, the Root Mean Square Error (RMSE) was computed on the test data as

$$RMSE = \sqrt{\frac{1}{n_{test}} \sum_{i=1}^{n_{test}} (y_i - \hat{y}_i)^2} \tag{2}$$

where y_i is the ith observation, i.e. the CHO counting error, of the test set, $\hat{y}_i$ is the estimate of the ith observation obtained with the model trained on the training set and n_{test} is the cardinality of the test set.

The estimate of the test error was obtained by computing the mean and 95% confidence interval of the RMSE obtained in the 50 iterations.

2.6 Incorporation in the T1D patient decision simulator

Both the linear and extended models are incorporated in the T1D patient decision simulator [20]. To take into account the variability of the CHO counting error generated by factors not included in the models we added a white Gaussian noise to each equation, with zero mean and constant variance calculated on the residuals obtained in the least-square identification of the final models. Then, we performed three simulations on 100 virtual subjects for 7 days with 3 meals per subject-day (i.e. breakfast, lunch and dinner) and therapy based on non-adjunctive CGM use. For each simulation, a total of 2100 meal CHO data were generated, with the corresponding CHO counting errors generated by different models. In the first simulation, the "old" model, that already present in the simulator, was used. This simple model consisted of a Student t probability density function fitted on the percentage CHO estimation error extracted from the Brazeau et al [10] dataset. In the second simulation, the "new" linear model, derived as in Section 2.3.1, was used while, in the third simulation, the CHO counting error was generated using the "new" extended model, derived as in Section 2.3.2. For each simulation, the meal CHO amount was generated by using a meal distribution similar to that of the real data produced by Brazeau et al. [10].

Lastly, the CHO counting error, generated by the "old" model used in [20] and the "new" models developed in this paper, were plotted vs. meal CHO in order to qualitatively assess whether the models could capture the relationship between the CHO counting error and the meal CHO observed in real data. In addition, a quantitative indicator of model accuracy

9

in capturing this relationship is provided by fitting regression lines against both real and simulated data and by comparing their parameters. The general equation of each regression line is as follows

$$\widehat{CHO} = \beta_0 + \beta_1 CHO \tag{3}$$

where

β_0 is the intercept and β_1 is the coefficient representing the correlation between CHO and the error (CHO)̆.

3. Results

3.1 Exploratory analysis of meals and CHO counting error

Panel A of Figure 1 shows the boxplots of meal CHO content in different types of meal. The average CHO content is 58.18 ± 26.06 g for breakfast, 77.97 ± 35.13 g for lunch, 80.57 ± 37.64 g for dinner and 36.53 ± 30.10 g for a snack. Hence, lunch and dinner seem, on average, to be larger meals than breakfast and snacks. Quantitatively similar boxplots were obtained for the other meal variables, i.e. energy, lipids, proteins and fiber (details are not reported here for reasons of space). These results are in line with the expectations generated by the fact that participants' food consumption was typical of the Canadian population's habits, where the smallest meal is breakfast and the largest is dinner [10].

To understand the extent to which individuals tend to eat the same amount of CHO at different meals (intra-subject variability) and what the difference is in the meal habits of the subjects (inter-subject variability), the CHO amount eaten by each subject for different types of meal is shown in Figure 2 (note that subjects are ordered according to the individual average daily CHO amount, shown in panel E). It is clear that there is some intra- and inter-subject variability. However, different behaviors in the CHO amount eaten by subjects can also be discerned. For example, patient #14 tends to eat the same amount of CHO for each of the meal types, i.e. about 50 g at breakfast, 50 g at lunch, 65 g at dinner and 25 g for snacks. On the other hand, other patients have much more irregular habits, e.g. patient #40 presents large variability at lunch and dinner, and small variability at breakfast and snacks. Moreover, the subjects' mean daily CHO amount varies widely between subjects, from a minimum of 100 g for subject #1, to about 350 g for subject #50. Nevertheless, the duration of the study (about three days) is, in all likelihood, too short to reveal the real daily habits of participants.

10

The boxplots of the CHO counting error (CHO)̆ in the different types of meal (Figure 1,panel B) show that the error is larger at lunch and dinner when compared with breakfast and snacks. Therefore, larger errors correspond to larger meals. Moreover, panel C of Figure 1 shows that after plotting the CHO counting error,(CHO,)̆ against the meal CHO amount, a specific trend appears: most of the meals are underestimated (about 63%),smaller errors correspond to meals with low CHO content and there is a tendency to underestimate large meals. The relationship between CHO and the error (CHO)̆ is quantitatively described by the regression line reported in the plot, with equation

$$\widehat{CHO} = 10.54 - 0.28\, CHO \qquad (4)$$

Similar behavior was observed after making separate scatterplots for the data on breakfasts, lunches, dinners and snacks (not reported for the sake of brevity). These results offer initial evidence that both meal type and CHO amount are factors that influence the CHO counting error. The behavior of T1D subjects in the CHO estimation found in this work is in line with that obtained by two recent works by Reiterer et al. [11,12]. Moreover, the tendency to underestimate large meals, and to overestimate small meals, has also been detected in other works in the literature [11,12,13,14]. Lastly, some parameters describing the distribution of CHO counting errors for different levels of meal CHO amount in greater detail, are reported in Table 1. It should be noted that increasing the amount of CHO results in a CHO counting error which becomes, on average, higher (in absolute value) and with a negative bias.

3.2 Model describing the CHO counting error

3.2.1 Selection of regressors

As a result of the steps explained in Section 2.4, we selected the following 10 variables, employed as regressors in the model: level of education (4 discrete values: 1 for primary school, 2 for secondary school, 3 for CEGEP and 4 for university), duration of insulin treatment (years), age (years), body weight (kg), CHO (g), amount of lipids (g), energy content (kcal), amount of proteins (g), fiber content (g), meal type (breakfast, lunch, dinner or snack). All the variables involved are quantitative, except for that representing the type of meal which is a categorical variable, with 4 levels, associated with each category, incorporated into the regression model using 3 dummy variables (or indicators). So, there are a total of 13 coefficients in the model (intercept included).

11

3.2.2 Linear model

Firstly, the full linear model is fitted. Estimates of the coefficients and the corresponding p-value related to the t-test with null hypothesis H_0: *the coefficient is equal to 0* are reported in Table 2. The coefficient of determination R^2 of the model is 0.311, while the *adjusted R^2* parameter is 0.299. The full linear model can explain 31% of the variance of the CHO counting error. This means either that other factors, together with meal and patient covariates, are needed to explain better the CHO estimation error or, that the majority of the variability of the CHO counting errors are due to random errors.

An F-test was performed to test the null hypothesis H_0: *all the slopes of the linear model are equal to 0*, which was rejected at the 5% level (F-statistic=25.4, p-value<0.00001). This means that the full model gives more information about the response than does the constant one; so, we have strong evidence that at least one of the covariates is important when predicting the CHO counting error. By looking at the t-test p-values reported in Table 2, we see that the only coefficients statistically significantly different from zero at the 5% significance level are those of the variables CHO (p<0.001), meal type snack (p<0.001) and body weight (p=0.05). In particular, the statistical significance of the meal type snack in the full model, highlights the fact that the error for snacks is significantly different, statistically speaking, from the error for the other meals (i.e. of breakfast, lunch and dinner) even when the same amounts of CHO are considered. This is probably due to the fact that, generally, it is easier to assess the CHO of snacks because they are often composed of a single ingredient. Furthermore, packaged snacks all have a label reporting nutrient content.

Since only a few coefficients show statistically significant differences from zero, the full model can be reduced in complexity by selecting the most important predictors to explain the response. To do this, a stepwise variable selection approach was applied to the full linear model. The results of the stepwise procedure, reported in Table 3, suggest that CHO and the type of meal are the most important determinants of any CHO counting error. Indeed, they are added to the model in the first two steps with very low p-values, almost explaining the total variance revealed by the full model (R^2=0.304 vs R^2=0.311). Indeed

12

CHO, alone, is able to explain 27% of the total variance of the response. At the third step, a patient-specific variable, body weight, was added to the model. However, body weight could be considered a borderline predictor: its p-value is very close to the threshold for entering and it results in only a small increase in R^2: from 0.304 to 0.308. No other variables were added to the model, and within this dataset no strong correlation was found between the other variables and the error.

The equation of the final linear model that includes only the predictors selected by stepwise variable selection is the following:

$$\widetilde{CHO} = 9.22 - 0.34\,CHO + 0.09\,body\,weight + 3.11\,meal_{lunch} + 0.68\,meal_{dinner}$$
$$- 7.05\,meal_{snack}$$

$$(5)$$

3.2.3 Extended model

Interaction and quadratic terms were added to the model in order to improve the description of the CHO counting error. At first, the model, fitted using the 10 manually selected predictors, had 85 coefficients because of the intercept, linear terms, quadratic terms and interaction terms between each pair of predictors.

The coefficient of determination R^2 of the full extended model is 0.523, while the *adjusted R^2* parameter is 0.456. Because, when the interaction and quadratic terms were added, the *adjusted R^2* increased from 0.299 to 0.456, we concluded that the addition of interactions and quadratic terms makes it possible to explain more information about the response, but does not offer any dramatic improvements.

An F-test was then performed in order to test the null hypothesis H_0: *all the coefficients of the linear model are equal to 0*. Since the resulting F-statistic was equal to 7.87 with p-value $p < 0.00001$, the null hypothesis was rejected.

The most important predictors that could explain the response were investigated using the stepwise variable selection approach (see Table 3). A quadratic term of *CHO* was added at step 2, an interaction term between *CHO* and *meal* at step 4 and, an interaction term between *fiber* and *meal* at step 6. Therefore, when the stepwise technique was performed on linear, interaction and quadratic terms, variable *body weight* was excluded from the

13

model while *fiber* was included. The importance of *CHO* in explaining the CHO counting error was highlighted by the addition of the CHO^2 term with a very low p-value. No strong relationship between the error and the other variables excluded from the model was found. This result is in line with the study by Meade et al. [16], in which no association was found between CHO counting accuracy and both the duration of diabetes and the level of education.

Lastly, since the *adjusted R^2* of the linear stepwise model is 0.302, while for the extended model it is 0.349, we concluded that the addition of interactions and quadratic terms makes it possible to explain the response better, but performance does not markedly improve.

The equation of the stepwise extended model obtained is as follows:

$$\widehat{CHO} = +3.38 - 0.08\, CHO - 0.0009\, CHO^2 + 9.57\, meal_{lunch} + 17.95\, meal_{dinner}$$
$$+ 0.47\, meal_{snack} + 0.004\, CHO{:}meal_{lunch} - 0.21\, CHO{:}meal_{dinner}$$
$$+ 0.03\, CHO{:}meal_{snack} + 0.19\, fiber - 1.38\, fiber{:}meal_{lunch}$$
$$- 0.45\, fiber{:}meal_{dinner}$$
$$- 2.00\, fiber{:}meal_{snack} \tag{6}$$

3.2.4 Test of collinearity

After developing the model, we performed a check on collinearity for both the linear and the extended models because of the "energy" variable, which is closely correlated with the other meal variables. To test for the possible presence of collinearity, we removed the variable, "energy", from the candidate predictors and performed stepwise variable selection on the remaining 9 predictors. The results obtained through this procedure, i.e. excluding "energy", were identical to those obtained when all the 10 predictor variables were involved (i.e. the results shown in Table 3). This check confirmed that the presence of this variable does not distort the results of the model.

3.2.5 Assessment of the generalizability of the models

The validation of the models developed was performed using cross-validation, as explained in Section 2.5. For the linear model, the RMSE mean is equal to 15.79 g, while the 95% confidence interval is [13.89, 17.69] g. Instead, for the extended model, the RMSE mean is equal to 15.41 g, while the 95% confidence interval is [13.43, 17.38] g. The confidence

intervals obtained are tight around the corresponding RMSE error means. Results suggest that the performance of the model using previously unseen data is satisfactory and comparable to that of the models of Equation (5) and Equation (6) for the entire dataset that was used for training the models (i.e. RMSE=16.61 g for the linear model and RMSE=15.97 g for the extended model) .

3.2.6 Incorporation in the T1D patient decision simulator

In order to be incorporated into the simulator, the extended CHO counting error model of Equation (6) needs to be reduced by excluding the terms containing the variable "fiber". The reason for doing this is that the variable "fiber" is not currently present in the virtual meals generated by the simulator. The resulting model:

$$\widetilde{CHO} = +3.56 - 0.07\, CHO - 0.0008\, CHO^2 + 6.77\, meal_{lunch} + 18.01\, meal_{dinner} - 0.49\, meal_{snack} - 0.08\, CHO\!:\!meal_{lunch} - 0.25\, CHO\!:\!meal_{dinner} - 0.06\, CHO\!:\!meal_{snack}$$

$$(7)$$

presents an R^2 of 0.34 and an *adjusted R^2* of 0.332. However, the models of Equation (5) and (7) are incorporated into the T1D patient decision simulator [20]. The residual variance of the models, through which the white Gaussian noise was generated, is equal to 278.26 g^2 and 266.37 g^2, respectively.

We then simulated CHO counting errors using the "new" developed models and the "old" one for 100 virtual subjects, 7 days, 3 meals per subject-day (i.e. breakfast, lunch and dinner) and therapy based on non-adjunctive CGM use, and then compared them with the real data of breakfast, lunch and dinner used in this work. Given that the current version of the simulator [20] does not yet include a snack model, only main meals were simulated. In Figure 3, one can see that the CHO counting error realizations, generated by the developed linear and extended models (panel C and D, respectively), describe the trend detected on real data (panel A) better when they are compared with the "old" model (panel B). In other words, the "old" model generated over- and under-estimations independently of the true CHO content while, in real life, subjects tend to underestimate large meals. This is

15

confirmed by comparing the parameters of the regression lines (see Equation 3) which fit the relationship between CHO and the error (CHO)ʳ for both real and simulated data.For the real data on breakfast,lunch and dinner,the β_0 coefficient of the regression line is equal to 17.87,while β_1 is -0.35.The regression lines fitted on the data when simulated with the newly developed models have parameters that are very close to those of the real data regression line (β_0=15.83 and β_1=-0.33 for the linear model,β_0=20.70 and β_1=-0.40 for the extended model).On the other hand,the regression line fitted on the data simulated using the "old" model has parameters β_0=0.74 and β_1=-0.08. Thus we can conclude that both the models developed in this work provide CHO counting error which, for the purpose of performing in silico clinical trials, are both more realistic and consonant with real data than are those offered by the "old" model.

4. Conclusions

In this work, adopting a dataset gathered on 50 T1D subjects, published in [10], we first demonstrated that both CHO amount and type of meal are the most important factors influencing any CHO counting errors made during T1D management. In particular, we noted that larger errors correspond to larger meals (i.e. lunch and dinner) and most of these large meals are underestimated.

Then, by applying multiple linear regression and a stepwise variable selection approach using 10 previously selected candidate predictors, we developed a model of these CHO counting errors. The final model only takes into account the CHO, the type of meal and the body weight variables. CHO and type of meal were shown to be the most important predictors useful for explaining CHO counting errors, with only a small portion of variance in the response explained by this model (R^2=0.308). A slight improvement in the performance of the model (R^2=0.360) was obtained by introducing a quadratic term of CHO, an interaction term between CHO and meal type and, finally, an interaction term between fiber and meal type.

Lastly, the linear model of Equation (5) and the extended model of Equation (7) were incorporated into the T1D patient decision simulator [20] which seeks to perform even more reliable in silico trials. Analysis of the CHO counting errors generated by the two new models demonstrated the credibility of their results, which were closer to the real data

16

than were those provided by the simpler model previously adopted in the T1D patient decision simulator [20]. Thus, the results reported in this paper could help to enable more realistic in silico clinical trials.

It is worth pointing out that, within the dataset used, no significant relationship was found between the CHO counting error and the subject's level of education, duration of insulin treatment, age, meal lipids, energy content and amount of protein. It is also worth mentioning that the same analysis conducted in this paper for the CHO counting error, $(CHO)^{\tilde{}}$, was also performed for the relative CHO counting error, i.e. the error $(CHO)^{\tilde{}}$ divided by the meal CHO amount ($(CHO)^{\tilde{}}$/CHO in the notation used in our equations). The results show that the relative error also depends on CHO and meal type, which was to be expected from the analysis conducted on the absolute error with quadratic and interaction terms that revealed the dependency of the absolute error on CHO^2 and CHO: meal. In order to make this paper clearer to read, the results obtained for the relative error are presented as supplementary material.

To conclude, some limitations of the used data should be highlighted. Of course, other factors, not included in our analysis because they were not in the dataset used, could influence the carb-counting error. For example, erroneous setting of the insulin: carbohydrate (IC) ratio parameter may influence the accuracy of CHO counting. Indeed, based on the subject's daily experience, the individual could voluntarily have made mistakes in CHO assessment in order to compensate for a wrong IC ratio parameter. However, the IC ratio parameter was not in the data used, so we were unable to analyze this aspect. Moreover, results may have been affected by the training and education level that patients had received, again, information which was not available in the dataset used.

It is also important to mention that the carb-counting error is only one of the factors that could affect post-prandial glycemic control. For example, the negative bias in the carb-counting error found in our data could afterwards be corrected by the diabetologist through appropriate adjustment of the IC ratio parameter. Moreover, the impact of food on glycemic control in T1D may be influenced by factors other than CHO, such as gastric emptying, fat and protein content, food glycemic index, courses of the meal and any previous physical activity. Although it would have been interesting to have examined these phenomena, unfortunately, there was no information in the dataset available. However,

17

our work does not intend to assess the impact of carb-counting on glycemic control, rather it seeks to develop a model of the carb-counting errors which makes it possible to simulate, with reasonable accuracy, the behavior of T1D subjects in carb-counting. The real novelty of this work is, in fact, the derived multivariable statistical model of the carb-counting error, which could be used to perform realistic in silico clinical trials. A carb-counting error model like the one proposed here has, so far, not been published in the literature.

Lastly, it is important to highlight that the carb-counting error model is only one of the aspects which make it possible to obtain a realistic patient behavior model. For example, further models, e.g. of the physical activity, would be required in order to obtain even more realistic T1D simulations.

While the leave-one-out cross-validation strategy used in this paper supports the generalizability of the predictive ability of the models, future work should include assessment of the performance of models proposed on independent datasets, should they become available, and would be possible because of the ease with which the proposed methodology could be applied to any new data. Other possible developments could also include the use of the CHO counting error models developed within the T1D patient decision simulator to quantitatively assess, in silico, the impact of CHO counting errors on the quality of diabetes management. Lastly, if appropriate data were to become available, it would be interesting to perform a comparison of the carb-counting error between groups of patients adopting different counting methods, e.g. comparing accuracy in CHO estimates when using apps, or other traditional counting methods, such as grams or exchanges.

5. Acknowledgments

6. Author Disclosure Statement

The authors have no conflicts of interest.

18

7. **Funding**

This study is part of Hypo-RESOLVE (Hypoglycemia – Redefining SOLutions for better liVEs) (http://hypo-resolve.eu/), a project currently ongoing (1.5.2018-30.4.2022) which has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 777460. The JU receives support from the European Union Horizon 2020 research and innovation programme and EFPIA and T1D Exchange, JDRF, International Diabetes Federation (IDF), The Leona M. and Harry B. Helmsley Charitable Trust.

8. References

1. De Witt DE, Hirsch IB: Outpatient Insulin Therapy in Type 1 and Type 2 Diabetes Mellitus. JAMA 2003; 289:2254-2264.

2. Schmidt S, Norgaard K: Bolus Calculator. J Diabetes Sci Technol 2014; 8:1035-1041.

3. American Diabetes Association: Nutrition Therapy Recommendations for the Management of Adults with Diabetes. Diabetes Care 2013; 36:3821-3842.

4. Vasiloglou MF, Mougiakakou S, Aubry E, et al: A Comparative Study on Carbohydrate Estimation: GoCARB vs Dietitians. Nutrients 2018; 10.

5. Rhyner D, Loher H, Dehais J, et al: Carbohydrate Estimation by a Mobile Phone-Based System Versus Self-Estimations of Individuals With Type 1 Diabetes Mellitus: A Comparative Study. J Med Internet Res 2016; 18.

6. Domhardt M, Tiefengrabner M, Dinic R, et al: Training of Carbohydrate Estimation for People with Diabetes Using Mobile Augmented Reality. J Diabetes Sci Technol 2015; 9:516-524.

7. Smart CE, Boss K, Edge JA, et al: Children and Adolescents on Intensive Insulin Therapy Maintain Postprandial Glycaemic Control Without Precise Carbohydrate Counting. Diabetic Medicine 2009; 26:279-285.

8. Bishop FK, Maahs DM, Spiegel G, et al: The Carbohydrate Counting in Adolescents with Type 1 Diabetes (CCAT) Study. Diabetes Spectrum 2009; 22:56-62.

9. Meade LT, Rushton WE: Accuracy of Carbohydrate Counting in Adults. Clin Diabetes Journal 2016; 34:142-147.

10. Brazeau AS, Mircescu H, Desjardins K, et al: Carbohydrate Counting Accuracy and Blood Glucose Variability in Adults With Type 1 Diabetes. Diabetes Res. Clin. Pract. 2013; 99:19-23.

11. Reiterer F, Freckmann G, Del Re L: Impact of Carbohydrate Counting Errors on Glycemic Control in Type 1 Diabetes. Science Direct 2018; 51:186-191.

12. Reiterer F, Freckmann G: Advanced Carbohydrate Counting: An Engineering Perspective. Annual Reviews in Control, https://doi.org/10.1016/j.arcontrol.2019.06.003

13. Kawamura T, Takamura C, Hirose M, et al: The Factors Affecting on Estimation of Carbohydrate Content of Meals in Carbohydrate Counting. Clin Pediatr Endocrinol 2015; 24:153-165.

14. Smart CE, Ross K, Edge JA, et al: Can Children with Type 1 Diabetes and Their Caregivers Estimate the Carbohydrate Content of Meals and Snacks? Diabetic Medicine 2010; 27:348-353.

15. Smart CE, King BR, McElduff P, et al: In Children Using Intensive Insulin Therapy, a 20-g Variation in Carbohydrate Amount Significantly Impacts on Postprandial Glycaemia. Diabetic Medicine 2012; 29:e21-24.

16. Laurenzi A, Bolla AM, Panigoni G, et al: Effects of Carbohydrate Counting on Glucose Control and Quality of Life Over 24 Weeks in Adult Patients With Type 1 Diabetes on Continuous Subcutaneous Insulin Infusion. Diabetes Care 2011; 34:823-827.

17. Deeb A, Al Hajeri A, Alhmoundi I, et al: Accurate Carbohydrate Counting Is an Important Determinant of Postprandial Glycemia in Children and Adolescents With Type I Diabetes on Insulin Pump Therapy. J Diabetes Sci Technol 2017; 11:753-758.

18. Sun Q, Jankovic MV, Mougiakakou SG: Impact of Errors in Carbohydrate Estimation on Control of Blood Glucose in Type 1 Diabetes. In 14th Symposium on Neural Networks and Applications (NEUREL), Belgrade, Serbia, Nov 2018.

19. Virdi NS, Mahoney JJ: Importance of Blood Glucose Meter and Carbohydrate Estimation Accuracy. J Diabetes Sci Technol 2012; 6:921-926.

20. Vettoretti M, Facchinetti A, Sparacino G, et al: Type 1 Diabetes Patient Decision Simulator for In Silico Testing Safety and Effectiveness of Insulin Treatments. IEEE Trans Biomed Eng 2018; 65:1281-1290.

21. Dalla Man C, Micheletto F, Lv D, et al: Type 1 Diabetes Simulator: New Features. J Diabetes Sci Technol 2014; 8:26-34.

22. Rawlings JO, Pantula SG, Dickey DA: Applied Regression Analysis – A Research tool. Second Edition, Springer, New York, 1998.

23. James G, Witten D, Hastie T, Tibshirani R: An Introduction to Statistical Learning: With Application in R. Springer, New York, 2013.

21

Table 1: The table shows mean (second column), standard deviation (third column), confidence interval at 95% (fourth column), minimum (fifth column) and maximum (last column) values of the CHO counting error for different levels of meal amount (expressed in grams).

	Mean (g)	SD (g)	95% CI (g)	Min (g)	Max (g)
$CHO \leq 20$	0.53	4.78	[-0.51 1.57]	-14.33	17
$20 < CHO \leq 40$	0.79	10.93	[-0.90 2.49]	-23.68	52.7
$40 < CHO \leq 60$	-0.39	11.95	[-2.29 1.51]	-36.08	69.93
$60 < CHO \leq 80$	-6.98	15.65	[-9.69 -4.27]	-55.76	43.17
$80 < CHO \leq 100$	-10.28	21.27	[-14.88 -5.67]	-63.25	55.69
$CHO > 100$	-29.37	33.40	[-36.27 -22.47]	-142.24	43.18

22

Table 2: Estimate and p-value of the coefficients of the full linear model.

Regressor	Estimate	p-value
Intercept	6.440	0.232
Education	0.739	0.344
Duration insulin treatment	0.003	0.954
Age	0.019	0.764
Body Weight	0.090	0.054
CHO	-0.316	**<0.001**
Lipids	0.001	0.993
Energy	-0.002	0.855
Proteins	0.033	0.686
Fiber	-0.273	0.143
Meal_lunch	3.254	0.112
Meal_dinner	0.766	0.729
Meal_snack	-7.526	**<0.001**

23

Table 3: Results of the stepwise procedure adoptive both for the linear model and for the model with interactions and quadratic terms. The variable added or removed at each step (first column), the value of the F-statistic (second column), the corresponding p-value (third column) and the value of the R^2 parameter (fourth column) for the current model are reported.

	Linear model		
	F-statistic	**p-value**	**R^2**
1. Add **CHO**	258.054	<0.00001	0.273
2. Add **meal**	9.840	<0.00001	0.304
3. Add **weight**	3.953	p=0.047	0.308

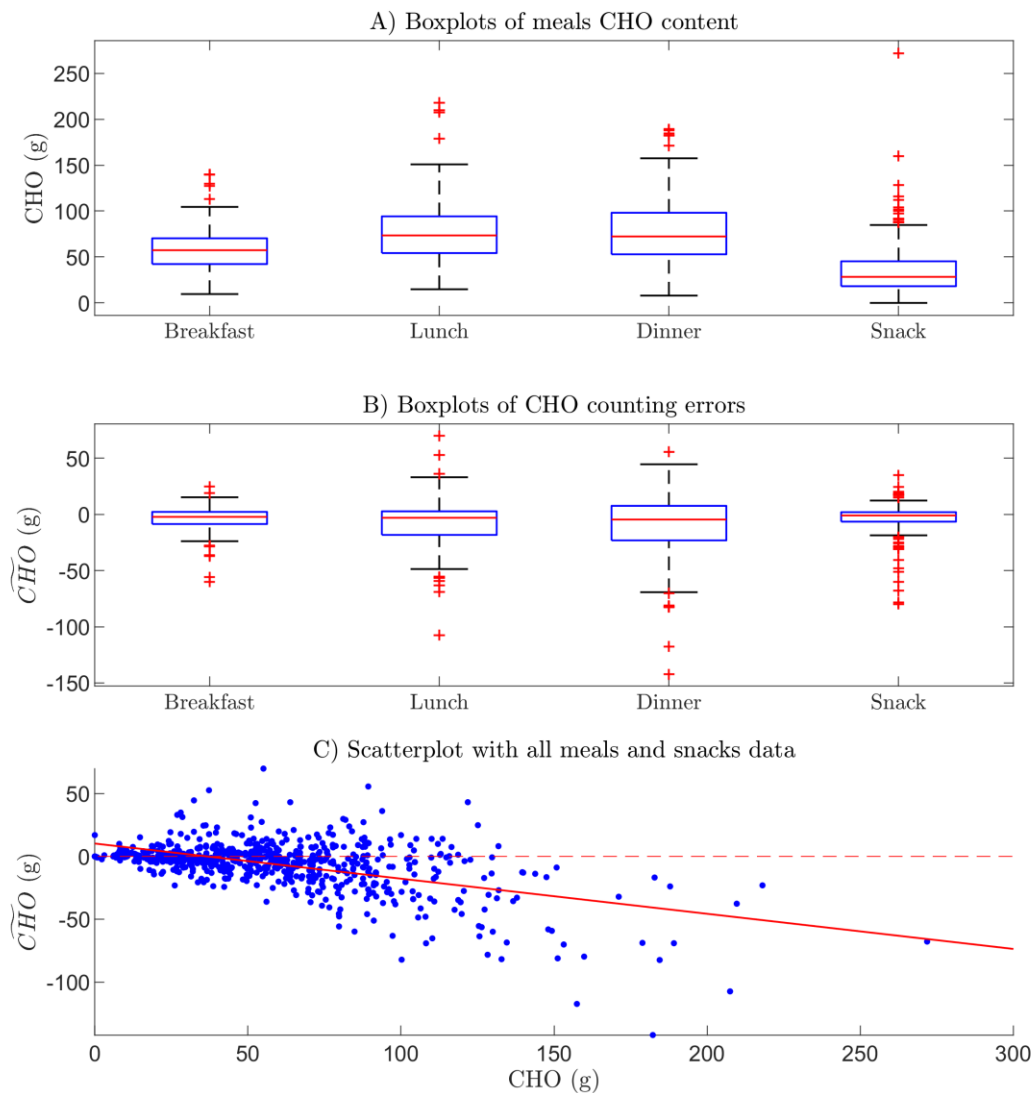|	Extended model		
	F-statistic	**p-value**	**R^2**
1. Add **CHO**	258.054	p<0.00001	0.273
2. Add **CHO^2**	31.700	p<0.00001	0.305
3. Add **meal**	4.655	p=0.003	0.319
4. Add **CHO:meal**	6.998	p=0.0001	0.340
5. Add **fiber**	5.115	p=0.024	0.345
6. Add **fiber:meal**	5.304	p=0.001	0.360

FIG. 1. Boxplots of CHO content (panel A) and CHO counting errors (panel B) of breakfast, lunch, dinner and snacks of all 50 patients. In each boxplot, the red horizontal line represents the median, the blue box marks the interquartile range, dashed lines are the whiskers and red crosses indicate outliers. Whiskers are drawn from the ends of the interquartile range to the adjacent values, which are the most extreme data values that are not outliers. By default, an outlier is a value that is more than 1.5 times the interquartile range away from the top or bottom of the box. The scatterplot of the CHO counting error against meal CHO amount, together with the corresponding regression line (in red), is reported in panel C.
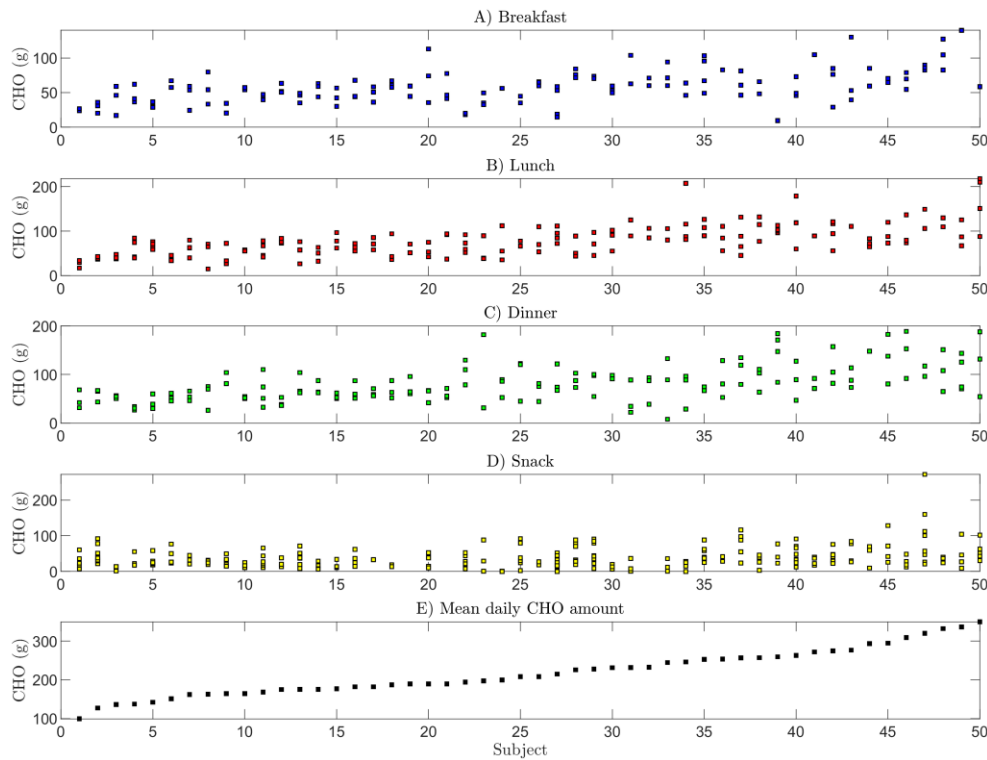
FIG. 2. Plots of meal CHO amounts eaten by each subject at breakfast (panel A), lunch (panel B), dinner (panel C) and snacks (panel D), and the mean daily CHO amount for each subject (panel E). Subjects are in ascending order with respect to their mean daily CHO dose. Note that since, on the four days of the study, the number of meals per day, per subject, differed from day to day, especially on the first and the last day when some meals were not registered, the mean daily dose for each subject was calculated as the sum of their mean breakfast dose, mean lunch dose, mean dinner dose and mean snacks dose.
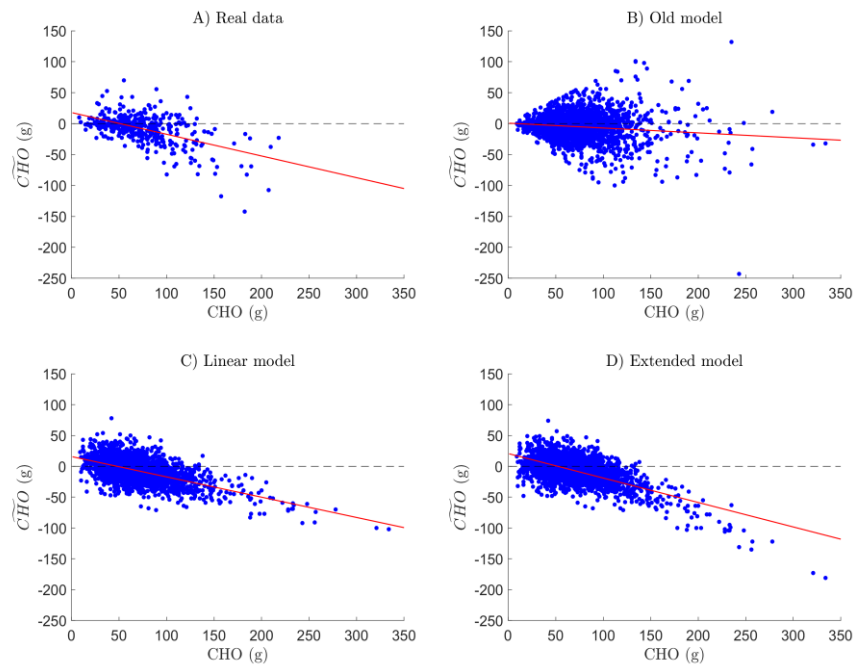
FIG. 3. Plots of the CHO counting error against meal CHO amount, together with the corresponding regression line (in red) for real data (panel A) and simulated data (panel B, C, D) of breakfast, lunch and dinner. Meal CHO amount was simulated with a distribution function similar to that of the real CHO data, while carb-counting errors were generated using the "old" (panel B), linear (panel C) and extended (D) models.

27

Table S1: Median and interquartile range (median [iqr]) of the macronutrient content for different type of meal, i.e. breakfast (first row), lunch (second row), dinner (third row) and snack (bottom row).

Type of meal	Energy (kcal)	Lipids (g)	Proteins (g)	Fiber (g)
Breakfast	382.04	10.99	12.28	3.98
	[287.92 498.83]	[5.64 18.73]	[8.58 18.43]	[2.05 6.27]
Lunch	603.88	23.30	26.75	5.79
	[458.85 851.44]	[11.49 39.33]	[16.01 43.06]	[3.86 8.80]
Dinner	722.77	26.21	31.67	5.85
	[536.04 934.44]	[14.90 42.17]	[22.84 48.54]	[3.52 8.62]
Snack	180	3.78	3.90	1.37
	[101.92 292.65]	[0.35 9.72]	[1.03 8.00]	[0.45 2.90]

28

Supplementary material

    I.   **Relative CHO counting error modeling**

    I.A Exploratory analysis of relative CHO counting error

The relative CHO counting error is defined as

$$\widehat{CHO}_{rel} = \frac{\widehat{CHO} - CHO}{CHO} \qquad\qquad (S1)$$

Where $\widehat{CHO}$ is the patient's CHO estimate and CHO is the reference CHO count. So, it represents the error in CHO counting expressed as a percentage of the reference meal CHO amount.

In Table S2, some parameters are reported describing, in more detail, the distribution of the relative CHO counting errors for different levels of meal CHO amount. It is evident that, when the amount of CHO is increased, the CHO counting error becomes, on average, higher (in absolute value) and has a negative bias.

    I.B Linear model

A model of relative CHO counting error is developed using the same methodology as that described in the main text and employed to develop the model of the absolute CHO counting error, i.e. multiple linear regression together with stepwise variable selection. The dependent variable of the model is the relative CHO counting error $\widehat{CHO}_{rel}$. The regressors are the 10 variables selected in Section 2.4, i.e. level of education (4 discrete values: 1 for primary school, 2 for secondary school, 3 for CEGEP and 4 for university), duration of insulin treatment (years), age (years), body weight (kg), CHO (g), amount of lipids (g), energy content (kcal), amount of proteins (g), fiber content (g), meal type (breakfast, lunch, dinner or snack). The full linear model, i.e. the model fitted using all the 10 predictors selected, presents a coefficient of determination R^2 of 0.104 and an adjusted R^2 of 0.087.

Table S3 includes the results of stepwise variable selection as applied to the full linear model. This reveals that the relative error depends not only on meal type, as would be expected, but also on meal amount. Moreover, an association has been found between relative error and fiber and age.

The equation of the final linear model including only the predictors selected by the stepwise variable selection is as follows:

29

$$\widetilde{CHO}_{rel} = +3.39 - 0.22\,CHO - 0.81\,fiber + 0.20\,age + 4.32\,meal_{lunch}$$
$$+ 4.37\,meal_{dinner} - 8.61\,meal_{snack}$$

(S2)

I.C Extended model

The linear model obtained in Section I.B is extended by introducing interaction and quadratic terms. First, the full extended model is fitted. The coefficient of determination of the full extended model R^2 is equal to 0.276, while the adjusted R^2 is equal to 0.175.

Then, the most important predictors adopted to explain the error are selected with stepwise variable selection (results included in Table S3). A quadratic term of fiber is added at step 4, while an interaction term between meal CHO and type are added at step 5.

The equation of the final extended model including only the predictors selected by stepwise variable selection is:

$$\widetilde{CHO}_{rel} = -0.71 - 0.09\,CHO + 10.20\,meal_{lunch} + 24.34\,meal_{dinner} - 8.73\,meal_{snack}$$
$$- 2.42\,fiber + 0.05\,fiber^2 - 0.07\,CHO{:}meal_{lunch}$$
$$- 0.25\,CHO{:}meal_{dinner} + 0.02\,CHO{:}meal_{snack} + 0.25\,age$$

(S3)

30

Table S2: The table shows the mean (second column), standard deviation (third column), confidence interval at 95% (fourth column), minimum (fifth column) and maximum (last column) values of the relative CHO counting error for different levels of meal amount (expressed in grams).

	Mean (%)	SD (%)	95% CI (%)	Min (%)	Max (%)
$CHO \leq 20g$	-1.28	43.64	[-10.97 8.40]	-100	125.56
$20g < CHO \leq 40g$	2.44	34.98	[-2.99 7.88]	-88.51	141.29
$40g < CHO \leq 60g$	-1.10	23.42	[-4.82 2.63]	-64.34	126.98
$60g < CHO \leq 80g$	-9.67	21.88	[-13.46 -5.88]	-69.91	67.63
$80g < CHO \leq 100g$	-11.38	23.89	[-16.55 -6.21]	-70.53	62.36
$CHO > 100g$	-21.32	23.63	[-26.20 -16.43]	-82.04	35.44

Table S3: Results of the stepwise procedure adopted both for the linear model and for the model with interactions and quadratic terms of the relative CHO counting error, the variable added or removed at each step (first column), the value of the F-statistic (second column), the corresponding p-value (third column) and the value of the R^2 parameter (fourth column) for the current model are reported.

	Linear model		
	F-statistic	**p-value**	**R^2**
1. Add **CHO**	44.77	<0.00001	0.062
2. Add **meal**	5.29	p=0.0013	0.083
3. Add **fiber**	6.26	p=0.013	0.091
4. Add **age**	4.16	p=0.042	0.097
	Extended model		
	F-statistic	**p-value**	**R^2**
1. Add **CHO**	44.77	p<0.00001	0.062
2. Add **meal**	5.28	p=0.001	0.082
3. Add **fiber**	6.26	p=0.012	0.091
4. Add **$fiber^2$**	9.10	p=0.003	0.103
5. Add **CHO:meal**	2.99	p=0.030	0.115
6. Add **age**	6.06	p=0.014	0.123