UNIVERSITY OF COPENHAGEN

Plasma metabolites and lipids predict insulin sensitivity improvement in obese, nondiabetic individuals after a 2-phase dietary intervention

Meyer, Antonin; Montastier, Emilie; Hager, Jörg; Saris, Wim H M; Astrup, Arne; Viguerie, Nathalie; Valsesia, Armand

Plasma metabolites and lipids predict insulin sensitivity improvement in obese, nondiabetic individuals after a 2-phase dietary intervention

Antonin Meyer,[1] Emilie Montastier,[2,3] Jörg Hager,[1] Wim HM Saris,[4] Arne Astrup,[5] Nathalie Viguerie,[2] and Armand Valsesia[1]

[1]Nestlé Institute of Health Sciences, Lausanne, Switzerland; [2]Institut National de la Santé et de la Recherche Médicale (INSERM) Unité 1048, Institute of Metabolic and Cardiovascular Diseases, University of Toulouse, Toulouse, France; [3]Department of Clinical Biochemistry and Nutrition, Toulouse University Hospital, Toulouse, France; [4]Department of Human Biology, School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre, Maastricht, Netherlands; and [5]Department of Nutrition, Exercise, and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark

ABSTRACT

Background: Weight loss in obese individuals aims to reduce the risk of type 2 diabetes by improving glycemic control. Yet, significant intersubject variability is observed, and the outcomes remain poorly predictable.

Objective: The aim of the study was to predict whether an individual will show improvements in insulin sensitivity above or below the median population change at 6 mo after a low-calorie-diet (LCD) intervention.

Design: With the use of plasma lipidomics and metabolomics for 433 subjects from the Diet, Obesity, and Genes (DiOGenes) Study, we attempted to predict good or poor Matsuda index improvements 6 mo after an 8-wk LCD intervention (800 kcal/d). Three independent analysis groups were defined: "training" ($n = 119$) for model construction, "testing" ($n = 162$) for model comparison, and "validation" ($n = 152$) to validate the final model.

Results: Initial modeling with baseline clinical variables (body mass index, Matsuda index, total lipid concentrations, sex, age) showed limited performance [area under the curve (AUC) on the "testing dataset" = 0.69; 95% CI: 0.61, 0.77]. Significantly better performance was achieved with an omics model based on 27 variables (AUC = 0.77; 95% CI: 0.70, 0.85; $P = 0.0297$). This model could be greatly simplified while keeping the same performance. The simplified model relied on baseline Matsuda index, proline, and phosphatidylcholine 0-34:1. It successfully replicated on the validation set (AUC = 0.75; 95% CI: 0.67, 0.83) with the following characteristics: specificity = 0.73, sensitivity = 0.68, negative predictive value = 0.60, and positive predictive value = 0.80. Marginally lower performance was obtained when replacing the Matsuda index with homeostasis model assessment of insulin resistance (AUC = 0.72; 95% CI: 0.64, 0.80; $P = 0.08$).

Conclusions: Our study proposes a model to predict insulin sensitivity improvements, 6 mo after LCD completion in a large population of overweight or obese nondiabetic subjects. It relies on baseline information from 3 variables, accessible from blood samples. This model may help clinicians assessing the large variability in dietary interventions and predict outcomes before an intervention. This trial was registered at www.clinicaltrials.gov as NCT00390637. *Am J Clin Nutr* 2018;108:13–23.

INTRODUCTION

Obesity is characterized by an excess of fat mass that affects adipocyte metabolism and is associated with comorbidities such as cardiovascular diseases, insulin resistance, type 2 diabetes (T2D), and cancer (1). Numerous studies have reported the causal link between obesity and T2D (2, 3). Dietary interventions aim to reduce fat mass, restore normal adipose tissue (AT) function, and improve dysfunctions linked to metabolic syndrome (4). Yet, high variability is observed in the capacity to lose and maintain weight (5). Within subjects who achieved weight loss >8% of initial body weight, only half had glycemic control improvement (defined as improvements in insulin sensitivity) (6). This stresses the need for clinical models to predict glycemic

13

improvement. Although the literature is extensive with regard to factors associated with the development of insulin resistance and T2D (7, 8), it remains sparse with regard to predictors of glycemic outcomes after weight loss. We previously proposed a model based on gene expression changes during a low-calorie diet (LCD) (9). This model achieved very good performance at predicting glycemic outcomes 6 mo after LCD completion (AUC = 0.80; 95% CI: 0.69, 0.92). However, this model may be difficult to implement in a routine clinical setting, because it requires intrusive subcutaneous AT biopsy samples taken at baseline and at LCD completion. Ideally, biomarkers measured from blood samples would be preferred. Both metabolomics and lipidomics data can be acquired from blood samples (plasma, serum) without the need for intrusive biopsies and, as such, hold promise as predictive markers. Several studies investigated the correlation between plasma metabolites and weight loss after bariatric surgery (10, 11) or dietary intervention (12). Yet, these studies only reported the association between individual metabolites and clinical outcomes (weight loss, glycemic control improvements) and did not construct predictive models. Therefore, it still remains unknown whether these plasma markers would enable the prediction of clinical outcomes with good performance.

In this report, we attempted to classify overweight or obese, nondiabetic subjects from the Diet, Obesity, and Genes (DiOGenes) Study into good or lesser insulin sensitivity improvers (as measured with the Matsuda index), 6 mo after an 8-wk LCD. We tested ∼170 models with the use of methodologies ranging from statistical to machine learning, variables from 3 types of data (clinical, plasma metabolomics, and lipidomics), and 3 different time points (baseline, at LCD completion, and fold-changes during LCD).

METHODS

Ethics

The study was performed according to the latest version of the Declaration of Helsinki and was approved by local ethics committees with informed consent obtained by all of the participants.

Clinical study design

The DiOGenes study is a multicenter (8 European countries) randomized, controlled, dietary intervention (clinicaltrials.gov; NCT00390637). The study has been described previously (13, 14) and a summary is shown in **Figure 1**A. Briefly, 938 overweight or obese nondiabetic adults underwent an 8-wk LCD. The LCD provided 800 kcal/d with a meal-replacement product (Modifast; Nutrition et Santé). During the LCD, 157 subjects (17%) withdrew from the study. This drop-out rate was extensively studied in Papadaki et al. (15). This rate is comparable (even slightly lower) than in other weight-loss interventions (16–19). In total, 781 subjects completed the LCD intervention, with only 8 who did not achieve 8% weight loss. The subjects who achieved >8% weight loss (n = 773) were randomly assigned to 1 of 5 weight-maintenance diets (WMDs) for 26 wk. The diets corresponded to a 2 × 2 factorial design: with high or low content in protein or glycemic index or a control diet according to local food customs (20). Of those 773 participants, 548 (71%)

completed this 6-mo weight-maintenance intervention. This drop-out rate (29%) has been previously discussed (5) and is in line with reports from other studies (21, 22).

Improvement in insulin sensitivity was expressed as the percentage of increase of the Matsuda index between clinical intervention day (CID) 1 (baseline) and CID3 (end of WMD). The responder group was defined as subjects who show an improvement in insulin sensitivity, as determined by an increase in the Matsuda index greater than the population median change [median = 40.36%, as defined by using the DiOGenes completers with 75-g oral-glucose-tolerance test (OGTT) data; n = 433]. The nonresponder class was composed of the remaining individuals.

Subjects were classified as having impaired fasting glucose if their fasting glucose concentrations were between 5.6 and 6.9 mmol/L as per American Diabetes Association recommendations. Similarly, subjects were classified as having impaired glucose tolerance if their 2-h OGTT glucose concentrations were between 7.8 and 11 mmol/L.

Definition of analyses data sets

Among the 433 DiOGenes completers with available OGTT data, 281 subjects had complete data (i.e., no missing values) for both clinical and all plasma "omics" variables (see Figure 1B). These subjects were used for discovery analyses. The remaining subjects (n = 152) were kept for validation of final models. In the discovery sample, statistical models were trained with subjects from the DiOGenes leading centers (Denmark and Netherlands; n = 119). An independent testing set was built with the individuals from the remaining 6 centers (n = 162) and was used to identify the best-trained model.

Blood sampling

Blood samples were taken after an overnight fasting period at baseline and upon completion of the LCD and WMD interventions.

Clinical variables

The following clinical variables were included in the study: age, sex, BMI, and fasting glucose and insulin concentrations. The HOMA-IR was calculated as fasting glucose (mM) × fasting insulin (μU/ mL)/22.5. The Matsuda index, an established measure of insulin sensitivity (23), was derived from OGTTs with measures at $t = 0, 30, 60, 90,$ and 120 min. Total lipid concentrations were measured from fasting blood samples and included total triglycerides, total cholesterol, HDL cholesterol. LDL-cholesterol concentrations were derived by using the Friedewald formula.

Lipidomics and metabolomics analyses

Liquid chromatography–mass spectrometry data generation was described previously (6). Metabolomics data generation (^{1}H nuclear magnetic resonance) was also described previously (24). In total, 125 lipids and 18 metabolites were quantified (**Supplemental Table 1**).
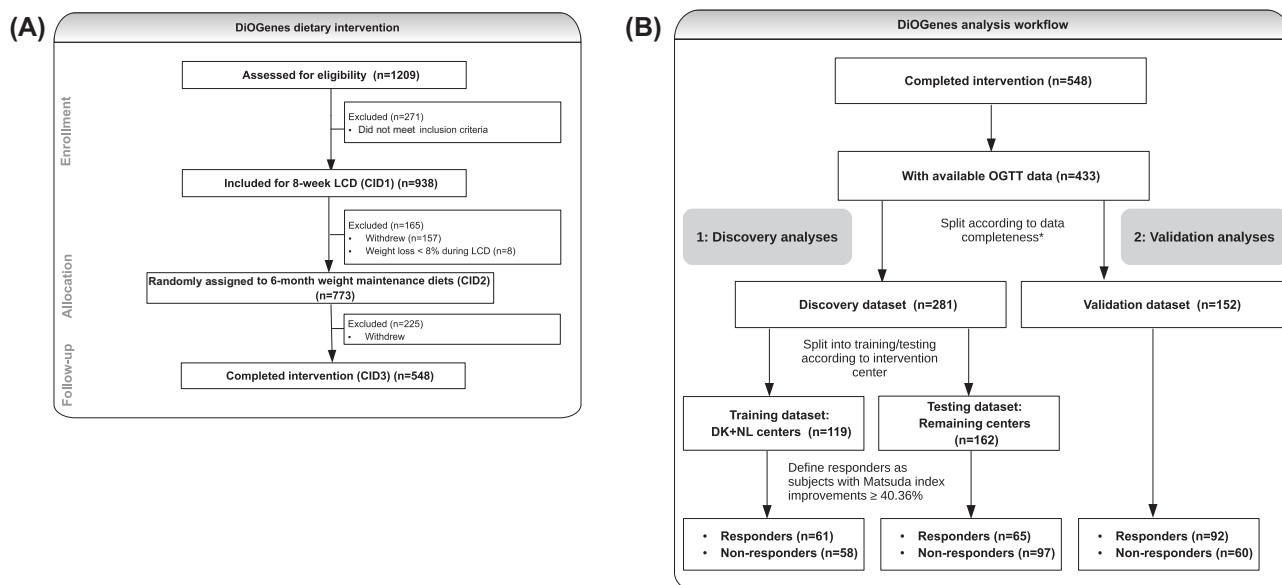
FIGURE 1 Study and analyses workflow. (A) DiOGenes dietary intervention; (B) analysis workflow and definition of the analysis data sets. The 2 leading centers (Netherlands and Denmark) provided most of the food to participants during the weight-maintenance intervention (following recommendations from dietitians and in accordance with the participant's randomly assigned diet). This enabled a better monitoring of patients' compliance during the WMD intervention. The definition of glycemic responders was based on Matsuda index improvements $\geq$40.36%, as estimated for all DiOGenes completers ($n = 433$ subjects with OGTT data). *The discovery data set was composed of subjects with complete data for all clinical variables (BMI, Matsuda index, total lipid concentrations from blood biochemistry, age, sex, fasting glucose and insulin concentrations, HOMA-IR) and all omics variables (125 lipids from liquid chromatography–mass spectrometry and 18 metabolites from nuclear magnetic resonance), and the validation data set was composed of subjects not included in the discovery analyses and who had complete data for the Matsuda index and plasma concentrations of proline and PC O-34:1. CID, clinical intervention day; DiOGenes, Diet, Obesity, and Genes; DK, Denmark; LCD, low-caloric diet; NL, Netherlands; OGTT, oral-glucose-tolerance test; PC O-31:1, phosphatidylcholine 0-34:1; WMD, weight-maintenance diet.

Statistical analysis

Overall model strategy

A combination of different sets of variables (clinical variables alone or in combination with omics data) was tested with different types of statistical models. This resulted in 169 different models (13 data sets × 13 types of models).

Sets of predictors (data sets)

Thirteen different groups of variables (data sets) were tested for model construction (Supplemental Table 1). The 13 data sets used either the values at baseline, LCD termination, or the fold-change during LCD.

Types of models

Thirteen different models from 4 distinct groups were tested, as follows:

- Linear classification models: logistic regression, linear discriminant analysis, and elastic nets
- Nonlinear models: neural network, support vector machine (SVM), and *k*-nearest neighbors
- Tree-based models: random forest; and 3 boosting approaches: adaptive boosting algorithm (ADA), stochastic gradient boosting [with Gradient Boosting Machine (GBM)], and C5.0
- Bayesian approaches: Bayesian generalized linear model, naive Bayes classifier, and Bayesian additive regression trees

Model training and evaluation

Model training was performed with the *caret* framework (25). Data preprocessing included variable standardization and removal of variables with near-zero variance. Model training was performed by using 10-fold cross-validation repeated 5 times. Models were optimized over a grid of variables (**Supplemental Table 2**). For ADA boosting, training was performed by using a single 10-fold cross-validation. The superiority of a single model was tested by using a 1-sided Delong's test (26). Comparison between 2 families of a model was tested with a Wilcoxon's Signed Rank test. All of the analyses were performed with the R statistical language (version 3.3.1; www.r-project.org), with all of the statistical models being available through the *caret* framework (25).

RESULTS

Overall clinical characteristics of the subjects

Clinical characteristics of the subjects are shown in **Table 1**. At baseline, subjects had a mean $\pm$ SD age of 42 $\pm$ 6 y, with a BMI (kg/m^2) of 35 $\pm$ 5, a Matsuda index of 4.99 $\pm$ 2.82, and an HOMA-IR of 3.05 $\pm$ 1.81.

After the LCD intervention, subjects had lost a mean $\pm$ SD of -11.30 ± 3.36 kg. At study termination, these subjects had a mean $\pm$ SD BMI of 31.36 $\pm$ 4.46, corresponding to a change of -3.65 ± 2.39. The Matsuda index also improved with the LCD ($+2.22 \pm 3.45$) and the WMD ($+1.57 \pm 3.08$) interventions.

Clinical characteristics of the 2 subject classes are also shown in Table 1. As expected, upon classification into "responders"

TABLE 1
Clinical characteristics of the subjects (discovery sample)[1]

Time point and variable	All subjects ($n = 281$)	Responders ($n = 126$)	Nonresponders ($n = 155$)	P
Baseline (CID1)				
Age, y	42.18 ± 6.33	42.82 ± 6.37	41.66 ± 6.28	0.1299
Female sex, %	65	63	67	0.4529
BMI, kg/m^2	34.98 ± 4.99	34.63 ± 4.54	35.27 ± 5.32	0.2788
Weight, kg	102.63 ± 18.18	103.05 ± 17.70	102.28 ± 18.61	0.722
Matsuda index	4.99 ± 2.83	3.83 ± 1.75	5.93 ± 3.18	2.26×10^{-11}
HOMA-IR	3.05 ± 1.81	3.73 ± 2.05	2.48 ± 1.34	1.95×10^{-8}
Fasting insulin, μU/mL	11.46 ± 6.49	13.84 ± 7.40	9.53 ± 4.87	6.15×10^{-8}
Fasting glucose, mmol/L	5.07 ± 0.63	5.17 ± 0.61	4.98 ± 0.63	0.0115
Impaired fasting glucose, %	14.23	19.84	9.67	0.0168
Impaired glucose tolerance, %	24.73	30.64	20	0.0504
End of LCD (CID2)				
BMI, kg/m^2	31.14 ± 4.49	30.74 ± 4.10	31.46 ± 4.76	0.1699
Weight, kg	91.33 ± 16.21	91.44 ± 15.75	91.23 ± 16.61	0.915
Matsuda index	7.21 ± 3.66	7.39 ± 3.94	7.06 ± 3.41	0.4639
HOMA-IR	2.04 ± 1.92	1.91 ± 1.21	2.15 ± 2.34	0.255
Fasting insulin, μU/mL	8.03 ± 6.37	7.70 ± 4.78	8.31 ± 7.42	0.4017
Fasting glucose, mmol/L	4.83 ± 0.49	4.78 ± 0.50	4.86 ± 0.49	0.2089
Impaired fasting glucose, %	5.34	4.76	5.80	0.7935
Impaired glucose tolerance, %	27.47	33.05	23.02	0.0763
End of WMD (CID3)				
BMI, kg/m^2	31.36 ± 4.46	30.72 ± 4.09	31.87 ± 4.69	0.0298
Weight, kg	91.99 ± 16.47	91.35 ± 15.96	92.50 ± 16.91	0.5615
Matsuda index	6.56 ± 3.36	7.74 ± 3.55	5.61 ± 2.87	1.21×10^{-7}
HOMA-IR	2.47 ± 2.29	1.89 ± 1.22	2.94 ± 2.80	4.16×10^{-5}
Fasting insulin, μU/mL	9.50 ± 8.42	7.57 ± 5.45	11.07 ± 9.97	0.0002
Fasting glucose, mmol/L	4.99 ± 0.51	4.88 ± 0.45	5.07 ± 0.55	0.0012
Impaired fasting glucose, %	8.89	5.55	11.61	0.0927
Impaired glucose tolerance, %	10.10	9.60	10.52	0.8436
Changes during LCD				
Change in BMI, kg/m^2	-3.85 ± 1.06	-3.90 ± 1.01	-3.81 ± 1.10	0.4787
Change in BMI, %	-10.98 ± 2.47	-11.24 ± 2.49	-10.78 ± 2.43	0.1194
Change in fasting glucose, mmol/L	-0.24 ± 0.53	-0.39 ± 0.46	-0.12 ± 0.56	1.92×10^{-5}
Change in fasting glucose, %	-3.77 ± 13.67	-7.02 ± 8.26	-1.14 ± 16.40	1.28×10^{-4}
Change in HOMA-IR	-1.00 ± 2.04	-1.82 ± 1.79	-0.32 ± 1.98	1.47×10^{-10}
Change in HOMA-IR, %	-24.79 ± 52.19	-42.92 ± 31.64	-9.87 ± 60.52	1.58×10^{-8}
Change in fasting insulin, μU/mL	-3.43 ± 6.64	-6.14 ± 6.18	-1.22 ± 6.18	1.74×10^{-10}
Change in fasting insulin, %	-22.59 ± 48.74	-38.90 ± 31.05	-9.33 ± 56.06	5.80×10^{-8}
Change in Matsuda index	2.22 ± 3.45	3.56 ± 3.49	1.13 ± 3.01	2.96×10^{-9}
Change in Matsuda index, %	67.65 ± 111.25	113.23 ± 142.76	30.61 ± 53.65	6.33×10^{-9}
Change in weight, kg	-11.30 ± 3.36	-11.61 ± 3.38	-11.05 ± 3.33	0.1605
Change in weight, %	-10.98 ± 2.47	-11.24 ± 2.49	-10.78 ± 2.43	0.1194
Changes during WMD				
Change in BMI, kg/m^2	-3.65 ± 2.39	-3.96 ± 2.24	-3.40 ± 2.47	0.0486
Change in BMI, %	-10.20 ± 6.01	-11.23 ± 5.79	-9.38 ± 6.08	0.0101
Change in fasting glucose, mmol/L	-0.08 ± 0.58	-0.29 ± 0.50	0.09 ± 0.58	7.97×10^{-9}
Change in fasting glucose, %	-0.51 ± 14.02	-5.00 ± 8.84	3.14 ± 16.25	2.09×10^{-7}
Change in HOMA-IR	-0.57 ± 2.36	-1.83 ± 1.54	0.47 ± 2.41	4.32×10^{-19}
Change in HOMA-IR, %	-6.90 ± 85.07	-46.90 ± 18.20	26.05 ± 102.69	5.38×10^{-15}
Change in fasting insulin, μU/mL	-1.96 ± 7.98	-6.27 ± 5.13	1.54 ± 8.18	2.08×10^{-19}
Change in fasting insulin, %	-8.04 ± 68.07	-44.09 ± 17.94	21.26 ± 78.97	7.63×10^{-19}
Change in Matsuda index	1.57 ± 3.08	3.91 ± 2.42	-0.33 ± 2.10	2.73×10^{-38}
Change in Matsuda index, %	49.64 ± 75.70	111.28 ± 70.35	-0.47 ± 27.68	7.26×10^{-37}
Change in weight, kg	-10.66 ± 6.97	-11.75 ± 6.73	-9.78 ± 7.05	0.0181
Change in weight, %	-10.20 ± 6.01	-11.23 ± 5.79	-9.38 ± 6.08	0.0101

[1] Values are means $\pm$ SDs unless otherwise indicated. Descriptive statistics are shown for the 281 subjects used in the discovery analyses, including the 126 responders and the 155 nonresponders. Clinical characteristics of these subjects are representative of the whole DiOGenes cohort. CID1 = baseline, CID2 = end of LCD intervention, and CID3 = end of WMD intervention. P values compare responders and nonresponders. For numerical variables, a 2-sided t test was used; for categorical variables (e.g., impaired glucose tolerance, impaired fasting glucose, sex), a Fisher's exact test was used. CID, clinical intervention day; DiOGenes, Diet, Obesity, and Genes; LCD, low-calorie diet; WMD, weight-maintenance diet.

and "nonresponders," some differences can be observed. Notably, these groups differed in baseline Matsuda index ($P = 2.26 \times 10^{-11}$). This observation was consistent with the use of fasting glucose and insulin concentrations as well as HOMA-IR. Significant differences were also found in the proportion of subjects with impaired fasting glucose (see definition in Methods) between responders and nonresponders (19.84% compared with 9.67%; $P = 0.0168$). There was a tendency for a difference in the proportion of subjects with impaired glucose tolerance (30.64% compared with 20%; $P = 0.0504$). This is in line with the above-mentioned observations using continuous measures of insulin sensitivity.

An investigation of other factors (age and sex) did not show significant differences between responders and nonresponders ($P > 0.05$). Finally, the 2 groups had similar baseline BMI ($P = 0.28$) and had no significant difference in the percentage of weight loss ($P = 0.12$) but displayed significant difference in weight maintenance ($P = 0.018$).

The strong differences at baseline in indexes of insulin sensitivity between responders and nonresponders suggested some potential to predict glycemic outcomes. Indeed, the baseline Matsuda index and the percentage of Matsuda changes between CID1 and CID3 were significantly correlated (Pearson's $r = -0.42$; 95% CI: -0.51, -0.32; **Supplemental Figure 1**). However, this correlation was not perfect, indicating that additional factors could affect the final glycemic outcome. In addition, although the proportion of subjects who could be considered insulin-resistant (baseline Matsuda index ≤ 2.5) was significantly higher in the responder group than in the nonresponder group (37.6% compared with 12.5%; Fisher's exact test, $P = 1.3 \times 10^{-9}$), this was not sufficient for prediction purposes. Indeed, a stratification model based on baseline insulin resistance status, sex, and age yielded a classification performance comparable to a random prediction (AUC on the testing set = 0.59; 95% CI: 0.49, 0.68). The use of other thresholds (Matsuda index ≤ 3) or other variables (HOMA-IR) led to similar performance. The addition of baseline BMI in these models also led to similar performance. This suggests that additional factors would need to be included in the model to achieve a better prediction.

Initial modeling of insulin sensitivity improvements

On the basis of the observations from Table 1, we attempted a simple model based on baseline clinical variables (BMI, Matsuda index, sex, age, and total lipid concentrations from biochemistry). Following best practices from the field, the data were split into a training set ($n = 119$, for model construction) and a testing set ($n = 162$, for evaluating performance). The 2 data sets were comparable: subjects had similar weight and insulin sensitivity during the LCD and after weight maintenance. We then used a logistic regression to classify responders and nonresponders. This preliminary model provided modest performance (AUC on testing set = 0.69; 95% CI: 0.61, 0.77). The same model substituting the Matsuda index with HOMA-IR led to a slightly lower performance (AUC = 0.66; 95% CI: 0.57, 0.74). The replacement of the HOMA-IR with fasting glucose and insulin concentrations resulted in the same performance.

Performance of 169 modeling approaches

To identify models with potentially better performance, we explored the combination from 13 different sets of predictors and 13 different modeling approaches, ranging from simple statistical approaches to more complex, machine-learning methods. Figure 2A summarizes the performance from all models. Approximately half of the models (89 of 169) were not better than a random predictor (with AUCs not significantly different from 0.50). In contrast, some models showed very good performance, with AUCs >0.75 and 95% CIs strictly greater than 0.50.

We investigated whether data from a specific time point (baseline, at LCD termination, or fold-change during LCD) would offer a better classification performance. Figure 2B shows the performance grouped by data set time point. Models based on data collected at LCD completion (CID2) had significantly lower performance than models built with baseline values (at CID1) or with fold-changes (Wilcoxon's Signed Rank test, $P < 5\%$). Approximately 70% of the models based on CID1 data had a performance better than a random classifier. This proportion was ~80% for data sets based on fold-changes, and it dropped to only 11% for data sets based on CID2 variables. Therefore, prediction of glycemic outcomes is best achieved by using either baseline information or the trajectories during weight loss.

Next, we searched whether specific statistical framework would lead to better performance than others. Figure 2C presents the performance grouped by type of model. It shows that elastic net models generally perform best (median AUC = 0.70; IQR = 0.16). The second best class of model pertains to boosting algorithms represented with the stochastic gradient boosting (GBM boosting; median AUC = 0.65; IQR = 0.15) and ADA boosting (median AUC = 0.64; IQR = 0.14). Elastic nets outperformed those 2 types of models ($P < 0.006$). Logistic regressions showed large variability in their performance and were outperformed by elastic nets ($P = 0.0017$), reflecting their limitations in the presence of numerous variables. Specific models, such as k-nearest neighbors, SVM, Naïve Bayes, and C5.0 boosting methods, had poor performance, likely due to their need for very large training data sets, which is not achievable in clinical studies.

Description of the top omics model

Our best omics model showed the following classification performance: AUC = 0.77 (95% CI: 0.70, 0.85). This model, based on an elastic net and baseline variables, retained 27 predictors with a non-zero coefficient (**Supplemental Table 3**); other variables were not considered informative by the model. The overall performance of the omics model is shown in Figure 3A. It outperformed the clinical model (Delong's $P = 0.0297$) and strong differences were observed. For example, by using a specificity cutoff at 80%, the omics model reached 65% sensitivity whereas the clinical model only reached 37%. By defining the optimal receiver operating characteristic (ROC) threshold with the use of the Youden index, the omics model obtained the following performance: 87% specificity, 61% sensitivity, and 77% and 75% negative and positive predictive values (NPVs and PPVs), respectively. By contrast, the clinical model only obtained 45% specificity, 86% sensitivity, and 83% and 51% NPVs and PPVs, respectively.
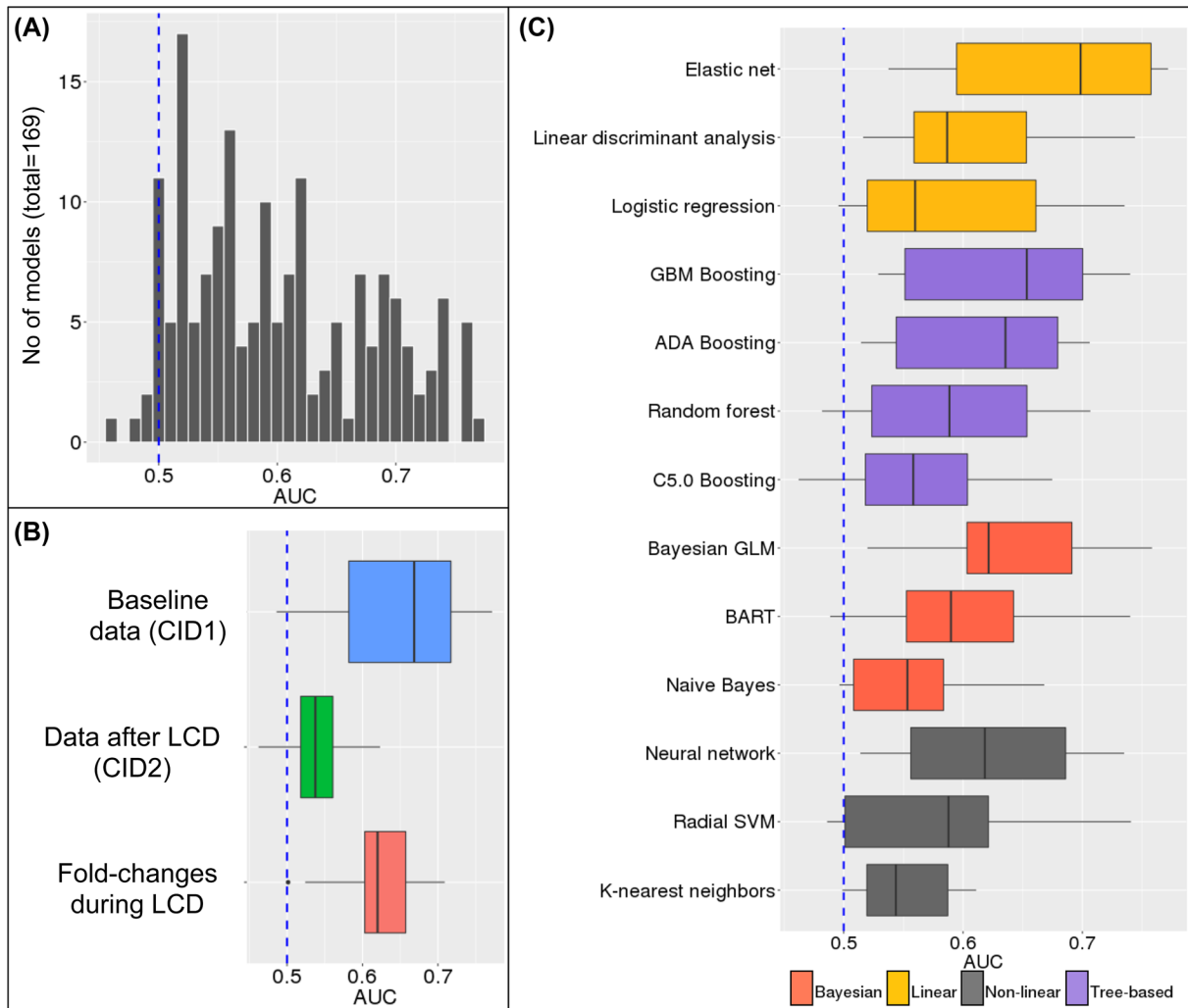
FIGURE 2 Performance of the 169 models. (A) Histogram of all 169 values of ROC AUCs. The dotted blue line indicates an AUC = 0.50, the performance of a random classifier. (B) Performances grouped by the time point of the variables involved in the models. (C) Performances grouped by the statistical approach used to construct the classification models. ADA, adaptive boosting algorithm; BART, Bayesian additive regression trees; CID, clinical intervention day; GBM, Gradiant Boosting Machine; GLM, generalized linear model; LCD, low-calorie diet; ROC, receiver operating characteristic; SVM, Support Vector Machine.

Next, we analyzed the importance of the variables in the omics model (Figure 3B). The Matsuda index was found to be the most informative variable for the prediction of its own evolution after weight maintenance, as expected. Several omics variables were selected by the elastic net and improved the model performance. The second most important variable, phosphatidylcholine (PC O-34:1) had a relative importance >60% (this variable was as informative as 60% of the information carried by the Matsuda index). Proline, the third most important variable, had a relative importance >50%. BMI was incorporated in the model, but only ranked as number 15. The model did not select sex and age, suggesting that these variables did not add significant information.
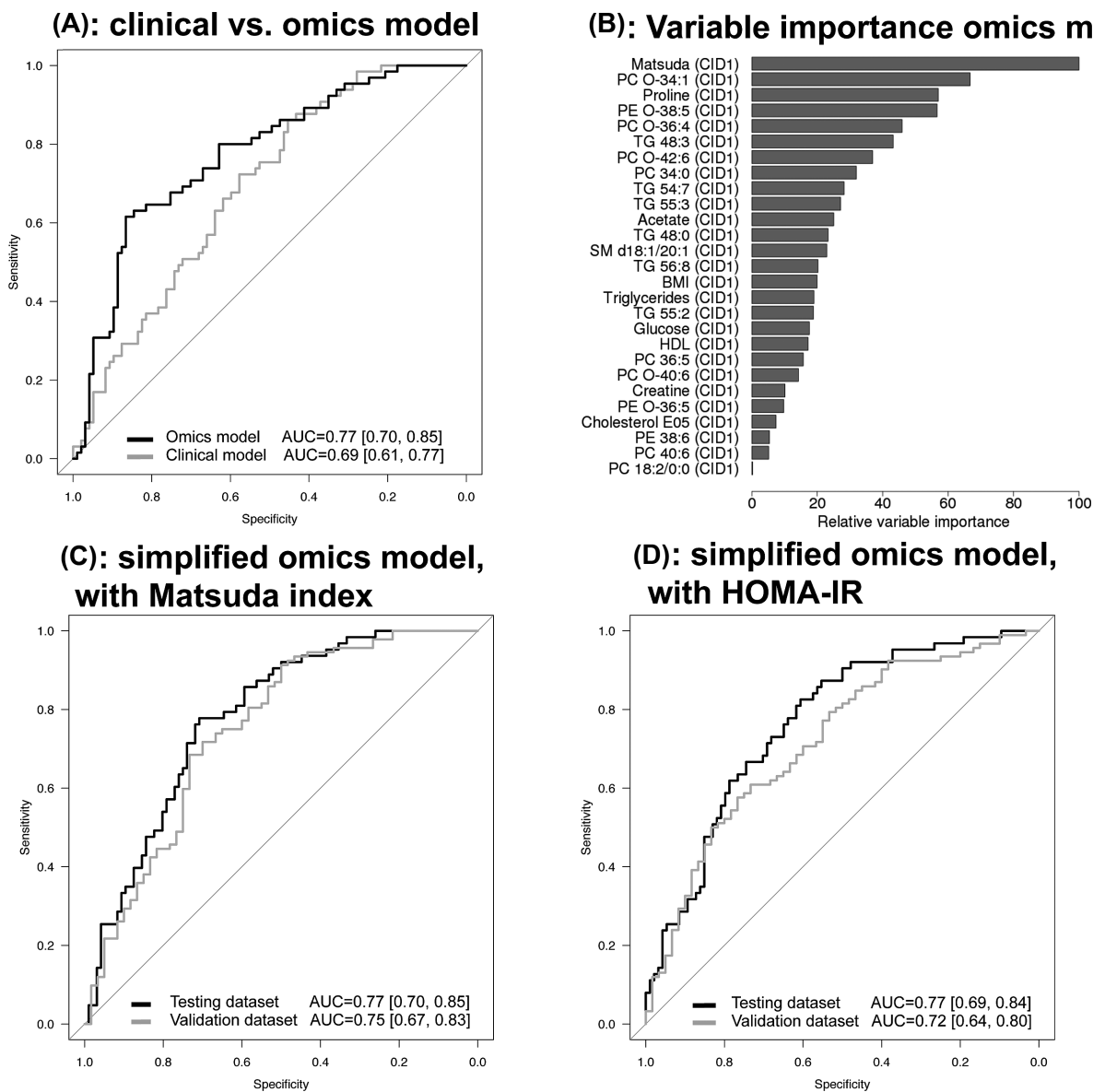
Effect of WMDs

We tested whether considering the WMD would improve the prediction. We modeled the WMD with the use of 3 different approaches. We first represented the diet as a categorical variable (representing each of the 5 possible diets). Second, we

expressed the diet into 2 categorical variables, one representing the protein content (high or low) and the other corresponding to the glycemic index (high or low). Finally, we used quantitative information about protein and carbohydrate macronutrient intakes, as derived from 3-d food diaries (collected within the first 2–4 wk of the WMD phase). In all approaches, the performance of previous models was not significantly improved ($P > 0.36$), suggesting that diet had little influence on the prediction of insulin sensitivity improvements when compared with the information already included in the omics model.

Testing additional model optimization

We tested whether adding the percentage of weight loss during the LCD intervention would help refine the performance of our omics model. However, the resulting AUC was identical: 0.76 (95% CI: 0.68, 0.83). Similarly, a model that used the percentage of Matsuda index improvement during the LCD did not improve the performance (AUC = 0.75; 95% CI: 0.67,

(A): clinical vs. omics model

(B): Variable importance omics model

(C): simplified omics model, with Matsuda index

(D): simplified omics model, with HOMA-IR

FIGURE 3 Performance of classification models. (A) ROC curves showing classification performance as obtained on the testing data set. The omics model outperforms the clinical model (Delong's $P = 0.0297$). The diagonal line indicates the performance from a random predictor. (B) Plot of relative variable importance from the top elastic net. Importance is relative to the top predictor (the one with the highest absolute coefficient). (C, D) ROC curves showing the performance on the testing set and validation set for the simplified omics models with the use of either the Matsuda index (C) or HOMA-IR (D), together with baseline proline and PC O-34:1 concentrations. CID, clinical intervention day; PC, phosphatidylcholine; PE, phosphatidylethanolamine; ROC, receiver operating characteristic; SM, sphyngomyelin; TG, triglyceride.

0.82). Models that used additional anthropometric variables did not yield better performance: adding waist circumference, hip circumference, or body fat mass all led to AUCs close to 0.63. Adding the waist-to-hip ratio resulted in poor performance (AUC = 0.58; 95% CI: 0.49, 0.68).

We also tested whether the Matsuda index could be replaced in our models by HOMA-IR. The resulting performance was slightly lower (AUC = 0.74; 95% CI: 0.67, 0.82), but this difference was not significant ($P = 0.27$). A model with glucose and insulin fasting variables instead of HOMA-IR led to a similar performance (0.73; 95% CI: 0.66, 0.81). This suggested that,

in the absence of OGTT data, HOMA-IR or fasting concentrations could be a suitable replacement with minimal loss of the model performance.

Replacing the Matsuda index with either the impaired glucose tolerance status or impaired fasting glucose status yielded significantly lesser performance ($P = 1.4 \times 10^{-4}$), with AUCs close to 0.66 (and 95% CIs ranging from ~0.58 to 0.75). This is expected because a model based on a continuous predictor would perform and be more robust than a model that uses a discretized version of this same variable (27).

Model simplification

In an attempt to simplify our model (27 variables), we constructed a new model based only on the top 3 variables (Matsuda index, proline, and PC O-34:1 at baseline). In the testing data set, this simplified model reached a performance (AUC = 0.77; 95% CI: 0.70, 0.85) similar to the full model ($P = 0.43$).

We further challenged the performance with the use of 152 subjects who had not been included in any previous analyses (see Figure 1B; clinical characteristics of this validation data set are shown in **Supplemental Table 4** and are representative of both the whole DiOGenes cohort and of the discovery data set). In this additional validation data set, the performance was confirmed, with an AUC = 0.75 (95% CI: 0.67, 0.83) (see Figure 3C), which shows the robustness of the model. Model coefficients and descriptive statistics for each variable are available in **Supplemental Tables 5** and **6**.

We also tested a simplified model based on the HOMA-IR instead of the Matsuda index. Here, performances were good: AUC = 0.77 (95% CI: 0.69, 0.84) in the testing set and a slightly lower AUC (0.72; 95% CI: 0.64, 0.80) in the validation set (Figure 3D). Compared with the Matsuda index–based model, this HOMA-IR model had marginally lower performance ($P = 0.08$).

Finally, we compared the performance in terms of sensitivity and specificity, as well as PPV and NPV (**Supplemental Table 7**). These metrics were derived from the ROC curve on the validation data set, and the best ROC threshold was obtained by using the Youden index. Both models provided similar and good specificity of >73% and a PPV >77%. However, the model that used the Matsuda index was more sensitive than the HOMA-IR–based model (sensitivity = 68.48% compared with 57.61%). Similar observations were made with the NPV (60.27% compared with 54.12%). The use of a different approach to find the optimal ROC threshold (specificity set at 80%) led to the same conclusions (Supplemental Table 7).

These simplified models can be interpreted as follows: non-responders include more subjects with good insulin sensitivity at baseline (high Matsuda index or low HOMA-IR; Figure 4A, B). Nonresponders also tend to have lower plasma PC O-34:1 concentrations and higher proline concentrations than responders (Figure 4C, D). Each of the individual variables shows some variability and overlap between responders and nonresponders. This shows that prediction cannot easily be achieved by using only 1 variable (as expected). Instead, the combination of 3 variables is required and leads to very good performance. This is expected because if a single variable had shown a clear separation between the groups, then our multivariate approach (combined with its intrinsic regularization and feature selection) would have kept only this variable in the final model.

DISCUSSION

This study explored approaches to predict glycemic outcomes 6 mo after an LCD in obese, nondiabetic subjects. We found significant differences in baseline glycemic control between responders and nonresponders. Specifically, subjects with better baseline insulin sensitivity had less room for improvement after the dietary intervention. However, a model based on baseline insulin-resistant status offered poor performance (AUC = 0.59; 95% CI: 0.48, 0.67). Prediction was improved by combining multiple variables (baseline BMI, Matsuda index, total lipid concentrations, sex, and age). This clinical model achieved better, yet modest, performance with an AUC = 0.69 (95% CI: 0.61, 0.77). Incorporating plasma metabolites and lipids enabled the model to significantly improve its performance (AUC = 0.77; 95% CI: 0.70, 0.85). This omics model corresponded to improvement in both sensitivity and specificity and outperformed the clinical model ($P = 0.0297$). Additional variables, such as the WMD arm, BMI, or glycemic changes during the LCD, did not improve the performance. This final performance was comparable to the one from our recent model based on transcriptomics analyses from AT biopsy sample (9), which reached an AUC = 0.80 (95% CI: 0.69, 0.92). Importantly, our new omics model, based on circulating markers, provides better translational potential because it does not depend on intrusive procedure (AT biopsy samples both at baseline and at LCD termination), while keeping a comparable performance. However, this omics model depends on 27 variables, which might challenge its application in a clinical context. Subsequent analyses greatly simplified this model. In particular, the use of only the 3 top variables (baseline Matsuda index and proline and PC O-34:1 concentrations) provided identical performance than the full omics model. This performance was confirmed on a second validation sample ($n = 152$), with an AUC = 0.75 (95% CI: 0.67, 0.83). Additional metrics reflecting the quality of our predictions can be found in Supplemental Table 7 (e.g., the simplified model has >73% specificity and >79% PPV).

This simplified model offers better translational potential than the full omics model, because it relies only on the Matsuda index and proline and PC O-34:1 concentrations. The Matsuda index can be measured from an OGTT. We found that, in the absence of such data, the HOMA-IR would represent a good compromise (with a model presenting an AUC = 0.72; 95% CI: 0.64, 0.80). Amino acid concentrations are already tested in clinical practice (28) [e.g., for the diagnosis of inborn error of metabolism in newborns (29)]. Lipids are becoming increasingly important for the prediction of clinical outcomes [e.g., for the prediction of cardiovascular events in subjects with T2D (30)]. Current advances in the field make this class of molecule easier to be acquired in a clinical context (e.g., using targeted approaches) (31, 32). Although not all clinics may have access to in-house omics facilities, both proline and PC O-34:1 markers are available on standard panels from external companies. Those commercial assays can be run on standard liquid chromatography–mass spectrometry machines (frequently available in universities or in large obesity clinics with translational research activities). The analysis can also be fully outsourced to the same commercial companies that provide these panels and to numerous other service companies.

The presence of proline and phosphatidylcholine in our model is not completely surprising, although they were not previously known as predictors of glycemic improvements (e.g., improvements in insulin sensitivity index). Proline plasma concentrations are known to be associated with obesity and insulin-resistance markers (BMI, HOMA-IR, and glycated hemoglobin, C-peptide, insulin, and leptin concentrations); specifically, higher proline concentrations are observed in hyperinsulinemic patients (33). Phosphatidylcholines are major components of the cellular membranes and play an important role in regulating lipid,
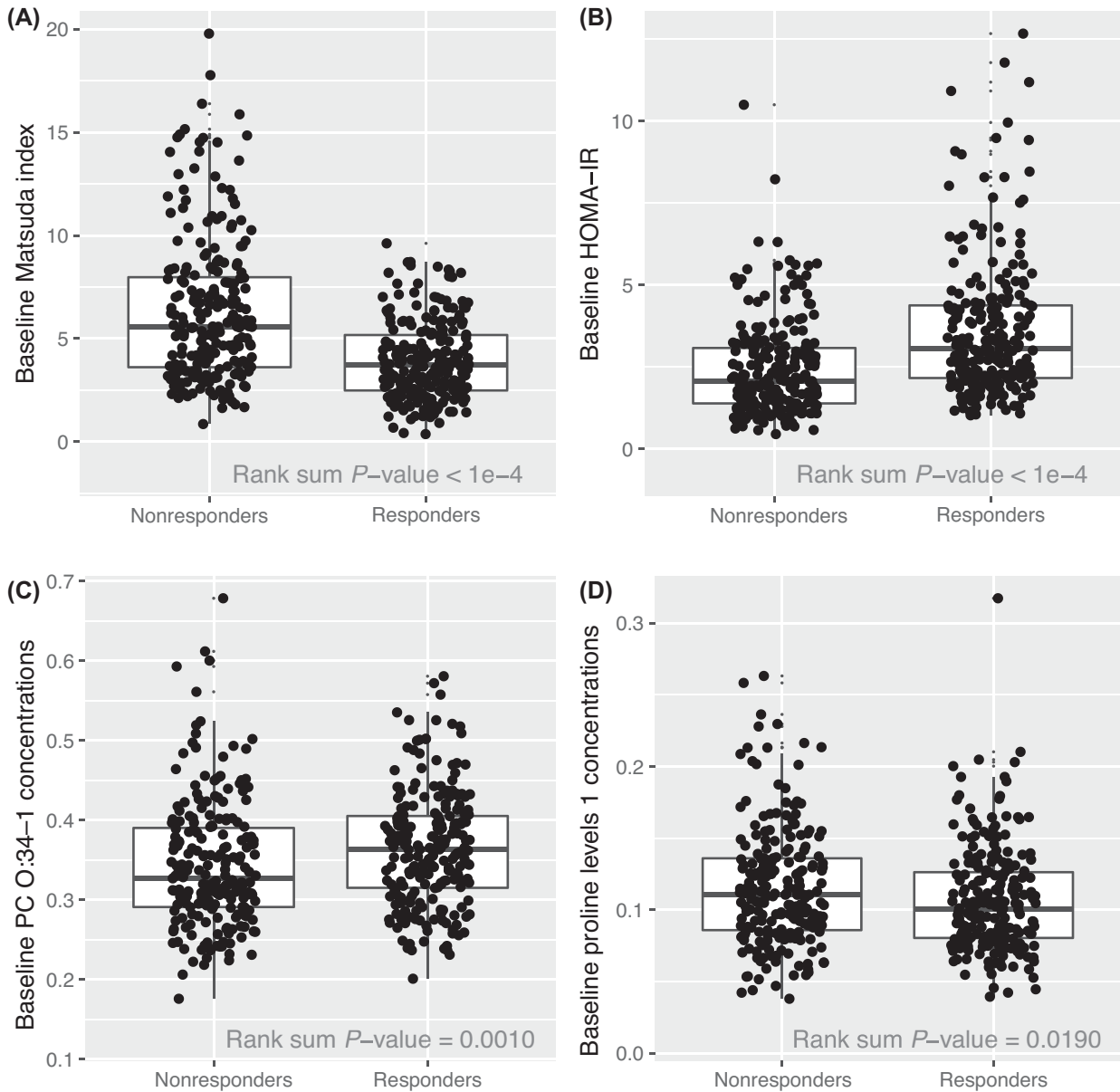
FIGURE 4 Boxplots of baseline values for the top (most predictive) variables from the simplified models, stratified by responders and nonresponders, for baseline Matsuda index (A), HOMA-IR (B), and concentrations of PC O-34:1 (C) and proline (D). PC, phosphatidylcholine.

lipoprotein, and whole-body energy metabolism (34). The content and alteration of PUFA-containing phospholipids in muscle are associated with insulin resistance (35). A recent metabolomics study (36) has shown the interplay between phosphatidylcholines, iron homeostasis, and glucose metabolism. Another study (37) found that changes in phosphatidylcholine concentrations after Roux-en-Y gastric bypass were associated with insulin sensitivity (HOMA-S), and that these changes were independent of weight loss. This is consistent with our results and the characteristics of the 2 patient groups (Figure 4)

Interestingly, branched-chain amino acids (BCAAs; isoleucine, valine, leucine) were not retained in our models. This suggests that, although BCAAs are prognostic markers of insulin resistance (7), these variables are not considered informative

for the prediction of glycemic outcomes. This is consistent with a recent analysis of the Preventing Overweight Using Novel Dietary Strategies Trial (POUNDS LOST) ($n = 774$) and Dietary Intervention Randomized Controlled Trial (DIRECT) ($n = 318$) studies, which that did not identify any significant association between change in BCAA concentrations and HOMA-IR improvements (12). Finally, adiposity (BMI and other measures such as waist circumference and body fat mass) had limited importance in our models. In the full omics model, baseline BMI only ranked as the 15th (out of 27) most-informative variable. Incorporating weight-loss trajectories during the LCD did not significantly improve the performance. Therefore, weight-related variables provide very limited information in predictive models because they fail to capture the metabolic state of the individual

(as opposed to the information carried by metabolites and lipids). This finding is in agreement with our recent lipidomics study (6) and stresses further the importance of lipid metabolism for the prediction of glycemic control improvements. The independence of the model with respect to weight-related variables also raises the question of whether it might be applicable to the prediction of glycemic improvements upon slight caloric-restriction or lifestyle intervention in a population of nonobese subjects. This would need to be formally tested in large cohorts such as the Look Ahead (38) or Diabetes Prevention Program (39). Our model may also be useful in cohorts of patients with more severe insulin resistance (e.g., cohorts of morbidly obese patients), to better study the variability in glycemic outcomes after an intervention (either LCD or surgery). This would also require formal testing and would likely require access to better tools (e.g., euglycemic clamp) to assess the degree of insulin resistance in such patients.

Finally, the recent study by Lean et al. (40) showed the superior efficacy of LCD studies over the current primary care practices (best practices by guidelines) in the management of obese subjects with T2D. Their LCD intervention yielded significant differences in terms of weight loss (with 24% of participants from the LCD group and 0% in the primary care group having >15 kg of weight loss at 1 y). Significant differences were also seen for the percentages of T2D remission (46% compared with 4% in the LCD and primary care groups, respectively). One would expect that LCD interventions will become increasingly important and present in the management of obese or prediabetic patients or those with T2D. Because the products used in LCDs provide similar daily caloric intake (800–900 kcal/d), our present results should apply to any such LCD study. In our study, we observed a drop-out rate of ∼17% during the LCD and we, de facto, could not include those subjects in our analyses due to lack of data. However, this number is in line with other multicenter weight-loss studies (16–19), and it does not challenge the generalization of our findings to other LCD studies that would face similar drop-out rates.

In conclusion, our study identified a model that enables the prediction of improvements in insulin sensitivity in a large population of overweight or obese subjects. This model relies on baseline information from only 3 variables that are accessible from blood samples. In addition, the performance is comparable to a more intrusive model that relies on molecular phenotyping from AT biopsy samples at 2 time points (baseline and after an LCD). To the best of our knowledge, there is no established model to predict glycemic outcomes; therefore, the proposed model defines a first standard that could serve as a benchmark for future models. This model already provides some potential for clinical practice and may help clinicians in understanding the large variability in dietary interventions. Finally, our modeling strategies might guide other biomarker studies that aim to predict clinical outcomes after an intervention.

REFERENCES

1. Haslam DW, James WPT. Obesity. Lancet 2005;366:1197–209.
2. Schienkiewitz A, Schulze MB, Hoffmann K, Kroke A, Boeing H. Body mass index history and risk of type 2 diabetes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)—Potsdam Study. Am J Clin Nutr 2006;84:427–33.
3. Perry IJ, Wannamethee SG, Walker MK, Thomson AG, Whincup PH, Shaper AG. Prospective study of risk factors for development of non-insulin dependent diabetes in middle aged British men. BMJ 1995;310:560.
4. Soare A, Weiss EP, Pozzilli P. Benefits of caloric restriction for cardiometabolic health, including type 2 diabetes mellitus risk. Diabetes Metab Res Rev 2014;30(Suppl 1):41–7.
5. Larsen TM, Dalskov S-M, van Baak M, Jebb SA, Papadaki A, Pfeiffer AFH, Martinez JA, Handjieva-Darlenska T, Kunešová M, Pihlsgård M, et al. Diets with high or low protein content and glycemic index for weight-loss maintenance. N Engl J Med 2010;363:2102–13.
6. Valsesia A, Saris WH, Astrup A, Hager J, Masoodi M. Distinct lipid profiles predict improved glycemic control in obese, nondiabetic patients after a low-caloric diet intervention: the Diet, Obesity and Genes randomized trial. Am J Clin Nutr 2016;104:566–75.
7. Würtz P, Soininen P, Kangas AJ, Rönnemaa T, Lehtimäki T, Kähönen M, Viikari JS, Raitakari OT, Ala-Korpela M. Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults. Diabetes Care 2013;36:648–55.
8. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, Lewis GD, Fox CS, Jacques PF, Fernandez C, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;17:448–53.
9. Armenise C, Lefebvre G, Carayol J, Bonnel S, Bolton J, Di Cara A, Gheldof N, Descombes P, Langin D, Saris W, et al. Transcriptome profiling from adipose tissue during low-caloric diet reveals predictors of weight and glycemic outcomes in obese, non-diabetic subjects. Am J Clin Nutr 2017;106(3):736–46.
10. Tan HC, Khoo CM, Tan MZ-W, Kovalik J-P, Ng ACM, Eng AKH, Lai OF, Ching JH, Tham KW, Pasupathy S. The effects of sleeve gastrectomy and gastric bypass on branched-chain amino acid metabolism 1 year after bariatric surgery. Obes Surg 2016;26:1830–5.
11. Magkos F, Bradley D, Schweitzer GG, Finck BN, Eagon JC, Ilkayeva O, Newgard CB, Klein S. Effect of Roux-en-Y gastric bypass and laparoscopic adjustable gastric banding on branched-chain amino acid metabolism. Diabetes 2013;62:2757–61.
12. Zheng Y, Ceglarek U, Huang T, Li L, Rood J, Ryan DH, Bray GA, Sacks FM, Schwarzfuchs D, Thiery J, et al. Weight-loss diets and 2-y changes in circulating amino acids in 2 randomized intervention trials. Am J Clin Nutr 2016;103:505–11.

13. Larsen TM, Dalskov S-M, van Baak M, Jebb SA, Papadaki A, Pfeiffer AFH, Martinez JA, Handjieva-Darlenska T, Kunešová M, Pihlsgård M, et al. Diets with high or low protein content and glycemic index for weight-loss maintenance. N Engl J Med 2010;363:2102–13.

14. Larsen TM, Dalskov S, van Baak M, Jebb S, Kafatos A, Pfeiffer A, Martinez JA, Handjieva-Darlenska T, Kunesová M, Holst C, et al. The Diet, Obesity and Genes (Diogenes) dietary study in eight European countries—a comprehensive design for long-term intervention. Obes Rev 2010;11:76–91.

15. Papadaki A, Linardakis M, Plada M, Larsen TM, van Baak MA, Lindroos AK, Pfeiffer AFH, Martinez JA, Handjieva-Darlenska T, Kunesova M, et al. A multicentre weight loss study using a low-calorie diet over 8 weeks: regional differences in efficacy across eight European cities. Swiss Med Wkly 2013;143:w13721.

16. Wadden TA, Foster GD, Letizia KA, Stunkard AJ. A multicenter evaluation of a proprietary weight reduction program for the treatment of marked obesity. Arch Intern Med 1992;152:961–6.

17. Hemmingsson E, Johansson K, Eriksson J, Sundström J, Neovius M, Marcus C. Weight loss and dropout during a commercial weight-loss program including a very-low-calorie diet, a low-calorie diet, or restricted normal food: observational cohort study. Am J Clin Nutr 2012;96:953–61.

18. Finley CE, Barlow CE, Greenway FL, Rock CL, Rolls BJ, Blair SN. Retention rates and weight loss in a commercial weight loss program. Int J Obes (Lond) 2006;31:292.

19. Tsai AG, Wadden TA. The evolution of very-low-calorie diets: an update and meta-analysis. Obesity 2006;14:1283–93.

20. Moore CS, Lindroos AK, Kreutzer M, Larsen TM, Astrup A, Van Baak MA, Handjieva-Darlenska T, Hlavaty P, Kafatos A, Kohl A, et al. Dietary strategy to manipulate ad libitum macronutrient intake, and glycaemic index, across eight European countries in the Diogenes Study. Obes Rev 2010;11:67–75.

21. Miller M, Beach V, Sorkin JD, Mangano C, Dobmeier C, Novacic D, Rhyne J, Vogel RA. Comparative effects of three popular diets on lipids, endothelial function, and C-reactive protein during weight maintenance. J Am Diet Assoc 2009;109:713–7.

22. Hession M, Rolland C, Kulkarni U, Wise A, Broom J. Systematic review of randomized controlled trials of low-carbohydrate vs. low-fat/low-calorie diets in the management of obesity and its comorbidities. Obes Rev 2009;10:36–50.

23. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462–70.

24. Stroeve JHM, Saccenti E, Bouwman J, Dane A, Strassburg K, Vervoort J, Hankemeier T, Astrup A, Smilde AK, van Ommen B, et al. Weight loss predictability by plasma metabolic signatures in adults with obesity and morbid obesity of the DiOGenes study. Obesity (Silver Spring) 2016;24:379–88.

25. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28:1–26.

26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.

27. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.

28. Shih VE. Amino acid analysis. In: Blau N, Duran M, Blaskovics ME, Gibson KM, editors. Physician's guide to the laboratory diagnosis of metabolic diseases [Internet]. Berlin, Heidelberg (Germany): Springer Berlin Heidelberg; 2003. p. 11–26. Available from: https://doi.org/10.1007/978-3-642-55878-8_2. Accessed on February 2018.

29. Vernon HJ. Inborn errors of metabolism: advances in diagnosis and therapy. JAMA Pediatr 2015;169:778–82.

30. Alshehry ZH, Mundra PA, Barlow CK, Mellett NA, Wong G, McConville MJ, Simes J, Tonkin AM, Sullivan DR, Barnes EH, et al. Plasma lipidomic profiles improve on traditional risk factors for the prediction of cardiovascular events in type 2 diabetes mellitus. Circulation 2016;134:1637–50.

31. Gonzalez-Covarrubias V, Beekman M, Uh H-W, Dane A, Troost J, Paliukhovich I, Kloet FM, Houwing-Duistermaat J, Vreeken RJ, Hankemeier T, et al. Lipidomics of familial longevity. Aging Cell 2013;12:426–34.

32. Alshehry ZH, Barlow CK, Weir JM, Zhou Y, McConville MJ, Meikle PJ. An efficient single phase method for the extraction of plasma lipids. Metabolites 2015;5:389–403.

33. Yamada C, Kondo M, Kishimoto N, Shibata T, Nagai Y, Imanishi T, Oroguchi T, Ishii N, Nishizaki Y. Association between insulin resistance and plasma amino acid profile in non-diabetic Japanese subjects. J Diabetes Investig 2015;6:408–15.

34. van der Veen JN, Kennelly JP, Wan S, Vance JE, Vance DE, Jacobs RL. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Membr Lipid Ther Drugs Target Biomembr 2017;1859:1558–72.

35. Borkman M, Storlien LH, Pan DA, Jenkins AB, Chisholm DJ, Campbell LV. The relation between insulin sensitivity and the fatty-acid composition of skeletal-muscle phospholipids. N Engl J Med 1993;328:238–44.

36. Stechemesser L, Eder SK, Wagner A, Patsch W, Feldman A, Strasser M, Auer S, Niederseer D, Huber-Schönauer U, Paulweber B, et al. Metabolomic profiling identifies potential pathways involved in the interaction of iron homeostasis with glucose metabolism. Mol Metab 2017;6:38–47.

37. Kayser BD, Lhomme M, Dao MC, Ichou F, Bouillot J-L, Prifti E, Kontush A, Chevallier J-M, Aron-Wisnewsky J, Dugail I, et al. Serum lipidomics reveals early differential effects of gastric bypass compared with banding on phospholipids and sphingolipids independent of differences in weight loss. Int J Obes (Lond) 2017;41:917–25.

38. Look AHEAD Research Group. Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the Look AHEAD Trial. Diabetes Care 2007;30:1374–83.

39. Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346:393–403.

40. Lean ME, Leslie WS, Barnes AC, Brosnahan N, Thom G, McCombie L, Peters C, Zhyzhneuskaya S, Al-Mrabeh A, Hollingsworth KG, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. Lancet [Internet]. [cited 2018 Jan 16]. Available from: http://dx.doi.org/10.1016/S0140-6736(17)33102-1.