

The advanced-DiaRem score improves prediction of diabetes remission 1 year post-Roux-en-Y gastric bypass

Judith Aron-Wisniewsky^{1,2} · Nataliya Sokolovska^{1,2} · Yuejun Liu^{1,2} · Doron S. Comaneshter³ · Shlomo Vinker^{3,4} · Tal Pecht⁵ · Christine Poitou^{1,2} · Jean-Michel Oppert¹ · Jean-Luc Bouillot⁶ · Laurent Genser⁷ · Dror Dicker^{4,8} · Jean-Daniel Zucker^{1,2,9} · Assaf Rudich⁵ · Karine Clément^{1,2}

Received: 2 March 2017 / Accepted: 8 June 2017 / Published online: 21 July 2017
© Springer-Verlag GmbH Germany 2017

Abstract

Aims/hypothesis Not all people with type 2 diabetes who undergo bariatric surgery achieve diabetes remission. Thus it is critical to develop methods for predicting outcomes that are applicable for clinical practice. The DiaRem score is relevant for predicting diabetes remission post-Roux-en-Y gastric bypass (RYGB), but it is not accurate for all individuals across the entire spectrum of scores. We aimed to develop an improved scoring system for predicting diabetes remission following RYGB (the Advanced-DiaRem [Ad-DiaRem]).
Methods We used a retrospective French cohort ($n = 1866$) that included 352 individuals with type 2 diabetes followed for 1 year post-RYGB. We developed the Ad-DiaRem in a test cohort ($n = 213$) and examined its accuracy in independent cohorts from France ($n = 134$) and Israel ($n = 99$).

Results Adding two clinical variables (diabetes duration and number of glucose-lowering agents) to the original DiaRem and modifying the penalties for each category led to improved predictive performance for Ad-DiaRem. Ad-DiaRem displayed improved area under the receiver operating characteristic curve and predictive accuracy compared with DiaRem (0.911 vs 0.856 and 0.841 vs 0.789, respectively; $p = 0.03$); thus correcting classification for 8% of those initially misclassified with DiaRem. With Ad-DiaRem, there were also fewer misclassifications of individuals with mid-range scores. This improved predictive performance was confirmed in independent cohorts.

Conclusions/interpretation We propose the Ad-DiaRem, which includes two additional clinical variables, as an optimised tool with improved accuracy to predict diabetes

Judith Aron-Wisniewsky, Nataliya Sokolovska and Yuejun Liu contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00125-017-4371-7) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

✉ Karine Clément
karine.clement@psl.aphp.fr

¹ Institute of Cardiometabolism and Nutrition (ICAN), Nutrition Department, Assistance Publique-Hôpitaux de Paris, Pitié-Salpêtrière Hospital, 46-83 Boulevard de l'Hôpital, F-75013 Paris, France

² Team 6 Nutriomics, UPMC Université Paris 06 and Inserm, UMR_S 1166, Sorbonne Universités, Paris, France

³ Central Headquarters, Clalit Health Services, Tel Aviv, Israel

⁴ Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

⁵ Department of Clinical Biochemistry and Pharmacology at the Faculty of Health Sciences and the National Institute of

Biotechnology in the Negev (NIBN), Ben-Gurion University of the Negev, Beer-Sheva, Israel

⁶ Surgery Department, Assistance Publique – Hôpitaux de Paris, Ambroise Paré Hospital, Boulogne-Billancourt, France

⁷ Hepato-biliary and Digestive Surgery Department, Assistance Publique – Hôpitaux de Paris, Pitié-Salpêtrière Hospital, Paris, France

⁸ Internal Medicine D and Obesity Clinic, Hasharon Hospital – Rabin Medical Center, Petach-Tikva, Israel

⁹ Unité Mixte Internationale Modélisation Math. et Info. des Systèmes Complexes, UMMISCO UMI 209 IRD/UPMC, Paris, France

remission 1 year post-RYGB. This tool might be helpful for personalised management of individuals with diabetes when considering bariatric surgery in routine care, ultimately contributing to precision medicine.

Keywords Bariatric surgery · Diabetes remission · Obese · Type 2 diabetes mellitus

Abbreviations

Ad-DiaRem	Advanced-DiaRem
AUROC	Area under the receiver operating characteristic
CHS	Clalit Health Services
CRP	C-reactive protein
DPP-IV	Dipeptidyl peptidase-4
DXA	Dual-energy X-ray absorptiometry scan
GLP-1	Glucagon-like peptide-1
PPV	Positive predictive value
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NPV	Negative predictive value
RYGB	Roux-en-Y gastric bypass
scAT	Subcutaneous adipose tissue
TZD	Thiazolidinedione

Introduction

Bariatric surgery has beneficial effects of major and sustained weight loss with improved metabolic comorbidities. Bariatric surgery improves glycaemic control and even induces diabetes remission, which can be complete or partial, defined by fasting glycaemia and HbA_{1c} normalisation without glucose-lowering treatment 1 year post-bariatric surgery [1]. These observations recently led to the revision of treatment guidelines, which now recommend bariatric surgery in the treatment of type 2 diabetes at any stage of obesity [2]. These guidelines are expected to substantially augment the already increasing number of bariatric surgery interventions worldwide [3]. However, despite the beneficial effects of bariatric surgery on metabolic conditions, there is significant inter-individual variability for individuals with type 2 diabetes. Outcomes are dependent on various factors including bariatric surgery procedure type and severity of type 2 diabetes before surgery.

A meta-analysis using an earlier definition of diabetes remission found that 78% of individuals with type 2 diabetes achieved diabetes remission post-bariatric surgery [4]. However, by applying the latest ADA proposed definitions [1] when considering all bariatric surgery procedures the proportion achieving diabetes remission decreased to 35%. When specifically focusing on the Roux-en-Y gastric bypass

(RYGB), 40–60% of individuals achieve diabetes remission within 1 year [5, 6]. This remission rate decreases to 37% 5 years post-RYGB, denoting an important prevalence of relapse [7]. Furthermore, although bariatric surgery has overall beneficial health outcomes, perioperative morbidity and mortality rates remain at 3.4% and 0.3%, respectively [4]. Deleterious effects, such as nutritional deficiency, are also observed with different types of bariatric surgery [8, 9]. Together, the anticipated increasing number of bariatric surgery procedures and the uncertainty in predicting clinical outcomes, both short- and long-term, emphasises the need to establish useful and clinically applicable tools to predict metabolic/bariatric surgery outcomes [2].

Current clinical predictors include preoperative clinical variables (e.g. young age, short diabetes duration, type 2 diabetes control [i.e. low HbA_{1c}], no insulin requirement), as well as post-bariatric surgery outcomes (e.g. significant post-bariatric surgery weight loss). Several scoring systems or statistical models based on these and other variables [10–13] currently help to predict diabetes remission post-bariatric surgery. Among them, the DiaRem, a scoring system based on preoperative age, HbA_{1c} and the use of some glucose-lowering treatments, has a predictive accuracy of 84% 1 year post-RYGB [14]. However, the use of the DiaRem across the entire score spectrum has limitations. Individuals with a mid-range DiaRem score (i.e. between 8 and 17) only show a 50% probability of diabetes remission [13]. Furthermore, one-third of patients with a high score (i.e. those predicted to not achieve diabetes remission) manage to attain diabetes remission [15]. Importantly, the current DiaRem does not take into account novel glucose-lowering agents, such as glucagon-like peptide-1 (GLP-1) analogues, dipeptidyl peptidase-4 (DPP-IV) inhibitors or sodium–glucose co-transporter-2 (SGLT2) inhibitors, which may also influence diabetes remission [13]. Collectively, these observations prompted us to examine how to optimise this scoring system in order to improve outcome prediction before bariatric surgery.

We aimed to develop an improved predictive scoring system (i.e. the advanced [Ad]-DiaRem) for diabetes remission post-bariatric surgery by adding easily accessible clinical variables and tested its predictive accuracy in a test cohort. We then validated this improved tool in two independent confirmation cohorts from France and Israel.

Methods

Study design and participants

We leveraged our ongoing cohort ('BARICAN' [BARIatric surgery cohort of the Institute of CARDiometabolism and Nutrition] recorded in CNIL [Commission nationale de l'informatique et des libertés] no. 1222666) followed in the

Pitié-Salpêtrière Hospital Nutrition Department (Paris, France), which consists of obese individuals who meet standard guideline recommendations for bariatric surgery intervention [16]. We selected individuals who had undergone RYGB, excluding revisional surgery, and who had a very detailed clinical dataset at 1 year follow-up. Intending to build this putative optimised scoring system, we identified individuals with type 2 diabetes with baseline bioclinical and anthropometric variables, obesity-related disease information and detailed information on treatment usage, blood metabolic and inflammatory variables, adipocyte size and liver histological diagnosis.

The test cohort, which consisted of 213 participants with type 2 diabetes with complete data for all of the abovementioned variables (patients' characteristics are displayed in Table 1 and ESM Table 1), enabled the development of two different scoring systems: (1) the Ad-DiaRem, which included additional simple clinical variables that significantly differed at baseline between participants who had achieved diabetes remission and those who had not. We further compared the Ad-DiaRem with the existing DiaRem, and (2) the Costly-DiaRem, which was constructed for individuals falling within the Ad-DiaRem middle score category in order to further improve prediction accuracy for this group.

A French confirmation cohort also from the 'BARICAN' cohort, consisted of 134 participants with type 2 diabetes with available information for each of the variables used in the Ad-DiaRem (Table 2). We further examined the Ad-DiaRem in another independent cohort from Israel, comprising 99 individuals with type 2 diabetes who had undergone RYGB as described previously [17]. The data for these participants were taken retrospectively from the electronic medical records of Clalit Health Services (CHS) and included individuals with type 2 diabetes who underwent bariatric surgery between 1999 and 2011, and had follow-up data until December 2014. Data from the CHS electronic database included the variables from the DiaRem and Ad-DiaRem. Figure 1 presents the study flow chart.

Ethics approval was obtained from the French Research Ethics Committee of CPP Ile de France-1 No. 13533 and the Rabin Medical Center Ethics Committee. All participants provided written informed consent.

Definition of diabetes and 1-year remission outcomes

Type 2 diabetes was defined according to ADA criteria [18]. In the French and Israeli confirmation cohorts, 1-year remission outcomes were defined according to the latest ADA definition [1] (i.e. partial remission: $HbA_{1c} < 6.5\%$, fasting plasma glucose [FPG] < 7.0 mmol/l and no use of glucose-lowering agents at 1 year; complete remission: $HbA_{1c} < 6.0\%$, FPG < 5.6 mmol/l and no use of glucose-lowering agents at 1 year). All participants who achieved complete or

partial remission were considered as the 'remission group' because they displayed blood glucose control normalisation without the use of glucose-lowering agents.

Test cohort: bioclinical, anthropological and histological variables

Baseline clinical information on diabetes duration (i.e. duration up to RYGB intervention), use of glucose-lowering agents, and obesity-related comorbidities and treatments (e.g. hypertension, obstructive sleep apnoea and dyslipidaemia) were collected as previously described [16]. Glucose-lowering medication groups were classified as follows: GLP-1 analogues, DPP-IV inhibitors, sulfonylureas, thiazolidinediones (TZDs), glinides, α -glucosidase inhibitors, metformin and insulin (basal and/or bolus). The number of glucose-lowering agents prescribed was considered the sum of the above drug categories.

Blood samples were collected at baseline after a 12 h overnight fast. Pancreatic beta cell function (insulin secretion) and insulin resistance were estimated using HOMA- β and HOMA-IR, respectively [19]. Body composition was evaluated by whole body dual-energy X-ray absorptiometry scan (DXA, Hologic Discovery W) [20].

Adipocyte diameter, which enabled the calculation of adipocyte morphology [21], was evaluated with Perfect Image (Clara Vision, Verrières le Busson, France) from subcutaneous adipose tissue (scAT) needle-aspirated biopsies after collagenase digestion as previously described [22]. Perioperative surgical liver biopsies were collected to assess non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH) using the steatosis activity fibrosis (SAF) score [23, 24].

The DiaRem score

The DiaRem was initially established to predict the probability of complete or partial diabetes remission following post-bariatric surgery. The DiaRem score, ranging from 0 to 22, was calculated for each participant using age, HbA_{1c} , and use of some glucose-lowering medications and insulin, each with a defined weight as described in [13] (see electronic supplementary material [ESM] Table 2).

Development of an optimised scoring system: Ad-DiaRem

We examined 43 baseline variables (11 clinical variables, 27 laboratory variables and five scAT and liver biopsy variables; Table 1) as potential variables that could improve the predictive power of the DiaRem. Multivariate logistic regression was performed to estimate the OR of potential predictors of remission. The variables whose ORs were significant (i.e. $p < 0.05$) were selected and included in the Ad-DiaRem

Table 1 Baseline characteristics of participants with type 2 diabetes before bariatric surgery according to remission status at 1 year post-surgery (test cohort)

Variable	Remission group (n = 137)	Non-remission group (n = 76)	p value	Adjusted p value
Male, n (%)	41 (30)	30 (40)	0.16	–
Age (years)	46 ± 10	53 ± 9	<0.01	–
BMI (kg/m ²)	48.1 ± 7.4	45.4 ± 7	0.01	0.09
Hypertension, n (%)	88 (65)	67 (89)	<0.01	–
Treated for hypertension, n (%)	83 (61)	66 (87)	<0.01	–
Obstructive sleep apnoea, n (%)	101 (75)	61 (81)	0.28	–
Treated with CPAP, n (%)	58 (43)	37 (50)	0.35	–
Diabetes characteristics				
Diabetes duration (years)	3.5 ± 3.8**	11.1 ± 7.6	<0.01	<0.01
Insulin therapy, n (%)	13 (9)	42 (55)	<0.01	–
Sulfonylureas or ISA other than metformin, n (%)	29 (21)	29 (38)	<0.01	–
Fasting glucose (mmol/l)	7.43 ± 2.32**	9.07 ± 0.3	0.01	<0.01
Fasting insulin (pmol/l)	170.14 ± 125.01	143.75 ± 117.71	0.26	0.89
HbA _{1c} (%)	7.0 ± 1.1**	8.4 ± 1.6	<0.01	<0.01
HbA _{1c} (mmol/mol)	53 ± 11.9	68 ± 17.8		
HOMA-IR	3.3 ± 2.3	2.9 ± 2.2	0.35	0.98
HOMA-β%	115.1 ± 61.7*	78.4 ± 53.1	<0.01	0.03
HOMA-S%	46.5 ± 41.8	54 ± 33.7	0.28	0.49
Body composition				
Fat mass (%)	47.9 ± 5.3**	45.3 ± 5.9	<0.01	<0.01
Fat-free mass (%)	49.9 ± 5.1**	52.4 ± 5.7	<0.01	<0.01
Fat mass/fat-free mass ratio	0.98 ± 0.20**	0.88 ± 0.20	<0.01	<0.01
Android fat mass (%)	66.2 ± 5.5	68.5 ± 5.3	<0.01	0.08
Gynoid fat mass (%)	32.2 ± 5.6*	29.5 ± 5.3	<0.01	0.02
Android/gynoid fat mass ratio	2.15 ± 0.53*	2.42 ± 0.61	<0.01	0.01
Adipokines				
Adiponectin (µg/ml)	4.9 ± 3.2	4.8 ± 2.9	0.97	0.35
Leptin (ng/ml)	52.0 ± 25.9	48.0 ± 32.6	0.37	0.21
Lipid variables				
Treated with lipid-lowering drugs, n (%)	48 (35)	55 (72)	<0.01	–
Total cholesterol (mmol/l)	4.88 ± 1.03*	4.39 ± 1.04	<0.01	0.01
Triacylglycerol (mmol/l)	1.90 ± 1.76	1.91 ± 1.13	0.98	0.64
HDL-cholesterol (mmol/l)	1.11 ± 0.32	1.18 ± 0.36	0.16	0.89
ApoA-1 (mmol/l)	1.36 ± 0.25	1.42 ± 0.29	0.19	0.81
ApoB (mmol/l)	0.99 ± 0.31	0.89 ± 0.26	0.02	0.09
Liver biology				
AST (µkat/l)	0.55 ± 0.30	0.53 ± 0.25	0.63	0.81
ALT (µkat/l)	0.78 ± 0.82	0.67 ± 0.40	0.19	0.93
γGT (µkat/l)	0.91 ± 0.72	1.12 ± 0.88	0.10	0.02
Inflammatory factors				
IL-6 (pg/ml)	4.2 ± 2.1	4.9 ± 3.5	0.14	0.08
hsCRP (mg/l)	10.8 ± 8.9	8.2 ± 9.5	0.06	0.60
Orosomucoid (g/l)	0.94 ± 0.22	0.90 ± 0.23	0.15	0.64
Adipose tissue needle aspirate				
Adipocyte diameter (µm)	121.1 ± 13.9	119.9 ± 9.8	0.53	0.46
Adipocyte morphology (pl)	46.7 ± 225.4	47.5 ± 216.1	0.98	0.57

Data are presented as mean ± SD unless otherwise stated

Baseline characteristics were compared using Student's *t* test for two groups, according to participants' 1-year type 2 diabetes remission outcome

Type 2 diabetes was defined according to ADA guidelines (i.e. FPG ≥7.0 mmol/l, 2 h plasma glucose ≥11.1 mmol/l when available, HbA_{1c} ≥6.5% or receiving any glucose-lowering agents); partial diabetes remission was defined as HbA_{1c} <6.5%, FPG <7.0 mmol/l and no use of glucose-lowering agents at 1 year; complete diabetes remission was defined as HbA_{1c} <6.0%, FPG <5.6 mmol/l and no use of glucose-lowering agents at 1 year

Continuous data were also adjusted for age

p* < 0.05, *p* < 0.01 vs non-remission

ALT, alanine transaminase; AST, aspartate transaminase; CPAP, continuous positive airway pressure therapy; γGT, gamma glutamyl transferase; hsCRP, high-sensitivity CRP; ISA, insulin sensitising agent

Table 2 Ad-DiaRem scoring system

Prediction factor	Score
Age (years)	
[15–41]	0
[42–52]	3
[53–69]	5
HbA _{1c} (%)	
[4.5–6.9]	0
[7.0–7.4]	2
[7.5–18.4]	4
Insulin	
No	0
Yes	3
Other glucose-lowering agents ^a	
No	0
Yes	1
Number of glucose-lowering agents ^b	
0	0
1	1
2	2
≥ 3	3
Diabetes duration (years)	
[0–6.9],	0
[7.0–13.9]	3
≥ 14	5
Ad-DiaRem overall score (sum of the above six components)	0–21

^a Includes sulfonylureas (glimepiride, glipizide and glibenclamide), ISAs other than metformin (pioglitazone and rosiglitazone)

^b Includes sulfonylureas, ISAs and GLP-1 analogues, DDP-IV inhibitors, insulin and other glucose-lowering agents

ISA, insulin sensitising agent

scoring system (i.e. all the variables included in the DiaRem plus two easily accessible clinical variables: the number of glucose-lowering agents and diabetes duration).

Statistical analyses

Categorical variables are expressed as numbers and percentages, continuous data as means ± SD. Categorical data were analysed using Fisher's exact test for two groups. Continuous data were analysed using the Student's *t* test. The analyses were adjusted by age. Two-tailed *p* values were considered significant at *p* < 0.05. All analyses were conducted using R software version 3.0.3 (<http://www.r-project.org>) and GraphPad Prism 6.0.

Learning ad-DiaRem A clinical scoring system should be able to select relevant clinical variables, propose interpretable clinical thresholds and estimate weights for corresponding bins. We applied a machine learning method that simultaneously learns the restricted set of informative variables to

retain. This method associates interpretable binning to map with each class variable (complete/partial diabetes remission or non-remission) and provides optimal weights to associate with these bins contributing to the score. For machine learning, we minimised empirical risk given the diabetes cohort and performed tenfold cross-validation to avoid possible overfitting. Specifically, as a classification algorithm, we used a sparse support vector machine. To optimise the problem of the score learning, we formulated it as a linear integer programming task and used the IBM ILOG CPLEX Optimization Studio (<http://www-03.ibm.com/software>), which is a state-of-the-art solver for constrained optimisation problems. We added integrity constraints to our task so that the resulting weights are integers. Also, constraints shrink similar variables to each other, creating bins and ordering them. The computations were done with R version 3.1.3 and the 'Rcplex' package, which is the interface to the IBM CPLEX Studio. The predictive performance of different scores were evaluated by the area under the receiver operating characteristic (AUROC) curve using the DeLong method.

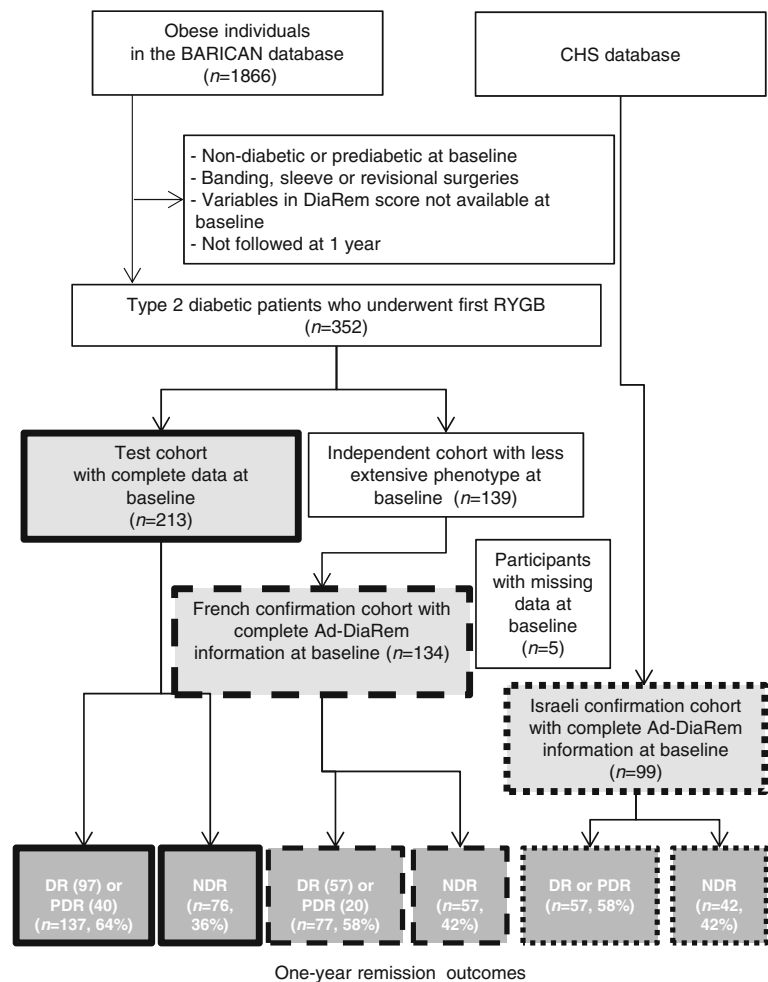
Results

Clinical variables associated with 1-year diabetes remission post-RYGB

In the test cohort, 64% of participants achieved diabetes remission within 1 year (Fig. 1), concordant with previous reports [25]. Participants who achieved diabetes remission were younger, had significantly lower FPG and HbA_{1c}, and were less likely to be treated with insulin or oral glucose-lowering agents other than metformin presurgery compared with those who did not achieve remission (Table 1, Fig. 2a). Those who achieved remission displayed a significantly higher BMI, higher DXA-evaluated fat mass and less abdominal fat distribution at baseline. Importantly, after adjustment for age, although differences in fat mass and its deleterious deposition (android/gynoid fat mass) remained significant, BMI did not. Participants who achieved remission also exhibited a shorter type 2 diabetes duration and potentially increased beta cell function as estimated by HOMA-β (Table 1). These differences remained significant after adjustment for age. The sex ratio was not significantly different between groups.

Although adipocyte diameter was increased in individuals with type 2 diabetes compared with those without diabetes (data not shown), it was not significantly different between the remission and non-remission groups at 1 year. Liver fibrosis scores (from liver biopsies) were more severe in participants who had not achieved remission compared with those who had (ESM Fig. 1), whereas other liver alterations (i.e. steatosis, inflammation activity, NAFLD/NASH scores) were similar between the groups. This exploration revealed that: (1) current DiaRem

Fig. 1 Study flow chart. The study included obese individuals with type 2 diabetes who underwent RYGB from a French cohort ($n = 352$) and an Israeli cohort ($n = 99$). The test cohort ($n = 213$) consisted of participants from the French cohort who had a complete set of data available at baseline. The French ($n = 134$) and Israeli ($n = 99$) confirmation cohorts were used for Ad-DiaRem external validation. DR, complete diabetes remission; PDR, partial diabetes remission; NDR, non-remission



variables differed between the remission and non-remission groups, and (2) additional factors (i.e. number of glucose-lowering agents, diabetes duration and body composition variables) also varied.

DiaRem score in the test cohort

When evaluating the DiaRem in our test cohort, we found an AUROC of 85% (Fig. 2b). Using the Youden method, the threshold for remission was calculated to be a score of 7 (i.e. participants with a DiaRem score <7 should achieve diabetes remission), confirming our previous findings in another independent group [14]. Although the overall predictive accuracy of DiaRem was 78.9% (Fig. 2b), the false-positive (remission was predicted in participants who failed to achieve remission, $n = 9$) and false-negative (non-remission was predicted in participants achieving remission, $n = 41$) rates were quite high. Positive predictive value (PPV) was high (0.91) but negative predictive value (NPV) was much lower (0.62).

Subsequently, participants were stratified into five groups according to their DiaRem score: 0–2 (highest probability of remission), 3–7, 8–12, 13–17 and 18–22 (lowest probability of

remission) (Fig. 2c). A high proportion of participants with low scores (i.e. 0–2 or 3–7) achieved remission, indicating a good predictive value of DiaRem for individuals in this range (Fig. 2c). However, only about half of the participants with scores ranging from 8–12 attained remission, demonstrating a poor predictive performance in this intermediate range. We highlighted a high degree of misclassification in this middle range (i.e. 27 participants (12.6%) with a DiaRem score of 8–17 still experienced remission) (Fig. 2c). Altogether, the majority of the individuals in either the remission or the non-remission groups were not readily separable by DiaRem, with an overlap between the score ranges that cumulatively included 80% of either group (Fig. 2d).

These results indicate a satisfactory predictive value of the DiaRem score for the extreme ranges, but a lot of participants remained incorrectly classified. This prompted us to evaluate the relevance of other variables in predictive accuracy.

Ad-DiaRem score improves prediction of diabetes remission 1 year post-RYGB

We examined baseline variables that significantly differed between the remission and non-remission groups (i.e. $p < 0.05$,

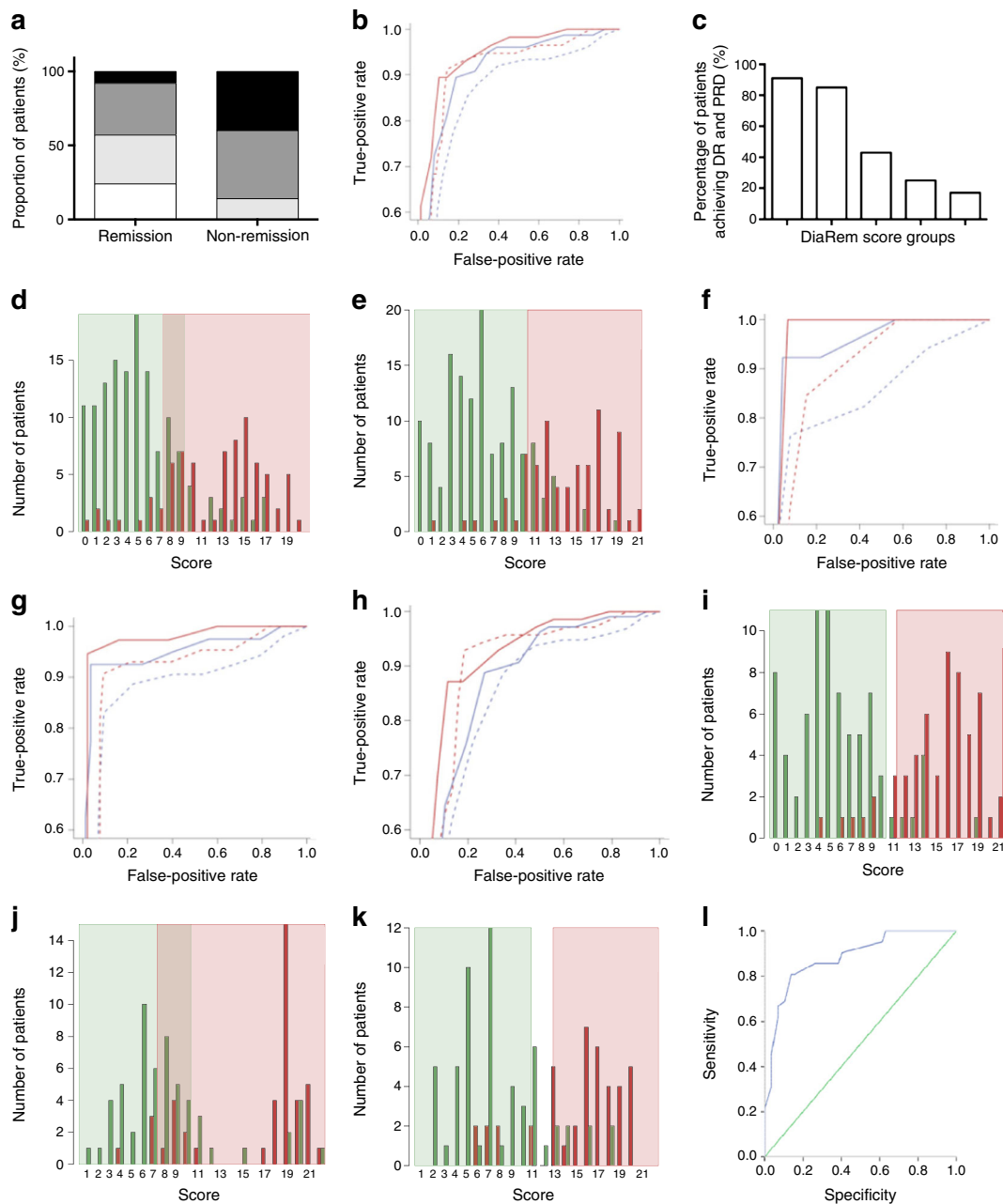


Table 1) to develop an improved tool for the prediction of remission (Ad-DiaRem; ESM Table 3). After adjustment for the four variables already present in the DiaRem, the ORs were significant for number of glucose-lowering agents, diabetes duration and DXA-evaluated body composition, but not BMI. Since DXA might not be easily accessible in all clinical settings, we first tested whether including only two additional clinical variables would be sufficient to improve the DiaRem accuracy.

The Ad-DiaRem (Table 2) led to a better classification of participants achieving remission with an improved AUROC and accuracy compared with the DiaRem (0.911 vs 0.856 and 0.841 vs 0.789, respectively; $p = 0.03$) (Fig. 2b, e). Compared with the DiaRem (Fig. 2d), the Ad-DiaRem created a better

separation of 80% of the participants that achieved remission vs those that did not (i.e. the majority [80%] of both groups did not overlap with the Ad-DiaRem) (Fig. 2e). Additionally, the Ad-DiaRem demonstrated better PPV and NPV (0.93 and 0.72, respectively) compared with the DiaRem (0.91 and 0.62, respectively); thus leading to improved classification of 16 (8%) participants who were initially misclassified. In total, the DiaRem correctly classified 164 of 213 participants from the test cohort, whereas 180 participants were correctly classified by the Ad-DiaRem. Using the Youden method, the threshold for remission was calculated to be below a score of 10 (i.e. participants with an Ad-DiaRem score <10 should achieve diabetes remission).

◀ **Fig. 2** (a) Number of glucose-lowering medications at baseline in the test cohort. Each bar represents the percentage of participants in the remission and non-remission groups receiving no glucose-lowering treatments (white), one glucose-lowering treatment (light grey), two glucose-lowering treatments (dark grey) or three or more glucose-lowering treatments (black). (b) Evaluation of DiaRem and Ad-DiaRem scores for the remission vs non-remission groups for the test cohort (DiaRem: AUROC 0.856, accuracy 0.789; Ad-DiaRem: AUROC 0.911, accuracy 0.841) and the French confirmation cohort (DiaRem: AUROC 0.893, accuracy 0.881; Ad-DiaRem: AUROC 0.939, accuracy 0.896). (c) The percentage of participants achieving remission according to DiaRem score in the test cohort (DR, complete diabetes remission; PDR, partial diabetes remission). (d) Distribution of participants according to each DiaRem score in the test cohort. (e) Distribution of participants according to Ad-DiaRem score in the test cohort in the remission vs the non-remission groups. (f) Evaluation of DiaRem and Ad-DiaRem scores in participants with low (0–2) compared with high scores (19–22 for DiaRem and 19–21 for Ad-DiaRem) for the remission vs the non-remission groups in the test (DiaRem test: AUROC 0.857, accuracy 0.873; Ad-DiaRem test: AUROC 0.955, accuracy 0.944) and French confirmation cohorts (DiaRem: AUROC 0.899, accuracy 0.846; Ad-DiaRem: AUROC 0.977, accuracy 0.96). (g) Evaluation of DiaRem and Ad-DiaRem scores in participants with low (0–5) compared with high scores (15–22 for DiaRem and 15–21 for the Ad-DiaRem) for the remission vs the non-remission groups in the test (DiaRem: AUROC 0.857, accuracy 0.887; Ad-DiaRem: AUROC 0.935, accuracy 0.965) and French confirmation cohorts (DiaRem: AUROC 0.891, accuracy 0.91; Ad-DiaRem: AUROC 0.964, accuracy 0.96). (h) DiaRem and Ad-DiaRem scores in participants with mid-range scores (8–14) for the remission vs non-remission groups in the test and confirmation cohorts. (i–k) Distribution of participants for the remission vs non-remission groups according to (i) Ad-DiaRem score in the French confirmation cohort, (j) DiaRem score in the Israeli confirmation cohort, and (k) Ad-DiaRem score in the Israeli confirmation cohort. (l) Evaluation of Ad-DiaRem in all participants in the Israeli confirmation cohort for the remission vs non-remission groups (AUROC 0.882). Diagonal segments are produced by ties. (d, e, i–k) Red bars, non-remission; green bars, remission; red background, 80% of participants with non-remission; green background, 80% of participants in remission. (b, f–h, l) Red lines, test cohort; blue lines, confirmation cohorts; dotted lines, DiaRem; solid lines, Ad-DiaRem

The predictive improvement was most noticeable for participants with low (0–5) or high (15–21) scores. As a consequence, the AUROC and accuracy calculation of Ad-DiaRem was better for the extreme ranges compared with the DiaRem, nearly reaching significance (Fig. 2f, g; $p = 0.06$ for comparison between scores 0–5 and 17–21/22 in the DiaRem and Ad-DiaRem).

For participants with mid-range scores, the Ad-DiaRem correctly reclassified 12 out of the 24 that the DiaRem incorrectly predicted as non-remission. Although AUROC and accuracy were increased with Ad-DiaRem for these participants (Fig. 2h), the difference did not reach statistical significance comparing the two scores.

Next we examined the Ad-DiaRem prediction accuracy in French and Israeli confirmation cohorts. In the French cohort, 58% of the participants achieved remission 1 year post-bariatric surgery (Fig. 1). Figure 2b and i show that, compared with the DiaRem, the Ad-DiaRem better classified participants in the French cohort, with an increased proportion with low scores

(0–2 and 3–5) achieving remission and a very high proportion with high scores (17–22) with non-remission. This improvement was retained in different scoring sub-categories (Fig. 2f–h). Compared with the DiaRem, the Ad-DiaRem score correctly reclassified ten (7.4%) participants, and the overall accuracy and AUROC of Ad-DiaRem in predicting remission (vs non-remission) was superior in the test and confirmation cohorts (Fig. 2b; $p = 0.03$). NPV also increased with Ad-DiaRem in this confirmation cohort compared with the DiaRem (0.82 vs 0.75, respectively). A similar added value of the Ad-DiaRem was found when comparing participants with complete diabetes remission vs non-remission in the test and confirmation cohorts (i.e. excluding participants with partial remission) (ESM Fig. 2).

Of the 99 participants from the Israeli cohort, 58% achieved diabetes remission. Similar to the observations made in the French cohort, Ad-DiaRem clearly separated the majority (80%) of those individuals who achieved remission from those who did not (Fig. 2k), whereas DiaRem exhibited an overlap between the groups (Fig. 2j). Furthermore, the AUROC increased from 0.825 with DiaRem to 0.882 with Ad-DiaRem (Fig. 2l). The Distribution of Ad-DiaRem scores among the remission and non-remission groups in the Israeli confirmation cohort are shown in ESM Table 4.

Added value of other bioclinical variables for predicting diabetes remission post-RYGB

To evaluate if we could further improve Ad-DiaRem performance for participants with mid-range scores (8–14), we tested the addition of other variables, e.g. DXA-measured fat mass, fat-free mass proportion, fat mass/fat-free mass ratio, serum C-reactive protein (CRP) and HOMA- β . At baseline, these variables significantly differed between participants who achieved remission and those who did not.

Using a binning method, we developed the Costly-DiaRem scoring system, which penalised for low fat mass (%), high fat mass/fat-free mass ratio, high android/gynoid fat mass ratio, high serum CRP and low HOMA- β (see ESM Table 5). Despite the inclusion of these additional bioclinical variables providing deeper phenotyping, the Costly-DiaRem did not perform better than the Ad-DiaRem in any scoring range (ESM Fig. 3). Furthermore, prediction accuracy was not improved by the simple addition of HOMA- β to the Ad-DiaRem (data not shown).

Discussion

Here, we show that the Ad-DiaRem scoring system improves predictive accuracy for diabetes remission 1 year post-RYGB compared with the currently proposed DiaRem in a population of severely obese individuals with type 2 diabetes. Of the 347 individuals from the French cohort (214 with complete/partial

diabetes remission), 16 were correctly reclassified using the Ad-DiaRem. This improved scoring system has two additional variables that are easily recordable in clinical practice (i.e. diabetes duration and number of glucose-lowering agents) and has modified scoring for each variable. The accuracy for predicting diabetes remission was significantly increased with Ad-DiaRem, as evidenced in French and Israeli confirmation cohorts. Developing an accurate scoring system to better predict the outcomes of bariatric surgery is becoming necessary as the number of bariatric surgery procedures being undertaken increases worldwide [3]. This increase is compounded by new guidelines for type 2 diabetes management now recommending a lower BMI cut-off for bariatric surgery in treatment algorithms [2]. Not all individuals experience the same beneficial outcomes following bariatric surgery, both in the extent of weight loss [26] and metabolic improvements [4]. Therefore, the development of reliable predictive tools will help routine care decision making and, in the future, to innovate personalisation of pre and postoperative care pathways.

The DiaRem, recently created using Cox regression analysis of 5-year follow-up data in 690 participants [13], demonstrated good predictive performance for 1-year remission, despite slightly lower accuracy in confirmation cohorts [13]. Here, although we confirmed the performance of the DiaRem in French and Israeli cohorts, a significant number of individuals remained misclassified [13, 15], primarily those in the medium score range (8–17), which comprised about one-third of participants. The Ad-DiaRem significantly decreased the predictive errors for the overall cohorts and for participants with mid-range scores. The Ad-DiaRem exhibited a PPV of 0.93 and NPV of 0.72 in predicting remission in the test cohort, thus improving the predictive accuracy of the previously published DiaRem.

The improved performance of Ad-DiaRem is likely to be due to multiple factors. First, the DiaRem score included age, a rather indirect marker of diabetes duration. Although increasing age is generally correlated with longer diabetes duration, it is known that, with the dramatic increase in obesity prevalence worldwide, type 2 diabetes now occurs earlier [27]. Therefore, the small penalty assigned for age below 40 years in the DiaRem might not be fully accurate anymore [13]. Diabetes duration is regarded as a consistent marker of disease progression and recognised as the best predictor of post-bariatric surgery diabetes remission [2, 28, 29]. Since this variable was not available in the Still et al database used for the DiaRem calculation, it could not be integrated [13]. Diabetes duration was integrated into another predictive tool for diabetes remission post-bariatric surgery, the ABCD [12]. However, the ABCD scoring system did not perform as well as the DiaRem [14] and might not be convenient for routine use as it relies on fasting C-peptide, an expensive serum

marker that is not easily available in routine care. Admittedly, diabetes duration is not absolutely accurate. It is usually self-reported and the true onset of disease is indolent. Frequently, type 2 diabetes is diagnosed long after beta cell function has declined [30]. Still, diabetes duration is easy to collect and its value is demonstrable in the Ad-DiaRem. Second, the DiaRem does not take into account currently available drugs for type 2 diabetes treatment, mainly DPP-IV inhibitors and GLP-1 analogues. This latter class is widely used in individuals with type 2 diabetes who are obese because it improves glucose control and decreases weight in some individuals [31]. We hypothesised that taking into account the overall number of drugs might be more reflective of disease progression during the preoperative stage. Thus, we integrated this information into the Ad-DiaRem (Table 2). Furthermore, since glucose-lowering agents are not standardised among countries [32, 33] and they are given according to tolerance and secondary effects, we believe that by adding the number of glucose-lowering agents into the score individual heterogeneity will be accounted for better.

By using this retrospective cohort of individuals with type 2 diabetes who have undergone RYGB that were extensively phenotyped at baseline, we also describe new clinical variables that were associated with non-remission. Compared with those individuals who achieved at least partial remission, those with non-remission had less adipose tissue (lower fat mass); however, these individuals displayed increased android fat mass repartition at baseline, which is recognised as detrimental for metabolic complications [34]. They also displayed liver fibrosis more frequently.

The Ad-DiaRem improved predictive accuracy compared with the DiaRem but did not fully solve misclassification of individuals with mid-range scores. Despite our effort to add other detailed phenotypic characteristics differing at baseline between participants who achieved remission and those who did not (i.e. body composition data and insulin secretion index) to the Ad-DiaRem, we were unable to further improve prediction accuracy. This opens up the possibility of testing other biological markers. For example, recent literature points to the importance of genomic variation (single nucleotide polymorphisms) related to insulin secretion in the prediction of diabetes remission post-bariatric surgery, suggesting that measures related to pancreas failure to (hyper-)secrete insulin might be of interest. The added value of genetic scoring must be examined in comparison with scores using clinical variables and other measurements linked to individuals' impaired metabolism. However, it is unknown whether adding more complex individual information derived from high throughput analysis, such as systemic proteomics, metabolomics or metagenomics [35, 36], or tissue alterations would be helpful in improving prediction, particularly in individuals with mid-range scores. As such, we previously described that adipose tissue fibrosis is associated with reduced weight loss post-

bariatric surgery [22, 37]. Whether adipose tissue scoring might be useful for predicting post-bariatric surgery outcomes needs to be investigated in the future.

Our study has some limitations. First, we focused on the achievement of complete/partial diabetes remission 1 year post-RYGB. Studies now demonstrate that remission is not sustained in all individuals [28]. For instance, 43% of participants who achieved 1-year remission displayed type 2 diabetes recurrence 5 years post-bariatric surgery [38]. This highlights the need to evaluate long-term glycaemic outcomes in type 2 diabetes and test the relevance of the Ad-DiaRem in the long term [39]. Indeed, type 2 diabetes is a progressive disease that worsens with time [40, 41] and bariatric surgery may only induce transient remission followed by resurgence or exacerbation. Despite this, while not all individuals achieve remission, they still improve their glycaemic control as seen with a reduction of the number of glucose-lowering agents and HbA_{1c} as observed in two long-term randomised control trials [7, 42]. When tested for prolonged remission 5 years post-RYGB, the DiaRem was not optimal for predicting remission in individuals with high scores [39]. Second, the Ad-DiaRem needs to be investigated for other types of bariatric surgery procedures, in particular post-sleeve gastrectomy, a procedure that is being used increasingly worldwide [3]. Finally, it should be noted that we tested the validity of the Ad-DiaRem solely in a population of severely obese individuals, which, to date, represents the majority of bariatric surgery candidates [43, 44]. However, the Ad-DiaRem should be further tested in diabetic individuals with less severe obesity as these will increasingly become candidates for bariatric surgery procedures based on recent recommendations [2].

Conclusions

Here, we described the development of the Ad-DiaRem for predicting diabetes remission following Roux-en-Y gastric bypass in obese individuals with type 2 diabetes. We demonstrated its ability to better separate between individuals predicted to achieve remission and those who will not, and its improved predictive performance over the original DiaRem scoring system. In the future, individuals with type 2 diabetes who are not expected to achieve remission might benefit from a care pathway with a more intensive follow-up and/or increased physical activity. These approaches need to be further tested and new guidelines proposed.

Acknowledgements We thank V. Lemoine (Clinical platform, ICAN, France) for patient follow-up, F. Marchelli (Clinical platform, Human Research Center on Nutrition [CRNH], France) for data management and R. Alili (NutriOmics Group, UPMC, France) for her contribution to bio-banking. T. Swartz (ICAN, France) performed English language editing. Some of the data were presented as an abstract at the European Congress on Obesity meeting in Porto, May 2017.

Data availability All data used for analyses in the current paper are available on request from the authors.

Funding This work was supported by funds obtained by the Campus programme (Maimonide ‘Franco-Israeli project’ and Ben-Gurion University). The clinical study was supported by the Assistance Publique – Hôpitaux de Paris and the Direction of Clinical Research (CRC) (PHRC 02076 to KC, CRC P050318 to CP and CRC-FIBROTA to JAW and KC), the National Agency of Research (ANR Adipofib, and the national programme ‘Investissements d’Avenir’ with the reference ANR-10-IAHU-05).

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement All authors provided substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; drafted the article and revised it critically for important intellectual content; gave final approval of the version to be published. Karine Clément is the guarantor of this work.

References

1. Buse JB, Caprio S, Cefalu WT et al (2009) How do we define cure of diabetes? *Diabetes Care* 32:2133–2135
2. Rubino F, Nathan DM, Eckel RH et al (2016) Metabolic surgery in the treatment algorithm for type 2 diabetes: a joint statement by international diabetes organizations. *Diabetes Care* 39:861–877
3. Angrisani L, Santonicola A, Iovino P et al (2015) Bariatric surgery worldwide 2013. *Obes Surg* 25:1822–1832
4. Schauer PR, Mingrone G, Ikramuddin S, Wolfe B (2016) Clinical outcomes of metabolic surgery: efficacy of glycemic control, weight loss, and remission of diabetes. *Diabetes Care* 39:902–911
5. Pournaras DJ, Aasheim ET, Søvik TT et al (2012) Effect of the definition of type II diabetes remission in the evaluation of bariatric surgery for metabolic disorders. *Br J Surg* 99:100–103
6. Davies SW, Efrid JT, Guidry CA et al (2014) Long-term diabetic response to gastric bypass. *J Surg Res* 190:498–503
7. Mingrone G, Panunzi S, De Gaetano A et al (2015) Bariatric-metabolic surgery versus conventional medical treatment in obese patients with type 2 diabetes: 5 year follow-up of an open-label, single-centre, randomised controlled trial. *Lancet* 386:964–973
8. Verger EO, Aron-Wisniewsky J, Dao MC et al (2015) Micronutrient and protein deficiencies after gastric bypass and sleeve gastrectomy: a 1-year follow-up. *Obes Surg* 26:785–796
9. Aron-Wisniewsky J, Verger EO, Bounaix C et al (2016) Nutritional and protein deficiencies in the short term following both gastric bypass and gastric banding. *PLoS One* 11:e0149588
10. Hayes MT, Hunt LA, Foo J et al (2011) A model for predicting the resolution of type 2 diabetes in severely obese subjects following Roux-en Y gastric bypass surgery. *Obes Surg* 21:910–916
11. Ramos-Levi AM, Matia P, Cabrerizo L et al (2014) Statistical models to predict type 2 diabetes remission after bariatric surgery. *J Diabetes* 6:472–477
12. Lee W-J, Hur KY, Lakadawala M et al (2013) Predicting success of metabolic surgery: age, body mass index, C-peptide, and duration score. *Surg Obes Relat Dis* 9:379–384
13. Still CD, Wood GC, Benotti P et al (2014) Preoperative prediction of type 2 diabetes remission after Roux-en-Y gastric bypass surgery: a retrospective cohort study. *Lancet Diabetes Endocrinol* 2: 38–45

14. Cotillard A, Poitou C, Duchâteau-Nguyen G et al (2015) Type 2 diabetes remission after gastric bypass: what is the best prediction tool for clinicians? *Obes Surg* 25:1128–1132
15. Lee W-J, Chong K, Chen S-C et al (2016) Preoperative prediction of type 2 diabetes remission after gastric bypass surgery: a comparison of DiaRem scores and ABCD scores. *Obes Surg* 26:2418–2424
16. Liu Y, Aron-Wisniewsky J, Marcelin G et al (2016) Accumulation and changes in composition of collagens in subcutaneous adipose tissue after bariatric surgery. *J Clin Endocrinol Metab* 101:293–304
17. Dicker D, Yahalom R, Comaneshter DS, Vinker S (2016) Long-term outcomes of three types of bariatric surgery on obesity and type 2 diabetes control and remission. *Obes Surg* 26:1814–1820
18. American Diabetes Association (2015) 2. Classification and diagnosis of diabetes. *Diabetes Care* 38:S8–S16
19. Matthews DR, Hosker JP, Rudenski AS et al (1985) Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419
20. Ciangura C, Bouillot J-L, Lloret-Linares C et al (2010) Dynamics of change in total and regional body composition after gastric bypass in obese patients. *Obesity (Silver Spring)* 18:760–765
21. Arner E, Westermark PO, Spalding KL et al (2010) Adipocyte turnover: relevance to human adipose tissue morphology. *Diabetes* 59:105–109
22. Divoux A, Tordjman J, Lacasa D et al (2010) Fibrosis in human adipose tissue: composition, distribution, and link with lipid metabolism and fat mass loss. *Diabetes* 59:2817–2825
23. Bedossa P, Poitou C, Veyrie N et al (2012) Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. *Hepatology* 56:1751–1759
24. Bedossa P, Tordjman J, Aron-Wisniewsky J et al (2016) Systematic review of bariatric surgery liver biopsies clarifies the natural history of liver disease in patients with severe obesity. *Gut*. doi:10.1136/gutjnl-2016-312238
25. Dixon JB, le Roux CW, Rubino F, Zimmet P (2012) Bariatric surgery for type 2 diabetes. *Lancet* 379:2300–2311
26. Courcoulas AP, Christian NJ, Belle SH, Longitudinal Assessment of Bariatric Surgery (LABS) Consortium et al (2013) Weight change and health outcomes at 3 years after bariatric surgery among individuals with severe obesity. *JAMA* 310:2416–2425
27. Amutha A, Mohan V (2016) Diabetes complications in childhood and adolescent onset type 2 diabetes—a review. *J Diabetes Complications* 30:951–957
28. Arterburn DE, Bogart A, Sherwood NE et al (2013) A multisite study of long-term remission and relapse of type 2 diabetes mellitus following gastric bypass. *Obes Surg* 23:93–102
29. Schauer PR, Bhatt DL, Kirwan JP, STAMPEDE Investigators et al (2014) Bariatric surgery versus intensive medical therapy for diabetes—3-year outcomes. *N Engl J Med* 370:2002–2013
30. Holman RR (1998) Assessing the potential for alpha-glucosidase inhibitors in prediabetic states. *Diabetes Res Clin Pract* 40(Suppl): S21–S25
31. Montanya E, Fonseca V, Colagiuri S et al (2016) Improvement in glycosylated haemoglobin evaluated by baseline body mass index: a meta-analysis of the liraglutide phase III clinical trial programme. *Diabetes Obes Metab* 18:707–710
32. American Diabetes Association (2017) Standards of medical care in diabetes—2017: summary of revisions. *Diabetes Care* 40(Suppl 1): S4–S5
33. Inzucchi SE, Bergenstal RM, Buse JB et al (2012) Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care* 35:1364–1379
34. Nazare J-A, Smith J, Borel A-L, INSPIRE ME IAA Investigators et al (2015) Usefulness of measuring both body mass index and waist circumference for the estimation of visceral adiposity and related cardiometabolic risk profile (from the INSPIRE ME IAA study). *Am J Cardiol* 115:307–315
35. Furet J-P, Kong L-C, Tap J et al (2010) Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: links with metabolic and low-grade inflammation markers. *Diabetes* 59:3049–3057
36. Aron-Wisniewsky J, Clement K (2014) The effects of gastrointestinal surgery on gut microbiota: potential contribution to improved insulin sensitivity. *Curr Atheroscler Rep* 16:454
37. Abdennour M, Reggio S, Le Naour G et al (2014) Association of adipose tissue and liver fibrosis with tissue stiffness in morbid obesity: links with diabetes and BMI loss after gastric bypass. *J Clin Endocrinol Metab* 99:898–907
38. Chikunguwo SM, Wolfe LG, Dodson P et al (2010) Analysis of factors associated with durable remission of diabetes after roux-en-Y gastric bypass. *Surg Obes Relat Dis* 6:254–259
39. Aminian A, Brethauer SA, Kashyap SR et al (2014) DiaRem score: external validation. *Lancet Diabetes Endocrinol* 2:12–13
40. Eliaschewitz FG, Tambascia MA (2012) Can we prevent beta cell apoptosis in type 2 diabetes? *Diabetes Metab Res Rev*. doi:10.1002/dmrr.2381
41. UK Prospective Diabetes Study Group (1995) UK prospective diabetes study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. *Diabetes* 44:1249–1258
42. Schauer PR, Bhatt DL, Kirwan JP, STAMPEDE Investigators et al (2017) Bariatric surgery versus intensive medical therapy for diabetes—5-year outcomes. *N Engl J Med* 376:641–651
43. Fried M, Yumuk V, Oppert JM, International Federation for Surgery of Obesity and Metabolic Disorders-European Chapter (IFSO-EC), European Association for the Study of Obesity (EASO), European Association for the Study of Obesity Obesity Management Task Force (EASO OMTF) et al (2014) Interdisciplinary European guidelines on metabolic and bariatric surgery. *Obes Surg* 24:42–55
44. Mechanick JL, Youdim A, Jones DB et al (2013) Clinical practice guidelines for the perioperative nutritional, metabolic, and nonsurgical support of the bariatric surgery patient – 2013 update: cosponsored by American Association of Clinical Endocrinologists, The Obesity Society, and American Society for Metabolic & bariatric surgery. *Obesity (Silver Spring)* 21(Suppl 1): S1–27