Supplemental Material - Type 2 diabetes remission after gastric bypass: what is the best prediction tool for clinicians?

Supplementary Information

Hayes et al. – Simple Logistic [1]

A subject is predicted as having type 2 diabetes (T2D) remission if :

$$\frac{e^{logit}}{1+e^{logit}} > 0.5 \text{, where } logit = 5.86 - 1.96 \times InsTT - 0.44 \times HbA1c$$

(InTT is 1 if the subject is under insulin therapy and 0 otherwise).

Hayes et al. – J48 decision trees [1]

T2D remission is predicted based on several bioclinical parameters : fasting glycemia, type of antidiabetic treatment, hypertension and HbA1c (**Supplementary Figure 2**).

Dixon et al. - Logistic regression for prediction of HbA1c ≤ 6 [2]

A subject is predicted as having HbA1c≤6 one year after surgery if :

$$\frac{e^{logit}}{1+e^{logit}} > 0.5 \text{, where } logit = -1.00 + 0.05 \times BMI - 2.70 \times \text{T2D Duration} + 0.38 \times \text{C - peptide}.$$

Dixon et al. - Logistic regression for prediction of HbA1c ≤ 7 [2]

A subject is predicted as having HbA1c≤7 one year after surgery if :

$$\frac{e^{logit}}{1+e^{logit}} > 0.5 \text{, where } logit = -0.065 - 0.15 \times \text{T2D Duration} + 1.1 \times \text{C - peptide}.$$

Lee et al. – Scoring system [3]

The score proposed by Lee et al. is defined as presented in **Supplementary Table 1**. In our cohort, figures had to be rounded to the right precision, especially BMI, on order to assign a score. Subjects with BMI 50 were assigned to a score of 3. C-peptide values below 0.9 were assigned to a score of 0.

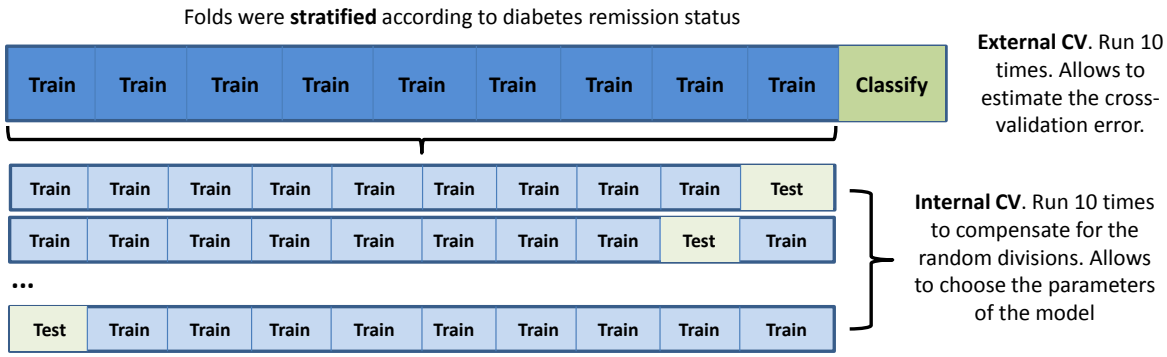
Still et al. – DiaRem score [4]

The score proposed by Still et al. is defined as presented in **Supplementary Table 2**.

Comment on predictive models

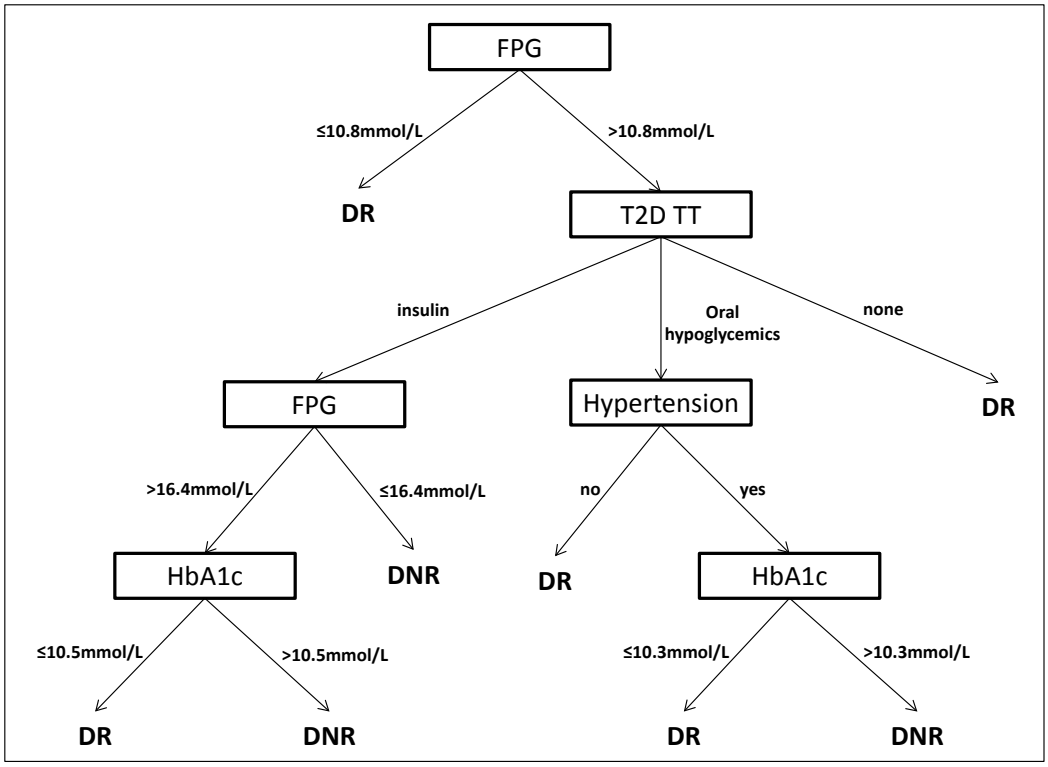
Decision trees separate the feature space into rectangles (for example: $x_1>15$ and $x_2<5$), while logistic regression is a generalized linear model and looks for a linear decision boundary (not necessarly parallel to the x or the y axis). Both methods are however prone to overfitting. The lasso regression used here is based on logistic regression but adds a penalty to the model which shrinks the coefficients of the variables and allows automatic variable selection. However, if two variables are correlated, it will only keep one of them. Elastic net is a combination of ridge regression (similar to lasso regression but with a different penalty not allowing variable selection) and lasso regression, and it allows to keep correlated variables in the model, but with smaller coefficients. The scores presented here were obtained from logistic regression (Lee at al. [3]) or Cox regression (Still et al. [4]) models. Cox regression models are time-to-event analysis, which can only be used if the time of T2D remission is known. The main advantage of this type of scoring is the simplicity of use.

Supplementary Figures



Supplementary Figure 1: Nested stratified 10×10 fold cross-validation

Adapted from Granholm et al. [5].



Supplementary Figure 2: J48 decision tree from Hayes et al. [1]

FPG: fasting plasma glucose; T2D TT: anti-diabetic treatment, DR: subject with T2D remission, DNR: subject with no T2D remission.

Supplementary Tables

Supplementary Table 1: Score proposed by Lee et al. [3]

Score	0	1	2	3
BMI (kg/m^2)	<30	30-39	40-49	>50
C-peptide (ng/mL)	0.9-1.9	2-9	4-6	>6
Duration of T2D (y)	>10	5-10	2-4.9	<2
Age (y)	≥40	<40		

Points from 0 to 3 are assigned for each variable depending on its value. These points are then summed up to obtain the final score. A higher score is associated with better chances of T2D remission.

Supplementary Table 2: Score proposed by Still et al. [4]

	Score
Age (y)	
<40	0
40-49	1
50-59	2
≥60	3
HbA1c (%)	
<6.5	0
6.5-6.9	2
7.0-8.9	4
≥9	6
Insulin therapy	
no	0
yes	10
Other diabetes drugs	
No sulfonylureas or insulin-sensitising agent other than metformin	0
Sulfonylureas and insulin-sensitising agent other than metformin	3

Points are assigned for each variable depending on its value. These points are then summed up to obtain the final score. A lower score is associated with better chances of T2D remission.

Supplementary References

1. Hayes MT, Hunt LA, Foo J, Tychinskaya Y, Stubbs RS. A model for predicting the resolution of type 2 diabetes in severely obese subjects following Roux-en Y gastric bypass surgery. Obes Surg. 2011;21:910‑6.

2. Dixon JB, Chuang L-M, Chong K, Chen S-C, Lambert GW, Straznicky NE, et al. Predicting the glycemic response to gastric bypass surgery in patients with type 2 diabetes. Diabetes Care. 2013;36:20‑6.

3. Lee W-J, Hur KY, Lakadawala M, Kasama K, Wong SKH, Chen S-C, et al. Predicting success of metabolic surgery: age, body mass index, C-peptide, and duration score. Surg Obes Relat Dis. 2013;9:379‑84.

4. Still CD, Wood GC, Benotti P, Petrick AT, Gabrielsen J, Strodel WE, et al. A probability score for preoperative prediction of type 2 diabetes remission following RYGB surgery. Lancet Diabetes Endocrinol. 2014;2:38‑45.

5. Granholm V, Noble W, Käll L. A cross-validation scheme for machine learning algorithms in shotgun proteomics. BMC Bioinformatics. 2012;13:S3.