876

EASD 2018
1–5 October
Berlin, Germany

A machine learning algorithm can identify clusters of patients with favorable glycemic outcomes in a pooled European Gla-300 studies (REALI): novel signposts for clinicians?

Insert QR code here

M. Rollot[1], M. Bonnemaire[2], C. Brulle-Wohlhueter[2], L. Pedrazzini[2], E. Boëlle-Le Corfec[2], G. Bigot[3], O. Hacman[1], A. Chiorean[1], R.C Bonadonna[4], P. Gourdy[5], D. Mauricio[6], D. Müller-Wieland[7], N. Freemantle[8]

[1] Quinten, Paris, France [2] Sanofi, Paris, France [3] IVIDATA, Paris, France [4]University of Parma School of Medicine and AOU of Parma, Italy [5]Metabolic Diseases and Nutrition, CHU Toulouse Hospital, France [6]Hospital de la Santa Creu i Sant Pau, Barcelona, Spain [7]University Clinic of the RWTH Aachen, Germany [8]University College London, UK

INTRODUCTION

The REALI project pools data from European (EU) study data bases and contains thousands of subjects with type 2 and type 1 diabetes mellitus (T2DM and T1DM) uncontrolled with their current antidiabetic therapy, initiated on Gla-300 and included in European Gla-300 clinical studies.

Machine learning approaches, together with traditional statistical methodologies, provides powerful tools for detecting potential patterns of response, and can lead to data-driven analysis identification of patient clusters of interest for clinicians.

OBJECTIVE

To undertake a proof-of-concept analysis of a subgroup discovery algorithm, leveraging data from the first study to be included in the REALI project. Patient-related variables from the Take Control study were used to identify clusters of patients who: 1. experienced HbA1c drop ≥ 0.5% from baseline to end of study (EoS); 2. experienced hypoglycemia event during the study (HEOS); 3. achieved the combined outcome of HbA1c < 7% at EoS without HEOS.

METHODS

Take Control, a multinational, randomized, 24-week interventional EU study of Gla-300 efficacy and safety with 2-parallel-group (patient- vs physician-managed titration) in T2DM patients, was the first study included in the pooled database.

Since Take Control was large and informed enough, it was an appropriate source of data for this proof-of-concept analysis. A total of 286 variables concerning the 631 patients enrolled in this study including baseline characteristics (concomitant medication, vital signs, medical history, laboratory tests, demographics) were analyzed.

Q-finder and machine learning process (Figure 1)

Q-finder was used to perform this analysis. It is a proprietary non-parametric supervised learning algorithm working with no particular assumptions regarding distributions of the outcome or explanatory variables. This algorithm explores the space of explanatory variables to identify areas where the outcome specified for the exploration shows higher or lower concentration than average.

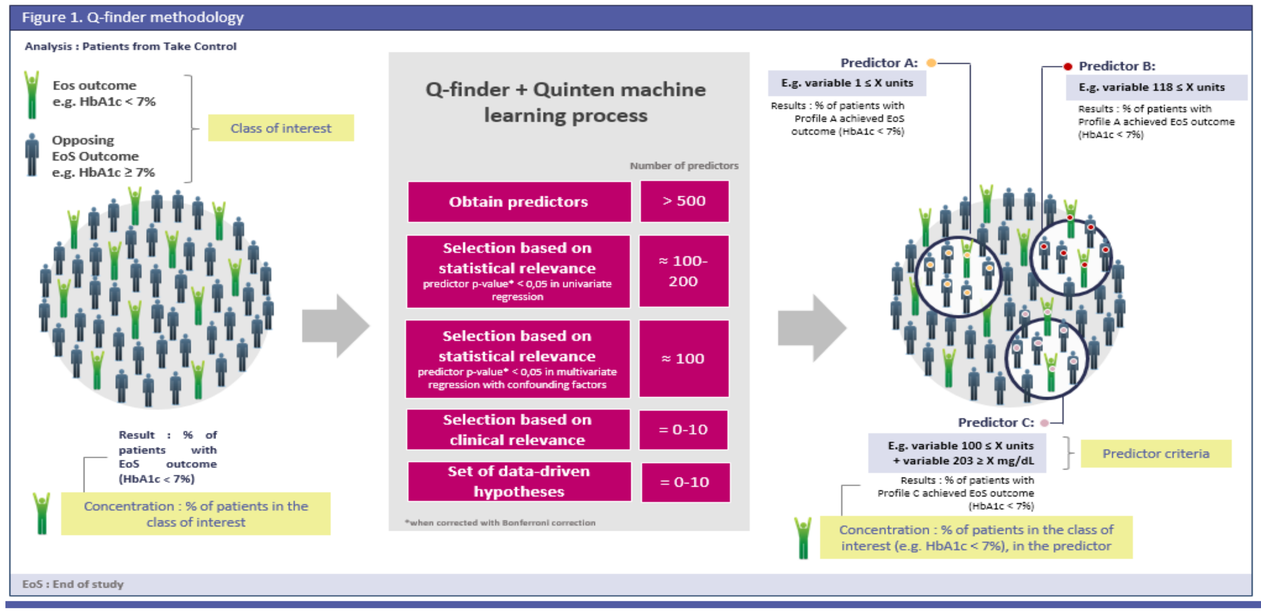
The output is a set of patient clusters, defined by combinations of variables and thresholds (predictors) with significantly higher/lower probability of experiencing an outcome of interest.

To account for multiple testing, predictors were considered statistically relevant if their p-value was < 0.05 in univariate regression when corrected–with Bonferroni correction, and in a multivariable logistic regression adjusting on the chosen confounding factors, i.e.:

- Baseline characteristics: HbA1c, age, BMI, prebreakfast SMPG, diabetes duration,
- Treatment group: physician or patient driven titration

Predictors were then verified with clinical experts for their relevance and assessed for their robustness with a bootstrap validation.

Figure 1. Q-finder methodology

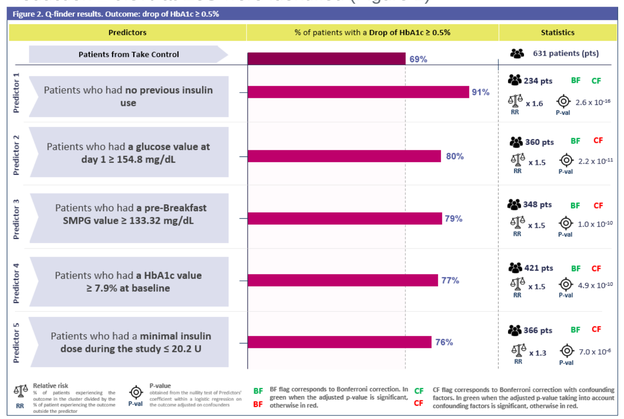


RESULTS

On thousands of queries performed by the algorithm, 32 were found statistically significant after Bonferroni adjustment in univariate and multivariable models adjusting for confounders, of which only those with clinical relevance are presented.

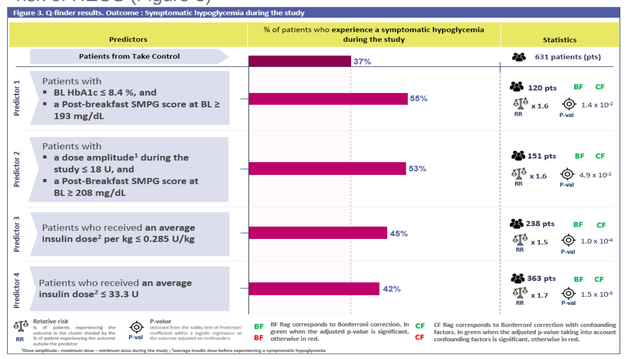
Five clusters of patients with higher probability of having an HbA1c reduction ≥ 0.5% at EoS were identified (Figure 2)

Figure 2. Q-finder results. Outcome: drop of HbA1c ≥ 0.5%



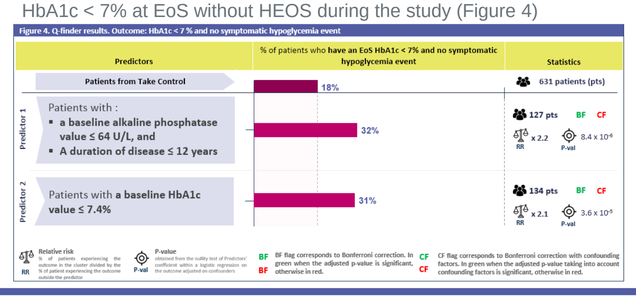
Additionally, our findings suggest four clusters of patients with a greater risk of HEOS (Figure 3)

Figure 3. Q-finder results. Outcome : Symptomatic hypoglycemia during the study



Two clusters of patients more likely to achieve the combined outcome of HbA1c < 7% at EoS without HEOS during the study (Figure 4)

Figure 4. Q-finder results. Outcome: HbA1c < 7 % and no symptomatic hypoglycemia event



DISCUSSION

- This first-stage analysis was a successful proof-of-concept of hypotheses generation using a machine learning algorithm conducted on the first study of the REALI project.
- This methodology can now be extended to the database pool in order to:
  - access to signals not detected due to small size of data sampling
  - assure a higher robustness of hypotheses by validating them into independent dataset not used for their generation
- Additionally, a few numbers of observations from this analysis emerged not sufficiently significant, but may become strong hypotheses when we will have the opportunity to test them on a larger dataset
- Following this proof-of-concept analysis, we will adjust the methodology more precisely in next analysis, especially regarding confounders, in order to assure stronger predictors.

CONCLUSIONS

The machine learning approach identified some patients' characteristics that may predict larger HbA1c reduction, higher risk of hypoglycemia or better chance to achieve glycemic control without hypoglycemia.

Machine learning can provide simple and efficient criteria for clinicians to identify clusters of patients who may benefit most from Gla-300 treatment.

Those criteria will be further validated by additional independent REALI datasets and new analysis will be performed in order to gain more robustness by accessing a larger number of observation and increase the likelihood to capture relevant signals.

Hence, this approach can inform individualized treatment and management of patients with T2DM.

DISCLOSURES AND ACKNOWLEDGEMENTS

M. Rollot — Quinten employee. M Bonnemaire – Sanofi employee. C. Brulle-Wohlhueter – Sanofi employee. L Pedrazzini – Sanofi employee. E. Boëlle-Le Corfec – Sanofi employee. G Bigot – IVIDATA employee on behalf of Sanofi O. Hacman — Quinten employee, A. Chiorean — Quinten employee. RC Bonadonna – **Speakers' bureau**: Sanofi, Merck, Sharp & Dohm, Bristol-Myers Squibb, AstraZeneca and Janssen; **Advisory panel**: Merck, Sharp & Dohm, Eli Lilly, Sanofi and Johnson & Johnson. P Gourdy – **Research**: AstraZeneca, Novo Nordisk and Sanofi; **Speakers' bureau and consultancy**: AstraZeneca, Bristol Myers Squibb, Boehringer Ingelheim, Eli Lilly, Janssen, Merck Sharp & Dohme, Novartis, Novo Nordisk, Sanofi, Servier and Takeda. D Mauricio – **Consultancy and/or speakers' bureau**: Almirall, Ascensia, AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Eli Lilly, Ferrer, Janssen, Menarini, Merck Sharp & Dohme, Novartis, Novo Nordisk and Sanofi. D Müller-Wieland – **Speakers' bureau and consultancy/advisory panel**: Amgen, AstraZeneca, Boehringer Ingelheim, MSD (Merck & Co.), Novartis, Novo Nordisk and Sanofi. N Freemantle – **Research, travel or consultancy**: Takeda, Pfizer, Biogen, Tesaro, Allergan, Ipsen and Sanofi.

The analysis presented in this poster was funded by Sanofi SA (France). Statistical data analysis was performed by Quinten.

The authors have received writing/editorial support from Dr T Rohban from Partner 4 Health (FR), funded by Sanofi SA (France)