Ibald-Mulli Angela (Orcid ID: 0000-0001-9006-9375)
Seufert Jochen (Orcid ID: 0000-0001-5654-7310)
Grimsmann Julia M (Orcid ID: 0000-0003-3627-1163)
Bramlage Peter (Orcid ID: 0000-0003-4970-2110)
Zhou Fang Liz (Orcid ID: 0000-0003-1353-6376)
Lanzinger Stefanie (Orcid ID: 0000-0002-3789-6459)

Identification of Predictive Factors of Diabetic Ketoacidosis in Type 1 Diabetes Using a Subgroup Discovery Algorithm

Running title: Predictive factors of DKA in T1D

Angela Ibald-Mulli[1], Jochen Seufert[2], Julia M. Grimsmann[3,4], Markus Laimer[5], Peter Bramlage[6], Alexandre Civet[7], Margot Blanchon[7], Simon Gosset[7], Alexandre Templier[7], W. Dieter Paar[8], Fang Liz Zhou[9], Stefanie Lanzinger[3,4]

[1]Real World Evidence and Clinical Outcome Generation, Sanofi, Paris, France

[2]Division of Endocrinology and Diabetology, Department of Medicine II, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[3]Institute of Epidemiology and Medical Biometry, ZIBMT, Ulm University, Ulm, Germany

[4]German Centre for Diabetes Research (DZD), München-Neuherberg, Germany

[5]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[6]Institute for Pharmacology and Preventive Medicine, Cloppenburg, Germany

[7]Quinten, Paris, France

[8]Sanofi-Aventis Deutschland GmbH, Berlin, Germany

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/dom.15039

[9]Sanofi, Bridgewater, USA

Corresponding author: Angela Ibald-Mulli, aibaldmulli@gmail.com, +491605823132

Journal: Diabetes, Obesity and Metabolism

Word count: 2454/ 3500 max

Abstract word count: 229 / 250 max

No. of references: 24 / 60 max

No. of figures/tables: 4 (2 tables, 2 figures) / 5 max

No. of supplementary figures/tables: 5 (4 tables, 1 figure)

Abstract

Aims

Diabetic ketoacidosis (DKA) is a serious and potentially fatal complication of type 1 diabetes and it is difficult to identify individuals at increased risk. The aim of this study was to identify predictive factors for DKA by retrospective analysis of registry data and use of a subgroup discovery algorithm.

Materials and Methods

Data from adults and children with type 1 diabetes and >2 diabetes-related visits were analyzed from the Diabetes Prospective Follow-up Registry. Q-Finder®, a supervised non-parametric proprietary subgroup discovery algorithm, was used to identify subgroups with clinical characteristics associated with increased DKA risk. DKA was defined as pH <7.3 during a hospitalization event.

Results

Data for 108,223 adults and children, of whom 5,609 (5.2%) had DKA, were studied. Q-Finder® analysis identified 11 profiles associated with increased risk of DKA: low body mass index standard deviation score; DKA at diagnosis; age 6–10 years; age 11–15 years; HbA_{1c} ≥8.87 [73 mmol/mol]; no fast-acting insulin intake; age <15 years and not using a continuous glucose monitoring system; physician diagnosis of nephrotic kidney disease; severe hypoglycemia; hypoglycemic coma; and autoimmune thyroiditis. Risk of DKA increased with number of risk profiles matching patients' characteristics.

Conclusions

Q-Finder® confirmed common risk profiles identified by conventional statistical methods and allowed the generation of new profiles that may help predict patients with type 1 diabetes who are at a greater risk of experiencing DKA.

Introduction

Diabetic ketoacidosis (DKA) is a serious and potentially fatal complication of type 1 diabetes, characterized by hyperglycemia, metabolic acidosis, and ketosis.[1] It is caused by relative or absolute insulin deficiency, leading to accelerated gluconeogenesis, glycogenolysis, impaired glucose utilization in peripheral tissues, ketogenesis, and lipolysis.[1, 2] Clinical management of DKA involves administration of fluids, electrolytes and insulin.[2]

Several risk factors for DKA, defined by pH <7.3 and/or DKA-related hospitalization, have been identified based on traditional analyses of the Diabetes Prospective Follow-up Registry (DPV), which includes data of predominantly German and Austrian patients with diabetes.[83] Identified risk factors for DKA include: age 11–15 years old at diabetes onset; lower mean body mass index-standard deviation score; higher HbA_{1c}; migration background (e.g. place of birth for ≥1 parent outside of Germany or Austria for the DPV); and higher mean number of hospitalisations.[3, 4]

Despite knowledge of these risk factors, DKA remains a significant cause of morbidity and mortality in people with diabetes, with hospital admissions in the US reaching 160 000 in 2017.[2] Annual incidence of DKA is reported as 8 episodes per 100 person-years in children (<18 years) and 2.5 per 100 person-years in adults,[4, 5] though it is possible that actual values may be higher given the difficulty in diagnosing DKA due to variability in symptoms, particularly in young children and those who have not been previously diagnosed with type 1 diabetes.[6, 7] Mortality rates in pediatric patients with DKA are between 0.15 and 0.31%, and cerebral edema accounts for up to 87% of all DKA-related deaths.[3]

Better knowledge of patient profiles associated with increased risk for experiencing DKA will provide useful information to physicians and could lead to improvements in the prevention and management of diabetes. Advanced analytics can supplement the conventional statistics described above to generate new insights on DKA risk factors using large diabetes cohorts such as the DPV. Q-Finder® (Quinten, Paris, France), is a subgroup discovery algorithm, which identifies patient profiles associated with an outcome of interest and the specific combinations of distinct characteristics for each profile.[9,10,11] The objective of this study was to identify predictive factors for DKA in type 1 diabetes by conducting a retrospective analysis of data from the DPV using a novel analytical approach leveraging the Q-Finder® algorithm.

Materials and Methods

Data Source

This was a retrospective cohort analysis of adults and children (<18 years) with type 1 diabetes in the DPV registry, from Germany and Austria. The analysis was performed in March 2019 and included patients with a diagnosis of type 1 diabetes with at least two diabetes-related visits in the full database. The index visit was defined as the first type 1 diabetes-related visit recorded in the DPV, since it was first established in 1995. The follow-up period was from day 10 after index date until a DKA event or the last recorded visit. Ethical approval was not required as this was a retrospective analysis of de-identified registry data; however, the overall DPV initiative was approved by the Ethics Committee of Ulm University and by the local Review Boards of participating centers.

Outcome for Subgroup Discovery

The outcome of interest for subgroup discovery was occurrence of any DKA (yes/no) defined as pH <7.3 during a hospitalization event throughout follow-up. A descriptive analysis of study variables and covariates at the index visit was performed for patients with and without DKA.

Q-Finder® Subgroup Discovery Analysis

The analysis was performed using Q-Finder® (Quinten, Paris, France), a supervised non-parametric proprietary subgroup discovery algorithm for which detailed methodology and algorithm principles have been described previously.[11] Prior to initiating the analysis, the full dataset was divided into a learning dataset (50%), test dataset (25%) and validation dataset (25%). The datasets were created by random sampling, stratified by DKA, age, gender, migration background (defined as a patient, or at least 1 parent of the patient, born outside of Germany, Austria, Switzerland or Luxembourg), HbA_{1c}, and follow-up duration variables. The learning dataset was first used to generate candidate profiles using 126 clinical variables (Supplementary Table 1) potentially associated with a higher frequency of DKA, based on the clinical and technical experience of a committee of 6 experts (from Ulm University and Sanofi). Identified profiles were then ranked according to statistical criteria; profiles that passed a higher number of these criteria, or scored higher on a given criteria (i.e. greater effect size or statistical significance) were described as more reliable and thus ranked higher.[11] With duplicates removed, profiles that passed these statistical criteria were tested on the test dataset, using the same criteria as the learning dataset, and an additional multiple test correction (Benjamini-Hochberg). Ranked profiles were then reviewed by the expert committee for clinical

relevance and then validated on the validation dataset. The criteria for this step-wise process are shown in Supplementary Figure 1.

Sensitivity analysis

Sensitivity analyses were performed to assess the robustness of the findings with DKA defined by the selected parameter (i.e. pH <7.3 during a hospitalization event), as compared with DKA defined by other parameters: (1) pH <7.3 or bicarbonate <15 mmol/L; (2) blood glucose >13.9 mmol/L (250 mg/dL) and pH <7.3 and bicarbonate <15 mmol/L and ketonemia and/or ketonuria (based on Deutsche Diabetes Gesellschaft/ German Diabetes Society [DDG] guidelines); and (3) blood glucose >11 mmol/L (200 mg/dL) and pH <7.3 and/or bicarbonate <15 mmol/L and ketonemia and/or ketonuria (based on International Society for Pediatric and Adolescent Diabetes [ISPAD] guidelines).

Output

The output comprised a set of patient profiles, each defined by either a criterion or a combination of criteria that characterized a subgroup with higher rates of DKA than the rest of the population (odds ratio [OR] >1; higher rate of DKA in patients defined by the profile versus patients with DKA not defined by the profile). Profiles were adjusted for the following confounding factors: age, gender, migration status and HbA1c level (within 3 months prior to a DKA event or at the end of the study for those who did not experience DKA), and follow-up duration.

Results

As of March 2019, the DPV included records for 516,091 people with diabetes. In total, 126,235 had a diagnosis of type 1 diabetes, of whom 83,905 were children (<18 years). The dataset analyzed comprised data for 108,223 patients with type 1

diabetes and >2 diabetes-related visits. Of the 108,223 adults and children in the dataset, 5.2% (5,609) experienced DKA.

Descriptive Analysis

Individuals with DKA were younger, had lower body mass index standard deviation score (BMI-SDS), higher HbA_{1c} level, an earlier onset of diabetes, and a shorter disease duration versus non-DKA patients (Table 1). Those that experienced DKA were also more likely to have a migration background, lower levels of supervised physical activity (e.g. club sport or attendance at fitness centre), and required more medical consultations versus those who did not experience DKA. For the full list of 126 clinical variables considered, refer to Supplementary Table 1.

Q-Finder® Analysis

The initial data-driven process generated approximately 120,000 candidate profiles, which were ranked using statistical measures to enable the top 45 profiles to be selected. The number of profiles was reduced to 15 during the testing phase, based on statistical factors, correction for multiple testing and the selection of relevant profiles by clinical experts. During the validation step, 11 profiles were confirmed to be associated with increased risk of DKA: low BMI-SDS (≤−1.69 adjusted for age); DKA at onset; children aged 6–10 years; children aged 11–15 years; HbA_{1c} ≥8.87 [73 mmol/mol] within 3 months prior to a DKA event (or by last visit for patients without DKA); no fast-acting insulin intake; age <15 years and not using a continuous glucose monitoring system (CGMS); patients diagnosed with nephrotic kidney disease; severe hypoglycemia; hypoglycemic coma between index visit and DKA event (or by last visit for patients without DKA); and autoimmune thyroiditis. The association of these profiles with the risk of DKA in the full dataset is presented in Figure 1.

Three profiles covered approximately 20% of the full dataset: HbA_{1c} ≥8.87% (23%), no fast-acting insulin (20%) and age between 11 and 15 years (19%). The profiles with the lowest coverage of ≤3% of the dataset were severe hypoglycemia (3%), diagnosis of nephrotic kidney disease (1%), and hypoglycemic coma (1%) (Figure 1). The findings with the selected clinical profiles were consistent across the learning, validation, and testing steps (Table 2).

Patients with a higher number of the profiles described above were at increased risk of having DKA (Figure 2); however, fewer patients were likely to have a higher number of profiles. For example, 64.7% of patients (n=70,018) had at least one profile, with a DKA rate of 7.1%. By comparison, less than 5% of patients had at least three profiles, but 19.5% of them had DKA. None of the patients had more than seven profiles.

Sensitivity analysis

Using the alternative definitions of DKA, 5.8% (6,266) of patients experienced DKA as defined by pH <7.3 or bicarbonate <15 mmol/L, 0.4% (448) of patients experienced DKA according to the DDG definition, and 2.7% (2,974) of patients experienced DKA according to the ISPAD definition. The profiles were generally comparable when using the different definitions (Supplementary tables 2–4). However, with the DDG definition, the profile of no short acting insulin is no longer associated with an increased risk of DKA, while the profiles for HbA_{1c} ≥8.87 [73 mmol/mol], hypoglycemia with coma, and severe hypoglycemia are no longer statistically significant (Supplementary table 3). With the ISPAD definition, the profile of severe hypoglycemia is no longer statistically significant (Supplementary Table 4).

Discussion

The descriptive statistics in this analysis provided characteristics for patients with and without DKA that were consistent with previously published data. Patients with DKA had higher HbA_{1c}, lower BMI, were younger, and were more likely to have a migration background.[4, 12, 13] The finding that DKA is associated with suboptimal glycemic control (higher HbA_{1c}) in proximity to the event is expected and has been reported previously.[4, 12, 13] The association with being underweight may reflect the link between having a higher BMI and greater residual β-cell function that has previously been reported in children with type 1 diabetes.[14] The higher likelihood of DKA in people with a migrant background may reflect the lower household incomes in this group, or issues around accessing appropriate diabetes educational materials.[4]

Using the Q-finder® analysis, 11 profiles were confirmed to be associated with an increased risk of DKA. The higher risk of DKA in the 11–15 years vs 6–10 years age groups may be explained by a combination of biological and social factors. Insulin resistance develops during puberty and may contribute to suboptimal glycemic control during adolescence.[15] Furthermore, diabetes-related stigma and fear of hypoglycemia have been associated with suboptimal glycemic control in adolescents.[16,17] As children get older and gain more autonomy in their own diabetes management,[18] they may engage in behaviors to avoid stigma and hypoglycemia such as reduced blood glucose monitoring and poor compliance to insulin therapy, which negatively impact glycemic control and may increase the risk of DKA.[19, 20]

The finding that autoimmune thyroiditis is associated with an increased risk of DKA may reflect findings that thyroiditis is a sign of more aggressive autoimmune responses in familial vs non-familial type 1 diabetes.[21] Familial type 1 diabetes tends

to manifest earlier and is associated with significantly higher prevalence of anti-thyroid antibodies and a closer link with thyroiditis versus non-familial type 1 diabetes.[21] Thyroiditis may therefore potentially be a marker for people with type 1 diabetes who have >1 autoimmune conditions and more aggressive type 1 diabetes. Additionally, there are links between DKA and thyroid function that may be related to a variety of factors including insufficient insulin secretion, highly fluctuating blood glucose, induction of pro-inflammatory cytokines, and cell hypoxia.[22]

The analysis also identified novel potential DKA risk factors that do not appear to have been described previously, such as nephrotic kidney disease, hypoglycemia with coma, severe hypoglycemia, no fast-acting insulin, and age <15 years with no CGMS experience. These newly identified risk factors provide valuable hypotheses that can be tested in further studies on independent datasets. Interestingly, these risk factors may also reflect sub-optimal glycemic control (nephrotic kidney disease, hypoglycemia with coma, severe hypoglycemia) or lack of experience with, or availability of tools to improve, glycemic control (age <15 years with no CGMS experience).

The fact that matching more than one of the selected profiles is associated with an increased risk of experiencing DKA supports their clinical relevance. The robustness of these results derives from the large multicenter population-based database that provided source data and the generation, testing, and validation of the clinical profiles in independent datasets. In addition, the analysis was controlled for confounding factors and the effects of multiple testing. In the current analysis, DKA was defined by pH <7.3, as per previous analyses of the DPV registry. Sensitivity analyses were performed using three other definitions of DKA based on updated guidelines that include parameters for bicarbonate levels, blood glucose, and

ketonemia and/or ketonuria. Results were generally consistent regardless of which definition of DKA was used; however, the DDG definition appeared to be too restrictive, identifying only 0.4% with DKA during follow-up.

This study is limited in some aspects. The results are limited to this specific population from Germany and Austria, therefore further validation of these results in an independent database with patients from a more diverse population would provide additional confidence in the relevance of the clinical profiles. A sub-analysis of the data would also be of interest to explore whether these identified profiles are more strongly associated with DKA risk in adults versus children. Another limitation is the need for human-processed data as well as the input from the committee of experts to be applied to the algorithm, meaning that some degree of subjectivity or bias cannot be discounted from the identification of the profiles. Additionally, the dataset analyzed lacked information on some factors that may be associated with DKA risk. For example, education level was not available in the dataset, and increasing educational attainment is associated with lower HbA_{1c} so may impact the risk of DKA.[23] Non-adherence to insulin treatment was also not captured in the dataset, which may highlight if patients were non-adherent to their therapy, such as insulin, which has been described as a risk factor for DKA in other studies.[12, 13]

In conclusion, this subgroup detection algorithm confirmed risk profiles for DKA that were identified previously using conventional statistical methods. Furthermore, the analysis allowed the generation of new profiles that may help healthcare professionals identify patients at greater risk of DKA. These profiles will need to be investigated further to better understand how they influence the development of DKA, which may help inform on how DKA risk can be managed in adults and children with type 1 diabetes.

Acknowledgments

Editorial assistance was provided by Arthur Holland, PhD, of Fishawack Communications Ltd, part of Fishawack Health, funded by Sanofi.

Some information in this article was presented at the EASD 2019 Virtual Meeting, 18 September 2019 (Oral Presentation # 117).

Conflict of Interest Statement

J.G. and S.L. have no conflicts of interest to report. J.S. is a consultant, speaker and principal trial investigator for Sanofi, on topics unrelated to the analysis reported in this article. P.B. is a consultant to Sanofi on topics unrelated to the analysis reported in this article. A.C., M.B., S.G., and A.T. are/were employees of Quinten France at the time of this analysis. W.D.P. is an employee of, and holds stocks in, Sanofi. A.I.-M. was an employee of and stockholder in Sanofi at the time of this analysis. F.L.Z. is an employee of and stockholder in Sanofi.

Data availability

Qualified researchers may request access to patient-level data and related documents. Patient-level data will be anonymized, and study documents will be redacted to protect the privacy of trial participants. Further details on Sanofi's data sharing criteria, eligible studies, and process for requesting access can be found at https://www.vivli.org

Author Contributions

Design: A. Ibald-Mulli, D. Paar, F.L. Zhou

Conduct/data collection: A. Civet, M. Blanchon, S. Gosset, and A. Tempelier

Analysis: All authors

Writing manuscript: All authors

References

1. Holt RIG, DeVries JH, Hess-Fischl A, et al. The management of type 1 diabetes in adults. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes care. 2021;44(11)

2. Karslioglu French E, Donihi AC, Korytkowski MT. Diabetic ketoacidosis and hyperosmolar hyperglycemic syndrome: review of acute decompensated diabetes in adult patients. BMJ. 2019;365:l1114.

3. Fritsch M, Rosenbauer J, Schober E, et al. Predictors of diabetic ketoacidosis in children and adolescents with type 1 diabetes. Experience from a large multicentre database. Pediatr Diabetes. 2011;12(4 Pt 1):307-312.

4. Kalscheuer H, Seufert J, Lanzinger S, et al. Event rates and risk factors for the development of diabetic ketoacidosis in adult patients with type 1 diabetes: analysis from the DPV Registry based on 46,966 patients. Diabetes care. 2019;42(3):e34-e36.

5. Baumer-Mouradian SH, Gray MP, Wolfgram PM, et al. Improving emergency department management of diabetic ketoacidosis in children. Pediatrics. 2019;144(4):e20182984.

6. Fayfman M, Pasquel FJ, Umpierrez GE. Management of hyperglycemic crises: diabetic ketoacidosis and hyperglycemic hyperosmolar state. Med Clin North Am. 2017;101(3):587-606.

7. Wolfsdorf J, Glaser N, Sperling MA, American Diabetes A. Diabetic ketoacidosis in infants, children, and adolescents: A consensus statement from the American Diabetes Association. Diabetes care. 2006;29(5):1150-1159.

8. Hofer SE, Schwandt A, Holl RW, Austrian/German DPVI. Standardized documentation in pediatric diabetology: experience from austria and germany. J Diabetes Sci Technol. 2016;10(5):1042-1049.

9. Zhou FL, Watada H, Tajima Y, et al. Identification of subgroups of patients with type 2 diabetes with differences in renal function preservation, comparing patients receiving sodium-glucose co-transporter-2 inhibitors with those receiving dipeptidyl peptidase-4 inhibitors, using a supervised machine-learning algorithm (PROFILE study): A retrospective analysis of a Japanese commercial medical database. Diabetes Obes Metab. 2019;21(8):1925-1934.

10. Dumontet C, Hulin C, Dimopoulos MA, et al. A predictive model for risk of early grade >/= 3 infection in patients with multiple myeloma not eligible for transplant: analysis of the FIRST trial. Leukemia. 2018;32(6):1404-1413.

11. Esnault C, Gadonna M-L, Queyrel M, Templier A, Zucker J-D. Q-Finder: an algorithm for credible subgroup discovery in clinical data analysis — An application to the International Diabetes Management Practice Study (IDMPS). Front Artif Intell. 2020;3:559927.

12. Del Degan S, Dube F, Gagnon C, Boulet G. Risk factors for recurrent diabetic ketoacidosis in adults with type 1 diabetes. Can J Diabetes. 2019;43(7):472-476 e471.

13. Al-Obaidi AH, Alidrisi HA, Abbas Ali Mansour AA. Precipitating factors for diabetic ketoacidosis among patients with type 1 diabetes mellitus: the effect of socioeconomic status. Int J Diabetes Metab 2019;25:52–60.

14. Redondo MJ, Rodriguez LM, Escalante M, O'Brian Smith E, Balasubramanyam A, Haymond MW. Beta cell function and BMI in ethnically diverse children with newly diagnosed autoimmune type 1 diabetes. Pediatr Diabetes. 2012;13(7):564-571.

15. Amiel SA, Sherwin RS, Simonson DC, Lauritano AA, Tamborlane WV. Impaired insulin action in puberty. A contributing factor to poor glycemic control in adolescents with diabetes. N Engl J Med. 1986;315(4):215-219.

16. Brazeau AS, Nakhla M, Wright M, et al. Stigma and its association with glycemic control and hypoglycemia in adolescents and young adults with type 1 diabetes: cross-sectional study. J Med Internet Res. 2018;20(4):e151.

17. Johnson SR, Cooper MN, Davis EA, Jones TW. Hypoglycaemia, fear of hypoglycaemia and quality of life in children with Type 1 diabetes and their parents. Diabet Med. 2013;30(9):1126-1131.

18. Markowitz JT, Garvey KC, Laffel LM. Developmental changes in the roles of patients and families in type 1 diabetes management. Curr Diabetes Rev. 2015;11(4):231-238.

19. Rausch JR, Hood KK, Delamater A, et al. Changes in treatment adherence and glycemic control during the transition to adolescence in type 1 diabetes. Diabetes care. 2012;35(6):1219-1224.

20. Morris AD, Boyle DI, McMahon AD, Greene SA, MacDonald TM, Newton RW. Adherence to insulin treatment, glycaemic control, and ketoacidosis in insulin-dependent diabetes mellitus. The DARTS/MEMO Collaboration. Diabetes Audit and Research in Tayside Scotland. Medicines Monitoring Unit. Lancet. 1997;350(9090):1505-1510.

21. Alyafei F, Soliman A, Alkhalaf F, et al. Clinical and biochemical characteristics of familial type 1 diabetes mellitus (FT1DM) compared to non-familial type 1 DM (NFT1DM). Acta Biomed. 2018;89(S5):27-31.

22. Xing Y, Chen J, Song G, Zhao L, Ma H. Impact of Diabetic Ketoacidosis on Thyroid Function in Patients with Diabetes Mellitus. Int J Endocrinol. 2021;2021:2421091.

23. Olesen K, AL FR, Joensen L, et al. Higher health literacy is associated with better glycemic control in adults with type 1 diabetes: a cohort study among 1399 Danes. BMJ Open Diabetes Res Care. 2017;5(1):e000437.

24. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289-300.

Tables

Table 1. Patient characteristics at the index visit according to occurrence of DKA events (Yes/No)

Patient Characteristic	DKA n=5,609 (5.2%)		No DKA n=102,614 (94.8%)	
	Mean	SD	Mean	SD
Age, years	14.7	9.0	28.5	18.5
BMI-SDS (AGA data)	0.1	1.1	0.3	1.2
HbA_{1c} (MOM-DCCT, [%])	9.3	2.0	8.0	1.8
Age at onset, years	8.8	6.4	15.3	14.1
Disease duration, years	5.9	5.9	10.4	10.8
Migration background, %	19.8	n/a	12.0	n/a
Supervised physical activity, units/wk	0.8	1.5	1.5	3.0
Medical visits, n	42	29	23	23.6

AGA, American Gastroenterological Association; BMI-SDS, body mass index-SD score; DKA, diabetic ketoacidosis; MOM-DCCT, multiple of the mean transformation

method - Diabetes Control and Complications Trial; n/a, not applicable; SD, standard deviation; wk, week.

Migration background defined as ≥1 parent of non-German or non-Austrian heritage.

Table 2. Statistical indicators of selected profiles during the learning, validation, and testing steps

Variable	Threshold	Learning (N=54,111) OR (95% CI)	Validation (N=27,056) OR (95% CI)	Testing (N=27,056) OR (95% CI)
Standardized HbA_{1c}, %	≥8.87 %	4.22 (2.74–6.55)	4.55 (2.43–8.68)	5.12 (2.83–9.46)
Age range, years	6–10	1.58 (1.35–1.84)	2.42 (1.97–2.97)	2.17 (1.76–2.65)
Age range, years	11–15	4.05 (3.65–4.50)	3.77 (3.25–4.38)	3.46 (2.99–4.01)
Diagnosis of nephrotic kidney disease	Yes	3.59 (2.64–4.81)	2.87 (1.88–4.26)	2.44 (1.59–3.63)
DKA at onset	Yes	1.57 (1.37–1.80)	1.99 (1.64–2.39)	1.53 (1.26–1.86)
Hypoglycemia with coma	≥1 event	6.36 (4.32–9.17)	7.69 (4.36–12.98)	6.48 (3.98–10.22)
Hypoglycemia no coma	≥1 event	2.35 (1.73–3.12)	3.52 (2.32–5.19)	2.65 (1.78–3.83)
Diagnosis of thyroiditis	Yes	1.99 (1.67–2.37)	1.81 (1.39–2.33)	1.35 (1.02–1.76)
Standardized BMI	≤-1.69	2.75 (2.10–3.56)	2.41 (1.60–3.52)	2.57 (1.77–3.64)
Fast-acting insulin intake	No	1.55 (1.39–1.74)	1.42 (1.21–1.67)	1.44 (1.23–1.69)
Age, years	<15.3	7.66 (6.13–9.62)	7.77 (5.59–10.95)	7.89 (5.68–11.07)

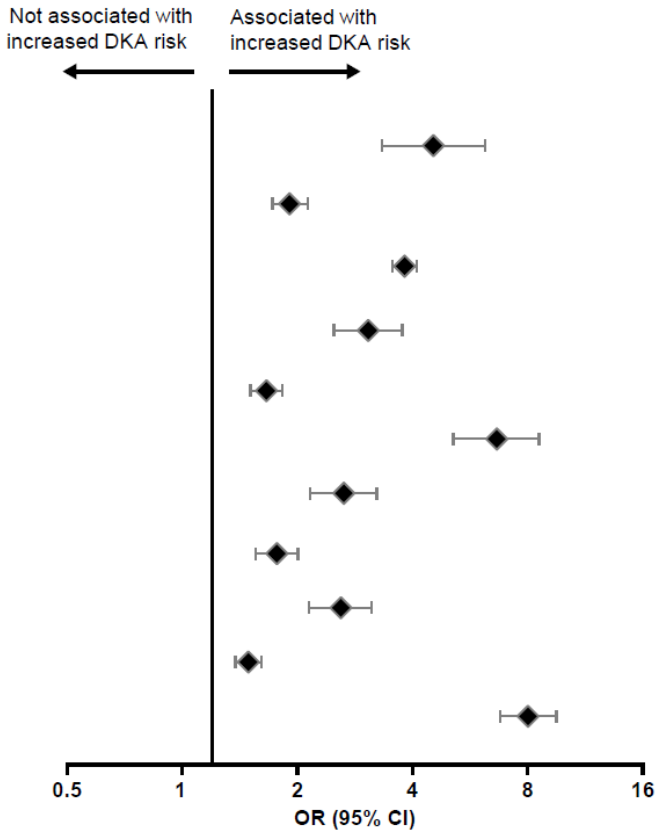
+ CGMS days	0			

BMI, body mass index; CGMS, continuous glucose monitoring system; CI, confidence interval; DKA, diabetic ketoacidosis; OR, odds ratio.

Figures

Figure 1. Association of validated profiles with the risk of DKA in the full dataset

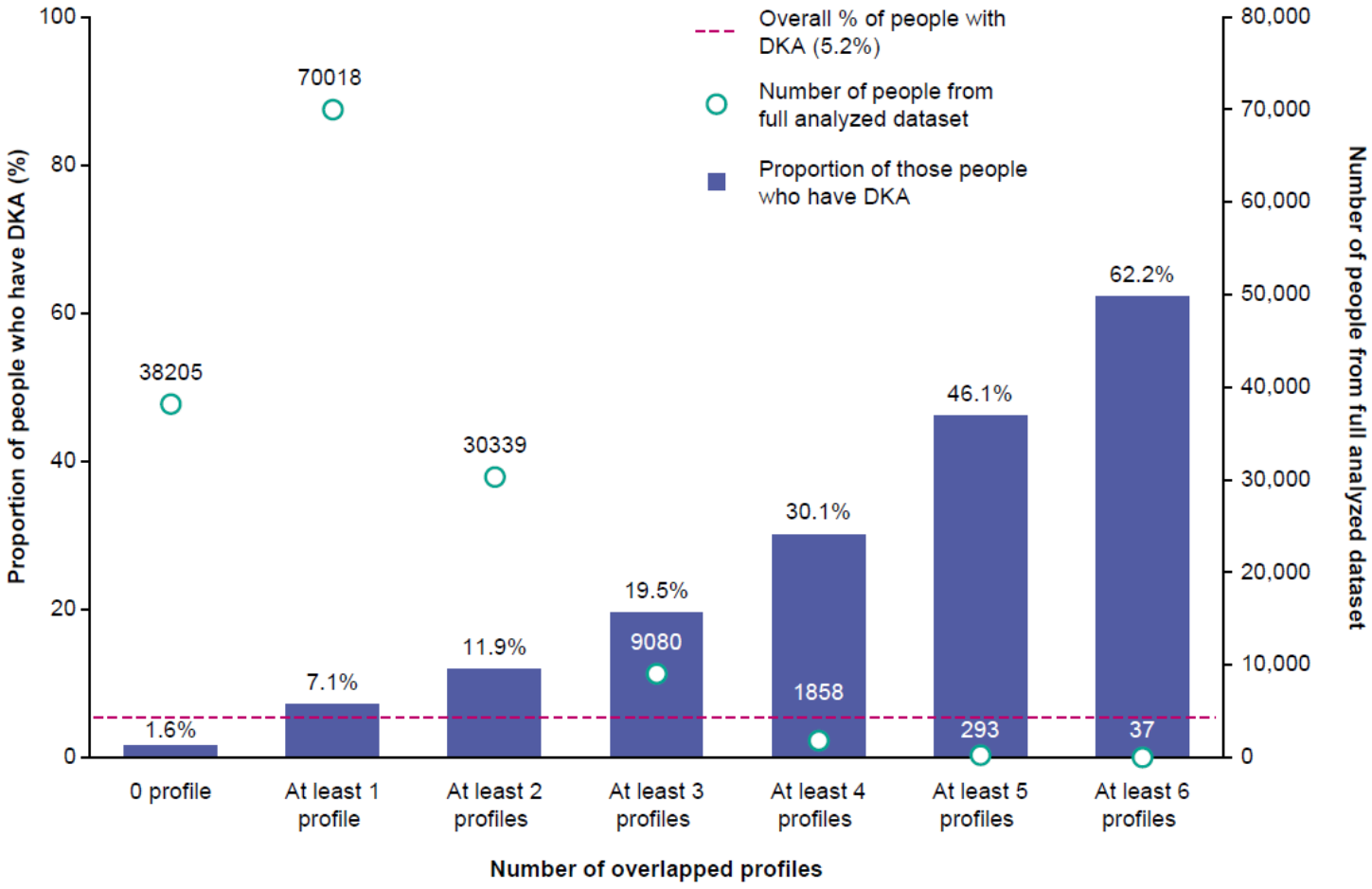
Profile	Size (n)	Coverage (%)	OR (95% CI)†	p-value‡	DKA in the profile (%)	DKA outside the profile (%)
Standardized HbA_{1c} >8.87 %	24,544	23	4.54 (3.34–6.19)	8.5E-22	6.9	2.2
Age range 6 to 10 years	8,656	8	1.91 (1.72–2.13)	6.0E-33	8.9	4.9
Age range 11 to 15 years	20,966	19	3.82 (3.55–4.11)	2.2E-276	13.2	3.3
Diagnosis of nephrotic kidney disease	1,467	1	3.07 (2.49–3.77)	7.4E-26	12.9	5.1
DKA at onset	8,709	8	1.66 (1.51–1.83)	1.5E-24	9.9	4.8
Hypoglycemia with coma, ≥1 event	1,425	1	6.65 (5.11–8.57)	1.6E-46	10.5	4.8
Severe hypoglycemia, ≥1 event	3,328	3	2.65 (2.16–3.23)	2.5E-21	6.7	4.8
Diagnosis of thyroiditis	6,353	6	1.77 (1.56–2.01)	2.2E-18	7.1	5.1
Standardized BMI ≤-1.69	4,862	4	2.60 (2.15–3.13)	1.7E-23	6.7	5.2
No fast-acting insulin	21,649	20	1.49 (1.38–1.61)	8.9E-23	9.6	5.7
Age <15.3 years and no CGMS	9,669	9	8.03 (6.79–9.52)	8.6E-129	16.3	2.5



BMI, body mass index; CGMS, continuous glucose monitoring system; CI, confidence interval; DKA, diabetic ketoacidosis; OR, odds ratio.

†Effect size criterion corrected for confounders; ‡Effect significance criterion corrected for confounders. Confounding factors: age, gender, migration status, last recorded HbA_{1c} within 3 months prior to a DKA event or at the end of the study for those who did not experience DKA, and follow-up duration.

Figure 2. Risk of DKA according to the number of overlapped profiles a patient belongs to



DKA, diabetic ketoacidosis.