



Universiteit
Leiden
The Netherlands

Potential value of identifying type 2 diabetes subgroups for guiding intensive treatment: a comparison of novel data-driven clustering with risk-driven subgroups

Li, X.Y.; Giessen, A. van; Altunkaya, J.; Slieker, R.C.; Beulens, J.W.J.; Hart, L.M. 't; ... ; Leal, J.

Citation

Li, X. Y., Giessen, A. van, Altunkaya, J., Slieker, R. C., Beulens, J. W. J., Hart, L. M. 't, ... Leal, J. (2023). Potential value of identifying type 2 diabetes subgroups for guiding intensive treatment: a comparison of novel data-driven clustering with risk-driven subgroups. *Diabetes Care*, 46(7), 1395-1403. doi:10.2337/dc22-2170

Version: Publisher's Version
License: [Leiden University Non-exclusive license](#)
Downloaded from: <https://hdl.handle.net/1887/3764306>

Note: To cite this publication please use the final published version (if applicable).



Potential Value of Identifying Type 2 Diabetes Subgroups for Guiding Intensive Treatment: A Comparison of Novel Data-Driven Clustering With Risk-Driven Subgroups

Diabetes Care 2023;46:1395–1403 | <https://doi.org/10.2337/dc22-2170>

Xinyu Li,¹ Anouk van Giessen,²
James Altunkaya,³ Roderick C. Sliker,^{4,5,6}
Joline W.J. Beulens,^{4,5,7}
Leen M. 't Hart,^{4,5,6,8} Ewan R. Pearson,⁹
Petra J.M. Elders,^{5,10}
Talitha L. Feenstra,^{1,2} and Jose Leal³

OBJECTIVE

To estimate the impact on lifetime health and economic outcomes of different methods of stratifying individuals with type 2 diabetes, followed by guideline-based treatment intensification targeting BMI and LDL in addition to HbA_{1c}.

RESEARCH DESIGN AND METHODS

We divided 2,935 newly diagnosed individuals from the Hoorn Diabetes Care System (DCS) cohort into five Risk Assessment and Progression of Diabetes (RHAPSODY) data-driven clustering subgroups (based on age, BMI, HbA_{1c}, C-peptide, and HDL) and four risk-driven subgroups by using fixed cutoffs for HbA_{1c} and risk of cardiovascular disease based on guidelines. The UK Prospective Diabetes Study Outcomes Model 2 estimated discounted expected lifetime complication costs and quality-adjusted life-years (QALYs) for each subgroup and across all individuals. Gains from treatment intensification were compared with care as usual as observed in DCS. A sensitivity analysis was conducted based on Ahlqvist subgroups.

RESULTS

Under care as usual, prognosis in the RHAPSODY data-driven subgroups ranged from 7.9 to 12.6 QALYs. Prognosis in the risk-driven subgroups ranged from 6.8 to 12.0 QALYs. Compared with homogenous type 2 diabetes, treatment for individuals in the high-risk subgroups could cost 22.0% and 25.3% more and still be cost effective for data-driven and risk-driven subgroups, respectively. Targeting BMI and LDL in addition to HbA_{1c} might deliver up to 10-fold increases in QALYs gained.

CONCLUSIONS

Risk-driven subgroups better discriminated prognosis. Both stratification methods supported stratified treatment intensification, with the risk-driven subgroups being somewhat better in identifying individuals with the most potential to benefit from intensive treatment. Irrespective of stratification approach, better cholesterol and weight control showed substantial potential for health gains.

To capture the heterogeneity and refine the current stratification of type 2 diabetes, a novel data-driven clustering analysis by Ahlqvist et al. (1) identified five subgroups, including severe autoimmune diabetes, severe insulin deficiency diabetes (SIDD), severe insulin resistance diabetes (SIRD), mild obesity-related diabetes (MOD), and mild

¹Faculty of Science and Engineering, Groningen Research Institute of Pharmacy, University of Groningen, Groningen, the Netherlands

²National Institute of Public Health and the Environment, Bilthoven, the Netherlands

³Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, U.K.

⁴Department of Epidemiology and Data Sciences, Amsterdam University Medical Center, Vrije Universiteit, Amsterdam, the Netherlands

⁵Amsterdam Cardiovascular Sciences, Amsterdam Public Health, Amsterdam, the Netherlands

⁶Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands

⁷Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

⁸Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

⁹Division of Population Health and Genomics, Ninewells Hospital and School of Medicine, University of Dundee, Dundee, U.K.

¹⁰Department of General Practice, Amsterdam University Medical Center, Vrije Universiteit, Amsterdam, the Netherlands

Corresponding author: Xinyu Li, li.xinyu@rug.nl

Received 8 November 2022 and accepted 11 April 2023

This article contains supplementary material online at <https://doi.org/10.2337/figshare.22619980>.

T.L.F. and J.L. contributed equally as senior authors.

© 2023 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

age-related diabetes (MARD), based on clinical parameters. These data-driven clustering methods have been replicated in many cohorts (2–6). However, questions remain concerning their clinical utility and cost-effectiveness. Soft clustering (7) or stratification based on predicted risk as estimated from continuous clinical features (2,8,9) might also identify type 2 diabetes phenotypes or predict outcomes for individuals, and it has been shown that using clinical measures in a regression model may outperform clustering for prediction of nephropathy risk and response to treatment (2). Nonetheless, data-driven clustering analysis might identify underlying phenotypic and pathologic subgroups and thus benefit medical decisions (6,10,11).

Alternatively, individuals could be classified based on clinically relevant risk thresholds as applied in diabetes and cardiovascular guidelines. European guidelines on cardiovascular disease prevention (12–14) recommend using the Systematic Coronary Risk Evaluation (SCORE) system (15) to inform intensity of care. U.S. and European guidelines for type 2 diabetes focus on HbA_{1c} values or goals to inform medical care (16,17).

In addition to uncertainty concerning the clinical utility of stratification approaches, it is unclear whether these approaches could potentially support a cost-effective use of health care resources. Allocating individuals into subgroups may help clinicians to make decisions about whether to treat individuals intensively because in some subgroups, individuals may benefit more from intensive treatment than the average or those in other subgroups (2). However, the potential benefit of this strategy to help decision making has not been explicitly evaluated. Hence, we used data from 2,935 contemporary individuals with type 2 diabetes from the Hoorn Diabetes Care System (DCS) to simulate the potential effect of their stratification (via data-driven clustering or using prespecified cutoffs for risk factor levels) and treatment intensification, relative to usual care, on predicted costs and (quality-adjusted) life expectancy. We further explored the potential gains from targeting cholesterol and weight, in addition to HbA_{1c}, in each subgroup and across all individuals.

To help decision making, we expressed our results as the maximum annual price in U.S. and U.K. settings that can be spent in the health care sector for identification and treatment of a certain subgroup while

remaining cost-effective. This straightforward indicator will inform clinicians and decision makers on whether intensifying treatment is beneficial and cost-effective.

RESEARCH DESIGN AND METHODS

Study Population

The DCS is a comprehensive dynamic prospective cohort of the natural course of type 2 diabetes from 103 general practitioners in the West Friesland region of the Netherlands (18). Laboratory measurements have been described in detail in previous studies (18,19).

The study population consisted of 2,935 individuals with newly diagnosed type 2 diabetes over the period 1998–2019 in the DCS cohort (Supplementary Appendix 1). Our inclusion criteria were age at diagnosis ≥ 35 years, clinical parameters available within 2 years after diagnosis, negative for GAD, complete data in clustering variables, and the presence of genome-wide association study data (19). The ethical review committee of VU University Medical Center approved the study, and informed consent was obtained from all participants.

Data-Driven Subgroups and Risk-Driven Subgroups

A recent study, as part of Risk Assessment and Progression of Diabetes project (RHAPSODY; <https://www.imi.europa.eu/projects-results/project-factsheets/rhapsody>), applied the data-driven clustering approach by Ahlqvist et al. (1) to participants with diabetes in three routine care cohorts, including the DCS. The RHAPSODY subgroups used clinical parameters available in routine care, replaced HOMA estimates in Ahlqvist's subgroups with C-peptide, and added HDL as an extra cluster indicator. This cluster replication in external data demonstrated a good concordance between cohorts and with the original clustering by Ahlqvist et al., while additionally refining the MARD into two subgroups (1,19,20).

Hence, as shown in Table 1, individuals in DCS were assigned to one of five RHAPSODY subgroups (19), including RHAPSODY SIDD (RHAP-SIDD), RHAP-SIRD, RHAP-MOD, RHAPSODY mild diabetes (RHAP-MD), and RHAPSODY mild diabetes with high HDL (RHAP-MDH), based on sex-specific k-means clustering by five scaled clustering indicators including age, BMI, HbA_{1c}, C-peptide, and HDL. The full details of the clustering methods and results have been published previously (1,19).

We also stratified individuals in DCS according to a combination of HbA_{1c} values and SCORE levels using prespecified thresholds (Table 1). The values were selected to reflect American Diabetes Association (ADA) (17) and European recommendations (16) on glucose goals (HbA_{1c} <7% [53 mmol/mol]) and European recommendations on cardiovascular risk management (with a SCORE of 5% discriminating between high or higher and moderate to lower cardiovascular risk categories) (14).

Care-as-Usual and Intensive Diabetes Management Strategies

The observed trajectories of risk factors such as HbA_{1c} and lipid levels captured care as usual in the contemporary DCS population. Intensive diabetes management interventions were simulated as guideline-based treat-to-target strategies because subgroup-specific treatment effects are unknown. We assumed that prespecified glycemic targets based on the ADA (17) and European guidelines (16) would be achieved (Supplementary Material Table 2.1). We followed European guidelines (14) for LDL and weight treatment targets. We analyzed a 5-year intensive intervention. Once intensive management interventions were discontinued, we assumed that risk factors would revert immediately to values observed under care as usual (base case).

Simulation

We used the UK Prospective Diabetes Study Outcomes Model version 2 (UKPDS-OM2) to simulate lifetime health outcomes and costs of the DCS cohort (21). The UKPDS-OM2 predicts an individual's absolute probability of experiencing any of eight diabetes complications (myocardial infarction, stroke, heart failure, ischemic heart disease, amputation, renal failure, blindness in one eye, foot ulcers) and death (21). These predictions depend on the individual's age, ethnicity, sex, and time-varying clinical risk factors (including diabetes duration, systolic blood pressure [SBP], HbA_{1c}, lipid levels, smoking status, and history of previous complications) (21). Model outputs include annual event probabilities, life expectancy, quality-adjusted life-years (QALYs), and lifetime costs.

The UKPDS-OM2 has been validated both internally and externally (21–23), and it has shown good performance in predicting macrovascular events in DCS (23). As our study focuses on the model's

Table 1—Subgroup characteristics and cutoffs

Subgroup	Characteristic or cutoff	n (%)
RHAPSODY data driven		
RHAP-SIDD	High HbA _{1c}	365 (12.44)
RHAP-SIRD	High C-peptide and age	637 (21.70)
RHAP-MOD	High BMI and C-peptide	520 (17.72)
RHAP-MD	Moderate in clustering indicators	860 (29.30)
RHAP-MDH	High HDL	553 (18.84)
Risk driven*		
H1S1	HbA _{1c} <7% and SCORE <5%	1,274 (43.41)
H1S2	HbA _{1c} <7% and SCORE ≥5%	542 (18.47)
H2S1	HbA _{1c} ≥7% and SCORE <5%	841 (28.65)
H2S2	HbA _{1c} ≥7% and SCORE ≥5%	278 (9.47)

H1S1, low HbA_{1c} and low SCORE level. *Additional information on the SCORE project can be found in Conroy et al. (15).

ability to capture differences between subgroups, we validated the relative risks of incidence of events for subgroups by testing whether simulated relative risks fell within the 95% CI of observed relative risks.

The model input variables are listed in Supplementary Appendix 2. We simulated an individual's lifetime outcomes for both care-as-usual and intensive diabetes management strategies. A 70-year simulation period was chosen to reflect a lifetime (study population minimum age 35 years).

After data cleaning (0.95% missing data) (Supplementary Appendix 3), baseline characteristics of each data-driven and risk-driven subgroup, as included in the simulation, were reported by frequency (percentage) for categorical variables or mean (SD) for continuous variables. An χ^2 test was applied to check for significant differences between subgroups within each stratification approach.

We used observed data until the end of the follow-up in the DCS cohort. For HbA_{1c}, LDL, BMI, and estimated glomerular filtration rate (eGFR) values after the end of follow-up, we extrapolated their progression using linear dynamic models fitted to DCS observations (Supplementary Appendix 4). As HDL and SBP remained relatively constant throughout the observation period (Supplementary Figs. 2.1 and 2.2), we extrapolated these by last observation carried forward.

A health care perspective was applied, and costs and utilities associated with diabetes management and diabetes-related complications were obtained for the U.S. and U.K. settings (Supplementary Table 2.3). Costs were expressed in 2019 values,

inflated to that year using a price index. Costs and QALYs were discounted at 3.0% in the U.S. setting (24) and 3.5% in the U.K. setting (25).

Simulated Outcomes and Standardization

Lifetime costs and QALYs for each subgroup under care as usual were simulated (mean and 95% CI). To remove the effect of unmodifiable risk factors (i.e., age and sex), we standardized the estimates to the average age for men and women separately in DCS (i.e., a 62-year-old man and a 63-year-old woman) by regressing the individual-level UKPDS-OM2-simulated outcomes on their age.

Maximum Annual Cost-Effective Price of Stratification and Intensive Management Interventions

Intensive management interventions were deemed cost-effective if the incremental cost-effectiveness ratio was below the threshold of \$100,000 and £20,000 per QALY in the U.S. and U.K., respectively (25,26). We estimated the maximum annual price for each strategy that would not exceed cost-effectiveness thresholds (equations in Supplementary Appendix 5) by subgroup and overall. A higher maximum annual price indicates that the subgroup can spend more on diabetes management costs while remaining cost-effective. The range (maximum – minimum) in maximum prices and in incremental QALYs among subgroups was used to indicate to what degree subgroups could distinguish between groups of individuals for whom intensive treatment was potentially more or less cost-effective.

Uncertainty

The analysis accounted for two types of uncertainty: Monte Carlo simulation error and parameter uncertainty. We reduced Monte Carlo simulation error by averaging 50,000 simulations per individual, and propagated parameter uncertainty by performing 400 random draws of different sets of model parameters derived from the UKPDS trial population (21). Maximum cost-effective prices of stratification and intensive treatments and further model outcomes were estimated for each of the 400 draws, and the 2.5% and 97.5% percentiles were used to present the level of uncertainty.

Sensitivity Analyses

To analyze the difference caused by different data-driven clustering approaches, individuals in DCS were also assigned to one of four subgroups following the method by Ahlqvist et al. (1), including SIDD, SIRD, MOD, and MARD based on sex-specific k-means clustering by five scaled clustering indicators, including age at diagnosis, BMI, HbA_{1c}, HOMA estimates (27) of β -cell function, and insulin resistance by C-peptide and fasting glucose. Because reaching treatment targets might be difficult, especially for weight loss, we analyzed a conservative 5% improvement scenario in which the values of care-as-usual risk factors will be improved by 5% based on the recommendation of achieving and maintaining ≥5% weight loss by ADA guidelines (28). We varied the duration of the intensive management interventions from 5 years to 10, 15, and 20 years. Moreover, we considered risk factors for returning to a care-as-usual trajectory gradually, rather than immediately, by introducing a scenario analysis in which the linear dynamic models for risk factor progression would inform the subsequent risk factor trajectories until they reached the observed care-as-usual values (scenario 1). Graphical representations of the scenario assumptions are presented in Supplementary Fig. 2.3.

Data and Resource Availability

The data are not publicly available but can be requested from VU University Medical Center. We accessed the data via a formal data request as a part of the RHAPSODY project.

RESULTS

Baseline Characteristics

We found significant differences in baseline characteristics in both data-driven subgroups and risk-driven subgroups (Table 2 and Supplementary Appendix 6). Of note, higher mean age was observed in the RHAP-SIRD, RHAP-MDH, and subgroups with high SCORE values compared with the remaining subgroups.

Lifetime Costs and Outcomes of Subgroups Under Care as Usual

Supplementary Figs. 7.1 and 7.2 show that simulated relative risks fit within the 95% CI of observed relative risks among subgroups, indicating that UKPDS-OM2 was able to reflect differences between subgroups in risks. Figure 1 and Supplementary Appendix 8 show the simulated lifetime costs and QALYs and their standardization to an average individual (62-year-old man or 63-year-old woman) for all data-driven and risk-driven subgroups and across all individuals with type 2 diabetes in DCS under care as usual (i.e., without intensive management intervention).

On average, an individual with type 2 diabetes in DCS was predicted to accrue 10.57 QALYs and \$165,000 in complication costs in their remaining lifetime (Supplementary Table 8.1). Both stratification methods showed significant differences in QALYs and complication costs among subgroups (Fig. 1). For data-driven subgroups, as expected, subgroups with older individuals had the worst simulated outcomes. The RHAP-SIRD subgroup had the lowest QALYs (7.90) and complication costs (\$125,000) and was predicted to have the highest diabetes-related macrovascular complication rates, explaining its low QALYs (Supplementary Fig. 8.2). For risk-driven subgroups, the high HbA_{1c} and high SCORE level (H2S2) subgroup had the lowest QALYs (6.83) and complication costs (\$114,000), with the highest simulated diabetes-related complication rates among all subgroups (Supplementary Fig. 8.3). Even at high rates of complication, complication costs were low when life expectancy was low.

After adjusting for sex and age, a standardized 62-year-old man and 63-year-old woman in DCS were predicted to accrue 9.98 and 11.12 QALYs and \$154,000 and \$176,000 in complication costs, respectively. For data-driven subgroups, the lowest standardized QALYs were seen in

Table 2—Selected baseline simulation characteristics of subgroups

	RHAPSODY data-driven subgroups					Risk-driven subgroups					P
	RHAP-SIDD	RHAP-SIRD	RHAP-MOD	RHAP-MD	RHAP-MDH	H1S1	H1S2	H2S1	H2S2	P	
n (%)	365 (12.44)	637 (21.70)	520 (17.72)	860 (29.30)	553 (18.84)	1,274 (43.41)	542 (18.47)	841 (28.65)	278 (9.47)	<0.001	
Age, years	61.39 (9.74)	70.74 (7.41)	55.90 (8.05)	57.56 (8.20)	68.79 (7.76)	59.08 (8.15)	72.75 (6.54)	58.18 (8.56)	73.58 (6.67)	<0.001	
Duration of diabetes, years	4.46 (3.29)	2.30 (2.74)	3.05 (3.21)	3.26 (3.40)	2.80 (3.18)	2.82 (3.22)	2.19 (2.69)	3.94 (3.38)	3.32 (3.21)	<0.001	
LDL-C, mmol/L	2.72 (0.91)	2.64 (0.89)	2.68 (0.90)	2.79 (0.94)	2.80 (0.91)	2.76 (0.93)	2.82 (0.90)	2.61 (0.89)	2.80 (0.92)	<0.001	
HDL-C, mmol/L	1.16 (0.31)	1.08 (0.22)	1.06 (0.26)	1.10 (0.23)	1.56 (0.34)	1.19 (0.32)	1.25 (0.33)	1.12 (0.31)	1.20 (0.33)	<0.001	
HbA _{1c} mmol/mol	61.51 (19.34)	47.87 (7.99)	51.17 (10.93)	49.47 (9.33)	46.90 (7.85)	45.87 (6.86)	45.76 (6.04)	58.21 (14.44)	56.92 (13.26)	<0.001	
%	7.78 (1.78)	6.53 (0.73)	6.83 (1.00)	6.68 (0.85)	6.44 (0.72)	6.35 (0.63)	6.34 (0.55)	7.48 (1.32)	7.36 (1.22)	<0.001	
eGFR,* mL/min/1.73 m ²	84.35 (18.02)	71.50 (15.94)	88.06 (16.84)	87.54 (16.12)	77.97 (14.83)	84.59 (16.25)	72.80 (15.50)	87.21 (17.37)	71.67 (16.13)	<0.001	
BMI, kg/m ²	29.50 (4.59)	29.89 (3.51)	37.82 (4.98)	28.90 (3.38)	27.15 (3.54)	30.87 (5.54)	29.10 (4.55)	31.06 (5.47)	29.24 (4.10)	<0.001	
SBP, mmHg	142.17 (19.84)	146.75 (19.99)	141.69 (17.87)	137.56 (17.56)	145.60 (18.57)	137.23 (16.16)	155.01 (19.70)	138.36 (17.36)	153.48 (18.33)	<0.001	
Male, n (%) [vs. female]	218 (59.7)	397 (62.3)	256 (49.2)	474 (55.1)	297 (53.7)	629 (49.4)	347 (64.0)	490 (58.3)	176 (63.3)	<0.001	
Smoking status [vs. never], n (%)										0.017	
Current	76 (21.1)	86 (13.9)	108 (21.1)	204 (24.8)	72 (13.4)	197 (16.1)	105 (20.0)	182 (21.8)	62 (22.9)		
Former	170 (47.1)	364 (58.9)	251 (48.9)	365 (44.3)	276 (51.3)	632 (51.6)	269 (51.2)	394 (47.3)	131 (48.3)		

Data are mean (1 SD) unless otherwise indicated. χ^2 test was applied to check for significant differences between subgroups. H1S1, low HbA_{1c} and low SCORE level; HDL-C, HDL cholesterol; LDL-C, LDL cholesterol. *Based on the Chronic Kidney Disease Epidemiology Collaboration equation.

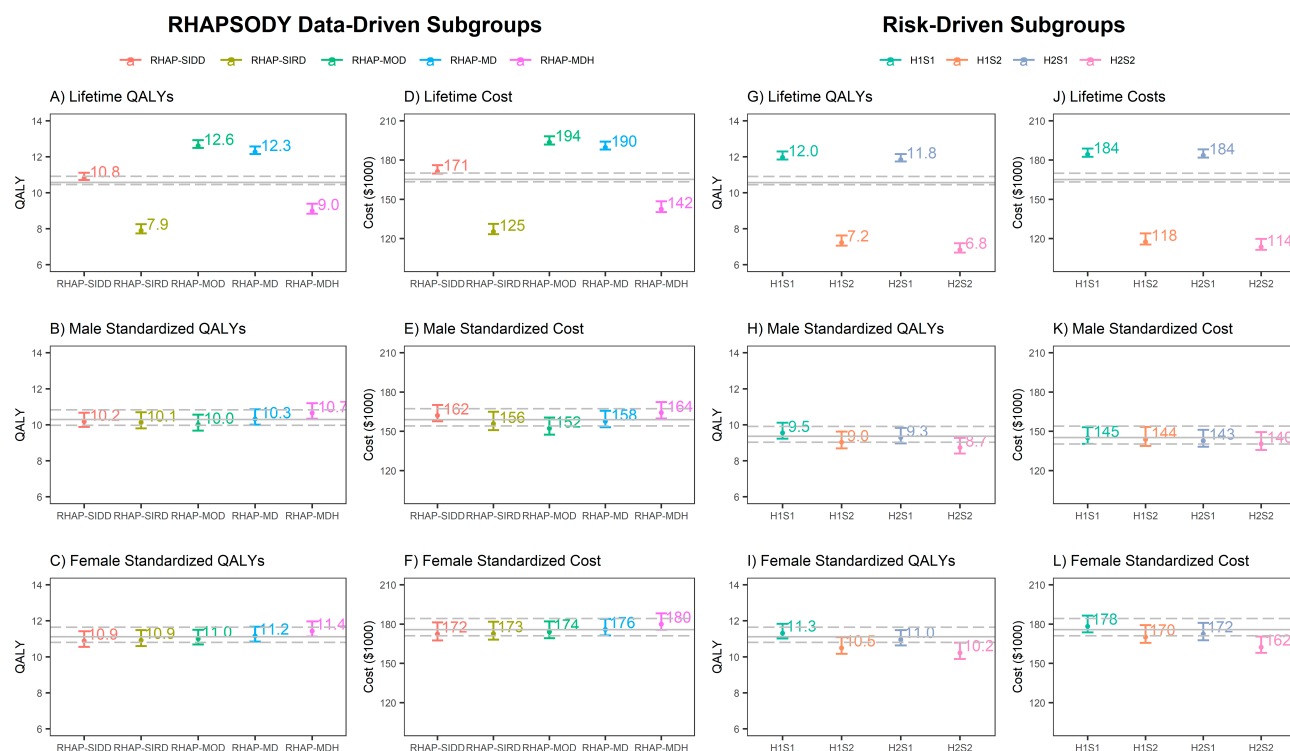


Figure 1—Nonstandardized and standardized mean simulated lifetime QALYs and costs in the U.S. setting for data-driven and risk-driven subgroups. The horizontal solid lines and dashed lines indicate the average value and its 95% CI. A, D, G, and J: Lifetime QALYs and costs. B, E, H, and K: Male standardized QALYs and costs. C, F, I, and L: Female standardized QALYs and costs. H1S1, low HbA_{1c} and low SCORE level.

RHAP-MOD for men (10.02) and RHAP-SIDD for women (10.88). For risk-driven subgroups, the ranking remained the same as before standardization, with the lowest standardized QALYs seen in H2S2 (men 8.73; women 10.22). The U.K. and U.S. settings featured similar outcomes, except the absolute values of the U.K. setting were lower because of higher discounting rates and lower complication costs (Supplementary Fig. 8.1 and Supplementary Table 8.2).

Maximum Annual Price of Stratification and Intensive Management

Table 3 shows the incremental complication costs, QALYs, and maximum prices of guideline-based treat-to-target strategy in the U.S. setting (threshold of \$100,000 per QALY). The outcomes of the remaining scenarios are provided in Supplementary Appendix 9.

Treat-to-target strategies led to an average reduction of 0.2% or 2.5 mmol/mol (2.7%) in HbA_{1c}, 0.5 mmol/L (14.7%) in LDL, and 5.0 kg/m² (15.0%) in BMI (14.9 kg in weight) (Supplementary Tables 2.4 and 2.5). In the base case, without stratification into subgroups, treat-to-target of HbA_{1c} could cost up to \$169 additionally per year while

remaining below the \$100,000 per QALY threshold. Furthermore, treating to the target of LDL and BMI in addition to HbA_{1c} could cost up to \$1,499 per year and remain cost-effective.

For RHAPSODY data-driven subgroups, intensive management interventions targeting HbA_{1c} resulted in the largest gains in QALYs (0.019) in the RHAP-SIDD subgroup and could cost up to \$368 per year and remain cost-effective. This finding indicates that individuals in the RHAP-SIDD subgroup can spend \$199 more on diabetes management than individuals with type 2 diabetes overall while remaining cost-effective.

Compared with focusing on HbA_{1c} only, treatment targeting HbA_{1c}, BMI, and LDL in combination achieved 10 times higher gains in QALYs and could cost substantially more per year while remaining cost-effective, ranging from 0.044 QALYs and \$799 per person in the RHAP-MD subgroup to 0.112 QALYs and \$1,973 per person in the RHAP-MOD subgroup. On average, for individuals in high-risk subgroups (RHAP-SIDD, RHAP-SIRD, and RHAP-MOD), the maximum annual price of intensive management in the U.S. setting could be 30.7% higher while remaining

cost-effective compared with the no stratification scenario.

For risk-driven subgroups, intensive management solely targeting HbA_{1c} resulted in the largest gains in QALYs in the subgroups with high HbA_{1c} levels (0.017 for high HbA_{1c} and low SCORE level [H2S1] and 0.012 for H2S2) and could cost up to \$323 and \$270 per year, respectively, while remaining cost-effective. Compared with solely targeting HbA_{1c}, treatment targeting BMI and LDL achieved more than 10 times the gains in QALYs and could cost substantially more at up to 0.114 QALYs and \$2,578 per person in the H2S2 subgroup. On average, for individuals in high-risk subgroups (H2S1, low HbA_{1c} and high SCORE level [H1S2], and H2S2) the maximum annual price of intensive management could be 31.2% higher, compared with a no stratification scenario, while remaining cost-effective in the U.S. setting.

Sensitivity Analyses

Replicating the current analyses by following the subgroups of Ahlqvist et al. (1) led to robust findings about discrimination (Supplementary Appendix 10). BMI (37.82 kg/m²) and C-peptide (1.43 nmol/L) values in RHAP-MOD were significantly higher than MOD (33.51 kg/m²

Table 3—Outcomes of 5-year guideline-based intensive management targeting HbA_{1c}, BMI, and LDL and targeting only HbA_{1c} compared with care as usual by subgroup in base case U.S. setting

Subgroup	Treat-to-target hypothetical intensive management			
	HbA _{1c}		HbA _{1c} + LDL + BMI	
	Maximum annual price of intervention (\$)	ΔQALY vs. care as usual	Maximum annual price of intervention (\$)	ΔQALY vs. care as usual
Overall*	169 (97–222)	0.008 (0.005–0.011)	1,499 (1,132–1,776)	0.073 (0.058–0.09)
RHAPSODY data driven				
RHAP-MOD	221 (150–296)	0.012 (0.008–0.015)	1,973 (1,444–2,603)	0.112 (0.083–0.146)
RHAP-MD	116 (67–167)	0.006 (0.004–0.009)	799 (666–966)	0.044 (0.036–0.052)
RHAP-SIDD	368 (248–477)	0.019 (0.013–0.024)	1,504 (1,233–1,779)	0.079 (0.065–0.092)
RHAP-MDH	58 (6–111)	0.003 (0–0.005)	1,267 (986–1,566)	0.061 (0.047–0.075)
RHAP-SIRD	96 (48–148)	0.004 (0.002–0.007)	1,902 (1,519–2,335)	0.087 (0.069–0.106)
Range†	309	0.016	1,174	0.068
Risk driven				
H1S1	82 (42–117)	0.004 (0.002–0.006)	930 (723–1,182)	0.052 (0.041–0.066)
H2S1	323 (235–416)	0.017 (0.012–0.021)	1,247 (990–1,546)	0.069 (0.055–0.084)
H1S2	69 (23–120)	0.003 (0–0.005)	2,356 (1,897–2,894)	0.105 (0.085–0.129)
H2S2	270 (164–396)	0.012 (0.007–0.017)	2,578 (2,080–3,100)	0.114 (0.093–0.137)
Range†	253	0.014	1,647	0.062

H1S1, low HbA_{1c} and low SCORE level. *Overall refers to a homogenous type 2 diabetes group. Results were generated based on extrapolations of subgroup-specific linear dynamic models and summarized by subgroup information. The overall result is summarized by the assumption that every individual was within this homogenous type 2 diabetes group. Each extrapolation from either RHAPSODY data-driven subgroups' or risk-driven subgroups' linear dynamic models led to an overall result, and the final overall result was taken as the average value. †Range is defined as the maximum – minimum of the mean maximum annual cost-effective price of intervention or incremental QALY.

and 1.04 nmol/L, respectively) (Supplementary Table 10.2). Although we observed RHAP-SIRD to have significantly higher BMI compared with other RHAPSODY subgroups except RHAP-MOD (Supplementary Fig. 6.2), this difference was less pronounced than the BMI difference observed between SIRD and SIDD or MARD by Ahlqvist et al. (Supplementary Fig. 10.2). MARD had the lowest absolute simulated QALYs, but after standardization, SIRD had the lowest QALYs (Supplementary Figs. 10.7 and 10.8). SIRD and MARD generally had the highest risk of complications, except for SIDD, which had the highest risk of amputation (Supplementary Fig. 10.9).

The scenario of a 5% improvement led to similar findings as the treat-to-target scenario, although with a less substantial reduction in risk factors (Supplementary Tables 2.4 and 2.5) and, therefore, less difference in results (Supplementary Tables 9.1–9.4). Overall, in considering both scenarios, compared with homogenous type 2 diabetes, treatment for individuals in high-risk subgroups could cost on average 22.0% and 25.3% more and still be cost-effective for data-driven and risk-driven subgroups, respectively.

A longer treatment period implied lower maximum annual prices of intensive management while remaining cost-effective (Supplementary Figs. 9.1–9.3). Allowing the

treatment effect to extend beyond the hypothetical treatment period (scenario 1) led to more incremental QALYs and higher maximum annual prices of intensive management among subgroups. In all scenarios, intensive management could cost significantly more in high-risk subgroups compared with no stratification and remain cost-effective.

CONCLUSIONS

The data-driven subgroups were able to stratify individuals with diverse prognosis, displaying significant differences in simulated lifetime QALYs and complication costs. However, the risk-driven subgroups showed somewhat larger differences between high- and low-risk subgroups compared with the data-driven subgroups. Both data-driven subgroups and risk-driven subgroups could support stratifying individuals for prioritizing treat-to-target strategies. For the individuals in high-risk subgroups, resources higher than average could be committed for treat-to-target strategies while remaining cost-effective. This difference in maximum annual prices indicates substantial financial incentives to identify individuals in high-risk groups and treat them more intensively.

About two-thirds of individuals with diabetes fail to achieve HbA_{1c} targets

(7%) (17,29), and we show the potential gains and value of targeting HbA_{1c} only. However, targeting LDL and BMI, in addition to HbA_{1c}, offered significant benefits in contemporary populations like the DCS. This finding is important when >90% of individuals with type 2 diabetes are overweight or obese (30) and less than one-half reach LDL targets (31). Our predicted gains may partly reflect that current targets for BMI and LDL are quite ambitious compared with actual risk factor levels observed in populations (32–34). Rather than treat-to-target, using 5% reductions of risk factor levels has produced similar findings but of smaller magnitude. Furthermore, the RHAP-SIRD, RHAP-SIDD, RHAP-MOD, and H2S2 subgroups benefited most from jointly targeting HbA_{1c}, LDL, and BMI. These subgroups had the largest simulated QALY gains from a combined intervention, highlighting an opportunity to target specific subgroups of individuals more intensively. Specifically, in a contemporary care-as-usual setting, the RHAP-SIRD and H2S2 subgroups had the lowest predicted lifetime QALYs and the highest risk of complications among all subgroups, partly driven by patient advanced age.

The findings regarding differences in baseline characteristics were in line with

previous studies (1,19). In addition, our article presents that, across all RHAPSODY data-driven subgroups, a guideline-based 5-year comprehensive intervention to lower HbA_{1c}, BMI, and LDL could cost up to \$799–\$1,973 per year in the U.S. and £196–£463 per year in the U.K. at \$100,000 per QALY and £20,000 per QALY cost-effectiveness thresholds, respectively. Thus, the costs of measuring any clustering indicators and intensifying treatment must be lower than these values for a subtype-specific treatment strategy to be cost-effective. For risk-driven subgroups, the intervention could cost up to \$930–\$2,578 per year in the U.S. and £230–£515 per year in the U.K. to be cost-effective. These ranges indicate financial incentives and potential benefits resulting from stratification of type 2 diabetes. The higher the range in annual prices, the more helpful stratification could be to inform treatment prioritization.

Comparing two stratification methods, risk-driven subgroups discriminated individuals better between mild and severe conditions than data-driven subgroups in the care-as-usual setting. Data-driven clustering better identified individuals who would benefit from more intensive glucose treatment alone. Risk-driven subgroups better identified individuals who would benefit from more intensive treatment targeting lipids, weight, and HbA_{1c} together. In general, also considering their more straightforward implementation, risk-driven subgroups seem better suited than data-driven subgroups for stratifying individuals with different risks and guiding comprehensive treatment.

Consistent with previous findings (19), RHAPSODY subgroups resembled those of Ahlqvist et al. (1), except that the RHAP-SIRD subgroup was older, less insulin resistant, and had a lower BMI than SIRD, while the RHAP-MOD subgroup had a higher BMI and was more insulin resistant than MOD. Although differences exist in their characteristics, using either of these two methods of data-driven clustering led to the same conclusion that classifying type 2 diabetes according to cutoffs for HbA_{1c} and cardiovascular risk might better identify individuals for treatment intensification compared with data-driven clustering. Furthermore, MOD is being recognized as a mild diabetes subgroup, but this recognition is highly influenced by the young age of individuals in that subgroup. In both RHAPSODY and

Ahlqvist's subgroups, after age standardization, the MOD subgroup had similar or even lower lifetime QALYs compared with other severe subgroups, including SIDD and SIRD, indicating that despite this group's mild designation, this population with high BMI still requires careful management.

This study had several limitations. First, despite the generally good fit of the linear dynamic models (Supplementary Appendix 4), they slightly underestimated eGFR, leading to overestimated kidney damage and underestimated QALYs. However, this likely had minimal impact on relative subgroup differences. Second, UKPDS-OM2 simulations predict complications using risk factor trajectories and preexisting complications. The prediction of risk factor trajectories was specified by subgroup based on subgroup-specific prediction models, while the prediction of complications was not specific to subgroups. The treatment intensification scenarios investigated were hypothetical and based on changes to risk factors to meet treatment targets. Our results provide a benchmark for stratified treatment strategies, allowing comparison of different stratification approaches. They warrant further research to investigate how to best reach treatment goals. Third, individuals with less favorable prognosis (e.g., those with a <5-year life expectancy) might fall under the HbA_{1c} <8% recommendation (17) rather than 7%, indicating lower incremental QALYs. However, our simulation cohort's average age (62.8 years) is ~18 years less than the mean life expectancy in the Netherlands (81 years as measured in 2020 [35]); therefore, we believe that our finding is relevant. Finally, two clustering indicators, namely C-peptide and fasting glucose, are not captured in the UKPDS-OM2, which might underestimate the discrimination ability of data-driven subgroups. However, C-peptide is found to be relatively stable over time (1), and HbA_{1c}, for which within-patient reproducibility is superior to that of fasting glucose (36), is included in the UKPDS-OM2. Therefore, we believe that our findings will not be largely affected.

This study suggests several potential directions for future research. We believe that cholesterol-lowering medicine and weight control interventions warrant further investigation for all individuals with diabetes (37,38), with special attention regarding their impact in specific subgroups (2,39). For example, as expected, treating to target of HbA_{1c} alone

is less cost-effective for individuals with SIRD than for most other subgroups, given their already low HbA_{1c} levels and the possibility that complications are primarily driven by hyperinsulinemia or insulin resistance (40). Treatment options targeting the latter are currently limited (39); while lifestyle programs may help to reduce insulin resistance through weight loss, long-term sustainability is challenging (39). The high maximum annual price (~\$2,000) we found in the combined intervention for SIRD suggests a significant potential return on investment, which could support the further development of therapeutic options specifically targeting them. Furthermore, future data-driven clustering of diabetes subtypes may benefit from incorporating some elements of the risk-driven approach, such as smoking status, SBP, and total cholesterol, and may help to refine the current clustering and indicate some etiologic pathways that might have remained unnoticed at the current clustering indicators.

In summary, stratification approaches examined in this article were successful in distinguishing among individuals with type 2 diabetes in terms of lifetime QALYs and costs. Both data-driven and risk-driven subgroup stratification methods suggest that research and investment in personalized care are attractive from an individual and economic perspective. Using a data-driven clustering approach, we estimated that the RHAP-SIDD, RHAP-SIRD, and RHAP-MOD subgroups would potentially benefit in a cost-effective way from treat-to-target strategies. However, a more straightforward stratification using risk-driven cutoff values for risk factors did slightly better than data-driven clustering in identifying priority groups of individuals. With maximum prices of up to \$3,786 or £815 per individual per year, strong economic incentives exist to research and identify the best ways to achieve established treatment targets, especially in high-risk individuals.

Acknowledgments. The authors thank Amber A. Van Der Heijden, PhD, Amsterdam University Medical Center, for providing DCS data and helping with various data-related questions; Sajad Emamipour, MSc, University of Groningen, for helping with the neuropathy and retinopathy data; Stefan R.A. Konings, MSc, University of Groningen, for helping with coding; and Junfeng Wang, PhD, Utrecht University, and Fang Li, MSc, University of Groningen, for scientific advice. The authors thank the anonymous

reviewers for insightful feedback, which contributed a lot to enhancing the quality and clarity of the article.

Funding. This project received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant 115881 (RHAPSODY). This joint undertaking receives support from the European Union's Horizon 2020 (H2020 Health) research and innovation program and the European Federation of Pharmaceutical Industries and Associations. This work was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0097-2.

The opinions expressed and arguments used herein do not necessarily reflect the official views of these funding bodies. The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Duality of Interest. No conflicts of interest relevant to this article were reported.

Author Contributions. X.L., T.L.F., and J.L. researched data, contributed to discussion, and wrote, reviewed, and edited the manuscript. A.V.G. and J.A. contributed to the discussion and wrote, reviewed, and edited the manuscript. R.C.S., J.W.J.B., L.M.H., E.R.P., and P.J.M.E. contributed to the discussion and reviewed and edited the manuscript. R.C.S., J.W.J.B., L.M.H., and P.J.M.E. contributed to DCS data gathering and delivery. All authors contributed to the critical revision of the manuscript for important intellectual content and approved the final version of the manuscript. X.L. and T.L.F. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this article were presented as a poster presentation at the 82nd Scientific Sessions of the American Diabetes Association in New Orleans, LA, 3–7 June 2022; as an oral presentation at the European Health Economics Association Conference 2022 in Oslo, Norway, 5–8 July 2022; and at the Mount Hood Diabetes Challenge Network Conference 2022 in Malmo, Sweden, 24 September 2022.

References

- Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018;6:361–369
- Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT. Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data. *Lancet Diabetes Endocrinol* 2019;7:442–451
- Sarría-Santamera A, Arazumbekova B, Maulenkul T, Gaipov A, Atageldiyeva K. The identification of diabetes mellitus subtypes applying cluster analysis techniques: a systematic review. *Int J Environ Res Public Health* 2020;17:9523
- Varghese JS, Narayan KMV. Ethnic differences between Asians and non-Asians in clustering-based phenotype classification of adult-onset diabetes mellitus: a systematic narrative review. *Prim Care Diabetes* 2022;16:853–856
- Tanabe H, Masuzaki H, Shimabukuro M. Novel strategies for glycaemic control and preventing diabetic complications applying the clustering-based classification of adult-onset diabetes mellitus: a perspective. *Diabetes Res Clin Pract* 2021;180:109067
- Herder C, Roden M. A novel diabetes typology: towards precision diabetology from pathogenesis to treatment. *Diabetologia* 2022;65:1770–1781
- Wesolowska-Andersen A, Brorsson CA, Bizzotto R, et al., IMI DIRECT Consortium. Four groups of type 2 diabetes contribute to the etiological and clinical heterogeneity in newly diagnosed individuals: an IMI DIRECT study. *Cell Rep Med* 2022;3:100477
- Lugner M, Gudbjörnsdóttir S, Sattar N, et al. Comparison between data-driven clusters and models based on clinical features to predict outcomes in type 2 diabetes: nationwide observational study. *Diabetologia* 2021;64:1973–1981
- van der Leeuw J, van Dieren S, Beulens JWJ, et al. The validation of cardiovascular risk scores for patients with type 2 diabetes mellitus. *Heart* 2015;101:222–229
- Ahlqvist E, Prasad RB, Groop L. Subtypes of type 2 diabetes determined from clinical parameters. *Diabetes* 2020;69:2086–2093
- Schrader S, Perflyev A, Ahlqvist E, et al. Novel subgroups of type 2 diabetes display different epigenetic patterns that associate with future diabetic complications. *Diabetes Care* 2022;45:1621–1630
- Perk J, De Backer G, Gohlke H, et al.; Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice; European Association for Cardiovascular Prevention and Rehabilitation. European guidelines on cardiovascular disease prevention in clinical practice (version 2012): the Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Atherosclerosis* 2012;223:1–68
- Piepoli MF, Hoes AW, Agewall S, et al. 2016 European guidelines on cardiovascular disease prevention in clinical practice: the Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts): developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur J Prev Cardiol* 2016;23:NP1–NP96
- Mach F, Baigent C, Catapano AL, et al.; ESC Committee for Practice Guidelines; ESC National Cardiac Societies. 2019 ESC/EAS guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Atherosclerosis* 2020;294:80–82
- Conroy RM, Pyörälä K, Fitzgerald AP, et al.; SCORE Project Group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24:987–1003
- Cosentino F, Grant PJ, Aboyans V, et al.; ESC Scientific Document Group. 2019 ESC guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. *Eur Heart J* 2020;41:255–323
- ElSayed NA, Aleppo G, Aroda VR, et al.; American Diabetes Association. 6. Glycemic targets: Standards of Care in Diabetes-2023. *Diabetes Care* 2023;46(Suppl. 1):S97–S110
- Van Der Heijden AAWA, Rauh SP, Dekker JM, et al. The Hoorn Diabetes Care System (DCS) cohort. A prospective cohort of persons with type 2 diabetes treated in primary care in the Netherlands. *BMJ Open* 2017;7:e015599
- Slieker RC, Donnelly LA, Fitipaldi H, et al. Replication and cross-validation of type 2 diabetes subtypes based on clinical variables: an IMI-RHAPSODY study. *Diabetologia* 2021;64:1982–1989
- Slieker RC, Donnelly LA, Fitipaldi H, et al. Distinct molecular signatures of clinical clusters in people with type 2 diabetes: an IMI-RHAPSODY study. *Diabetes* 2021;70:2683–2693
- Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82. *Diabetologia* 2013;56:1925–1933
- Si L, Willis MS, Asseburg C, et al. Evaluating the ability of economic models of diabetes to simulate new cardiovascular outcomes trials: a report on the Ninth Mount Hood Diabetes Challenge. *Value Health* 2020;23:1163–1170
- Pagano E, Konings SRA, Di Cuozzo D, et al. Prediction of mortality and major cardiovascular complications in type 2 diabetes: external validation of UK Prospective Diabetes Study outcomes model version 2 in two European observational cohorts. *Diabetes Obes Metab* 2021;23:1084–1091
- Sanders GD, Neumann PJ, Basu A, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 2016;316:1093–1103
- National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. Accessed 26 April 2023. Available from <https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781>
- Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness—the curious resilience of the \$50,000-per-QALY threshold. *N Engl J Med* 2014;371:796–797
- Holman R, Hines G, Kennedy I, Stevens R, Matthews D, Levy J. A calculator for HOMA. *Diabetologia* 2004;47:A222
- ElSayed NA, Aleppo G, Aroda VR, et al.; American Diabetes Association. 8. Obesity and weight management for the prevention and treatment of type 2 diabetes: Standards of Care in Diabetes-2023. *Diabetes Care* 2023;46(Suppl. 1):S128–S139
- Mannucci E, Monami M, Dicembrini I, Piselli A, Porta M. Achieving HbA1c targets in clinical trials and in the real world: a systematic review and meta-analysis. *J Endocrinol Invest* 2014;37:477–495
- Bramante CT, Lee CJ, Gudzone KA. Treatment of obesity in patients with diabetes. *Diabetes Spectr* 2017;30:237–243
- Morieri ML, Avogaro A; the DARWIN-T2D Network of the Italian Diabetes Society. Cholesterol lowering therapies and achievement of targets for primary and secondary cardiovascular prevention in type 2 diabetes: unmet needs in a large population of outpatients at specialist clinics. *Cardiovasc Diabetol* 2020;19:190
- Breuker C, Clement F, Mura T, et al. Non-achievement of LDL-cholesterol targets in patients with diabetes at very-high cardiovascular risk

receiving statin treatment: incidence and risk factors. *Int J Cardiol* 2018;268:195–199

33. Presta V, Figliuzzi I, Miceli F, et al.; EFFECTUS Steering Committee. Achievement of low density lipoprotein (LDL) cholesterol targets in primary and secondary prevention: analysis of a large real practice database in Italy. *Atherosclerosis* 2019;285:40–48
34. Van Gaal L, Scheen A. Weight management in type 2 diabetes: current and emerging approaches to treatment. *Diabetes Care* 2015;38:1161–1172
35. The World Bank. Life expectancy at birth, total(years) – Netherlands. Accessed 23 December

2022. Available from <https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=NL>

36. Gonzalez A, Deng Y, Lane AN, et al. Impact of mismatches in HbA_{1c} vs glucose values on the diagnostic classification of diabetes and prediabetes. *Diabet Med* 2020;37:689–696
37. Athyros VG, Doumas M, Imprialos KP, et al. Diabetes and lipid metabolism. *Hormones (Athens)* 2018;17:61–67
38. Magkos F, Hjorth MF, Astrup A. Diet and exercise in the prevention and treatment of type 2 diabetes mellitus. *Nat Rev Endocrinol* 2020;16:545–555

39. Veelen A, Erazo-Tapia E, Oscarsson J, Schrauwen P. Type 2 diabetes subgroups and potential medication strategies in relation to effects on insulin resistance and beta-cell function: a step toward personalised diabetes treatment? *Mol Metab* 2020;46:101158

40. Christensen DH, Nicolaisen SK, Ahlqvist E, et al. Type 2 diabetes classification: a data-driven cluster study of the Danish Centre for Strategic Research in Type 2 Diabetes (DD2) cohort. *BMJ Open Diabetes Res Care* 2022;10:e002731