

Supplementary Materials for

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

Xinyu Li, Mphil¹, Anoukh van Giessen, PhD², James Altunkaya, MSc³, Roderick C. Slieker, PhD^{4,5,6}, Joline W.J. Beulens, PhD^{4,5,7}, Leen M. 't Hart, PhD^{4,5,6,8}, Ewan R. Pearson, PhD⁹, Petra J. M. Elders, PhD^{5,10}, Talitha L. Feenstra, PhD^{1,2*} & Jose Leal, Dphil^{3*}

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

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

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

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

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

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

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

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

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

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

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

Contact information of the corresponding author

Xinyu Li, Mphil

Groningen Research Institute of Pharmacy, A. Deusinglaan1, 9713AV, Groningen, Netherlands.

Email: li.xinyu@rug.nl

Phone: (+31) 627574221

The supplementary material provides additional information on the methods used in our study as well as additional supporting results.

List of contents

Appendix 1	The selection process of study population	5
Appendix 2	Model input.....	7
Appendix 3	Data cleaning	16
Appendix 4	The method of imputing risk-factor progression beyond observed values	18
Appendix 5	The method of calculating annual maximum cost-effective price	32
Appendix 6	Baseline characteristics.....	33
Appendix 7	Model validation.....	37
Appendix 8	Model outcomes under care-as-usual.....	39
Appendix 9	Model outcomes under hypothetical treatment strategies	44
Appendix 10	The sensitivity analysis of applying Ahlqvist’s subgroups	57

List of Tables

Table 2.1	Treatment targets used in scenario analyses for treat-to-target strategies.....	7
Table 2.2	The input of United Kingdom Prospective Diabetes Study Outcomes Model version 2 (UKPDS-OM2)	8
Table 2.3	Quality of life mean estimates and costs* used in the scenario’s analyses. (US costs expressed in \$2019 prices and UK costs expressed in £2019 prices).....	9
Table 2.4	The effect of hypothetical treatment on risk factors of each RHAPSODY data-driven subgroup	11
Table 2.5	The effect of hypothetical treatment on risk factors of each risk-driven subgroup	12
Table 3.1	Characteristics of pre-existing events in UKPDS OM2 input for DCS-dataset.....	17
Table 4.1	Missing data.....	19
Table 4.2	Random effect models for HbA1c over time by subgroups.....	20
Table 4.3	Random effect models for weight over time by subgroups	21
Table 4.4	Random effect models for eGFR over time by subgroups.....	22
Table 4.5	Random effect models for LDL over time by subgroup	23
Table 6.1	Baseline simulation characteristics of data-driven subgroups and risk-driven subgroups.....	33
Table 8.1	Absolute and standardized simulated outcome (mean and 95% CI) by RHAPSODY data-driven subgroups and risk-driven subgroups in the U.S. setting.....	39
Table 8.2	Absolute and standardized simulated outcome (mean and 95% CI) by RHAPSODY data-driven subgroups and risk-driven subgroups in the U.K. setting	40

Table 9.1 Outcomes of five-year 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in base case U.S. setting	44
Table 9.2 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in Scenario 1 U.S. setting	45
Table 9.3 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in Base case U.K. setting.....	46
Table 9.4 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in Scenario 1 U.K. setting.....	47
Table 10.1 Baseline simulation characteristics of Ahlqvist's data-driven subgroups.....	57
Table 10.2 Comparison between simulation baseline characteristics of corresponding subgroups in two data-driven approaches	58
Table 10.3 Absolute and standardized simulated outcome (mean and 95% CI) by Ahlqvist's data-driven subgroups in the U.S. setting	59
Table 10.4 Absolute and standardized simulated outcome (mean and 95% CI) by Ahlqvist's data-driven subgroups in the U.K. setting	60
Table 10.5 The effect of hypothetical treatment on risk factors of each Ahlqvist's data-driven subgroup	61
Table 10.6 Random effect models for HbA1c and weight over time by Ahlqvist's subgroups.....	62
Table 10.7 Random effect models for eGFR and LDL over time by Ahlqvist's subgroups.....	63
Table 10.8 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in U.S. setting	64
Table 10.9 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in U.K. setting	65

List of Figures

Figure 1.1 Flow chart depicting merging of different datasets	5
Figure 1.2 Flow diagram of adjusting diabetes duration	6
Figure 2.1 Progression of clinical parameters by RHAPSODY data-driven subgroups' mean over time since diagnosis.....	13
Figure 2.2 Progression of clinical parameters by risk-driven subgroups' mean over time since diagnosis	14
Figure 2.3 Graphical representations of scenario assumptions	15
Figure 4.1 The graphs for goodness-of-fit over time for HbA1c of each RHAPSODY subgroup	24
Figure 4.2 The graphs for goodness-of-fit over time for weight of each RHAPSODY subgroup	25
Figure 4.3 The graphs for goodness-of-fit over time for eGFR of each RHAPSODY subgroup	26
Figure 4.4 The graphs for goodness-of-fit over time for LDL of each RHAPSODY subgroup	27

Figure 4.5 The graphs for goodness-of-fit over time for HbA1c of each risk-driven subgroup	28
Figure 4.6 The graphs for goodness-of-fit over time for weight of each risk-driven subgroup	29
Figure 4.7 The graphs for goodness-of-fit over time for eGFR of each risk-driven subgroup	30
Figure 4.8 The graphs for goodness-of-fit over time for LDL of each risk-driven subgroup	31
Figure 6.1 Violin graphs and pairwise comparisons of baseline clustering indicators of risk-driven subgroups .	35
Figure 6.2 Violin graphs and pairwise comparisons of baseline clustering indicators of RHAPSODY data-driven subgroups.....	36
Figure 7.1 Model validation of subgroup-wise differences of relative risks in 5 years of simulation for RHAPSODY data-driven subgroups	37
Figure 7.2 Model validation of subgroup-wise differences of relative risks in 5 years of simulation for risk-driven subgroups.....	38
Figure 8.1 Non-standardized and standardized mean simulated lifetime QALYs and cost with 95% CI (in U.K. setting) for RHAPSODY data-driven and risk-driven subgroups.....	41
Figure 8.2 Simulated Kaplan-Meier plots of RHAPSODY data-driven subgroups.....	42
Figure 8.3 Simulated Kaplan-Meier plots of risk-driven subgroups.....	43
Figure 9.1 Maximum annual cost-effective price, incremental QALYs and complication costs of multiple-year guideline-based intensive management targeting HbA1c, BMI and LDL by RHAPSODY data-driven subgroups (in U.S. setting), for a range of treatment durations	48
Figure 9.2 Maximum annual cost-effective price, incremental QALYs and complication costs of multiple-year guideline-based intensive management targeting HbA1c, BMI and LDL by risk-driven subgroups (in U.S. setting), for a range of treatment durations	49
Figure 9.3 The results of maximum annual cost-effective prices of multiple scenarios, for a range of treatment duration.....	50
Figure 10.1 Comparisons among subgroups defined by different methodology	66
Figure 10.2 Violin graphs and pairwise comparisons of baseline clustering indicators of Ahlqvist's data-driven subgroups.....	67
Figure 10.3 The graphs for goodness-of-fit over time for HbA1c of each Ahlqvist's data-driven subgroup	68
Figure 10.4 The graphs for goodness-of-fit over time for weight of each Ahlqvist's data-driven subgroup.....	69
Figure 10.5 The graphs for goodness-of-fit over time for eGFR of each Ahlqvist's data-driven subgroup	70
Figure 10.6 The graphs for goodness-of-fit over time for LDL of each Ahlqvist's data-driven subgroup.....	71
Figure 10.7 Non-standardized and standardized mean simulated lifetime QALYs and cost with 95% CI (in U.S. setting) for Ahlqvist's data-driven clustering subgroups.....	72
Figure 10.8 Non-standardized and standardized mean simulated lifetime QALYs and cost with 95% CI (in U.K. setting) for Ahlqvist's data-driven clustering subgroups.....	73
Figure 10.9 Simulated Kaplan-Meier plots of Ahlqvist's data-driven subgroups	74
Figure 10.10 Maximum annual cost-effective price, incremental QALYs and complication costs of multiple-year guideline-based intensive management targeting HbA1c, BMI and LDL by Ahlqvist's data-driven subgroups (in U.S. setting), for a range of treatment durations	75

Appendix 1 The selection process of study population

- 1) Merge datasets to obtain most recent information
Three DCS dataset (A dataset with subgroup indicators as the output of previous published work (1) [DCS_Clusters_JC], an updated DCS dataset with observations over the period 1998 to 2019 [DCS_1998_2019_28042020e], and a dataset containing C peptide information [DCS_C peptide]) were merged and filtered, see Figure 1.1.
- 2) Adjust diabetes duration
The diabetes duration was adjusted according to age at diagnosis and current age, see Figure 1.2.

Figure 1.1 Flow chart depicting merging of different datasets

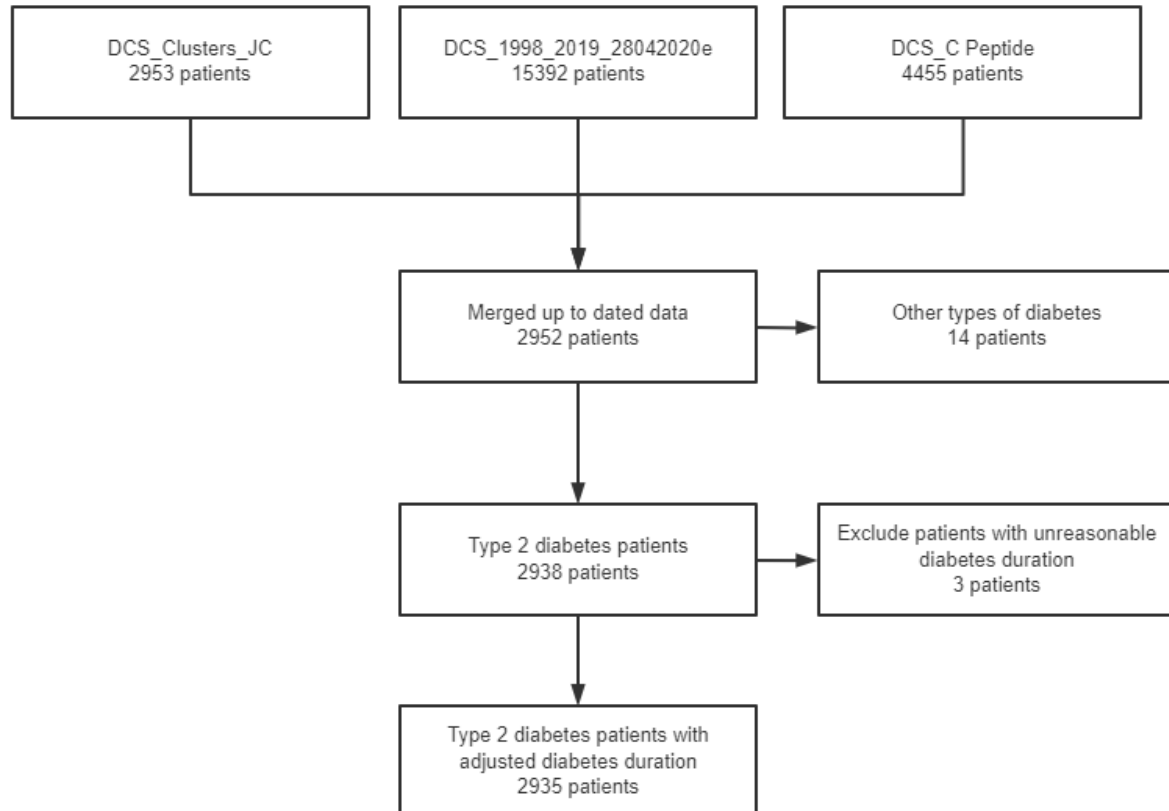


Figure 1.2 Flow diagram of adjusting diabetes duration

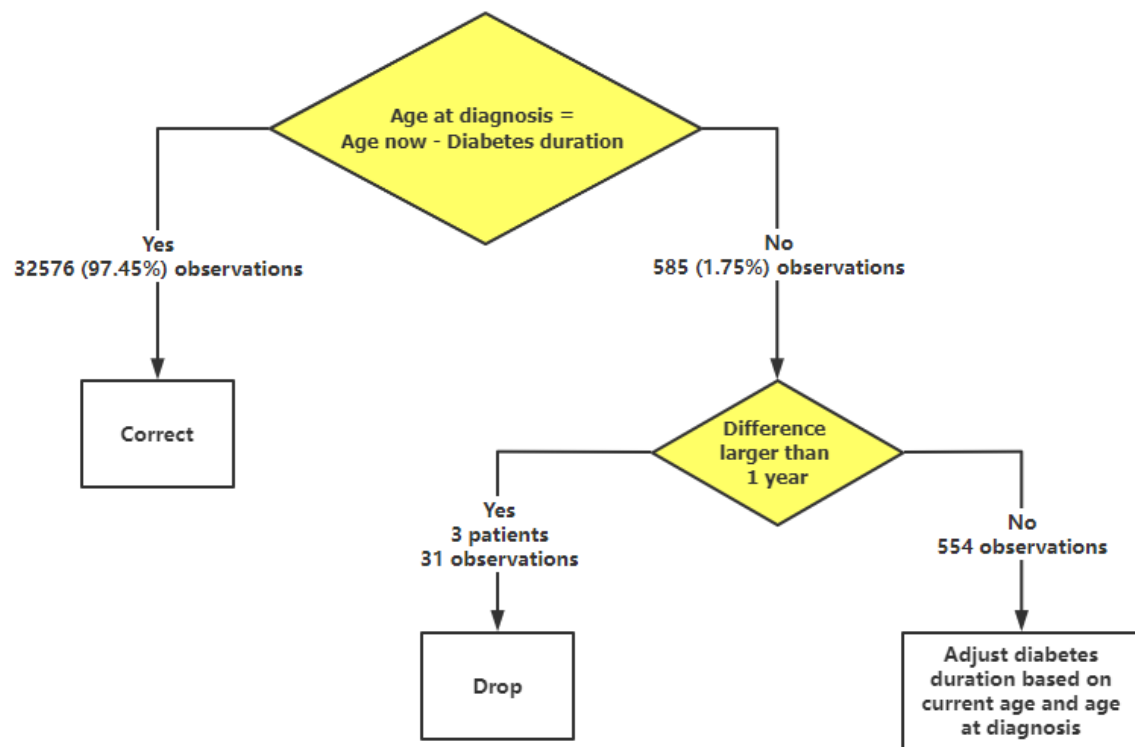


Table 2.1 Treatment targets used in scenario analyses for treat-to-target strategies

Guideline Based Treatment Target			
HbA1c (%, mmol/mol)	BMI (kg/m ²)	LDL (mmol/liter)*	
7% (53 mmol/mol)	25 (kg/m ²)	Very high risk (SCORE $\geq$ 10%)	1.4 mmol/L
		High risk (10% $>$ SCORE $\geq$ 5%)	1.8 mmol/L
		Moderate risk (5% $>$ SCORE $\geq$ 1%)	2.6 mmol/L
		Low risk (1% $>$ SCORE)	3 mmol/L

* An individualized target was applied for LDL based on each patient's cardiovascular risk calculated by Systematic Coronary Risk Estimation (SCORE) using age, sex, smoking status, total cholesterol and SBP (2; 3).

Table 2.2 The input of United Kingdom Prospective Diabetes Study Outcomes Model version 2 (UKPDS-OM2)

Required input in UKPDS-OM2			
	Baseline	Follow up	Extrapolation over life time
Demographic characteristics	Ethnicity		
	Gender		
	Age now		
	Duration of diabetes		
	Weight	Weight	Linear dynamic model
	Height		
Risk factor values	AF (Y/N)*	AF (Y/N)*	Regression model* and carry forward
	PVD (Y/N)*	PVD (Y/N)*	Regression model* and carry forward
	Current smoker (Y/N)	Current smoker (Y/N)	Carry forward
	Albuminuria (Y/N)	Albuminuria (Y/N)	Carry forward
	HDL	HDL	Carry forward
	LDL	LDL	Linear dynamic model
	Systolic BP	Systolic BP	Carry forward
	HbA1c	HbA1c	Linear dynamic model
	Heart rate*	Heart rate*	Regression model* and carry forward
	White blood cells*	White blood cells*	Regression model* and carry forward
	Haemoglobin*	Haemoglobin*	Regression model* and carry forward
	eGFR	eGFR	Linear dynamic model
Pre-existing events [†]	IHD		
	Heart failure		
	Amputation		
	Blindness in one eye		
	Renal failure		
	Stroke		
	MI		
	Ulcer		

AF, Atrial Fibrillation; PVD, Peripheral Vascular Disease; eGFR, estimated Glomerular Filtration Rate; IHD, Ischemic Heart Disease; MI, Myocardial Infarction

* The value is completely missing in our dataset. We used available risk factors (e.g. BMI, HbA1c) to impute them. For more information, please refer to Appendix 3.

[†] Details related to ICD and other characteristics of pre-existing events were listed in Table 3.1. The duration of pre-existing event (years) was measured as (baseline measurement date - event date) (days)/365, and rounded to the nearest integer.

Table 2.3 Quality of life mean estimates and costs* used in the scenario's analyses. (US costs expressed in \$2019 prices and UK costs expressed in £2019 prices)

Category	US†		UK†	
	Mean estimate	Source	Mean estimate	Source
Initial utility	0.807	Alva 2014(4)	0.807	Alva 2014(4)
Quality of life decrement				
IHD	0.000	Alva 2014(4)	0.000	Alva 2014(4)
MI	-0.065	Alva 2014(4)	-0.065	Alva 2014(4)
Heart failure	-0.101	Alva 2014(4)	-0.101	Alva 2014(4)
Stroke	-0.165	Alva 2014(4)	-0.165	Alva 2014(4)
Amputation	-0.172	Alva 2014(4)	-0.172	Alva 2014(4)
Blindness in one eye	0.000	Alva 2014(4)	0.000	Alva 2014(4)
Renal failure	-0.330	Lung 2011 (5)	-0.330	Lung 2011 (5)
Ulcer	-0.210	Lung 2011 (5)	-0.210	Lung 2011 (5)
Costs				
Diabetes management costs‡	\$12,904	ADA 2018 (6)	£875-£2,956*	Alva 2015(7)
<i>In the year of non-fatal event§</i>				
IHD	\$25,716	Ward 2014 (8)	£10,879-£19,888*	Alva 2015(7)
MI	\$67,811	Ward 2014 (8)	£7,692-£13,871*	Alva 2015(7)
Heart failure	\$28,542	Ward 2014 (8)	£4,030-£6,704*	Alva 2015(7)
Stroke	\$0	Ward 2014 (8)	£7,509-£12,631*	Alva 2015(7)
Amputation	\$10,861	Ward 2014 (8)	£13,033-£20,029*	Alva 2015(7)
Blindness in one eye	\$3,438	Ward 2014 (8)	£2,078-£5,993*	Alva 2015(7)
Renal failure	\$86,154	Ward 2014 (8)	£23,300	NHS Blood and Transplant 2009 (9)
Ulcer	\$2,579	Ward 2014 (8)	£7,779	Kerr 2014 (10)
<i>In the year of fatal event</i>				
IHD	\$25,716	Ward 2014 (8)	£4,715-£7,559*	Alva 2015(7)
MI	\$67,811	Ward 2014 (8)	£2,341-£8,474*	Alva 2015(7)
Heart failure	\$14,271	50% of non-fatal cost	£2,015-£3,352*	50% of non-fatal cost
Stroke	\$25,716	Ward 2014 (8)	£5,255-£8,124*	Alva 2015(7)
Amputation	\$10,861	Non-fatal minus 50% of cost subsequent years	£11,428-£16,561*	Non-fatal minus 50% of cost subsequent years
Renal failure	\$43,077	50% of non-fatal cost	£11,650	50% of non-fatal cost
<i>In subsequent years§</i>				
IHD	\$2,287	Ward 2014 (8)	£1,616-£5,050*	Alva 2015(7)
MI	\$2,287	Ward 2014 (8)	£1,586-£4,843*	Alva 2015(7)
Heart failure	\$2,287	Ward 2014 (8)	£2,197-£5,864*	Alva 2015(7)
Stroke	\$18,670	Ward 2014 (8)	£1,659-£5,034*	Alva 2015(7)
Amputation	\$0	Ward 2014 (8)	£3,209-£6,934*	Alva 2015(7)
Blindness in one eye	\$3,438	Ward 2014 (8)	£1,098-£3,177*	Alva 2015(7)
Renal failure	\$86,154	Ward 2014 (8)	£23,300	NHS Blood and Transplant 2009 (9)
Ulcer	\$2,579	Ward 2014 (8)	£1,179	Kerr 2014 (10)

* Costs varied by sex and age group (up to 50, 51-60, 61-70, 71-80, 81+).

† US costs include medications, outpatient consultations, emergency department visits and inpatient stays(8). UK costs include non-inpatient contacts (e.g. general practitioner, nurse, hospital eye clinic, etc.) and inpatient stay(7).

‡ US diabetes management costs obtained from ADA 2018 excluding costs of complications. UK diabetes management costs obtained from UKPDS and concern costs in the absence of complications(7). IGT costs estimated by applying ratio of IGT and diabetes costs per individual from Khan 2007(11) for the US setting (0.74) and DPP(12) for the UK setting (0.77).

§ In the US analysis, IGT/diabetes management costs (\$6,762/\$9,158) were added to costs of non-fatal events in the year of the event and in subsequent years. In the UK analysis, management costs were already included in the costs of complications and no further costs were added to them.

Table 2.4 The effect of hypothetical treatment on risk factors of each RHAPSODY data-driven subgroup

		Year 1	Year 2	Year 3	Year 4	Year 5	
Treat-to-target scenario	HbA1c	RHAP-SIDD	0.45 (5.24%)	0.5 (5.74%)	0.51 (5.75%)	0.5 (5.78%)	0.5 (5.89%)
		RHAP-SIRD	0.08 (0.95%)	0.1 (1.24%)	0.11 (1.35%)	0.11 (1.35%)	0.13 (1.59%)
		RHAP-MOD	0.24 (2.84%)	0.28 (3.3%)	0.27 (3.1%)	0.27 (3.15%)	0.32 (3.64%)
		RHAP-MD	0.16 (1.99%)	0.22 (2.59%)	0.2 (2.34%)	0.21 (2.49%)	0.24 (2.87%)
		RHAP-MDH	0.06 (0.7%)	0.09 (1.07%)	0.11 (1.26%)	0.1 (1.28%)	0.1 (1.2%)
	LDL	RHAP-SIDD	0.4 (11.62%)	0.33 (9.99%)	0.32 (9.16%)	0.31 (8.98%)	0.29 (8.67%)
		RHAP-SIRD	0.65 (20.86%)	0.58 (19.07%)	0.52 (17.78%)	0.52 (17.64%)	0.51 (17.5%)
		RHAP-MOD	0.31 (8.57%)	0.26 (7.27%)	0.24 (6.66%)	0.21 (5.98%)	0.22 (6.19%)
		RHAP-MD	0.37 (9.78%)	0.3 (8.05%)	0.27 (7.24%)	0.27 (7.4%)	0.23 (6.4%)
		RHAP-MDH	0.66 (20.59%)	0.63 (19.5%)	0.59 (18.37%)	0.56 (17.67%)	0.55 (17.42%)
	Weight	RHAP-SIDD	13.23 (13.67%)	13.44 (13.84%)	13.3 (13.76%)	13.04 (13.53%)	12.65 (13.09%)
		RHAP-SIRD	13.42 (14.57%)	13.33 (14.45%)	12.87 (13.99%)	12.13 (13.25%)	11.49 (12.62%)
		RHAP-MOD	36.18 (31.83%)	35.85 (31.66%)	35.24 (31.19%)	34.4 (30.67%)	33.29 (29.91%)
		RHAP-MD	11.61 (12.5%)	11.71 (12.56%)	11.52 (12.38%)	11.06 (11.91%)	10.89 (11.74%)
		RHAP-MDH	7.29 (8.35%)	7.24 (8.26%)	7.14 (8.12%)	6.97 (7.94%)	6.79 (7.73%)
5%-Improvement-based scenario	HbA1c	RHAP-SIDD	0.36 (5%)	0.36 (5%)	0.36 (5%)	0.36 (5%)	0.36 (5%)
		RHAP-SIRD	0.32 (5%)	0.33 (5%)	0.33 (5%)	0.33 (5%)	0.33 (5%)
		RHAP-MOD	0.34 (5%)	0.34 (5%)	0.34 (5%)	0.34 (5%)	0.34 (5%)
		RHAP-MD	0.33 (5%)	0.34 (5%)	0.33 (5%)	0.34 (5%)	0.34 (5%)
		RHAP-MDH	0.32 (5%)	0.32 (5%)	0.32 (5%)	0.32 (5%)	0.32 (5%)
	LDL	RHAP-SIDD	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		RHAP-SIRD	0.12 (5%)	0.12 (5%)	0.11 (5%)	0.11 (5%)	0.11 (5%)
		RHAP-MOD	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		RHAP-MD	0.13 (5%)	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		RHAP-MDH	0.13 (5%)	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
	Weight	RHAP-SIDD	4.37 (5%)	4.38 (5%)	4.37 (5%)	4.36 (5%)	4.33 (5%)
		RHAP-SIRD	4.36 (5%)	4.35 (5%)	4.32 (5%)	4.28 (5%)	4.24 (5%)
		RHAP-MOD	5.53 (5%)	5.51 (5%)	5.48 (5%)	5.44 (5%)	5.38 (5%)
		RHAP-MD	4.31 (5%)	4.31 (5%)	4.3 (5%)	4.27 (5%)	4.26 (5%)
		RHAP-MDH	3.92 (5%)	3.91 (5%)	3.91 (5%)	3.89 (5%)	3.87 (5%)

Table 2.5 The effect of hypothetical treatment on risk factors of each risk-driven subgroup

			Year 1	Year 2	Year 3	Year 4	Year 5
Treat-to-target scenario	HbA1c	H1S1	0.1 (1.17%)	0.11 (1.39%)	0.11 (1.35%)	0.12 (1.44%)	0.15 (1.74%)
		H1S2	0.06 (0.77%)	0.07 (0.82%)	0.06 (0.79%)	0.07 (0.97%)	0.09 (1.14%)
		H2S1	0.45 (4.98%)	0.46 (5.36%)	0.45 (5.12%)	0.45 (5.22%)	0.46 (5.41%)
		H2S2	0.34 (3.79%)	0.22 (2.72%)	0.25 (3.01%)	0.23 (2.76%)	0.26 (3.17%)
	LDL	H1S1	0.31 (8.09%)	0.27 (6.95%)	0.24 (6.21%)	0.23 (5.99%)	0.21 (5.34%)
		H1S2	1.03 (33.06%)	0.93 (30.64%)	0.85 (28.47%)	0.84 (28.08%)	0.81 (27.58%)
		H2S1	0.23 (6.09%)	0.19 (4.96%)	0.18 (4.69%)	0.18 (4.5%)	0.17 (4.33%)
		H2S2	0.95 (31.38%)	0.8 (27.56%)	0.77 (26.04%)	0.74 (25.55%)	0.76 (26.06%)
	Weight	H1S1	17.11 (16.85%)	17.07 (16.86%)	16.88 (16.64%)	16.51 (16.31%)	16.18 (16.03%)
		H1S2	11.48 (12.25%)	11.26 (12%)	10.74 (11.52%)	10.27 (11.01%)	9.47 (10.17%)
		H2S1	17.72 (17.22%)	17.81 (17.27%)	17.48 (17.04%)	16.87 (16.55%)	16.31 (16.12%)
		H2S2	11.6 (12.68%)	11.65 (12.66%)	11.28 (12.29%)	10.38 (11.42%)	10 (11.02%)
5%-Improvement-based scenario	HbA1c	H1S1	0.32 (5%)	0.33 (5%)	0.32 (5%)	0.33 (5%)	0.33 (5%)
		H1S2	0.32 (5%)	0.32 (5%)	0.32 (5%)	0.32 (5%)	0.32 (5%)
		H2S1	0.36 (5%)	0.36 (5%)	0.36 (5%)	0.36 (5%)	0.36 (5%)
		H2S2	0.35 (5%)	0.34 (5%)	0.35 (5%)	0.35 (5%)	0.35 (5%)
	LDL	H1S1	0.13 (5%)	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		H1S2	0.13 (5%)	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		H2S1	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.11 (5%)
		H2S2	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
	Weight	H1S1	4.53 (5%)	4.53 (5%)	4.52 (5%)	4.5 (5%)	4.48 (5%)
		H1S2	4.2 (5%)	4.18 (5%)	4.15 (5%)	4.11 (5%)	4.06 (5%)
		H2S1	4.62 (5%)	4.63 (5%)	4.61 (5%)	4.58 (5%)	4.55 (5%)
		H2S2	4.21 (5%)	4.21 (5%)	4.19 (5%)	4.13 (5%)	4.1 (5%)

Figure 2.1 Progression of clinical parameters by RHAPSODY data-driven subgroups' mean over time since diagnosis

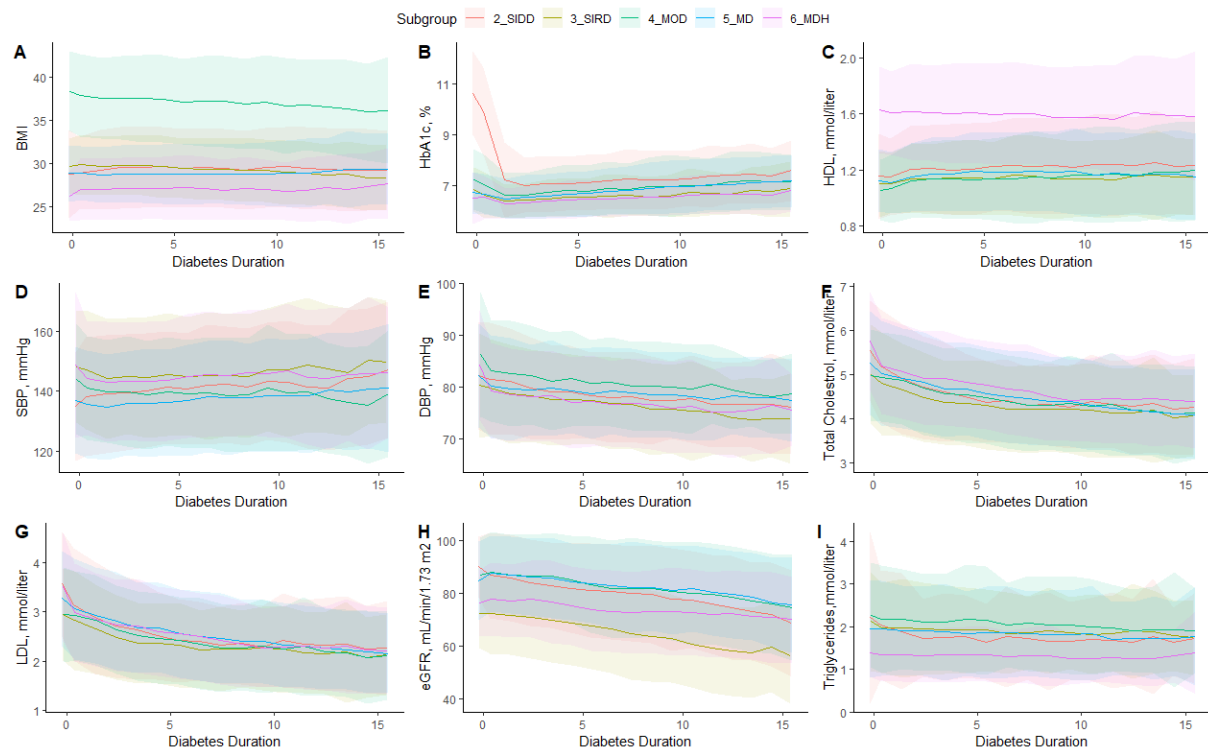
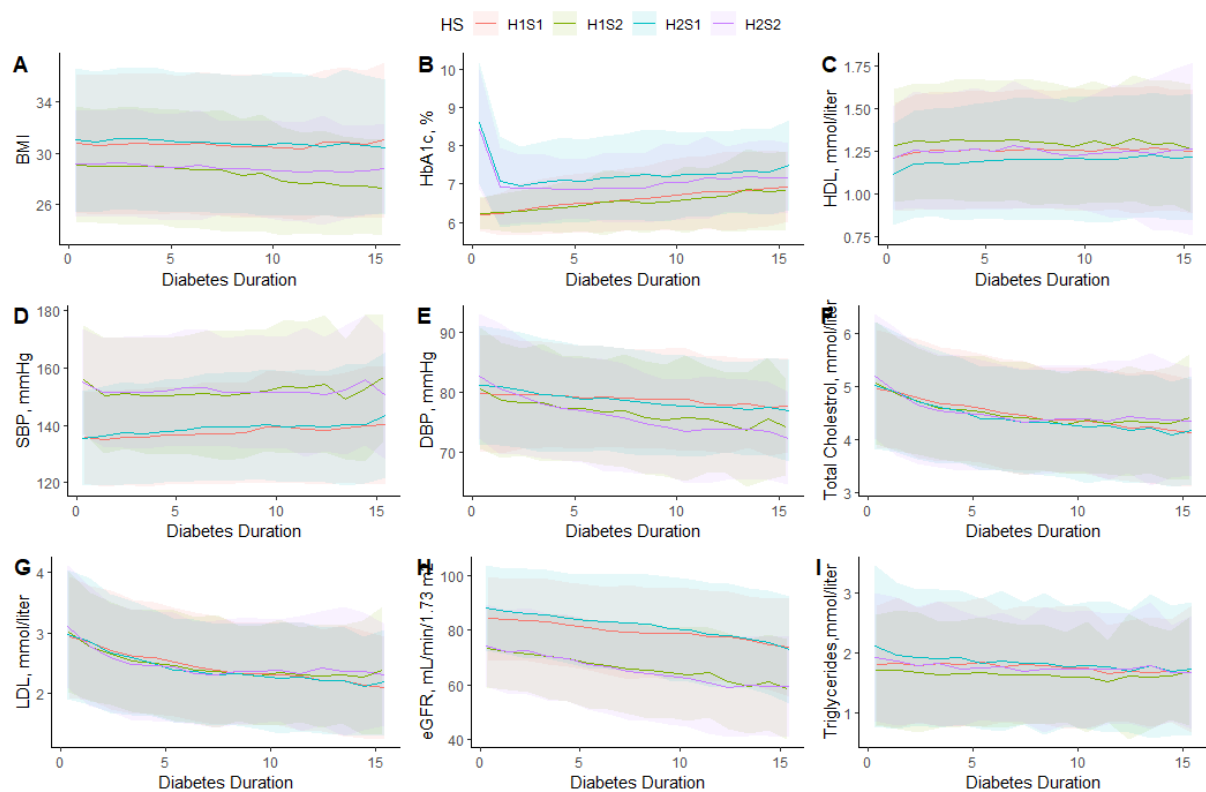
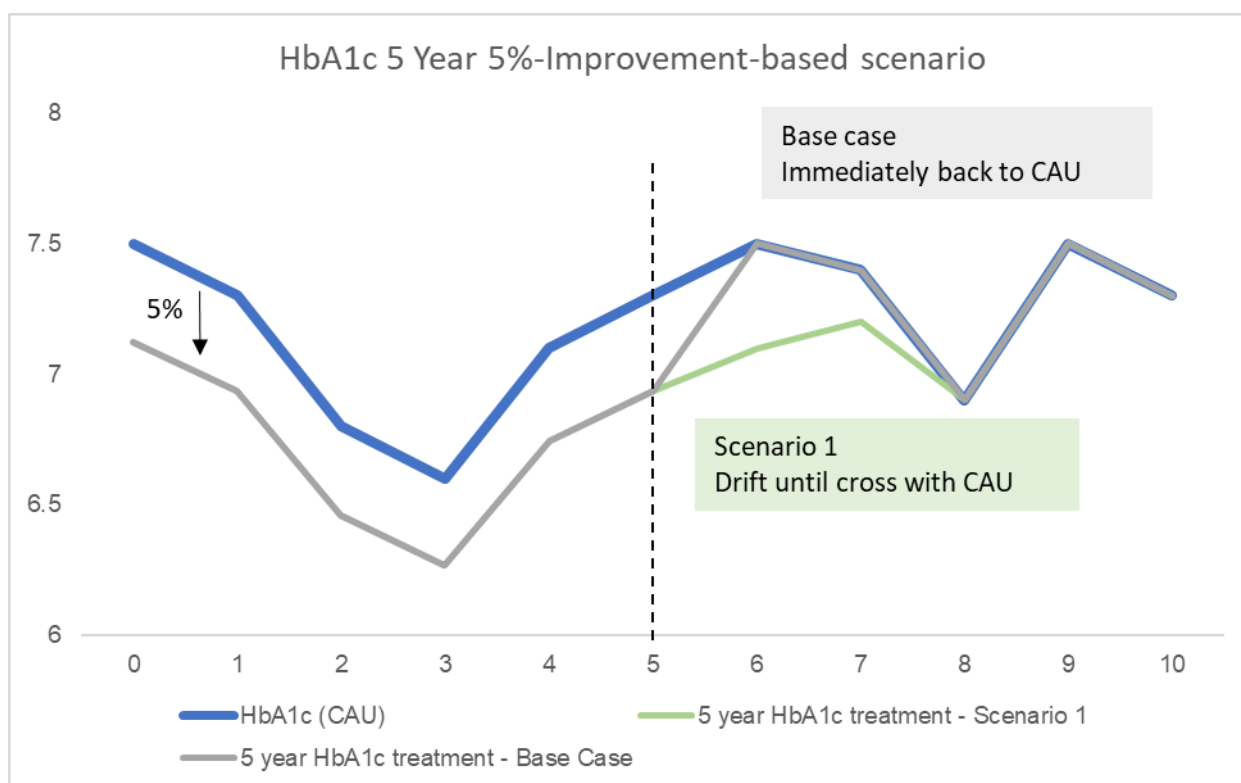
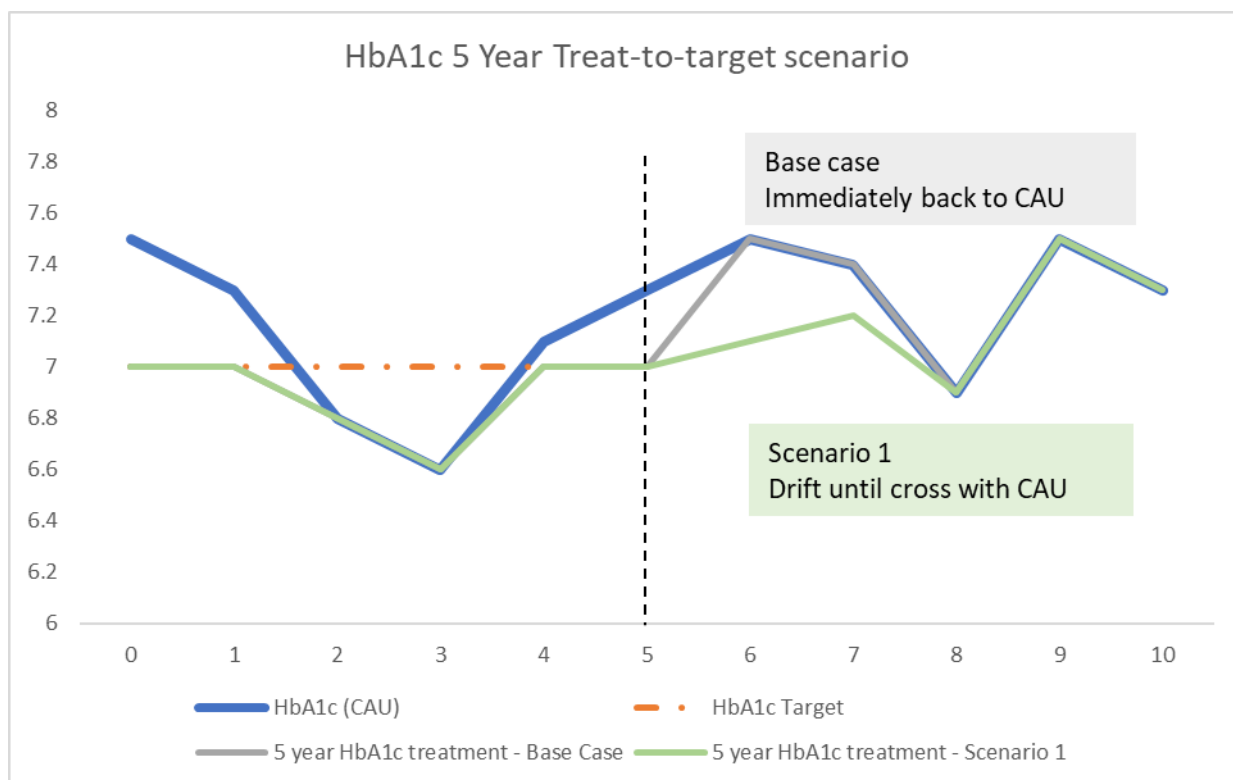


Figure 2.2 Progression of clinical parameters by risk-driven subgroups' mean over time since diagnosis



The data are represented as annually diabetes duration mean values (solid line) $\pm$ standard deviation (shaded areas) of each risk-driven subgroup.

Figure 2.3 Graphical representations of scenario assumptions



CAU, Care-as-usual

Appendix 3 Data cleaning

- 1) To avoid survival bias, as the baseline we used observed patient-level data from each individual's first observation after 2008 (the time of obtaining informed consent).
- 2) Missing ethnicities (4.97%) was set to "White", reflecting the most frequently observed ethnicity (92.13%).
- 3) Micro or macroalbuminuria was defined as urinary microalbumin (mg/liter) >50 (mg/liter).
- 4) Missing diabetes duration and age values were interpolated between two known values over time, and extrapolated by adding 1 to the last observed value annually.
- 5) Missing values of first observation of HbA1c, LDL, weight, eGFR, atrial fibrillation, peripheral vascular disease, smoking status, albuminuria, HDL, SBP, heart rate, white blood cells, and hemoglobin were imputed by the mean at first observation of each subgroup.
- 6) Missing values for risk factors over time were imputed as described in Table 2.2 by linear dynamic models (for HbA1c, LDL, weight, and eGFR; see details in Appendix 4) or carried forward from last observed value (for atrial fibrillation, peripheral vascular disease, smoking status, albuminuria, HDL, SBP, heart rate, white blood cells, and hemoglobin).
- 7) For individuals with multiple records of the same variable within a single year, the first observation was applied.
- 8) Variables which were not recorded in the registry (Table 2.2) but were required by the UKPDS-OM2 were imputed based on existing DCS data and published equations (13).
- 9) Events were coded according to the International Stratification of Diseases, Injuries and Causes of Death, ninth revision (ICD9) (14) and defined by self-report, which was partly verified by the electronic patient records from the regional hospital and GP (15). Furthermore, Sims (16), and retinopathy grades from EURODIAB scale (17) were also applied as in Table 3.1. Amputation was identified for patients identified as Sims=3. The duration of pre-existing event (years) was measured by (baseline measurement date - event date) (days)/365, and rounded to the nearest integer.

Table 3.1 Characteristics of pre-existing events in UKPDS OM2 input for DCS-dataset

	ICD 9	Additional information
Ischemic heart disease	411-414	
Heart failure	428	
Amputation	84.10, 84.11-84.14, 84.15-84.18, 84.3, 84.4; except: 800-959 all diagnosis fields	Sims=3 (Exclude patients identified as Ulcer)
Blindness in one eye	369	Score=5
Renal failure	586	
Stroke	430-434 and 436 (non-fatal) & Fatal 430-438	
Myocardial infarction	410&798	
Ulcer	707.10, 707.14	

Appendix 4 The method of imputing risk-factor progression beyond observed values

Linear dynamic models were estimated on the time paths of HbA1c, LDL, weight and eGFR for each subgroup, to allow for imputation of missing data and extrapolation of risk factor trajectories where data were censored. Following recent research (18), these models included the value of the risk factor in the previous year, the number of years since diagnosis (in log form for improved fit), value of the risk factor at baseline, sex, an indicator for first year of follow-up (to capture treatment effect following diagnosis), and an interaction term for first year of follow-up and risk factor at baseline (to capture different treatment effects conditional on initial risk factor value). Missing data (Table 4.1) were imputed by Multivariate Imputation by Chained Equations when estimating the linear dynamic models, and results were combined in 25 imputed data sets using the packages *mice* and *mitools* in R 4.1.0 (19-21). Models were estimated for each clustered subgroup and each risk-driven subgroup. Results were listed in Table 4.2-Table 4.5 and Figure 4.1-Figure 4.8.

Table 4.1 Missing data

Data	Number of missing	Proportion of missing
Length	1	0.00%
BMI	207	0.62%
Weight	209	0.63%
SBP	262	0.78%
DBP	264	0.79%
Total cholesterol	300	0.90%
Triglyceride	304	0.91%
Creatinine	316	0.95%
Cholesterol ratio	316	0.95%
HDL	317	0.95%
HbA1c	323	0.97%
eGFR	324	0.97%
LDL	510	1.53%
Fast glucose	532	1.59%
Smoking status	593	1.78%

Table 4.2 Random effect models for HbA1c over time by subgroups

	Overall	RHAP-SIDD	RHAP-SIRD	RHAP-MOD	RHAP-MOD	RHAP-MDH	H1S1	H1S2	H2S1	H2S2
No. of observations	28130	4090	5433	4899	8688	5020	12197	8928	4436	2569
No. of individuals	2930	364	637	519	859	551	1272	541	840	277
Adjusted R ²	0.448	0.357	0.401	0.395	0.468	0.517	0.448	0.357	0.401	0.395
Parameters	Estimate of coefficient (standard error)									
Constant	1.819*** (0.036)	2.655*** (0.146)	1.725*** (0.096)	2.081*** (0.107)	1.352*** (0.076)	1.327*** (0.083)	0.851** * (0.087)	2.405*** (0.085)	0.939** * (0.15)	2.843*** (0.157)
Male	0.018* (0.009)	0.002 (0.029)	0.003 (0.017)	0.041. (0.024)	0.035* (0.015)	0.018 (0.015)	0.03** (0.011)	0.043* (0.019)	0.025 (0.018)	-0.063* (0.032)
Value of Y in previous year	0.645*** (0.005)	0.598*** (0.014)	0.623*** (0.013)	0.63*** (0.012)	0.638*** (0.009)	0.631*** (0.015)	0.656** * (0.009)	0.617*** (0.009)	0.629** * (0.015)	0.551*** (0.019)
Log(diabetes duration)	0.278*** (0.018)	0.427*** (0.059)	0.2*** (0.037)	0.262*** (0.049)	0.298*** (0.032)	0.197*** (0.032)	0.259** * (0.024)	0.306*** (0.038)	0.263** * (0.039)	0.329*** (0.064)
Baseline of Y	0.054*** (0.003)	-0.009 (0.01)	0.095*** (0.011)	0.038*** (0.011)	0.133*** (0.01)	0.143*** (0.011)	0.198** * (0.014)	0.01 (0.006)	0.206** * (0.025)	0.009 (0.01)
Dummy variable of Year 1	3.184*** (0.08)	3.215*** (0.399)	2.282*** (0.234)	3.185*** (0.278)	2.084*** (0.24)	2.25*** (0.217)	1.144** * (0.302)	3.957*** (0.21)	1.431** (0.453)	3.054*** (0.337)
Baseline of Y*Dummy variable of Year 1	- 0.499*** (0.011)	-0.491*** (0.04)	-0.367*** (0.035)	-0.49*** (0.039)	-0.328*** (0.036)	-0.368*** (0.033)	- 0.18*** (0.049)	- 0.576*** (0.024)	- 0.219** (0.072)	- 0.454*** (0.041)

Significance is indicated as 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 4.3 Random effect models for weight over time by subgroups

	Overall	RHAP-SIDD	RHAP-SIRD	RHAP-MOD	RHAP-MOD	RHAP-MDH	H1S1	H1S2	H2S1	H2S2
No. of observations	28130	4090	5433	4899	8688	5020	12197	8928	4436	2569
No. of individuals	2930	364	637	519	859	551	1272	541	840	277
Adjusted R ²	0.939	0.933	0.903	0.866	0.935	0.930	0.939	0.933	0.903	0.866
Parameters	Estimate of coefficient (standard error)									
Constant	2.216** *	2.448***	2.342***	5.641***	1.565***	1.648***	1.892** *	3.156** *	2.828** *	2.436** *
	(0.174)	(0.478)	(0.53)	(0.767)	(0.316)	(0.392)	(0.256)	(0.337)	(0.509)	(0.711)
Male	0.273** *	0.493***	0.163	0.322.	0.29**	0.267*	0.144.	0.591** *	0.309*	0.489**
	(0.055)	(0.142)	(0.118)	(0.191)	(0.097)	(0.106)	(0.082)	(0.106)	(0.136)	(0.191)
Value of Y in previous year	0.876** *	0.892***	0.857***	0.84***	0.895***	0.893***	0.879** *	0.88***	0.803** *	0.843** *
	(0.006)	(0.014)	(0.021)	(0.012)	(0.011)	(0.017)	(0.008)	(0.009)	(0.028)	(0.03)
Log(diabetes duration)	- 0.789** *	-1.143***	-1.249***	-1.551***	-0.296.	-0.782***	- 0.617** *	- 0.972** *	- 1.475** *	- -1.07**
	(0.108)	(0.27)	(0.24)	(0.378)	(0.163)	(0.207)	(0.163)	(0.204)	(0.286)	(0.345)
Baseline of Y	0.102** *	0.087***	0.119***	0.112***	0.087***	0.089***	0.102** *	0.087** *	0.166** *	0.129** *
	(0.006)	(0.014)	(0.021)	(0.011)	(0.01)	(0.016)	(0.008)	(0.008)	(0.027)	(0.028)
Dummy variable of Year 1	1.295* (0.53)	-1.411 (1.452)	1.306 (1.667)	-1.117 (2.386)	4.364*** (1.146)	0.932 (1.123)	1.593* (0.785)	0.877 (0.966)	0.608 (1.403)	1.246 (1.861)
Baseline of Y*Dummy variable of Year 1	- 0.023** *	0.015	-0.026	-0.005	-0.06***	-0.019	-0.025**	-0.02*	-0.017	-0.026
	(0.006)	(0.016)	(0.019)	(0.021)	(0.013)	(0.014)	(0.008)	(0.01)	(0.016)	(0.021)

Significance is indicated as 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 4.4 Random effect models for eGFR over time by subgroups

	Overall	RHAP-SIDD	RHAP-SIRD	RHAP-MOD	RHAP-MOD	RHAP-MDH	H1S1	H1S2	H2S1	H2S2
No. of observations	28130	4090	5433	4899	8688	5020	12197	8928	4436	2569
No. of individuals	2930	364	637	519	859	551	1272	541	840	277
Adjusted R^2	0.786	0.792	0.767	0.751	0.752	0.746	0.786	0.792	0.767	0.751
Parameters	Estimate of coefficient (standard error)									
Constant	4.345** *	4.447***	5.473***	6.355***	5.986***	5.231***	6.599***	5.604***	4.006** *	6.76***
	(0.351)	(0.954)	(0.787)	(0.96)	(0.748)	(0.818)	(0.542)	(0.708)	(0.831)	(1.296)
Male	-0.125	-0.106	-0.219	0.025	0.35.	-0.38	0.217	-0.077	-0.069	0.208
	(0.102)	(0.271)	(0.235)	(0.265)	(0.18)	(0.233)	(0.148)	(0.188)	(0.261)	(0.382)
Value of Y in previous year	0.794** *	0.839***	0.783***	0.765***	0.77***	0.785***	0.756***	0.809***	0.784** *	0.792** *
	(0.006)	(0.013)	(0.014)	(0.013)	(0.011)	(0.012)	(0.008)	(0.012)	(0.014)	(0.023)
Log(diabetes duration)	- 3.312** *	-2.282***	-4.336***	-4.653***	-3.425***	-3.069***	- 3.689***	- 3.251***	- 3.754** *	- 5.062** *
	(0.219)	(0.566)	(0.542)	(0.569)	(0.382)	(0.484)	(0.322)	(0.39)	(0.571)	(0.864)
Baseline of Y	0.162** *	0.106***	0.156***	0.181***	0.171***	0.159***	0.178***	0.137***	0.167** *	0.13***
	(0.006)	(0.013)	(0.014)	(0.014)	(0.012)	(0.013)	(0.008)	(0.011)	(0.016)	(0.022)
Dummy variable of Year 1	8.519** *	12.593***	4.448*	13.791***	11.2***	6.231*	11.358** *	11.603** *	5.137*	1.002
	(1.082)	(2.885)	(1.955)	(2.771)	(2.567)	(2.518)	(1.59)	(2.357)	(2.261)	(3.435)
Baseline of Y*Dummy variable of Year 1	- 0.115** *	-0.158***	-0.076**	-0.178***	-0.139***	-0.087**	- 0.144***	- 0.147***	-0.081**	-0.051
	(0.013)	(0.032)	(0.026)	(0.031)	(0.029)	(0.032)	(0.019)	(0.027)	(0.03)	(0.045)

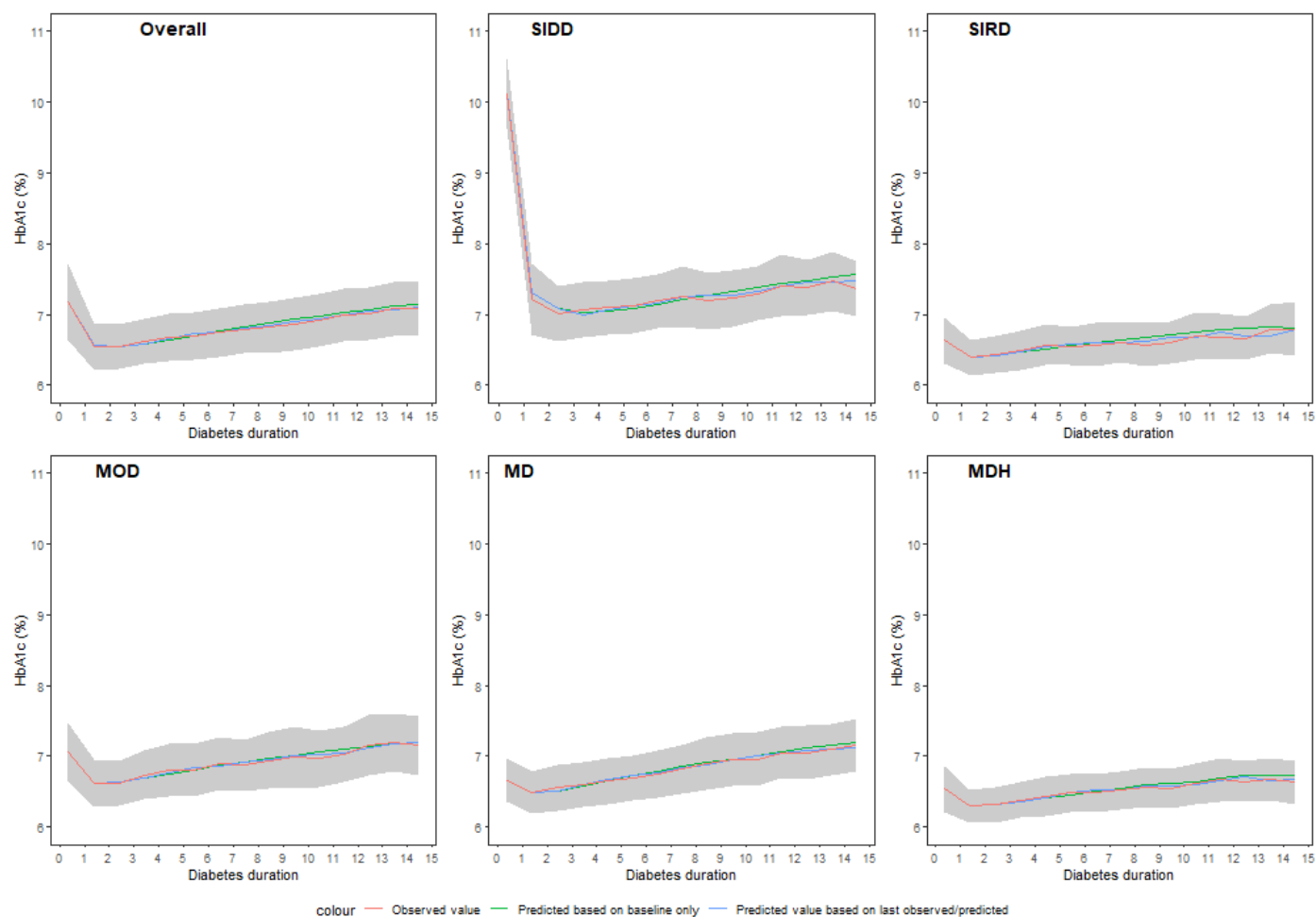
Significance is indicated as 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1.

Table 4.5 Random effect models for LDL over time by subgroup

	Overall	RHAP-SIDD	RHAP-SIRD	RHAP-MOD	RHAP-MOD	RHAP-MDH	H1S1	H1S2	H2S1	H2S2
No. of observations	28130	4090	5433	4899	8688	5020	12197	8928	4436	2569
No. of individuals	2930	364	637	519	859	551	1272	541	840	277
Adjusted R ²	0.502	0.497	0.469	0.520	0.484	0.544	0.502	0.497	0.469	0.520
Parameters	Estimate of coefficient (standard error)									
Constant	0.717** *	0.729***	0.687***	0.599***	0.854***	0.678***	0.714** *	0.746** *	0.752** *	0.578** *
	(0.174)	(0.478)	(0.53)	(0.767)	(0.316)	(0.392)	(0.256)	(0.337)	(0.509)	(0.711)
Male	- 0.063** *	-0.021	-0.039*	-0.09***	-0.083***	-0.057***	- 0.057** *	- 0.071** *	-0.062**	-0.048.
	(0.055)	(0.142)	(0.118)	(0.191)	(0.097)	(0.106)	(0.082)	(0.106)	(0.136)	(0.191)
Value of Y in previous year	0.659** *	0.632***	0.636***	0.684***	0.643***	0.69***	0.675** *	0.628** *	0.66***	0.665** *
	(0.006)	(0.014)	(0.021)	(0.012)	(0.011)	(0.017)	(0.008)	(0.009)	(0.028)	(0.03)
Log(diabetes duration)	- 0.163** *	-0.161***	-0.104**	-0.109**	-0.222***	-0.197***	- 0.194** *	- 0.181** *	- 0.141** *	0.014
	(0.108)	(0.27)	(0.24)	(0.378)	(0.163)	(0.207)	(0.163)	(0.204)	(0.286)	(0.345)
Baseline of Y	0.074** *	0.08***	0.073***	0.076***	0.068***	0.077***	0.072** *	0.091** *	0.058** *	0.063** *
	(0.006)	(0.014)	(0.021)	(0.011)	(0.01)	(0.016)	(0.008)	(0.008)	(0.027)	(0.028)
Dummy variable of Year 1	0.47***	0.05	0.43***	0.577***	0.605***	0.527***	0.529** *	0.371** *	0.507** *	0.385*
	(0.53)	(1.452)	(1.667)	(2.386)	(1.146)	(1.123)	(0.785)	(0.966)	(1.403)	(1.861)
Baseline of Y*Dummy variable of Year 1	- 0.181** *	-0.061	-0.164***	-0.194***	-0.222***	-0.217***	- 0.199** *	- 0.136** *	- 0.215** *	- 0.148**
	(0.006)	(0.016)	(0.019)	(0.021)	(0.013)	(0.014)	(0.008)	(0.01)	(0.016)	(0.021)

Significance is indicated as 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Figure 4.1 The graphs for goodness-of-fit over time for HbA1c of each RHAPSODY subgroup



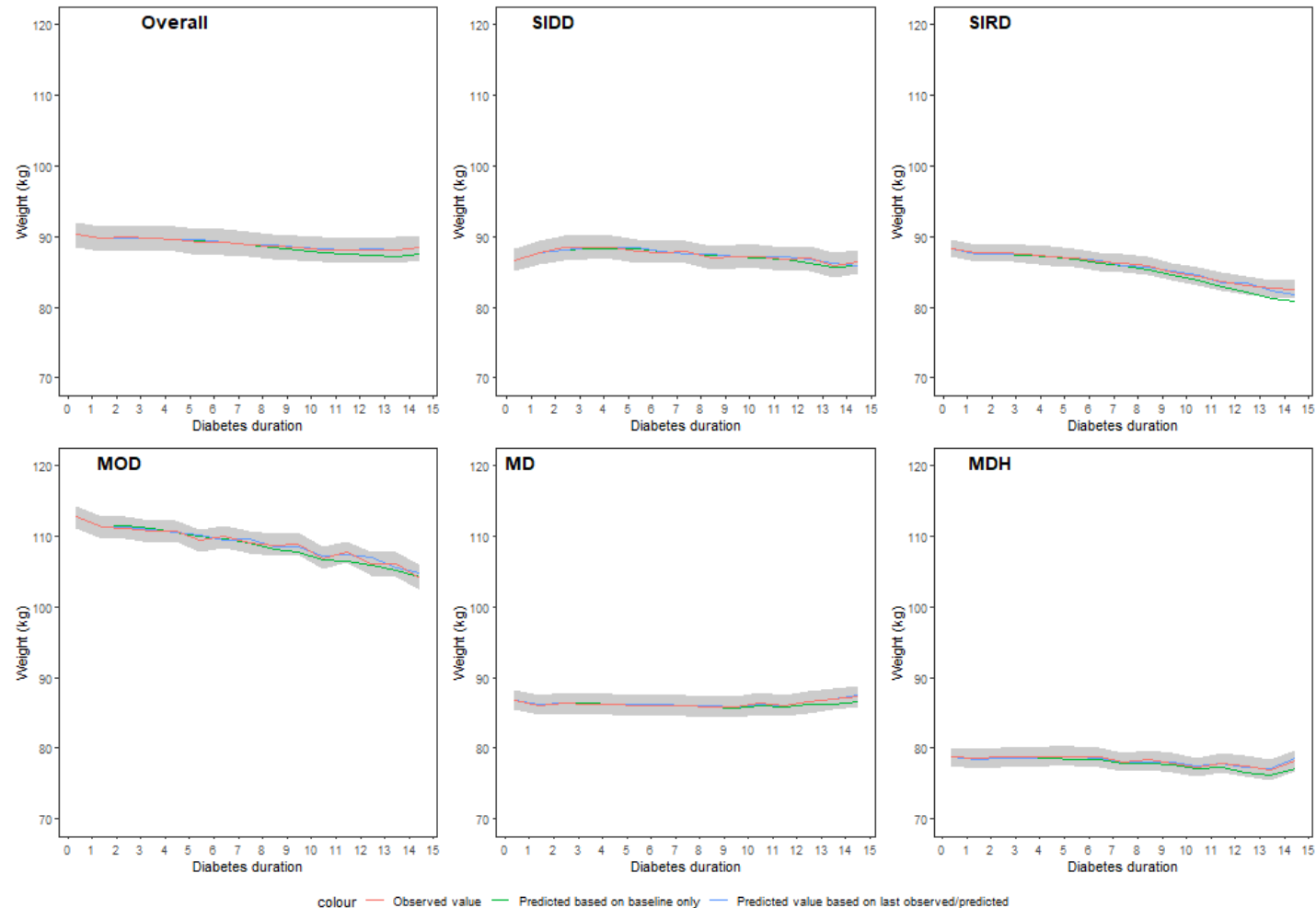
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.2 The graphs for goodness-of-fit over time for weight of each RHAPSODY subgroup



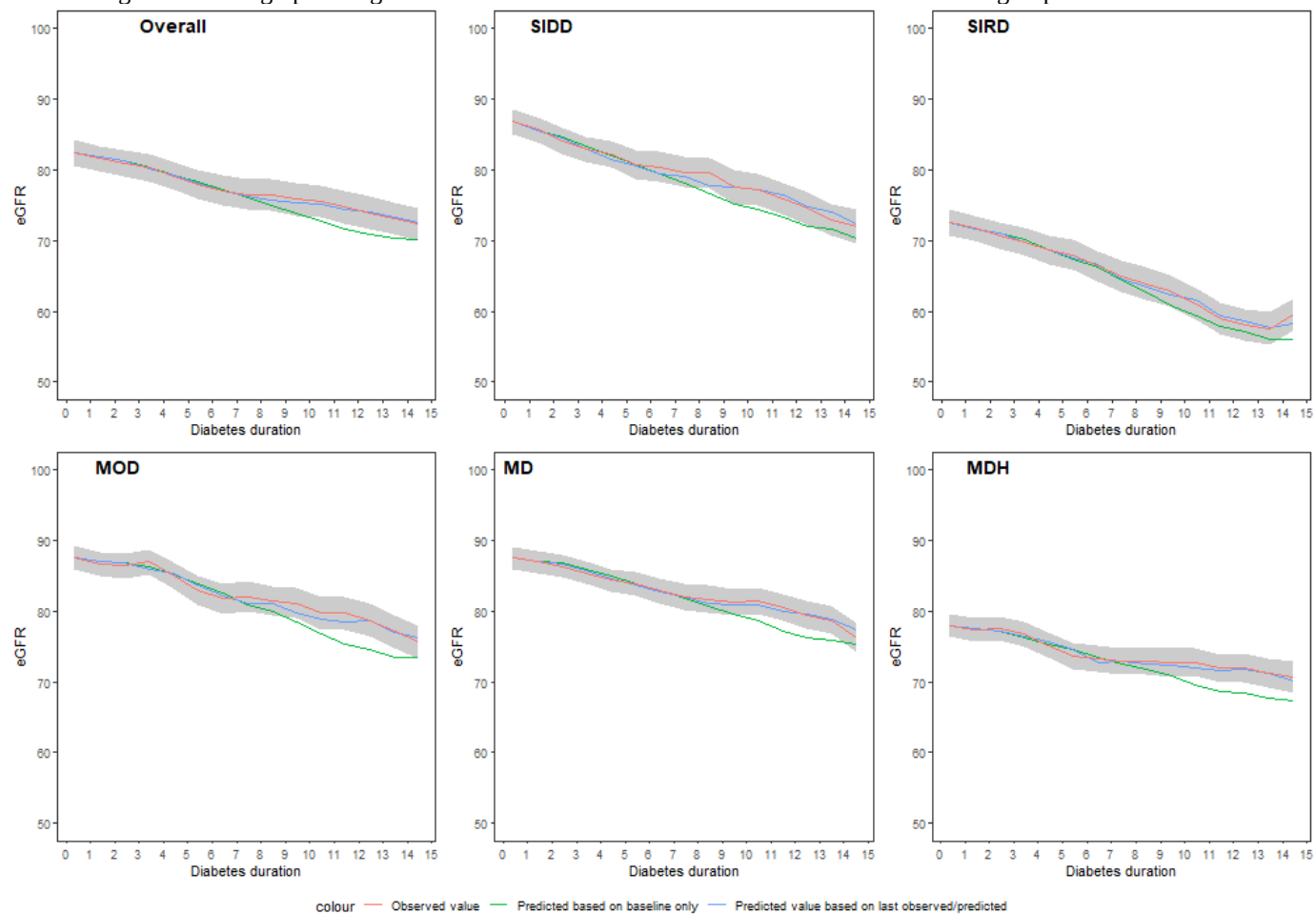
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.3 The graphs for goodness-of-fit over time for eGFR of each RHAPSODY subgroup



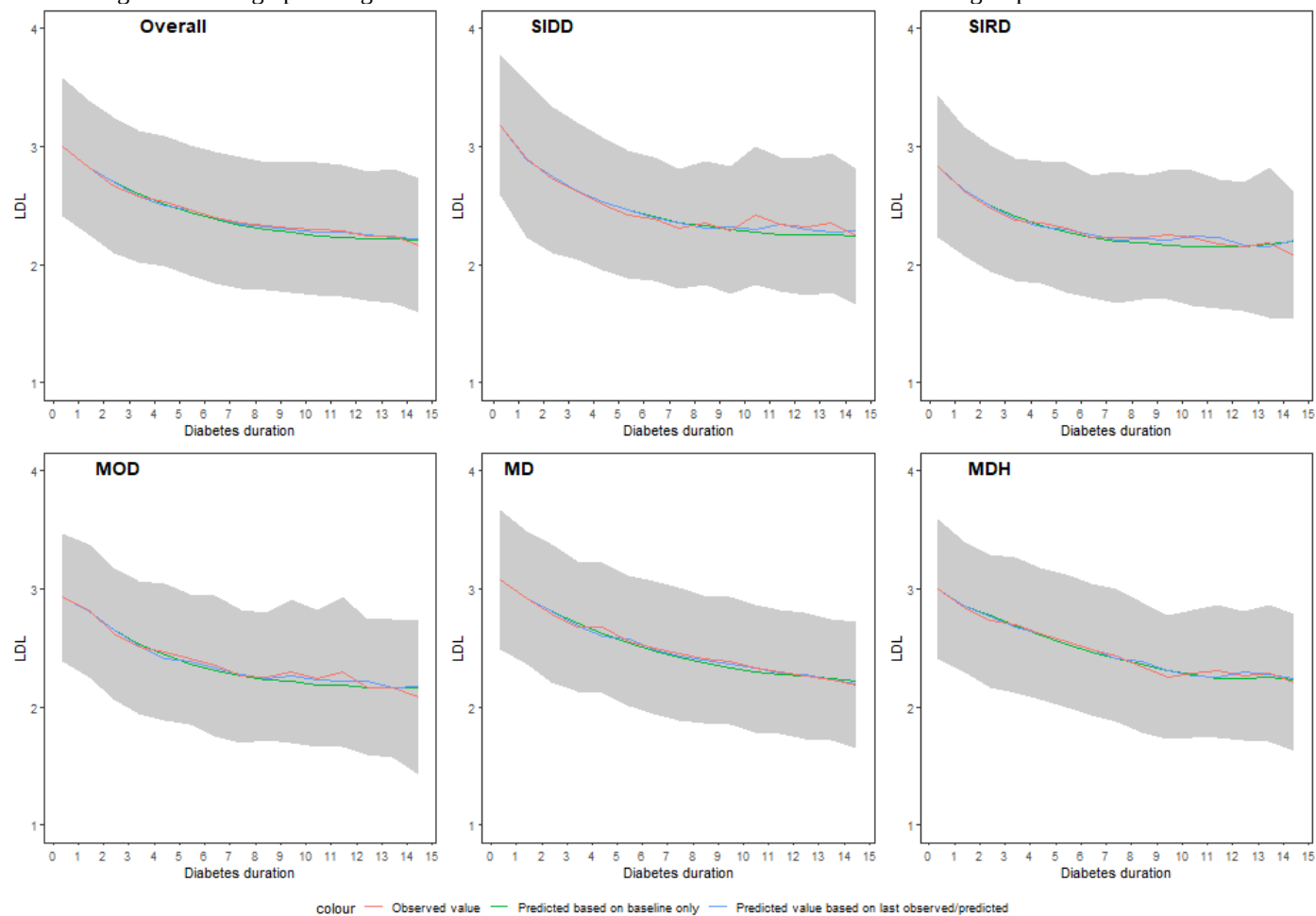
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.4 The graphs for goodness-of-fit over time for LDL of each RHAPSODY subgroup



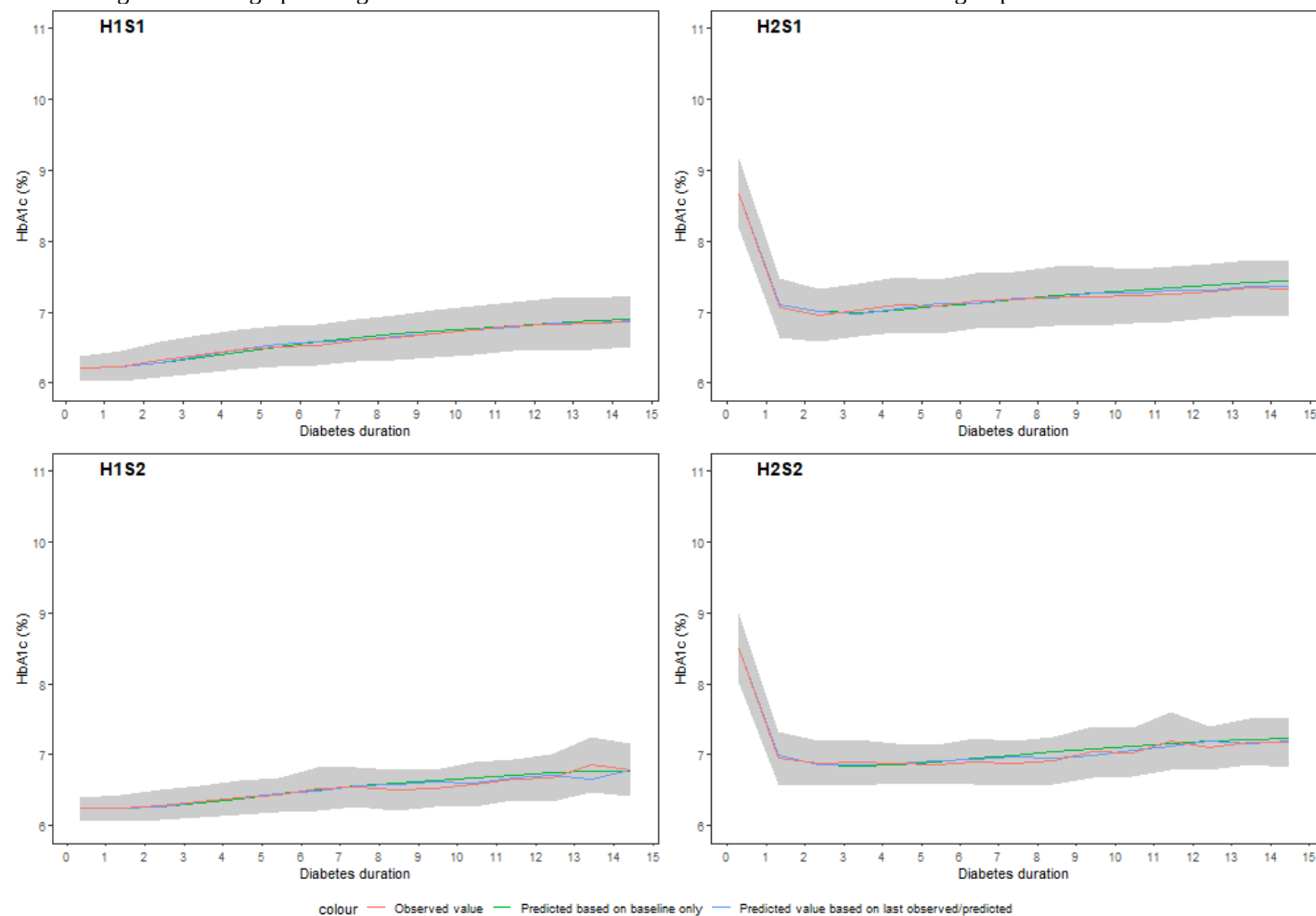
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.5 The graphs for goodness-of-fit over time for HbA1c of each risk-driven subgroup



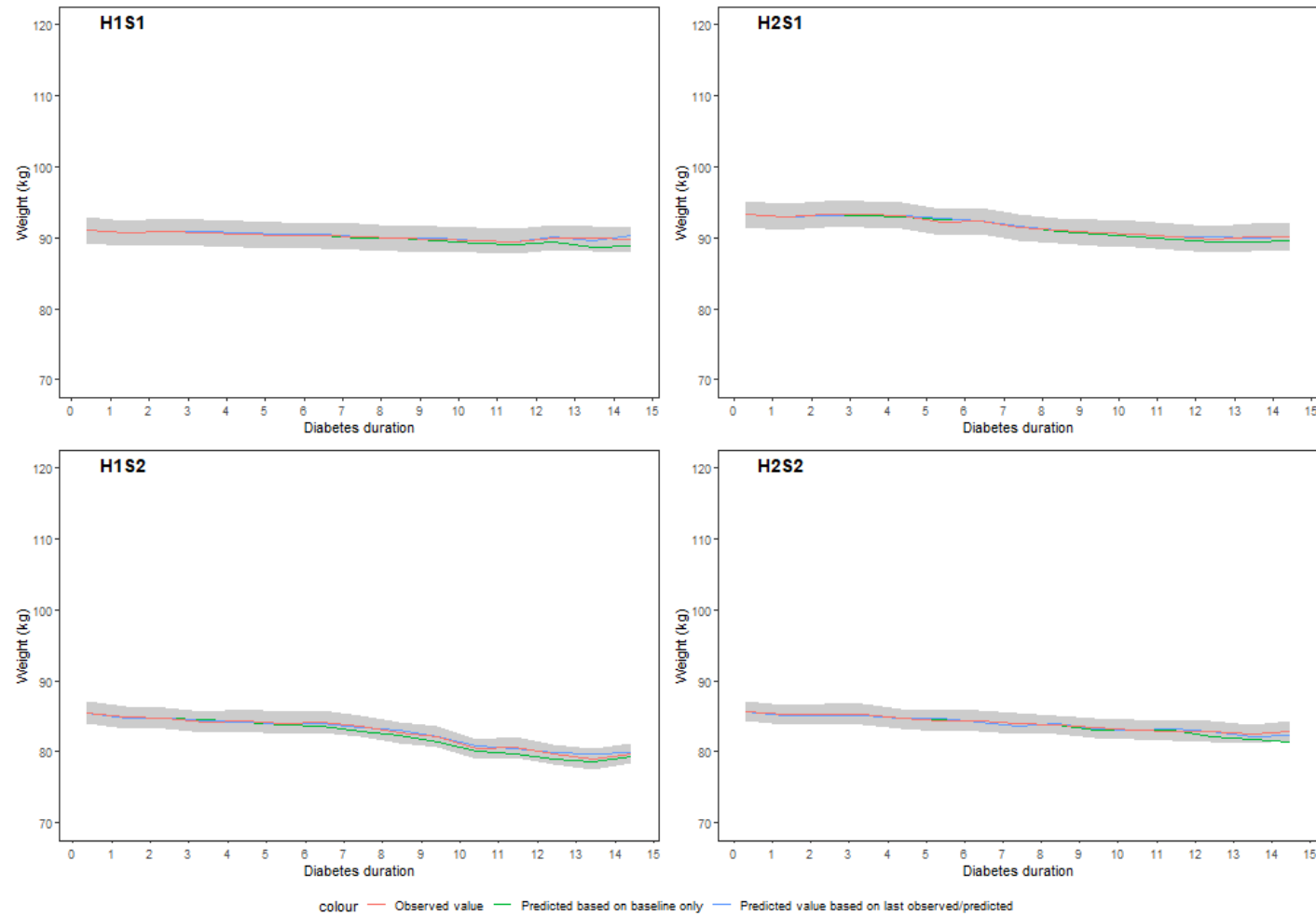
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.6 The graphs for goodness-of-fit over time for weight of each risk-driven subgroup



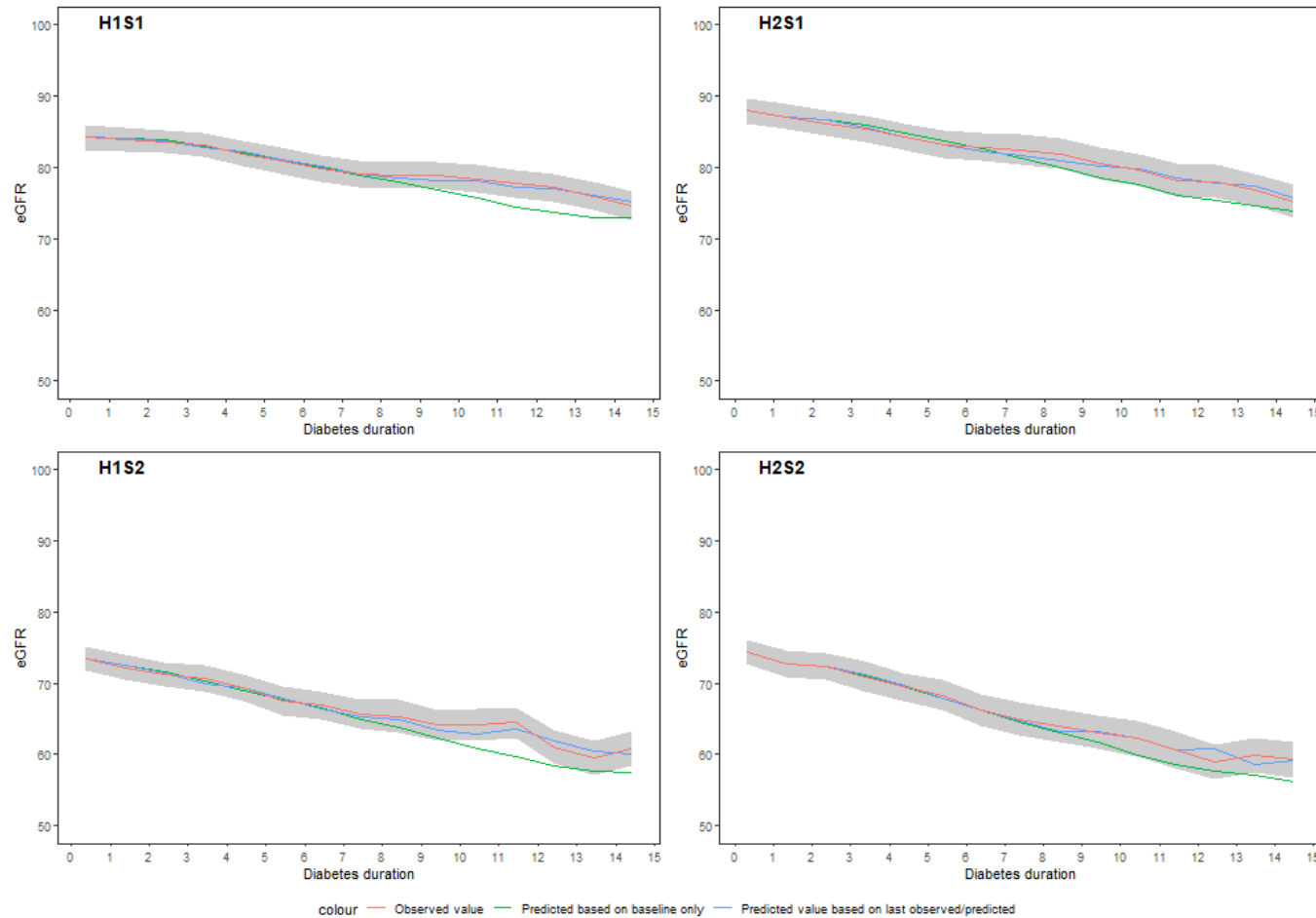
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.7 The graphs for goodness-of-fit over time for eGFR of each risk-driven subgroup



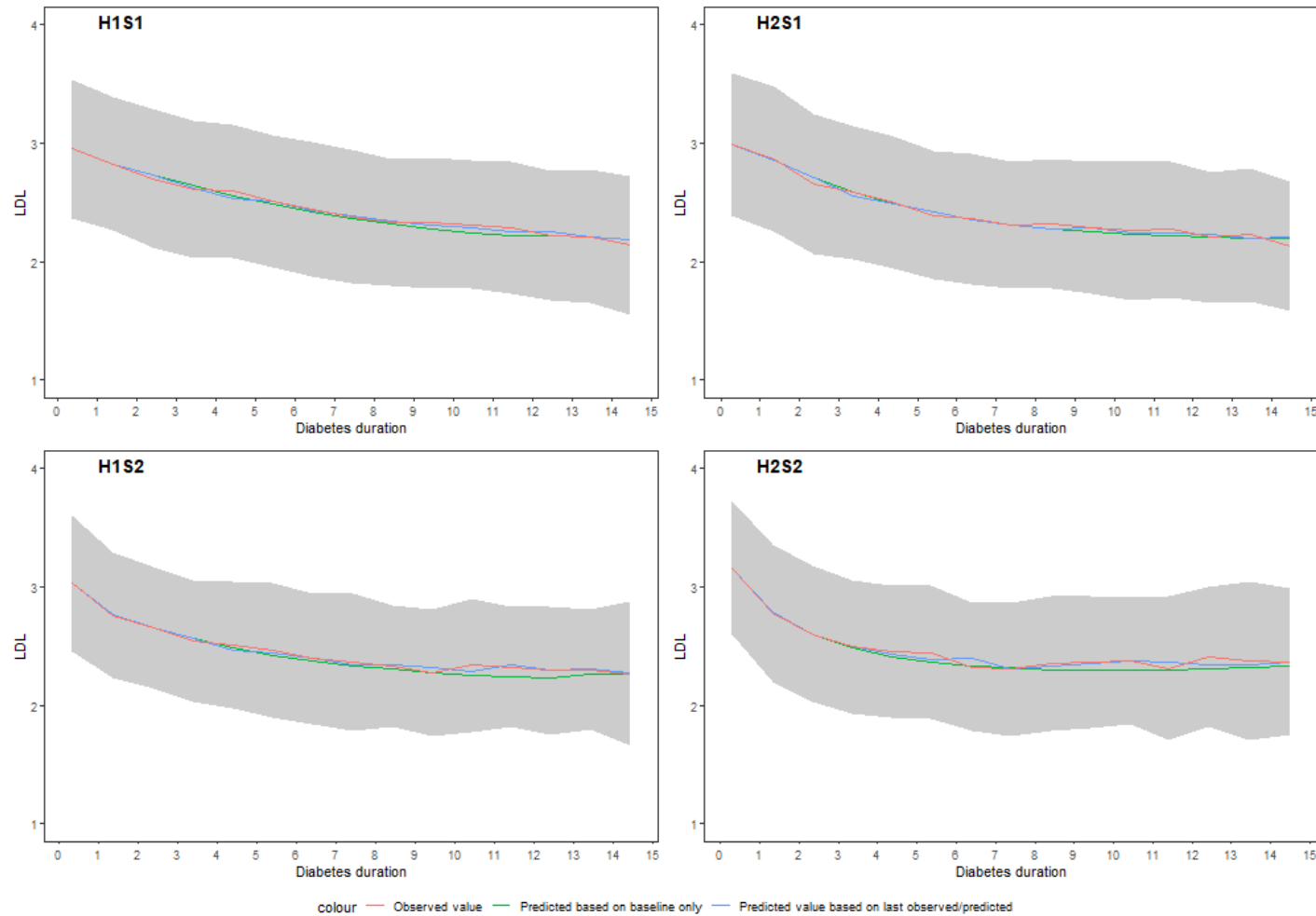
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 4.8 The graphs for goodness-of-fit over time for LDL of each risk-driven subgroup



Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Appendix 5 The method of calculating annual maximum cost-effective price

First, note that the definition of cost-effectiveness thresholds implies that

$$\frac{\Delta \text{Complication cost} + \text{Maximum discounted intervention price}}{\Delta \text{QALY}} = \text{Cost} - \text{effective threshold} \quad (1)$$

Also note that the maximum discounted intervention price can be calculated as:

$$\text{Maximum discounted intervention price} = \sum_{k=1}^i \frac{\text{Average year } k \text{ survival rate} * \text{Maximum annual price}}{(1 + \text{discounting factor})^{k-1}} \quad (2)$$

Now, combine equation (1) and (2) to obtain (3) for an i year hypothetical treatment:

$$\text{Maximum annual price} = \frac{\Delta \text{QALY} * \text{Cost} - \text{effective threshold} - \Delta \text{Complication cost}}{\sum_{k=1}^i \frac{\text{Average year } k \text{ survival rate}}{(1 + \text{discounting factor})^{k-1}}} \quad (3)$$

where

$$\text{Average year } k \text{ survival rate} = \text{mean}((1 - \sum_{j=1}^k \text{death rate}_{\text{year } j}) + 0.5 * \text{death rate}_{\text{year } k})$$

Note:

Δ denotes increments between hypothetical treatment scenario and care-as-usual scenario.

i denotes the duration of hypothetical treatment scenario (in years), which is five-year in the base case.

QALYs, quality-adjusted life years.

Discounting factor is either 3% in U.S. setting or 3.5% in U.K. setting.

Mean() is averaging across all patients.

Appendix 6 Baseline characteristics

Table 6.1 Baseline simulation characteristics of data-driven subgroups and risk-driven subgroups

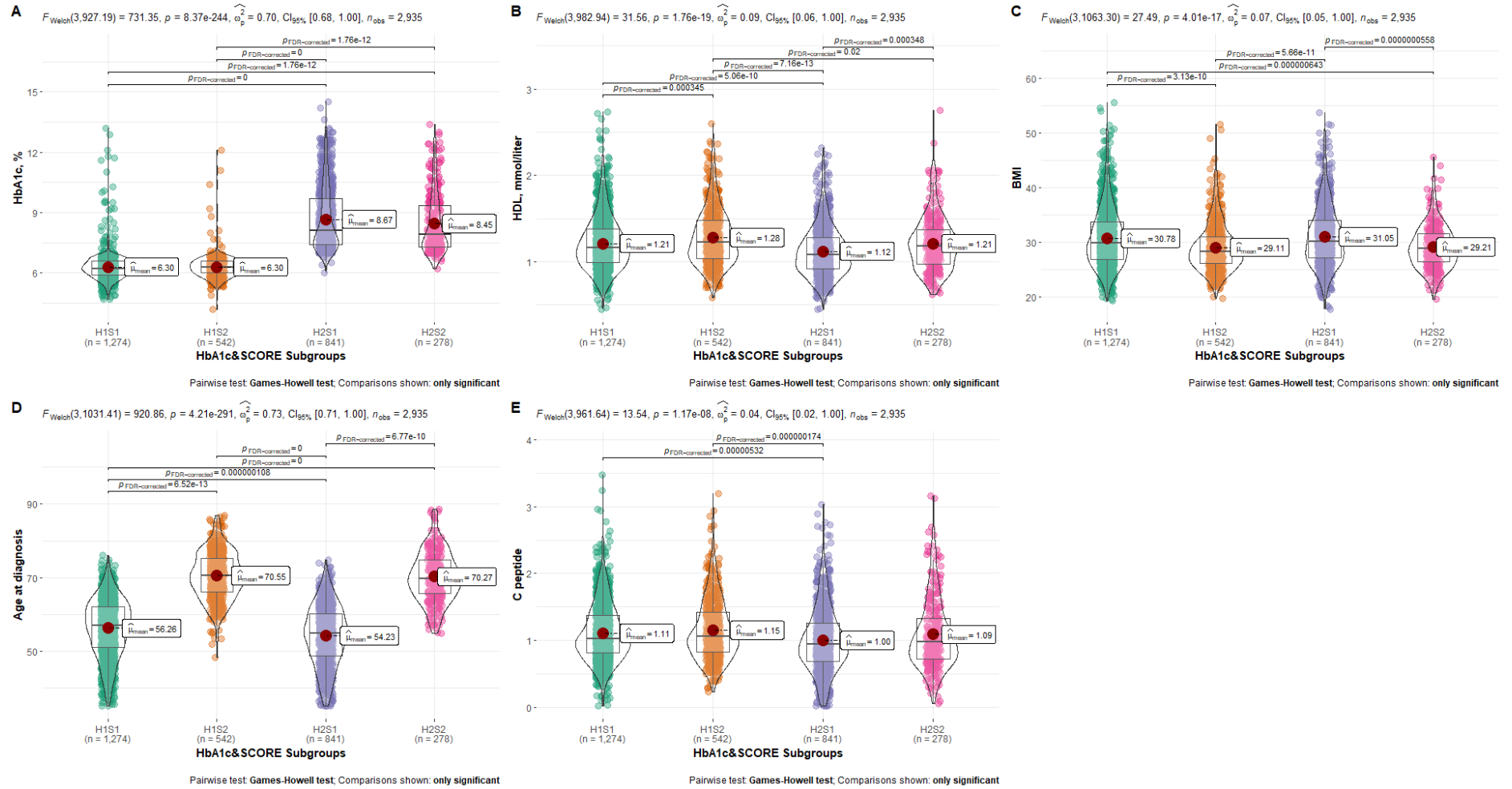
	RHPSODY Data-driven subgroups					Risk-driven subgroups					
	RHAP-SIDD	RHAP-SIRD	RHAP-MOD	RHAP-MD	RHAP-MDH	p	H1S1	H1S2	H2S1	H2S2	p
n	365	637	520	860	553		1274	542	841	278	
Age, years	61.39 (9.74)	70.74 (7.41)	55.90 (8.05)	57.56 (8.20)	68.79 (7.76)	<0.001	59.08 (8.15)	72.75 (6.54)	58.18 (8.56)	73.58 (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)	<0.001	2.82 (3.22)	2.19 (2.69)	3.94 (3.38)	3.32 (3.21)	<0.001
Age at diabetes diagnosis, years	56.93 (9.28)	68.45 (7.22)	52.85 (7.78)	54.31 (7.74)	66.00 (7.83)	<0.001	56.26 (7.84)	70.55 (6.39)	54.23 (8.14)	70.27 (6.83)	<0.001
LDL cholesterol, mmol/liter	2.72 (0.91)	2.64 (0.89)	2.68 (0.90)	2.79 (0.94)	2.80 (0.91)	0.005	2.76 (0.93)	2.82 (0.90)	2.61 (0.89)	2.80 (0.92)	<0.001
Total cholesterol, mmol/liter	4.65 (1.05)	4.59 (1.04)	4.68 (1.03)	4.72 (1.11)	4.96 (1.06)	<0.001	4.75 (1.08)	4.85 (1.06)	4.57 (1.04)	4.83 (1.09)	<0.001
HDL cholesterol, mmol/liter	1.16 (0.31)	1.08 (0.22)	1.06 (0.26)	1.10 (0.23)	1.56 (0.34)	<0.001	1.19 (0.32)	1.25 (0.33)	1.12 (0.31)	1.20 (0.33)	<0.001
Total cholesterol/HDL ratio	4.24 (1.28)	4.38 (1.23)	4.61 (1.34)	4.43 (1.23)	3.27 (0.82)	<0.001	4.18 (1.24)	4.04 (1.17)	4.34 (1.38)	4.24 (1.30)	<0.001
HbA1c, %	7.78 (1.78)	6.53 (0.73)	6.83 (1.00)	6.68 (0.85)	6.44 (0.72)	<0.001	6.35 (0.63)	6.34 (0.55)	7.48 (1.32)	7.36 (1.22)	<0.001
HbA1c, mmol/mol	61.51 (19.34)	47.87 (7.99)	51.17 (10.93)	49.47 (9.33)	46.90 (7.85)	<0.001	45.87 (6.86)	45.76 (6.04)	58.21 (14.44)	56.92 (13.26)	<0.001
Fasting blood glucose, mmol/liter	8.82 (2.83)	7.69 (1.35)	8.02 (1.85)	7.85 (1.77)	7.42 (1.23)	<0.001	7.45 (1.37)	7.41 (1.13)	8.67 (2.39)	8.39 (1.91)	<0.001
Triglycerides, mmol/liter	1.77 (1.00)	1.94 (0.99)	2.15 (1.29)	1.89 (1.05)	1.32 (0.59)	<0.001	1.80 (0.99)	1.72 (0.87)	1.92 (1.21)	1.85 (1.02)	0.003
Urinary microalbumin (mg/liter)	23.22 (67.38)	28.50 (150.02)	30.59 (130.71)	15.36 (62.49)	14.48 (75.73)	0.017	16.55 (74.85)	35.61 (192.91)	19.84 (59.64)	24.10 (62.52)	0.005
Urinary creatinine (umol/liter)	9.18 (5.15)	8.88 (7.21)	10.74 (5.83)	10.24 (5.80)	8.12 (5.37)	<0.001	9.98 (6.24)	8.41 (6.36)	10.01 (5.95)	7.79 (3.95)	<0.001
Albumin creatinine ratio (mg/mmol)	3.18 (10.43)	3.93 (19.70)	3.11 (14.25)	1.65 (7.19)	1.53 (5.78)	0.002	1.76 (7.87)	4.28 (22.33)	2.25 (8.17)	3.94 (13.36)	<0.001
Blood creatinine (umol/liter)	78.65 (21.97)	87.95 (23.76)	75.88 (17.38)	76.15 (16.12)	79.36 (17.52)	<0.001	77.41 (17.37)	85.33 (20.78)	76.91 (19.55)	86.48 (24.40)	<0.001

eGFR based on CKD-EPI, 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)	<0.001	84.59 (16.25)	72.80 (15.50)	87.21 (17.37)	71.67 (16.13)	<0.001
Length, cm	172.72 (9.68)	171.61 (8.89)	171.97 (9.71)	172.70 (10.17)	170.13 (9.60)	<0.001	171.72 (9.68)	170.84 (9.20)	173.03 (10.10)	170.88 (9.10)	<0.001
Weight, kg	88.11 (15.93)	88.06 (11.94)	111.79 (16.97)	86.48 (13.85)	78.67 (12.52)	<0.001	91.18 (18.39)	85.00 (14.72)	93.12 (18.52)	85.35 (13.55)	<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)	<0.001	30.87 (5.54)	29.10 (4.55)	31.06 (5.47)	29.24 (4.10)	<0.001
Systolic blood pressure, mmHg	142.17 (19.84)	146.75 (19.99)	141.69 (17.87)	137.56 (17.56)	145.60 (18.57)	<0.001	137.23 (16.16)	155.01 (19.70)	138.36 (17.36)	153.48 (18.33)	<0.001
Diastolic blood pressure, mmHg	78.18 (9.43)	77.72 (9.85)	79.85 (9.30)	77.91 (9.34)	77.73 (8.73)	0.001	78.13 (8.97)	78.85 (10.47)	77.93 (9.06)	78.23 (9.86)	0.342
Male, n (%)	218 (59.7)	397 (62.3)	256 (49.2)	474 (55.1)	297 (53.7)	<0.001	629 (49.4)	347 (64.0)	490 (58.3)	176 (63.3)	<0.001
Smoking status						<0.001					0.017
Current, n (%)	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, n (%)	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)	
Never, n (%)	115 (31.9)	168 (27.2)	154 (30.0)	255 (30.9)	190 (35.3)		396 (32.3)	151 (28.8)	257 (30.9)	78 (28.8)	

Frequencies (percentage) for categorical variables, or mean (standard deviation) for continuous variables. Chi-square tests were applied to check for significant differences between groups.

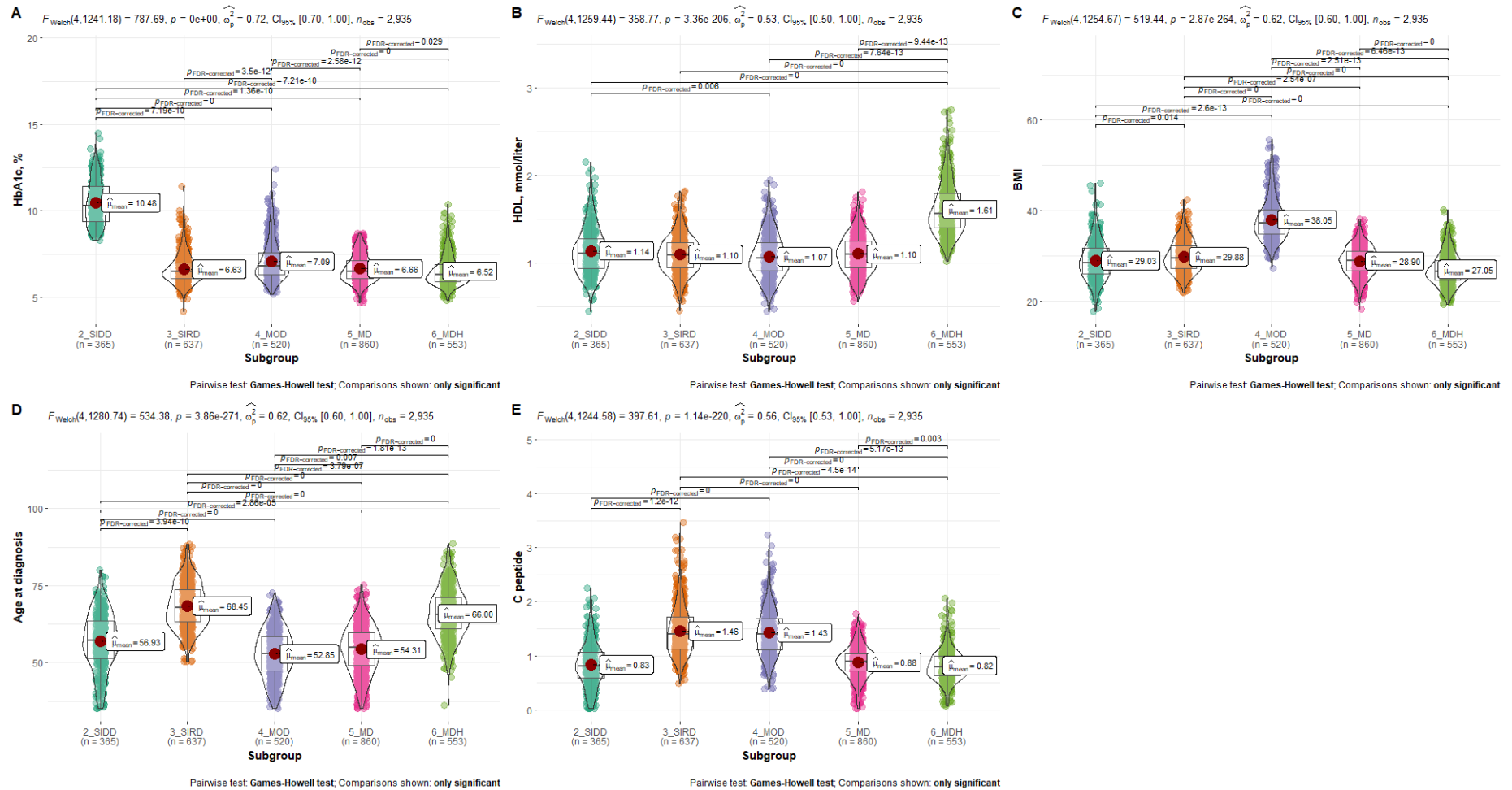
Simulation baseline for each individual was set as the landmark of the date of participation in the biobank sample around 2008 to avoid survival bias.

Figure 6.1 Violin graphs and pairwise comparisons of baseline clustering indicators of risk-driven subgroups



The Games-Howell test was applied to enable pairwise comparison and p values were adjusted by using Benjamini-Hochberg correction to account for false discovery rate. P-values below 0.05 were considered statistically significant. These are the baseline characteristics at diagnosis, rather than the baseline simulation characteristics.

Figure 6.2 Violin graphs and pairwise comparisons of baseline clustering indicators of RHAPSODY data-driven subgroups

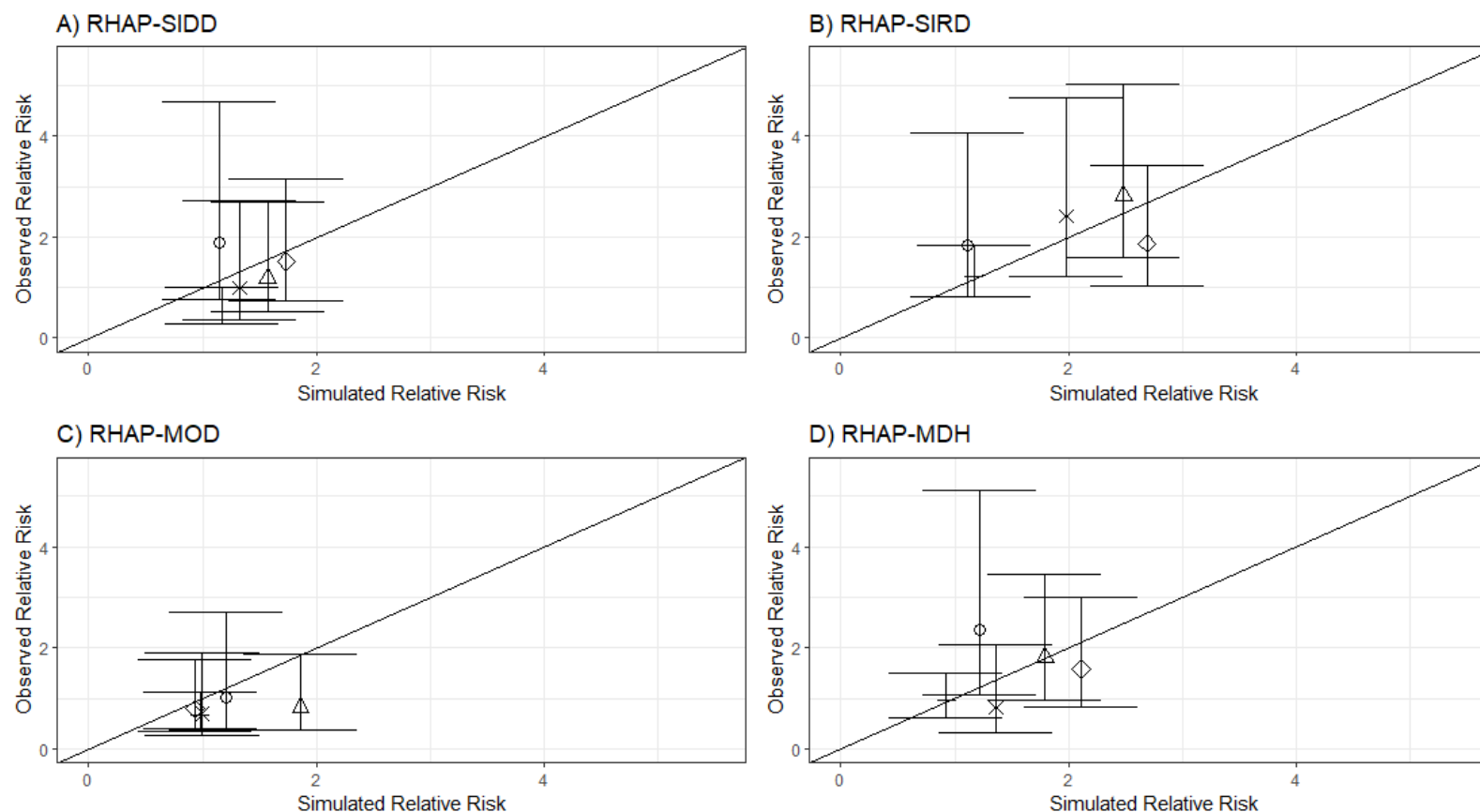


The Games-Howell test was applied to enable pairwise comparison and p values were adjusted by using Benjamini-Hochberg correction to account for false discovery rate. P-values below 0.05 were considered statistically significant. These are the baseline characteristics at diagnosis, rather than the baseline simulation characteristics.

Figure 7.1 Model validation of subgroup-wise differences of relative risks in 5 years of simulation for RHAPSODY data-driven subgroups

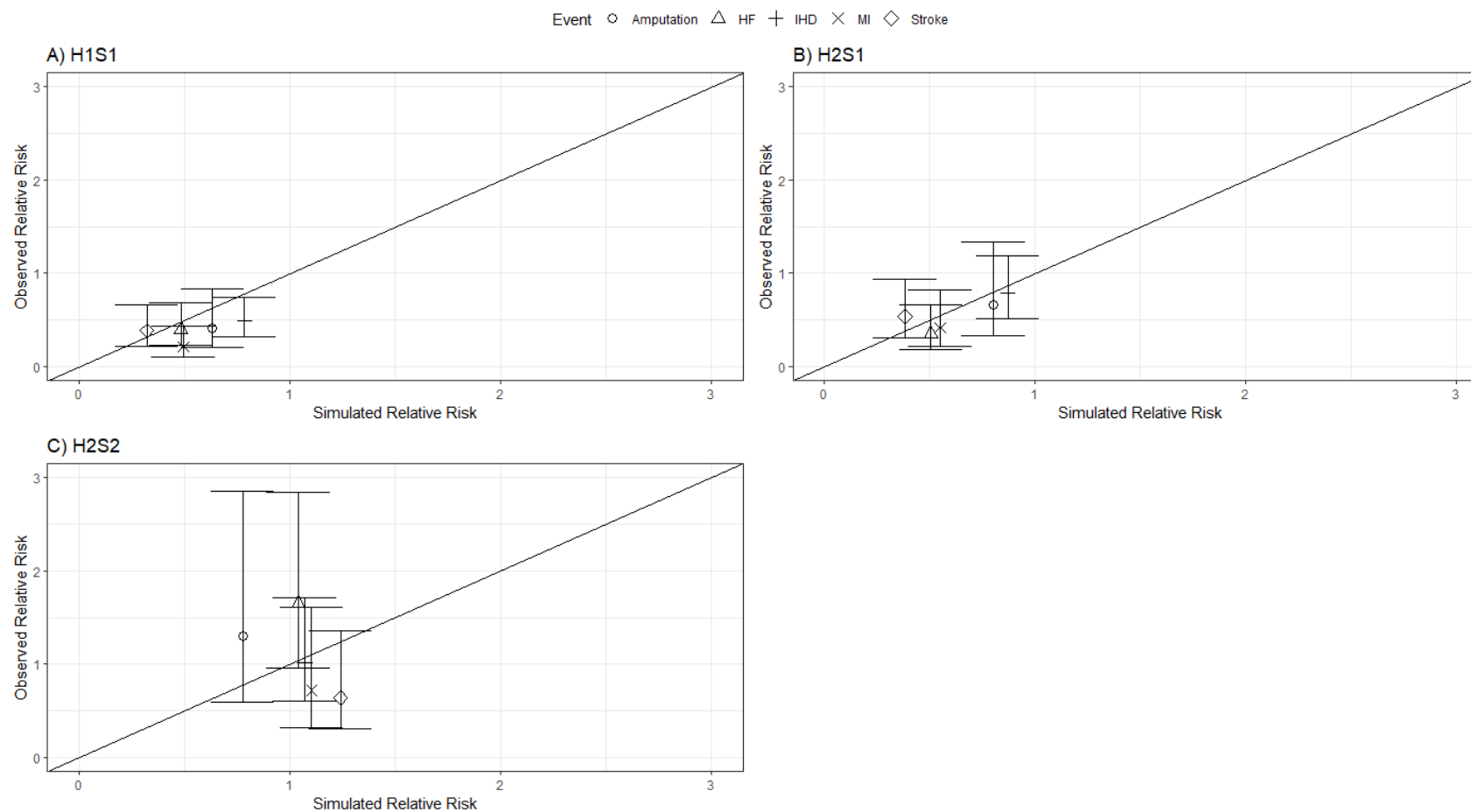
The comparison of relative risks in the 5 years of simulation (Reference group is RHAP-MD)

Event ○ Amputation △ HF + IHD × MI ◇ Stroke



The 45-degree line indicates perfect calibration (observations equals to simulations). The closer the point is to the 45-degree line, the better the model's prediction. The error bar indicates the 95% confidence interval of observed relative risks, and if simulated relative risks fit within the interval (i.e., the error bar crossed with the 45-degree line), we considered it as a valid prediction.

Figure 7.2 Model validation of subgroup-wise differences of relative risks in 5 years of simulation for risk-driven subgroups
The comparison of relative risks in the 5 years of simulation (Reference group is H1S2)



The 45-degree line indicates perfect calibration (observations equals to simulations). The closer the point is to the 45-degree line, the better the model's prediction. The error bar indicates the 95% confidence interval of observed relative risks, and if simulated relative risks fit within the interval (i.e., the error bar crossed with the 45-degree line), we considered it as a valid prediction.

Table 8.1 Absolute and standardized simulated outcome (mean and 95% CI) by RHAPSODY data-driven subgroups and risk-driven subgroups in the U.S. setting

	Lifetime QALYs	Standardized QALYs - Male	Lifetime QALYs - Female	Standardized QALYs - Female	Lifetime Complication (100,000\$)	cost	Standardized complication (100,000\$) - Male	cost	Standardized complication (100,000\$) - Female	cost
Overall*	10.57 (10.46-10.91)	9.98 (9.65-10.52)	11.12 (10.8-11.65)	11.12 (10.8-11.65)	1.65 (1.63-1.7)		1.54 (1.49-1.63)		1.76 (1.71-1.84)	
	RHAPSODY data-driven clustering subgroups									
RHAP-MOD	12.61 (12.49-12.92)	10.02 (9.67-10.56)	11 (10.68-11.51)	11 (10.68-11.51)	1.94 (1.92-1.98)		1.52 (1.47-1.61)		1.74 (1.69-1.82)	
RHAP-MD	12.26 (12.15-12.57)	10.32 (10.01-10.86)	11.15 (10.85-11.67)	11.15 (10.85-11.67)	1.9 (1.88-1.94)		1.58 (1.53-1.66)		1.76 (1.71-1.84)	
RHAP-SIDD	10.81 (10.7-11.11)	10.18 (9.88-10.67)	10.88 (10.55-11.42)	10.88 (10.55-11.42)	1.71 (1.7-1.76)		1.62 (1.58-1.7)		1.72 (1.68-1.81)	
RHAP-MDH	8.99 (8.83-9.39)	10.65 (10.34-11.2)	11.43 (11.13-11.97)	11.43 (11.13-11.97)	1.42 (1.4-1.48)		1.64 (1.6-1.72)		1.8 (1.76-1.88)	
RHAP-SIRD	7.9 (7.74-8.25)	10.13 (9.8-10.7)	10.92 (10.6-11.49)	10.92 (10.6-11.49)	1.25 (1.23-1.31)		1.56 (1.51-1.65)		1.73 (1.68-1.82)	
Range†	4.71	0.63	0.55	0.55	0.69		0.12		0.08	
	HbA1c and SCORE risk-driven subgroups									
H1S1	11.96 (11.85-12.3)	9.55 (9.23-10.11)	11.31 (11.01-11.83)	11.31 (11.01-11.83)	1.84 (1.82-1.89)		1.45 (1.4-1.53)		1.78 (1.74-1.86)	
H2S1	11.84 (11.74-12.15)	9.29 (8.97-9.83)	10.95 (10.63-11.49)	10.95 (10.63-11.49)	1.84 (1.82-1.88)		1.43 (1.38-1.51)		1.72 (1.68-1.81)	
H1S2	7.23 (7.06-7.62)	9.03 (8.69-9.62)	10.49 (10.16-11.09)	10.49 (10.16-11.09)	1.18 (1.15-1.24)		1.44 (1.39-1.53)		1.7 (1.66-1.79)	
H2S2	6.83 (6.68-7.19)	8.73 (8.4-9.28)	10.22 (9.88-10.78)	10.22 (9.88-10.78)	1.14 (1.11-1.2)		1.4 (1.36-1.49)		1.62 (1.58-1.7)	
Range†	5.13	0.82	1.09	1.09	0.7		0.05		0.16	

CAU, Care as usual

* Overall was defined as the average result, indicating the result of calculating group-wise simulation outcome altogether as a homogenous type 2 diabetes group. (Average values of results based on three extrapolations of linear dynamic models estimated from Ahlqvist's and RHAPSODY data-driven subgroups and risk-driven subgroups were applied.)

†Range was defined as the maximum – minimum of the mean maximum annual cost-effective price of intervention or incremental QALYs. To compare the value of RHAPSODY data-driven subgroups versus risk-driven subgroups, the range between subgroups was considered, with a wider range indicating a greater ability to differentiate between subgroups.

Table 8.2 Absolute and standardized simulated outcome (mean and 95% CI) by RHAPSODY data-driven subgroups and risk-driven subgroups in the U.K. setting

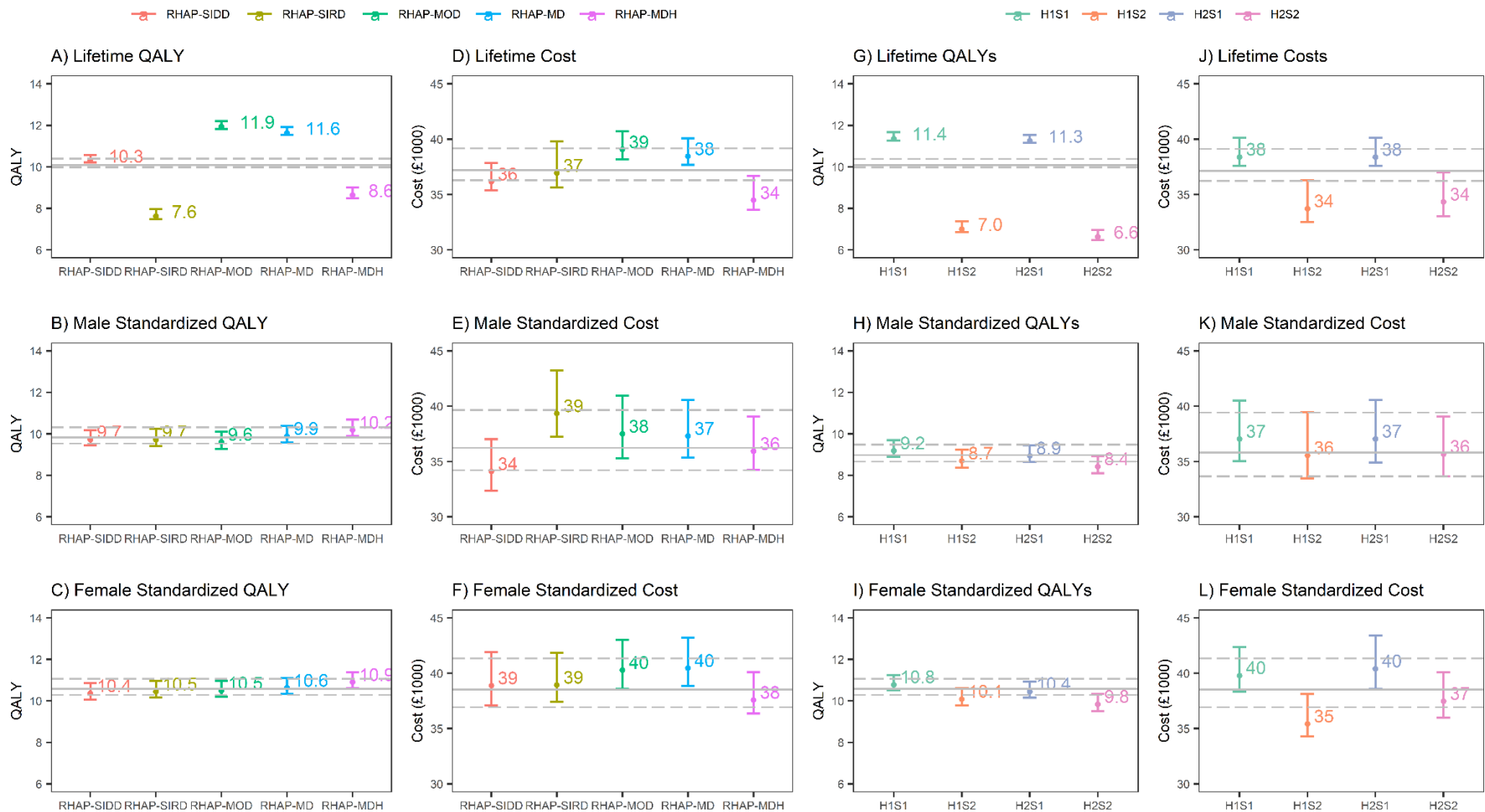
	Lifetime QALYs	Standardized QALYs - Male	Lifetime QALYs - Female	Standardized QALYs - Female	Lifetime Complication (100,000£)	cost	Standardized complication (100,000£) - Male	cost	Standardized complication (100,000£) - Female	cost
Overall*	10.09 (9.98-10.39)	9.54 (9.24-10.04)	10.58 (10.29-11.06)	0.37 (0.36-0.39)	0.36 (0.34-0.4)	0.36 (0.34-0.4)	0.39 (0.37-0.41)			
	RHAPSODY data-driven clustering subgroups									
RHAP-MOD	11.94 (11.84-12.22)	9.61 (9.29-10.1)	10.5 (10.2-10.95)	0.39 (0.38-0.41)	0.38 (0.35-0.41)	0.38 (0.35-0.41)	0.4 (0.39-0.43)			
RHAP-MD	11.64 (11.53-11.91)	9.89 (9.6-10.38)	10.63 (10.35-11.1)	0.38 (0.38-0.4)	0.37 (0.35-0.41)	0.37 (0.35-0.41)	0.4 (0.39-0.43)			
RHAP-SIDD	10.3 (10.21-10.57)	9.73 (9.45-10.18)	10.37 (10.06-10.85)	0.36 (0.35-0.38)	0.34 (0.32-0.37)	0.34 (0.32-0.37)	0.39 (0.37-0.42)			
RHAP-MDH	8.64 (8.5-9.01)	10.19 (9.9-10.68)	10.9 (10.63-11.39)	0.34 (0.34-0.37)	0.36 (0.34-0.39)	0.36 (0.34-0.39)	0.38 (0.36-0.4)			
RHAP-SIRD	7.62 (7.48-7.96)	9.72 (9.42-10.24)	10.45 (10.16-10.97)	0.37 (0.36-0.4)	0.39 (0.37-0.43)	0.39 (0.37-0.43)	0.39 (0.37-0.42)			
Range†	4.32	0.58	0.53	0.05	0.05	0.05	0.02			
	HbA1c and SCORE risk-driven subgroups									
H1S1	11.36 (11.26-11.66)	9.19 (8.89-9.71)	10.78 (10.51-11.25)	0.38 (0.38-0.4)	0.37 (0.35-0.41)	0.37 (0.35-0.41)	0.4 (0.38-0.42)			
H2S1	11.25 (11.16-11.53)	8.94 (8.64-9.44)	10.45 (10.15-10.93)	0.38 (0.38-0.4)	0.37 (0.35-0.41)	0.37 (0.35-0.41)	0.4 (0.39-0.43)			
H1S2	7.01 (6.85-7.37)	8.7 (8.38-9.25)	10.08 (9.78-10.62)	0.34 (0.33-0.36)	0.36 (0.34-0.39)	0.36 (0.34-0.39)	0.35 (0.34-0.38)			
H2S2	6.63 (6.48-6.97)	8.42 (8.11-8.93)	9.84 (9.52-10.35)	0.34 (0.33-0.37)	0.36 (0.34-0.39)	0.36 (0.34-0.39)	0.37 (0.36-0.4)			
Range†	4.73	0.77	0.94	0.04	0.01	0.01	0.05			

CAU, Care as usual

* Overall was defined as the average result, indicating the result of calculating group-wise simulation outcome altogether as a homogenous type 2 diabetes group. (Average values of results based on three extrapolations of linear dynamic models estimated from Ahlqvist's and RHAPSODY data-driven subgroups and risk-driven subgroups were applied.)

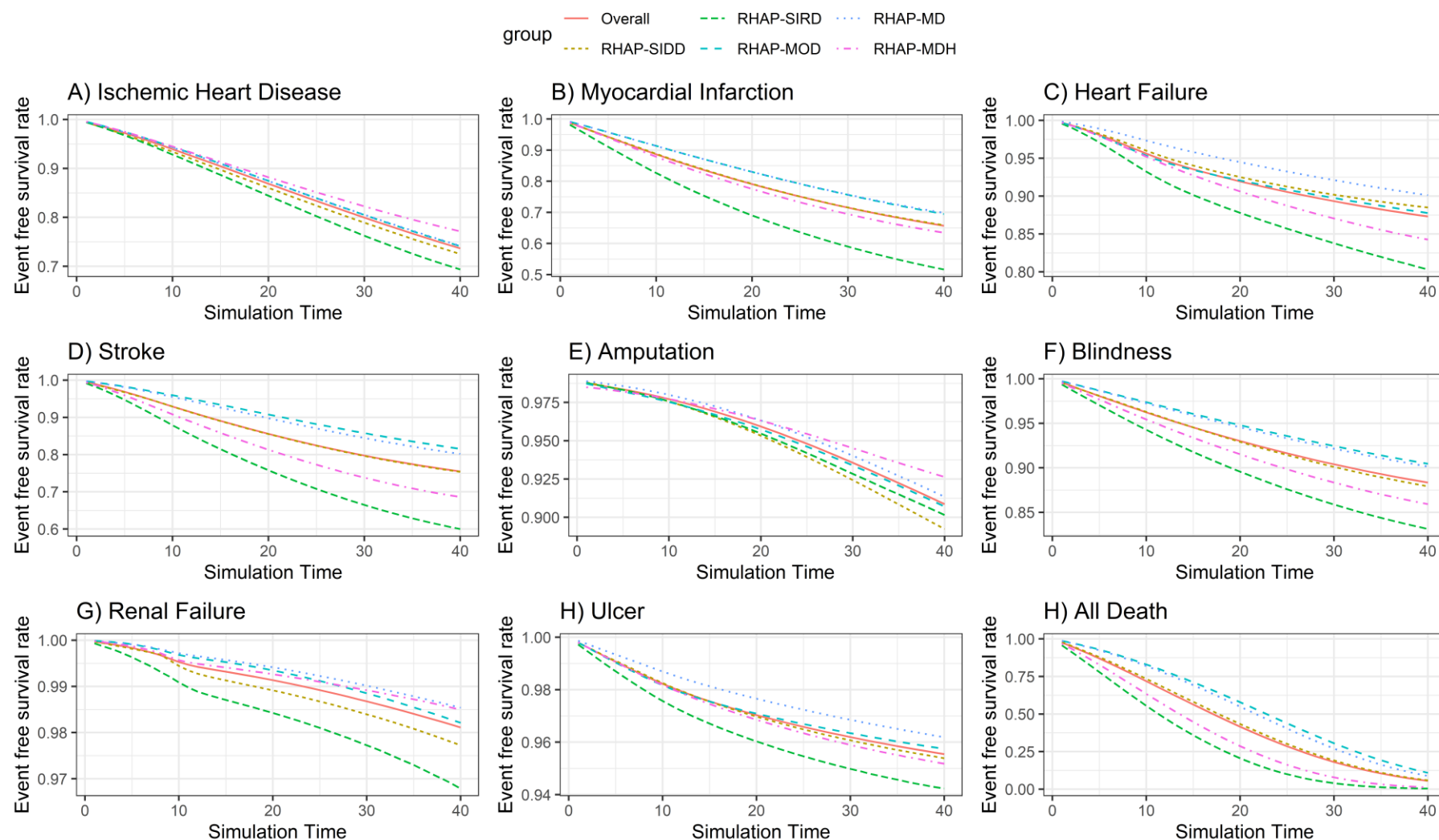
†Range was defined as the maximum – minimum of the mean maximum annual cost-effective price of intervention or incremental QALYs. To compare the value of RHAPSODY data-driven subgroups versus risk-driven subgroups, the range between subgroups was considered, with a wider range indicating a greater ability to differentiate between subgroups.

Figure 8.1 Non-standardized and standardized mean simulated lifetime QALYs and cost with 95% CI (in U.K. setting) for RHAPSODY data-driven and risk-driven subgroups



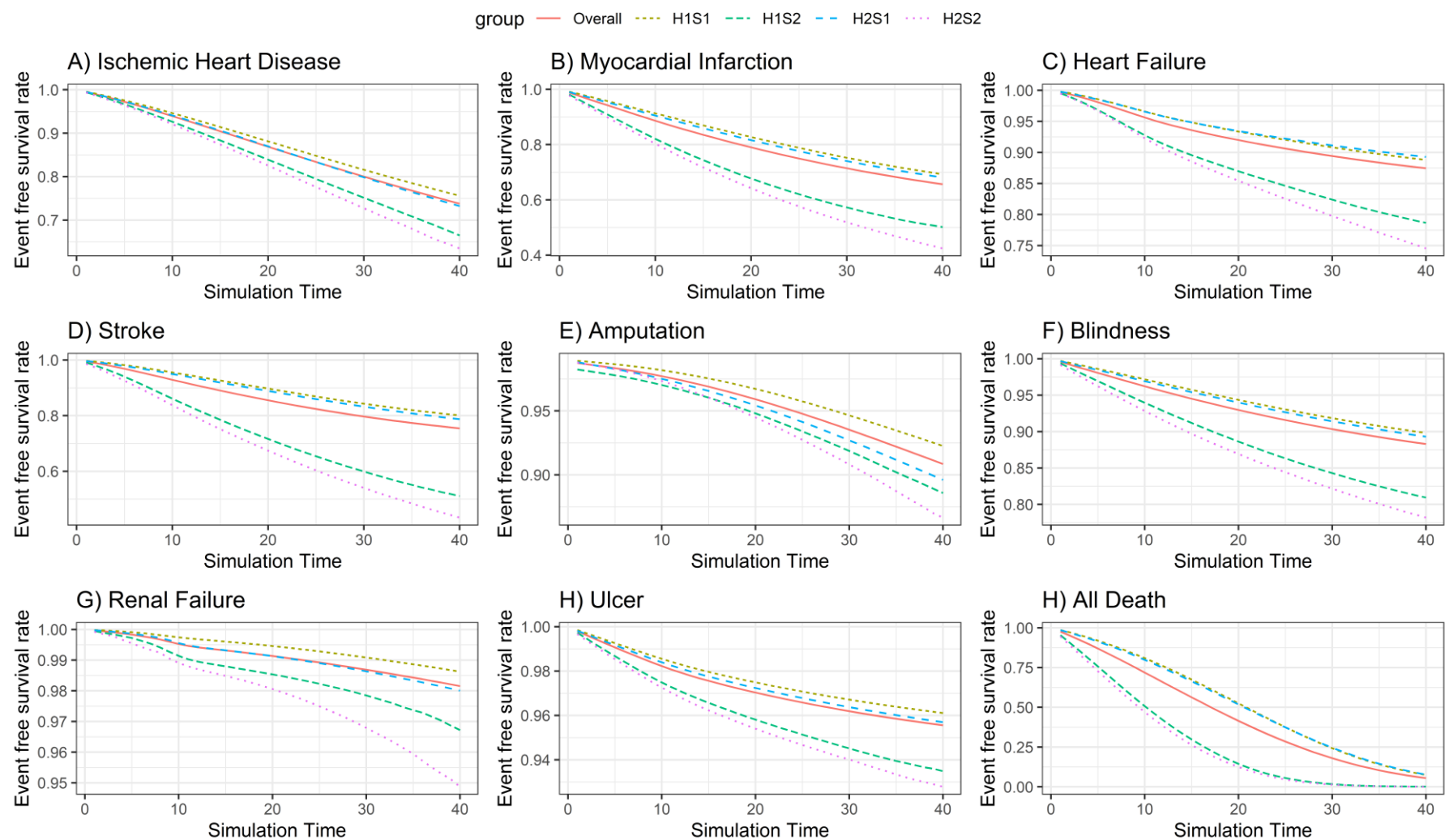
QALYs, Quality-Adjusted Life-Years

Figure 8.2 Simulated Kaplan-Meier plots of RHAPSODY data-driven subgroups



The UKPDS-OM2 generated Kaplan-Meier survival plots for each event, stratified by year of simulation and subgroup. A higher value on the plot indicates a lower risk of experiencing the event for an individual within that subgroup. Blindness refers to blindness in one eye.

Figure 8.3 Simulated Kaplan-Meier plots of risk-driven subgroups



The UKPDS-OM2 generated Kaplan-Meier survival plots for each event, stratified by year of simulation and subgroup. A higher value on the plot indicates a lower risk of experiencing the event for an individual within that subgroup. Blindness refers to blindness in one eye.

Appendix 9 Model outcomes under hypothetical treatment strategies

Table 9.1 Outcomes of five-year 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in base case U.S. setting

		5%-Improvement-based hypothetical intensive management			
		HbA1c		HbA1c+LDL+BMI	
		Max annual price of intervention (\$)	Δ QALY vs CAU	Max annual price of intervention (\$)	Δ QALY vs CAU
RHAPSODY data-driven subgroups	Overall*	237 (150-302)	0.011 (0.008-0.015)	582 (436-673)	0.028 (0.022-0.034)
	RHAP-MOD	208 (144-274)	0.011 (0.008-0.014)	589 (455-746)	0.032 (0.025-0.041)
	RHAP-MD	154 (94-215)	0.008 (0.005-0.011)	415 (335-493)	0.023 (0.019-0.026)
	RHAP-SIDD	285 (192-378)	0.015 (0.01-0.019)	637 (519-746)	0.033 (0.027-0.039)
	RHAP-MDH	205 (121-286)	0.01 (0.006-0.014)	407 (278-555)	0.019 (0.013-0.027)
	RHAP-SIRD	269 (179-358)	0.012 (0.008-0.016)	730 (578-888)	0.033 (0.026-0.04)
	Range†	131	0.006	323	0.014
Risk-driven subgroups	H1S1	165 (107-222)	0.009 (0.006-0.011)	456 (365-544)	0.024 (0.02-0.029)
	H2S1	278 (208-348)	0.014 (0.011-0.018)	558 (446-666)	0.03 (0.024-0.035)
	H1S2	287 (186-391)	0.012 (0.008-0.017)	680 (528-834)	0.03 (0.023-0.037)
	H2S2	339 (223-470)	0.015 (0.01-0.02)	857 (697-1037)	0.038 (0.031-0.046)
	Range†	174	0.006	401	0.014

*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.

Table 9.2 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in Scenario 1 U.S. setting

		Treat-to-target hypothetical intensive management				5%-Improvement-based hypothetical intensive management			
		HbA1c		HbA1c+LDL+BMI		HbA1c		HbA1c+LDL+BMI	
		Max annual cost of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU
	Overall*	335 (214-417)	0.016 (0.011-0.02)	2742 (2066-3204)	0.131 (0.105-0.161)	405 (267-503)	0.019 (0.013-0.024)	1348 (1044-1502)	0.063 (0.052-0.074)
Data-driven subgroups	RHAP-MOD	347 (237-458)	0.018 (0.013-0.024)	3611 (2700-4719)	0.199 (0.149-0.256)	330 (231-441)	0.017 (0.012-0.023)	1497 (1253-1775)	0.078 (0.066-0.093)
	RHAP-MD	262 (176-356)	0.013 (0.009-0.018)	1818 (1507-2136)	0.097 (0.08-0.113)	301 (205-395)	0.016 (0.011-0.02)	1111 (937-1291)	0.058 (0.049-0.067)
	RHAP-SIDD	717 (525-923)	0.036 (0.027-0.046)	2793 (2226-3330)	0.145 (0.12-0.171)	585 (415-755)	0.029 (0.021-0.037)	1485 (1241-1726)	0.075 (0.063-0.087)
	RHAP-MDH	177 (103-252)	0.008 (0.005-0.012)	2042 (1599-2508)	0.097 (0.076-0.118)	325 (209-446)	0.015 (0.01-0.021)	897 (651-1153)	0.042 (0.03-0.055)
	RHAP-SIRD	185 (106-269)	0.008 (0.004-0.011)	3050 (2398-3746)	0.14 (0.11-0.172)	436 (311-567)	0.019 (0.014-0.025)	1427 (1131-1699)	0.064 (0.051-0.076)
	Range†	540	0.028	1794	0.102	284	0.014	600	0.037
Risk-driven subgroups	H1S1	238 (169-301)	0.012 (0.009-0.015)	2108 (1693-2619)	0.114 (0.092-0.14)	333 (235-430)	0.017 (0.012-0.022)	1177 (985-1357)	0.061 (0.051-0.07)
	H2S1	534 (384-677)	0.028 (0.02-0.035)	2531 (1985-3116)	0.136 (0.11-0.166)	441 (323-559)	0.023 (0.017-0.028)	1417 (1196-1631)	0.073 (0.061-0.084)
	H1S2	176 (101-252)	0.007 (0.004-0.011)	3469 (2802-4238)	0.154 (0.124-0.188)	400 (258-544)	0.017 (0.011-0.023)	1261 (993-1537)	0.054 (0.043-0.066)
	H2S2	422 (267-592)	0.018 (0.011-0.025)	3786 (3077-4583)	0.167 (0.135-0.201)	489 (332-673)	0.021 (0.014-0.028)	1508 (1212-1810)	0.066 (0.053-0.079)
	Range†	358	0.02	1678	0.053	157	0.006	331	0.018

*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.

Table 9.3 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in Base case U.K. setting

		Treat-to-target hypothetical intensive management				5%-Improvement-based hypothetical intensive management			
		HbA1c		HbA1c+LDL+BMI		HbA1c		HbA1c+LDL+BMI	
		Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU
	Overall*	41 (26-50)	0.007 (0.005-0.01)	338 (255-393)	0.069 (0.055-0.085)	55 (37-67)	0.01 (0.007-0.014)	137 (105-154)	0.026 (0.021-0.032)
Data-driven subgroups	RHAP-MOD	53 (37-68)	0.011 (0.007-0.014)	463 (345-603)	0.104 (0.078-0.136)	49 (36-63)	0.01 (0.007-0.013)	147 (116-182)	0.03 (0.024-0.038)
	RHAP-MD	30 (20-41)	0.006 (0.003-0.008)	196 (165-231)	0.041 (0.034-0.049)	38 (26-51)	0.008 (0.005-0.011)	105 (88-121)	0.021 (0.017-0.025)
	RHAP-SIDD	87 (64-110)	0.018 (0.012-0.023)	338 (282-393)	0.074 (0.061-0.087)	65 (47-83)	0.014 (0.009-0.018)	145 (120-166)	0.031 (0.026-0.037)
	RHAP-MDH	15 (5-24)	0.003 (0-0.005)	285 (227-349)	0.058 (0.045-0.072)	47 (31-63)	0.009 (0.006-0.013)	105 (80-131)	0.018 (0.012-0.025)
	RHAP-SIRD	21 (12-30)	0.004 (0.002-0.006)	380 (300-467)	0.083 (0.066-0.101)	59 (43-77)	0.011 (0.008-0.015)	152 (123-183)	0.032 (0.025-0.038)
	Range†	72	0.015	267	0.063	27	0.006	47	0.013
Risk-driven subgroups	H1S1	20 (13-27)	0.004 (0.002-0.006)	230 (183-285)	0.049 (0.039-0.062)	42 (30-53)	0.008 (0.005-0.011)	114 (94-133)	0.023 (0.018-0.027)
	H2S1	78 (59-97)	0.016 (0.012-0.02)	287 (232-348)	0.064 (0.051-0.079)	62 (48-77)	0.013 (0.01-0.017)	133 (109-155)	0.028 (0.023-0.033)
	H1S2	17 (8-26)	0.003 (0-0.005)	494 (394-608)	0.101 (0.081-0.123)	62 (43-81)	0.012 (0.008-0.016)	151 (120-183)	0.029 (0.022-0.035)
	H2S2	56 (35-80)	0.011 (0.007-0.017)	515 (418-621)	0.109 (0.089-0.131)	71 (49-96)	0.014 (0.009-0.02)	172 (141-205)	0.037 (0.03-0.044)
	Range†	61	0.013	285	0.061	30	0.006	59	0.014

*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.

Table 9.4 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in Scenario 1 U.K. setting

		Treat-to-target hypothetical intensive management				5%-Improvement-based hypothetical intensive management			
		HbA1c		HbA1c+LDL+BMI		HbA1c		HbA1c+LDL+BMI	
		Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU
	Overall*	78 (52-94)	0.014 (0.01-0.019)	592 (450-676)	0.123 (0.098-0.15)	94 (65-111)	0.017 (0.012-0.023)	299 (234-327)	0.058 (0.048-0.068)
Data-driven subgroups	RHAP-MOD	85 (62-110)	0.017 (0.012-0.022)	815 (629-1036)	0.184 (0.138-0.236)	82 (60-105)	0.016 (0.011-0.021)	345 (292-407)	0.072 (0.06-0.085)
	RHAP-MD	63 (45-83)	0.012 (0.009-0.017)	409 (344-472)	0.089 (0.074-0.104)	74 (53-93)	0.015 (0.01-0.019)	254 (216-291)	0.053 (0.045-0.061)
	RHAP-SIDD	166 (126-207)	0.034 (0.025-0.043)	601 (499-695)	0.135 (0.112-0.159)	133 (100-165)	0.027 (0.02-0.035)	324 (275-373)	0.07 (0.059-0.08)
	RHAP-MDH	38 (25-52)	0.008 (0.005-0.011)	448 (356-550)	0.091 (0.072-0.112)	72 (50-94)	0.014 (0.009-0.02)	211 (166-258)	0.039 (0.028-0.051)
	RHAP-SIRD	40 (25-56)	0.007 (0.004-0.011)	591 (469-721)	0.132 (0.104-0.163)	92 (67-116)	0.018 (0.013-0.023)	283 (230-332)	0.06 (0.048-0.071)
	Range†	128	0.026	407	0.095	61	0.013	135	0.033
Risk-driven subgroups	H1S1	55 (40-68)	0.011 (0.008-0.014)	476 (388-573)	0.105 (0.085-0.13)	80 (60-100)	0.016 (0.011-0.02)	270 (226-309)	0.056 (0.047-0.065)
	H2S1	127 (96-160)	0.026 (0.019-0.032)	550 (446-660)	0.126 (0.102-0.154)	102 (78-126)	0.021 (0.016-0.026)	314 (268-357)	0.067 (0.057-0.077)
	H1S2	37 (23-51)	0.007 (0.004-0.01)	704 (561-866)	0.147 (0.119-0.178)	86 (58-112)	0.016 (0.01-0.022)	263 (212-318)	0.051 (0.04-0.063)
	H2S2	90 (60-122)	0.017 (0.011-0.024)	742 (594-902)	0.159 (0.129-0.191)	105 (75-138)	0.02 (0.013-0.027)	297 (240-354)	0.063 (0.05-0.075)
	Range†	91	0.019	267	0.054	24	0.005	51	0.015

*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.

Figure 9.1 Maximum annual cost-effective price, incremental QALYs and complication costs of multiple-year guideline-based intensive management targeting HbA1c, BMI and LDL by RHAPSODY data-driven subgroups (in U.S. setting), for a range of treatment durations

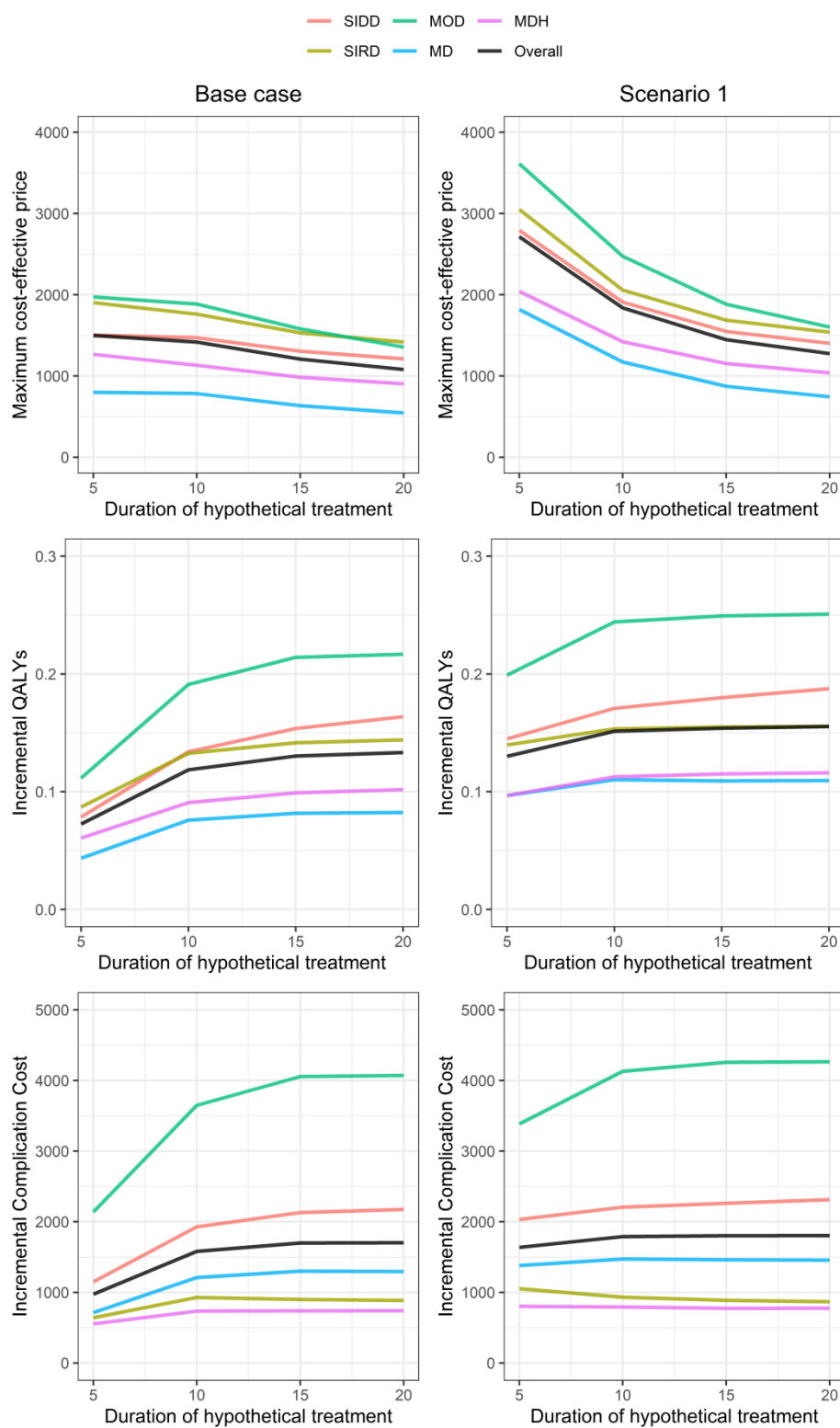


Figure 9.2 Maximum annual cost-effective price, incremental QALYs and complication costs of multiple-year guideline-based intensive management targeting HbA1c, BMI and LDL by risk-driven subgroups (in U.S. setting), for a range of treatment durations

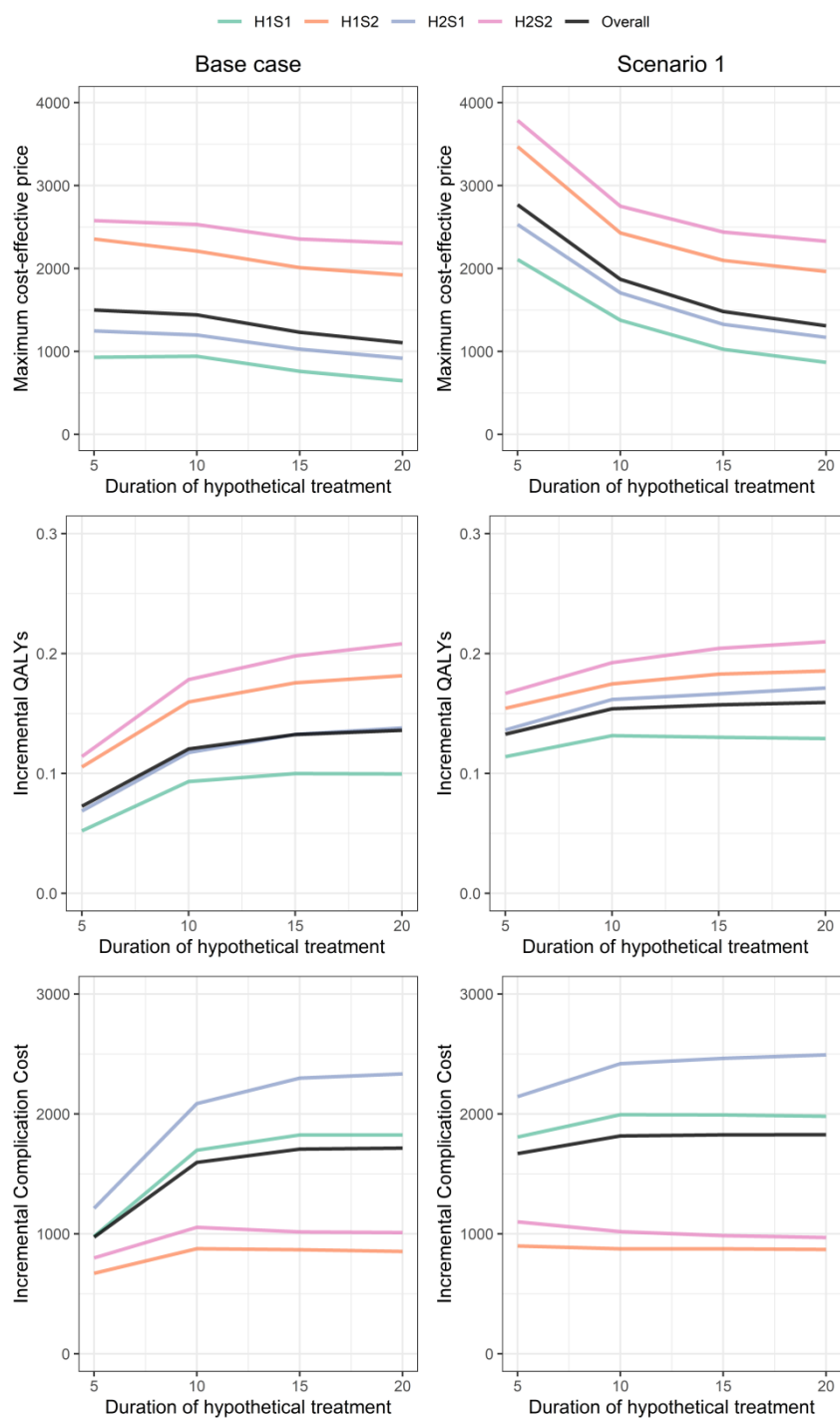
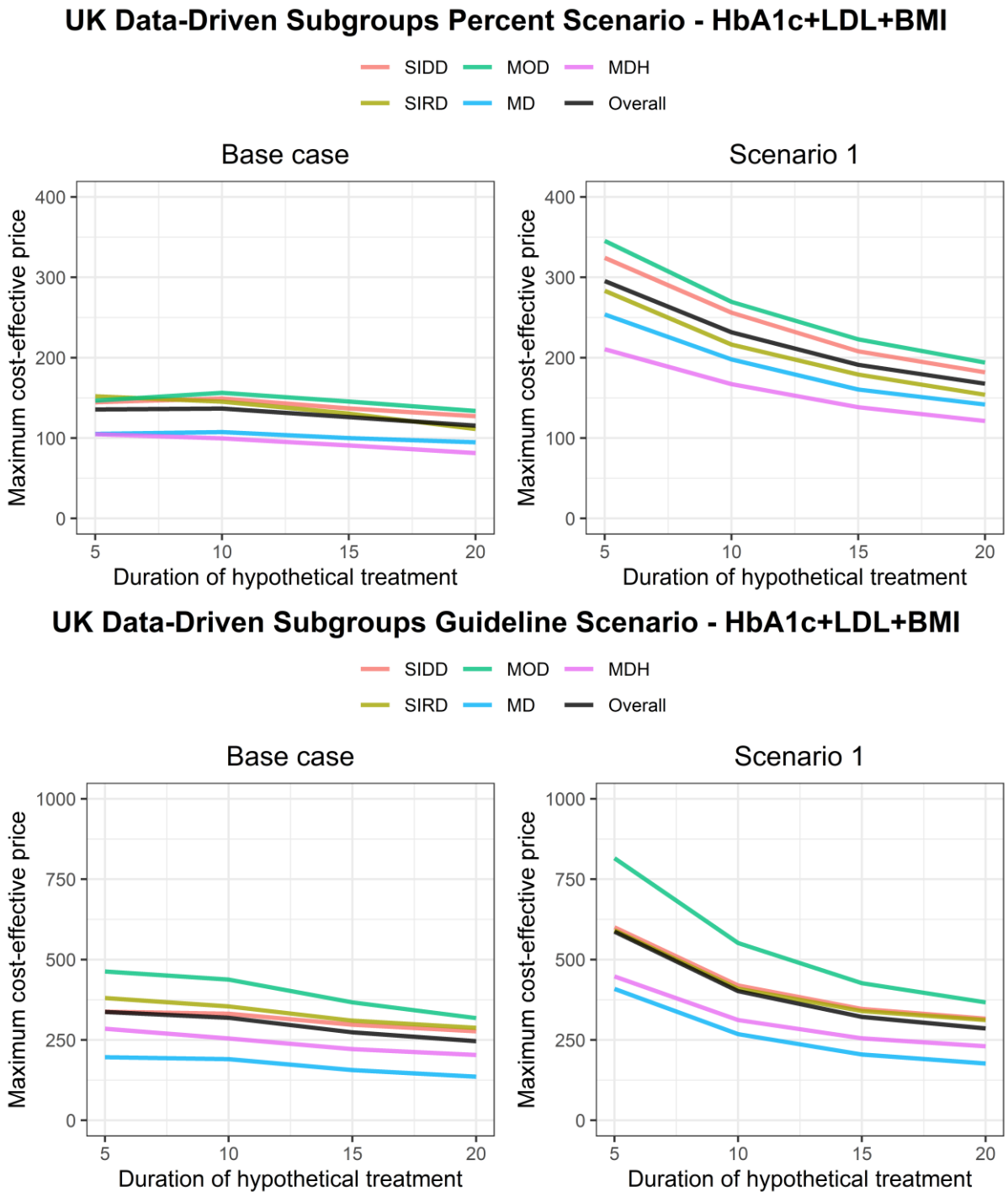
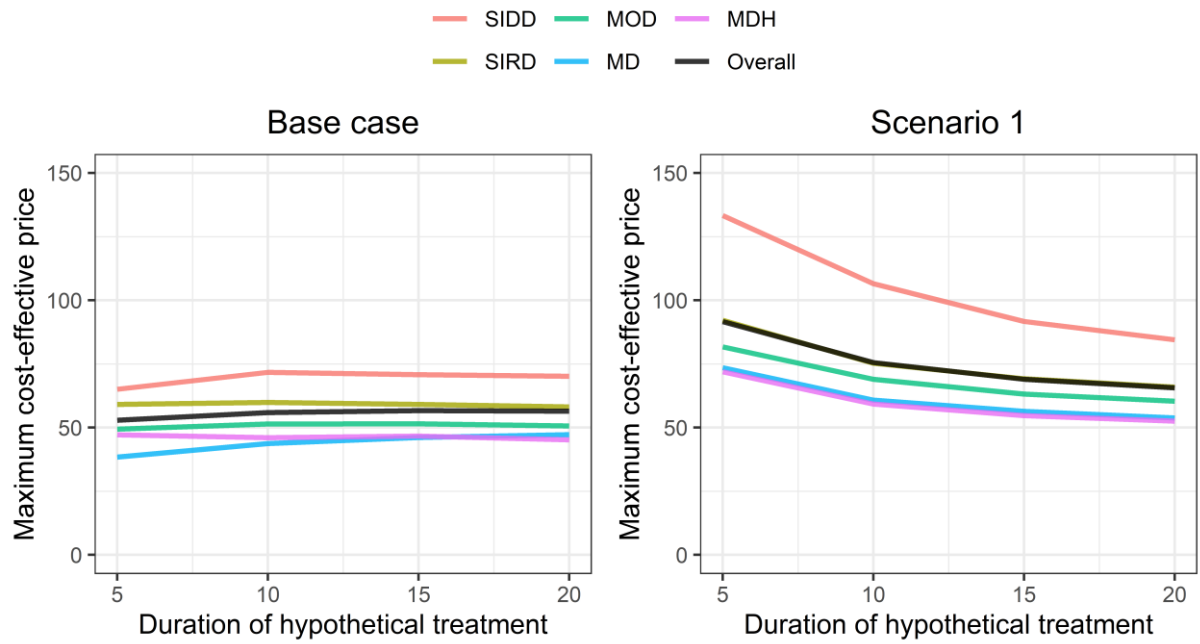


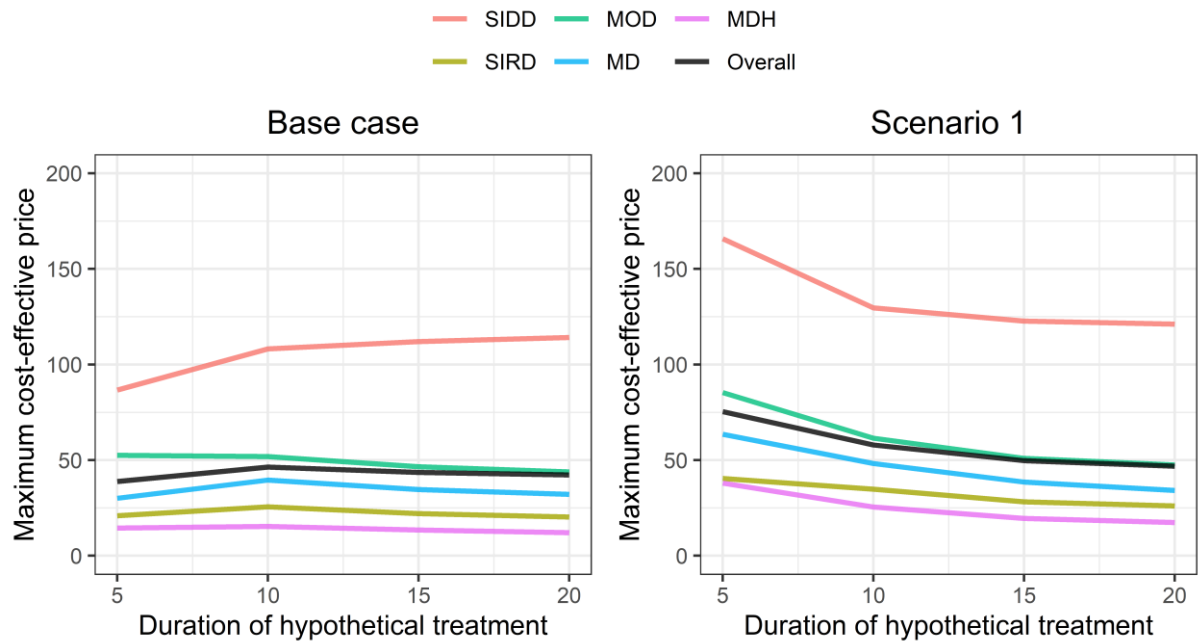
Figure 9.3 The results of maximum annual cost-effective prices of multiple scenarios, for a range of treatment duration



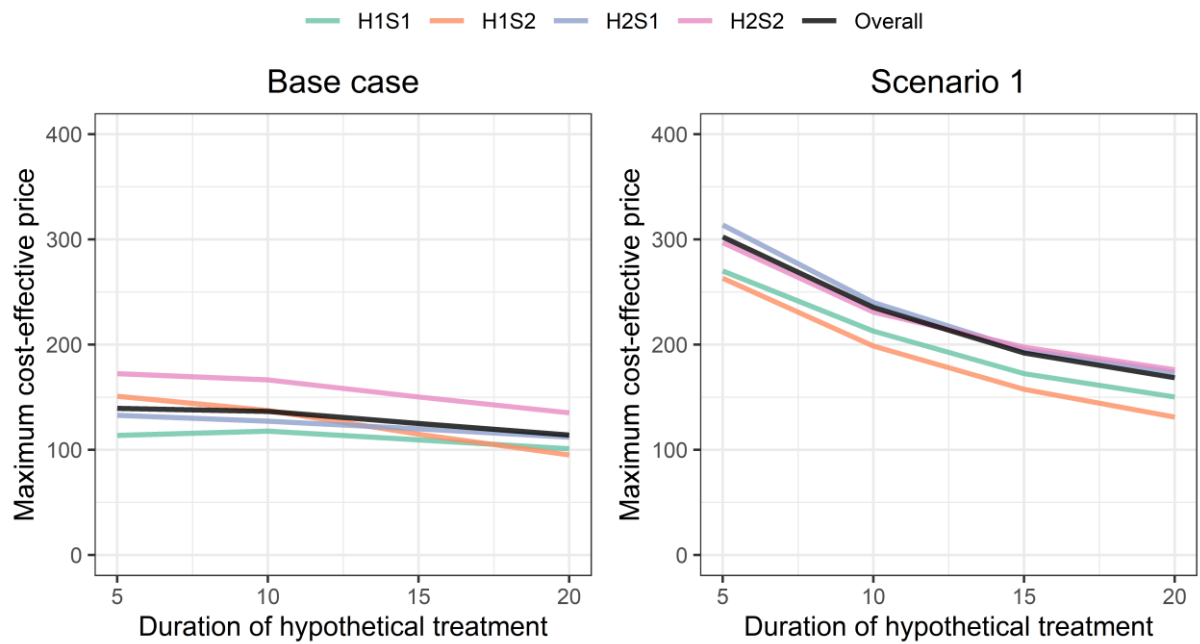
UK Data-Driven Subgroups Percent Scenario - HbA1c



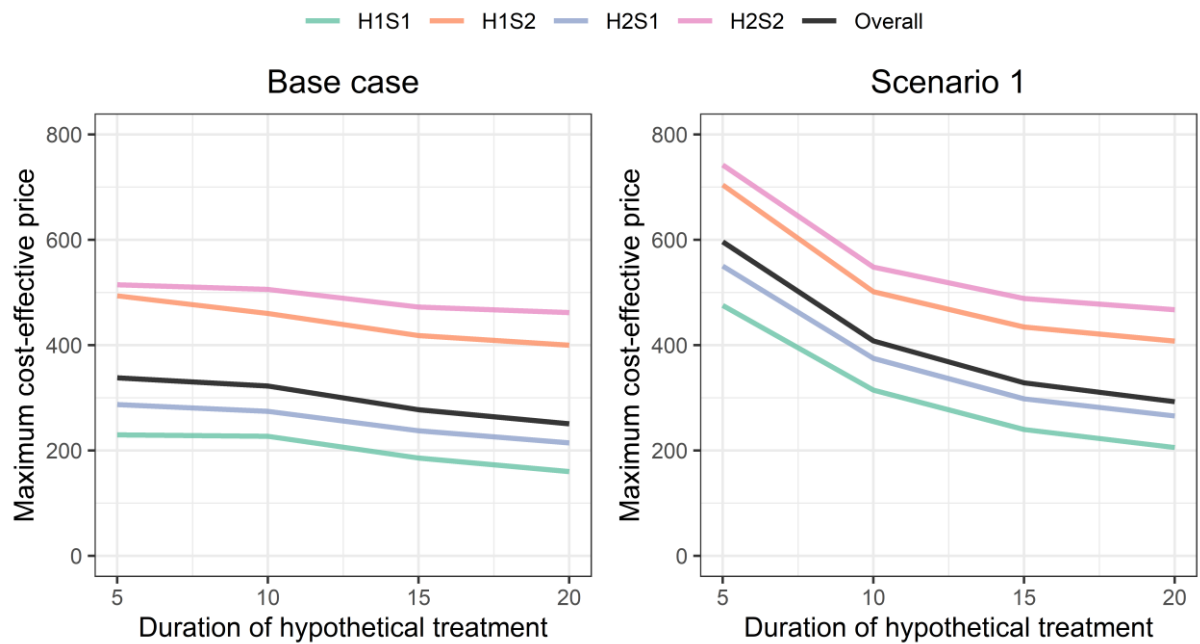
UK Data-Driven Subgroups Guideline Scenario - HbA1c



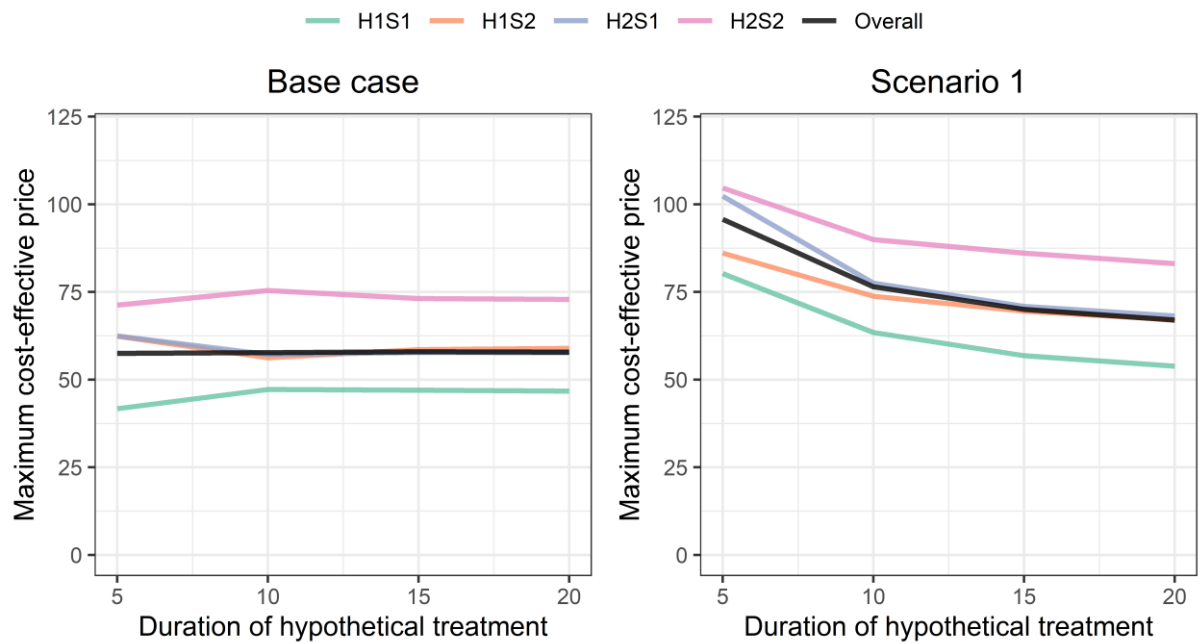
UK Risk-Driven Subgroups Percent Scenario - HbA1c+LDL+BMI



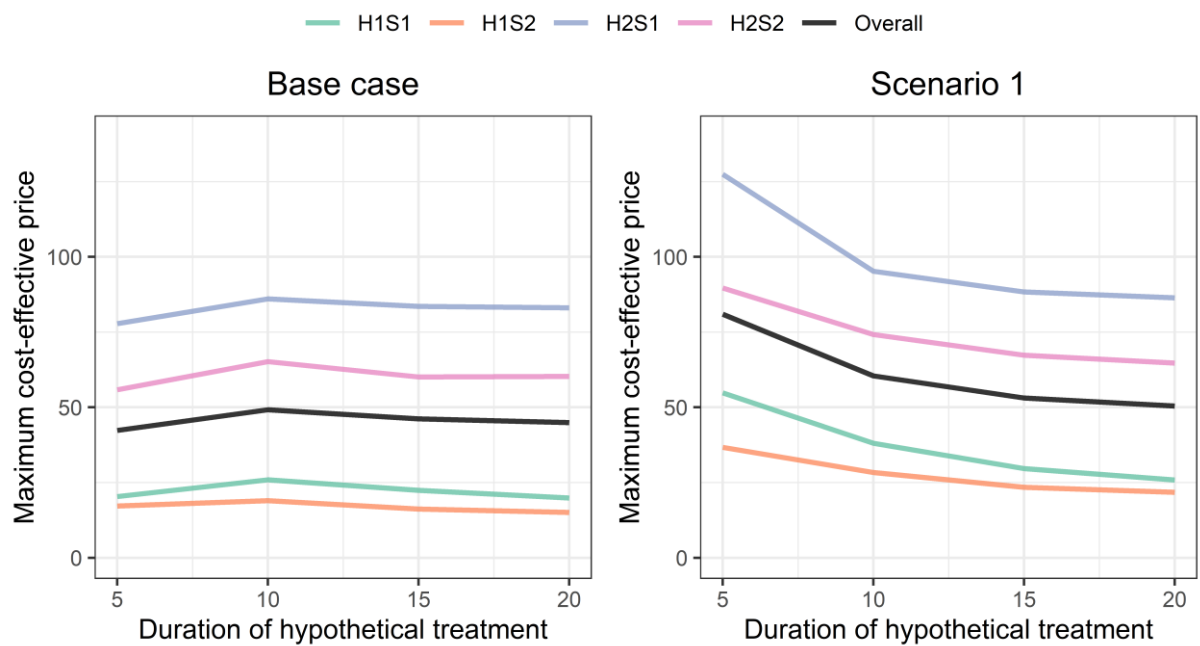
UK Risk-Driven Subgroups Guideline Scenario - HbA1c+LDL+BMI



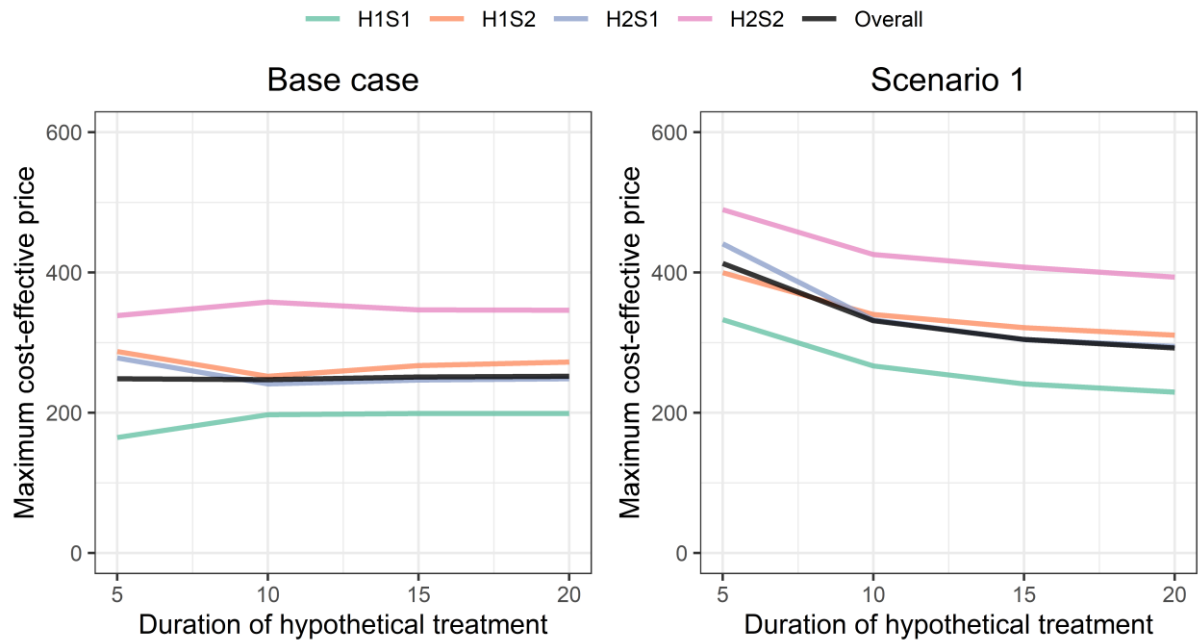
UK Risk-Driven Subgroups Percent Scenario - HbA1c



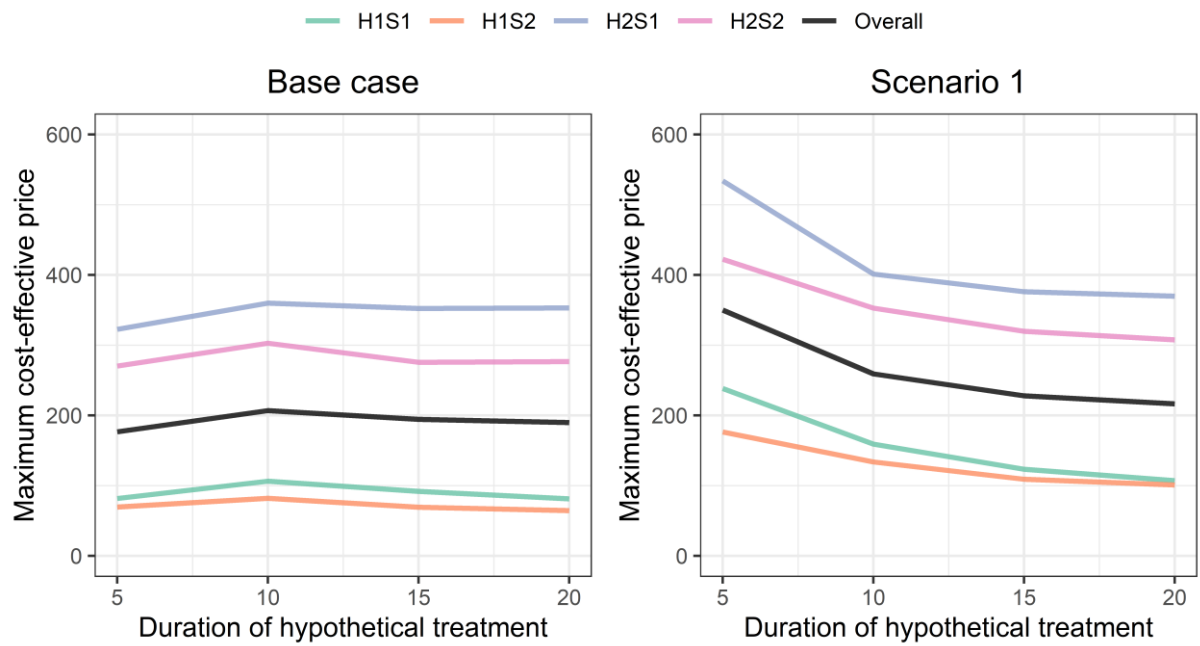
UK Risk-Driven Subgroups Guideline Scenario - HbA1c



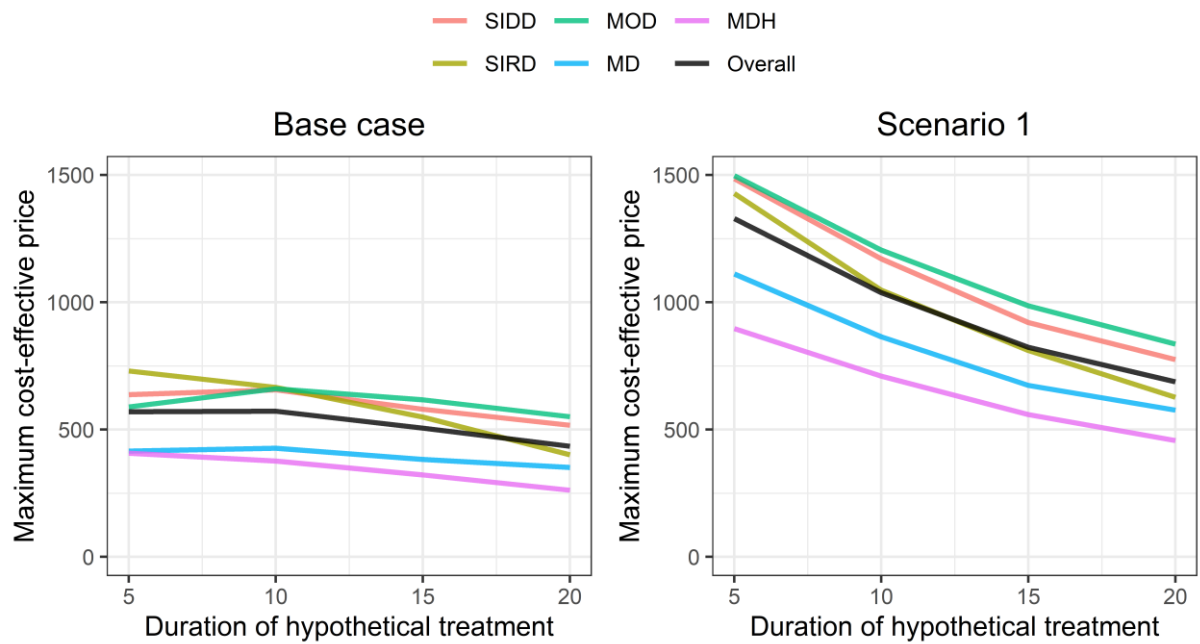
US Risk-Driven Subgroups Percent Scenario - HbA1c



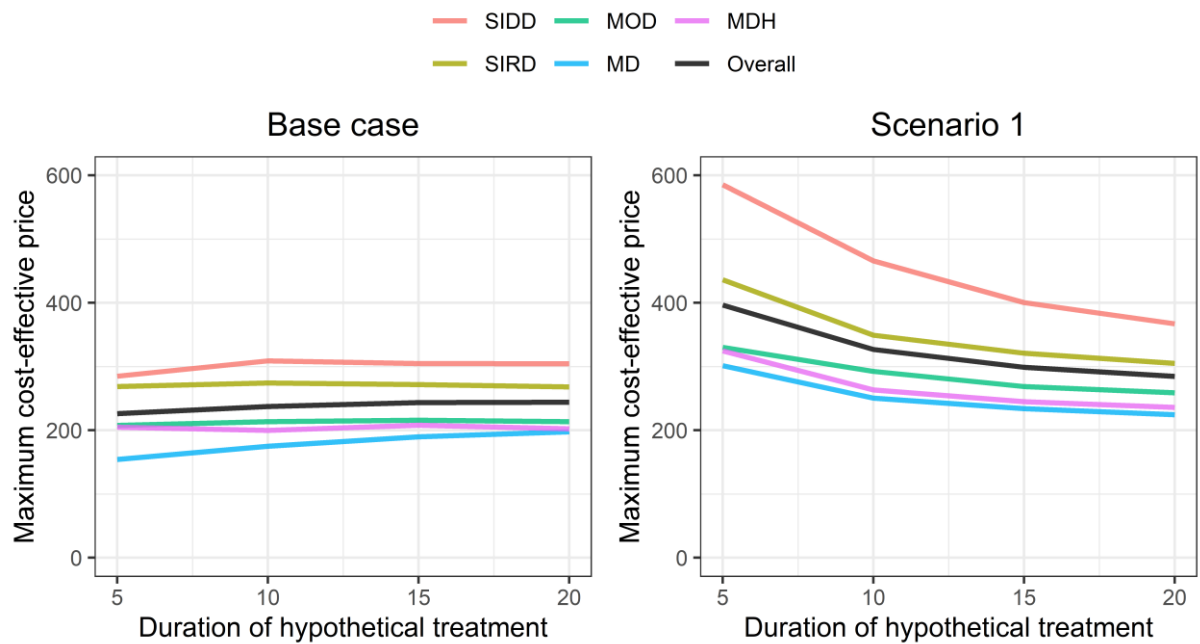
US Risk-Driven Subgroups Guideline Scenario - HbA1c



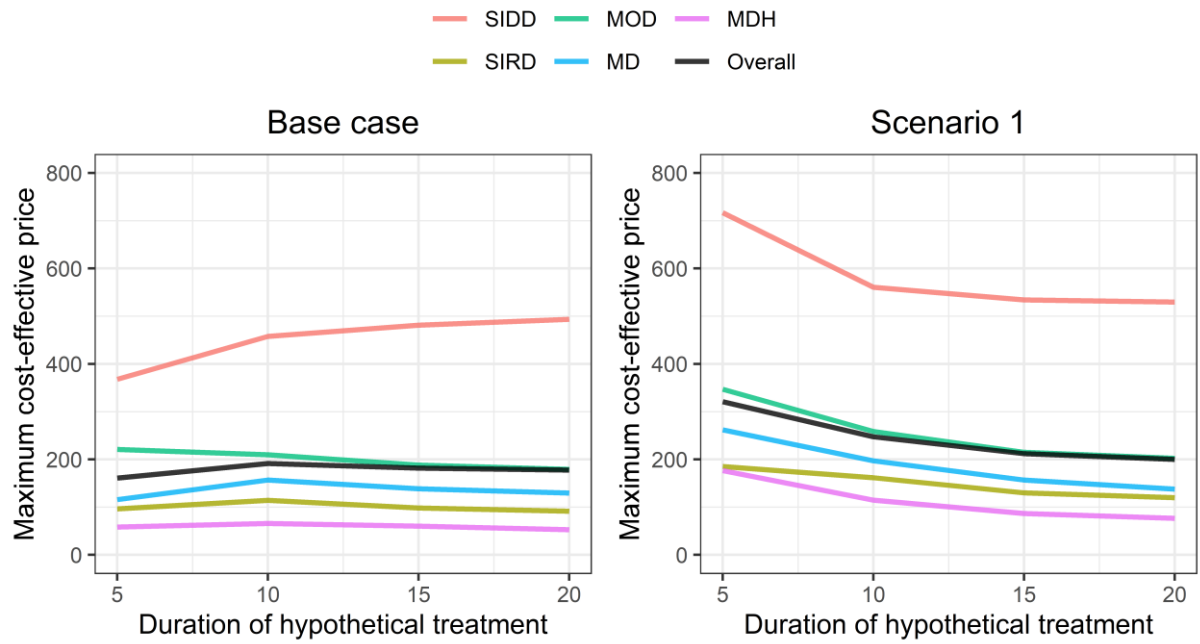
US Data-Driven Subgroups Percent Scenario - HbA1c+LDL+BMI



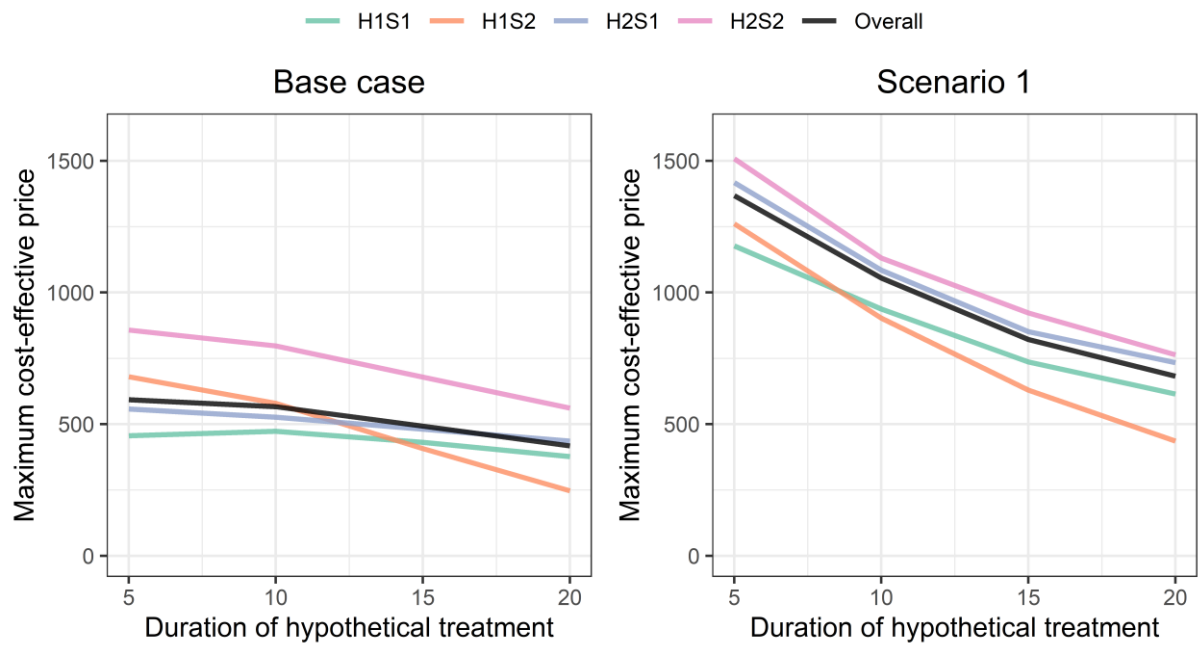
US Data-Driven Subgroups Percent Scenario - HbA1c



US Data-Driven Subgroups Guideline Scenario - HbA1c



US Risk-Driven Subgroups Percent Scenario - HbA1c+LDL+BMI



Appendix 10 The sensitivity analysis of applying Ahlqvist's subgroups

Table 10.1 Baseline simulation characteristics of Ahlqvist's data-driven subgroups

	Ahlqvist's Data-driven subgroups				p
	SIDD	SIRD	MOD	MARD	
n	447	580	774	1096	
Age, years	60.93 (9.58)	65.09 (9.38)	54.42 (8.04)	68.05 (7.93)	<0.001
Duration of diabetes, years	4.41 (3.40)	2.33 (2.76)	3.48 (3.38)	2.64 (3.10)	<0.001
Age at diabetes diagnosis, years	56.52 (9.38)	62.76 (9.22)	50.94 (7.19)	65.41 (7.87)	<0.001
LDL cholesterol, mmol/liter	2.68 (0.93)	2.68 (0.88)	2.73 (0.89)	2.79 (0.94)	0.063
Total cholesterol, mmol/liter	4.63 (1.10)	4.71 (1.01)	4.72 (1.04)	4.77 (1.11)	0.143
HDL cholesterol, mmol/liter	1.16 (0.32)	1.10 (0.28)	1.11 (0.28)	1.29 (0.35)	<0.001
Total cholesterol/HDL ratio	4.22 (1.35)	4.48 (1.30)	4.46 (1.31)	3.88 (1.15)	<0.001
HbA1c, %	7.70 (1.52)	6.52 (0.79)	6.79 (1.00)	6.49 (0.77)	<0.001
HbA1c, mmol/mol	60.65 (16.58)	47.78 (8.64)	50.73 (10.96)	47.43 (8.43)	<0.001
Fasting blood glucose, mmol/liter	8.77 (2.59)	7.56 (1.44)	8.16 (1.98)	7.51 (1.29)	<0.001
Triglycerides, mmol/liter	1.81 (1.13)	2.12 (1.11)	2.02 (1.22)	1.54 (0.70)	<0.001
Urinary microalbumin (mg/liter)	22.67 (67.65)	42.23 (194.26)	19.45 (69.33)	12.51 (57.71)	<0.001
Urinary creatinine (umol/liter)	9.27 (4.88)	9.92 (7.84)	11.22 (6.17)	8.17 (4.98)	<0.001
Albumin creatinine ratio (mg/mmol)	3.20 (11.58)	5.33 (24.04)	1.76 (5.75)	1.47 (5.42)	<0.001
Blood creatinine (umol/liter)	77.36 (19.10)	86.39 (24.91)	75.52 (15.97)	79.85 (18.51)	<0.001
eGFR based on CKD-EPI, mL/min/1.73 m2	85.94 (16.58)	75.04 (18.21)	90.91 (15.60)	77.64 (15.64)	<0.001
Length, cm	172.80 (9.75)	171.10 (9.18)	174.00 (10.14)	170.39 (9.33)	<0.001
Weight, kg	87.31 (15.58)	96.84 (17.05)	101.43 (17.31)	79.61 (11.64)	<0.001
BMI, kg/m2	29.21 (4.47)	33.09 (5.54)	33.51 (5.15)	27.40 (3.20)	<0.001
Systolic blood pressure, mmHg	141.88 (19.12)	144.68 (19.67)	137.36 (16.93)	144.87 (19.19)	<0.001
Diastolic blood pressure, mmHg	78.24 (9.20)	78.72 (9.39)	78.94 (9.35)	77.43 (9.36)	0.003
Male, n (%)	280 (62.6)	316 (54.5)	466 (60.2)	562 (51.3)	<0.001
Smoking status					<0.001
Current, n (%)	97 (21.8)	105 (18.4)	179 (23.8)	157 (14.9)	
Former, n (%)	206 (46.4)	304 (53.3)	350 (46.6)	549 (52.1)	
Never, n (%)	141 (31.8)	161 (28.2)	222 (29.6)	348 (33.0)	

Table 10.2 Comparison between simulation baseline characteristics of corresponding subgroups in two data-driven approaches

	SIDD			SIRD			MOD		
	Ahlqvist*	RHAP†	p	Ahlqvist*	RHAP†	p	Ahlqvist*	RHAP†	p
n	447	365		580	637		774	520	
Age, years	60.93 (9.58)	61.39 (9.74)	0.499	65.09 (9.38)	70.74 (7.41)	<0.001	54.42 (8.04)	55.90 (8.05)	0.001
Duration of diabetes, years	4.41 (3.40)	4.46 (3.29)	0.821	2.33 (2.76)	2.30 (2.74)	0.804	3.48 (3.38)	3.05 (3.21)	0.022
Age at diabetes diagnosis, years	56.52 (9.38)	56.93 (9.28)	0.537	62.76 (9.22)	68.45 (7.22)	<0.001	50.94 (7.19)	52.85 (7.78)	<0.001
LDL cholesterol, mmol/liter	2.68 (0.93)	2.72 (0.91)	0.563	2.68 (0.88)	2.64 (0.89)	0.495	2.73 (0.89)	2.68 (0.90)	0.328
Total cholesterol, mmol/liter	4.63 (1.10)	4.65 (1.05)	0.797	4.71 (1.01)	4.59 (1.04)	0.045	4.72 (1.04)	4.68 (1.03)	0.586
HDL cholesterol, mmol/liter	1.16 (0.32)	1.16 (0.31)	0.783	1.10 (0.28)	1.08 (0.22)	0.205	1.11 (0.28)	1.06 (0.26)	0.003
Total cholesterol/HDL ratio	4.22 (1.35)	4.24 (1.28)	0.825	4.48 (1.30)	4.38 (1.23)	0.18	4.46 (1.31)	4.61 (1.34)	0.05
HbA1c, %	7.70 (1.52)	7.78 (1.78)	0.496	6.52 (0.79)	6.53 (0.73)	0.863	6.79 (1.00)	6.83 (1.00)	0.501
HbA1c, mmol/mol		61.51					50.73	51.17	
	60.65 (16.58)	(19.34)	0.496	47.78 (8.64)	47.87 (7.99)	0.854	(10.96)	(10.93)	0.483
Fasting blood glucose, mmol/liter	8.77 (2.59)	8.82 (2.83)	0.791	7.56 (1.44)	7.69 (1.35)	0.1	8.16 (1.98)	8.02 (1.85)	0.205
Triglycerides, mmol/liter	1.81 (1.13)	1.77 (1.00)	0.618	2.12 (1.11)	1.94 (0.99)	0.003	2.02 (1.22)	2.15 (1.29)	0.072
Urinary microalbumin (mg/liter)		23.22		42.23	28.50		19.45	30.59	
	22.67 (67.65)	(67.38)	0.91	(194.26)	(150.02)	0.176	(69.33)	(130.71)	0.051
Urinary creatinine (umol/liter)	9.27 (4.88)	9.18 (5.15)	0.79	9.92 (7.84)	8.88 (7.21)	0.02	11.22 (6.17)	10.74 (5.83)	0.178
Albumin creatinine ratio (mg/mmol)		3.18							
	3.20 (11.58)	(10.43)	0.982	5.33 (24.04)	3.93 (19.70)	0.278	1.76 (5.75)	3.11 (14.25)	0.02
Blood creatinine (umol/liter)		78.65		86.39	87.95		75.52	75.88	
	77.36 (19.10)	(21.97)	0.368	(24.91)	(23.76)	0.267	(15.97)	(17.38)	0.703
eGFR based on CKD-EPI, mL/min/1.73 m2		84.35		75.04	71.50		90.91	88.06	
	85.94 (16.58)	(18.02)	0.193	(18.21)	(15.94)	<0.001	(15.60)	(16.84)	0.002
BMI, kg/m2	29.21 (4.47)	29.50 (4.59)	0.364	33.09 (5.54)	29.89 (3.51)	<0.001	33.51 (5.15)	37.82 (4.98)	<0.001
Systolic blood pressure, mmHg		142.17		144.68	146.75		137.36	141.69	
	141.88 (19.12)	(19.84)	0.831	(19.67)	(19.99)	0.071	(16.93)	(17.87)	<0.001
Diastolic blood pressure, mmHg	78.24 (9.20)	78.18 (9.43)	0.93	78.72 (9.39)	77.72 (9.85)	0.073	78.94 (9.35)	79.85 (9.30)	0.086
C-Peptide, nmol/liter	0.82 (0.42)	0.83 (0.40)	0.63	1.75 (0.38)	1.46 (0.45)	<0.001	1.04 (0.30)	1.43 (0.44)	<0.001
Male, n (%)	280 (62.6)	218 (59.7)	0.438	316 (54.5)	397 (62.3)	0.007	466 (60.2)	256 (49.2)	<0.001
Smoking status			0.961			0.064			0.495
Current, n (%)	97 (21.8)	76 (21.1)		105 (18.4)	86 (13.9)		179 (23.8)	108 (21.1)	
Former, n (%)	206 (46.4)	170 (47.1)		304 (53.3)	364 (58.9)		350 (46.6)	251 (48.9)	
Never, n (%)	141 (31.8)	115 (31.9)		161 (28.2)	168 (27.2)		222 (29.6)	154 (30.0)	

*Ahlqvist indicates the subgroups defined according to Ahlqvist's data-driven clustering subgroups (22).

†RHAP indicates the subgroups defined according to RHAPSODY data-driven subgroups (1).

Table 10.3 Absolute and standardized simulated outcome (mean and 95% CI) by Ahlqvist's data-driven subgroups in the U.S. setting

	Lifetime QALYs	Standardized Lifetime QALYs - Male	Standardized Lifetime QALYs - Female	Complication cost (100,000\$)	Standardized complication cost (100,000\$) - Male	Standardized complication cost (100,000\$) - Female
Overall*	10.57 (10.46-10.91)	9.98 (9.65-10.52)	11.12 (10.8-11.65)	1.65 (1.63-1.7)	1.54 (1.49-1.63)	1.76 (1.71-1.84)
Ahlqvist's Data-driven clustering subgroups						
MARD	9.15 (9-9.51)	10.6 (10.3-11.13)	11.31 (11.02-11.83)	1.44 (1.42-1.5)	1.62 (1.58-1.7)	1.79 (1.74-1.87)
SIRD	9.61 (9.48-9.96)	9.99 (9.62-10.61)	10.92 (10.58-11.48)	1.5 (1.48-1.56)	1.53 (1.47-1.62)	1.73 (1.68-1.82)
SIDD	10.91 (10.81-11.22)	10.13 (9.82-10.63)	11 (10.67-11.52)	1.72 (1.7-1.77)	1.6 (1.56-1.68)	1.74 (1.69-1.83)
MOD	13.09 (12.98-13.4)	10.22 (9.89-10.76)	11.05 (10.73-11.56)	2.01 (1.99-2.05)	1.55 (1.51-1.63)	1.72 (1.67-1.8)
Range†	3.94	0.61	0.39	0.57	0.09	0.07

* Overall was defined as the average result, indicating the result of calculating group-wise simulation outcome altogether as a homogenous type 2 diabetes group. (Average values of results based on three extrapolations of linear dynamic models estimated from Ahlqvist's and RHAPSODY data-driven subgroups and risk-driven subgroups were applied.)

†Range was defined as the maximum – minimum of the mean maximum annual cost-effective price of intervention or incremental QALYs. To compare the value of RHAPSODY data-driven subgroups versus risk-driven subgroups, the range between subgroups was considered, with a wider range indicating a greater ability to differentiate between subgroups.

Table 10.4 Absolute and standardized simulated outcome (mean and 95% CI) by Ahlqvist's data-driven subgroups in the U.K. setting

	Lifetime QALYs	Standardized Lifetime QALYs - Male	Standardized Lifetime QALYs - Female	Complication cost (100,000£)	Standardized complication cost (100,000£) - Male	Standardized complication cost (100,000£) - Female
Overall*	10.09 (9.98-10.39)	9.54 (9.24-10.04)	10.58 (10.29-11.06)	0.37 (0.36-0.39)	0.36 (0.34-0.4)	0.39 (0.37-0.41)
Ahlqvist's Data-driven clustering subgroups						
MARD	8.79 (8.65-9.12)	10.14 (9.86-10.61)	10.79 (10.52-11.27)	0.36 (0.35-0.39)	0.37 (0.35-0.4)	0.38 (0.37-0.41)
SIRD	9.2 (9.08-9.52)	9.56 (9.21-10.12)	10.41 (10.09-10.92)	0.38 (0.37-0.41)	0.38 (0.36-0.43)	0.4 (0.38-0.43)
SIDD	10.39 (10.3-10.67)	9.68 (9.39-10.14)	10.48 (10.17-10.95)	0.37 (0.36-0.39)	0.35 (0.33-0.38)	0.39 (0.37-0.42)
MOD	12.37 (12.27-12.64)	9.8 (9.49-10.29)	10.54 (10.24-11.01)	0.38 (0.37-0.4)	0.37 (0.35-0.41)	0.41 (0.4-0.44)
Range†	3.58	0.58	0.38	0.02	0.03	0.03

* Overall was defined as the average result, indicating the result of calculating group-wise simulation outcome altogether as a homogenous type 2 diabetes group. (Average values of results based on three extrapolations of linear dynamic models estimated from Ahlqvist's and RHAPSODY data-driven subgroups and risk-driven subgroups were applied.)

†Range was defined as the maximum – minimum of the mean maximum annual cost-effective price of intervention or incremental QALYs. To compare the value of RHAPSODY data-driven subgroups versus risk-driven subgroups, the range between subgroups was considered, with a wider range indicating a greater ability to differentiate between subgroups.

Table 10.5 The effect of hypothetical treatment on risk factors of each Ahlqvist's data-driven subgroup

		Year 1	Year 2	Year 3	Year 4	Year 5	
Treat-to-target scenario	HbA1c	SIDD	0.44 (5.11%)	0.52 (5.98%)	0.55 (6.11%)	0.52 (6.05%)	0.53 (6.18%)
		SIRD	0.11 (1.23%)	0.1 (1.25%)	0.12 (1.4%)	0.13 (1.57%)	0.16 (1.97%)
		MOD	0.23 (2.74%)	0.27 (3.21%)	0.24 (2.9%)	0.26 (3.05%)	0.29 (3.37%)
		MARD	0.07 (0.91%)	0.11 (1.33%)	0.1 (1.29%)	0.11 (1.31%)	0.13 (1.53%)
	LDL	SIDD	0.42 (11.96%)	0.33 (9.83%)	0.32 (9.3%)	0.32 (9.21%)	0.3 (8.95%)
		SIRD	0.51 (15.76%)	0.47 (14.84%)	0.4 (13.05%)	0.42 (13.48%)	0.4 (12.94%)
		MOD	0.29 (7.77%)	0.25 (6.72%)	0.24 (6.39%)	0.22 (5.87%)	0.2 (5.44%)
		MARD	0.61 (18.64%)	0.55 (16.89%)	0.49 (15.62%)	0.48 (15.26%)	0.46 (14.86%)
	Weight	SIDD	12.65 (13.13%)	12.87 (13.32%)	12.67 (13.17%)	12.24 (12.77%)	11.89 (12.41%)
		SIRD	22.4 (21.72%)	22.22 (21.61%)	21.58 (21.16%)	20.67 (20.36%)	19.83 (19.64%)
		MOD	24.28 (22.8%)	24.22 (22.73%)	23.83 (22.33%)	23.29 (21.95%)	22.75 (21.56%)
		MARD	7.45 (8.63%)	7.5 (8.65%)	7.37 (8.5%)	7.08 (8.18%)	6.85 (7.91%)
5%-Improvement-based scenario	HbA1c	SIDD	0.36 (5%)	0.36 (5%)	0.36 (5%)	0.36 (5%)	0.36 (5%)
		SIRD	0.32 (5%)	0.33 (5%)	0.33 (5%)	0.33 (5%)	0.33 (5%)
		MOD	0.34 (5%)	0.34 (5%)	0.34 (5%)	0.34 (5%)	0.34 (5%)
		MARD	0.32 (5%)	0.33 (5%)	0.32 (5%)	0.33 (5%)	0.33 (5%)
	LDL	SIDD	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		SIRD	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		MOD	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
		MARD	0.13 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)	0.12 (5%)
	Weight	SIDD	4.34 (5%)	4.35 (5%)	4.34 (5%)	4.31 (5%)	4.29 (5%)
		SIRD	4.79 (5%)	4.78 (5%)	4.75 (5%)	4.7 (5%)	4.65 (5%)
		MOD	5.02 (5%)	5.01 (5%)	4.99 (5%)	4.96 (5%)	4.94 (5%)
		MARD	3.96 (5%)	3.96 (5%)	3.95 (5%)	3.93 (5%)	3.91 (5%)

Table 10.6 Random effect models for HbA1c and weight over time by Ahlqvist's subgroups

	HbA1c (%)				Weight (kg)			
	SIDD	SIRD	MOD	MARD	SIDD	SIRD	MOD	MARD
No. of observations	4918	5002	7747	10102	4918	5002	7747	10102
No. of individuals	446	580	773	1093	446	580	773	1093
Adjusted R ²	0.356	0.379	0.423	0.480	0.936	0.918	0.910	0.936
Parameters	Estimate of coefficient (standard error)							
Constant	2.719*** (0.138)	1.995*** (0.097)	1.815*** (0.084)	1.535*** (0.058)	2.159*** (0.413)	3.501*** (0.526)	2.969*** (0.448)	1.569*** (0.262)
Male	0.027 (0.027)	0.008 (0.019)	0.01 (0.018)	0.005 (0.011)	0.491*** (0.126)	0.266. (0.143)	0.311* (0.124)	0.114. (0.066)
Value of Y in previous year	0.598*** (0.013)	0.624*** (0.014)	0.649*** (0.01)	0.68*** (0.009)	0.905*** (0.011)	0.803*** (0.019)	0.879*** (0.009)	0.923*** (0.008)
Log(diabetes duration)	0.374*** (0.053)	0.198*** (0.041)	0.271*** (0.037)	0.212*** (0.024)	-0.814*** (0.234)	-1.527*** (0.306)	-0.659** (0.241)	-0.725*** (0.129)
Baseline of Y	-0.012 (0.01)	0.056*** (0.009)	0.059*** (0.009)	0.065*** (0.006)	0.073*** (0.011)	0.166*** (0.019)	0.091*** (0.009)	0.061*** (0.008)
Dummy variable of Year 1	3.233*** (0.369)	3.139*** (0.208)	2.902*** (0.216)	3.329*** (0.135)	0.33 (1.276)	-0.775 (1.476)	3.01* (1.363)	1.017 (0.765)
Baseline of Y*Dummy variable of Year 1	-0.486*** (0.039)	-0.504*** (0.031)	-0.451*** (0.031)	-0.53*** (0.02)	-0.006 (0.014)	-0.001 (0.015)	-0.043*** (0.013)	-0.022* (0.009)

Significance is indicated as 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 '' 1.

Table 10.7 Random effect models for eGFR and LDL over time by Ahlqvist's subgroups

	eGFR (mL/min/1.73 m2)				LDL (mmol/liter)			
	SIDD	SIRD	MOD	MARD	SIDD	SIRD	MOD	MARD
No. of observations	4918	5002	7747	10102	4918	5002	7747	10102
No. of individuals	446	580	773	1093	446	580	773	1093
Adjusted R ²	0.789	0.803	0.717	0.764	0.505	0.467	0.511	0.509
Parameters	Estimate of coefficient (standard error)							
Constant	5.131*** (0.873)	4.035*** (0.752)	9.245*** (0.905)	4.315*** (0.589)	0.7*** (0.413)	0.726*** (0.526)	0.728*** (0.448)	0.725*** (0.262)
Male	0.053 (0.244)	-0.522* (0.243)	0.291 (0.204)	-0.364* (0.163)	-0.015 (0.126)	-0.063*** (0.143)	-0.1*** (0.124)	-0.058*** (0.066)
Value of Y in previous year	0.834*** (0.012)	0.814*** (0.013)	0.736*** (0.011)	0.784*** (0.011)	0.64*** (0.011)	0.64*** (0.019)	0.657*** (0.009)	0.671*** (0.008)
Log(diabetes duration)	-3.051*** (0.502)	-3.398*** (0.542)	-4.175*** (0.431)	-3.541*** (0.36)	-0.152*** (0.234)	-0.122** (0.306)	-0.2*** (0.241)	-0.172*** (0.129)
Baseline of Y	0.111*** (0.012)	0.142*** (0.013)	0.177*** (0.013)	0.173*** (0.011)	0.08*** (0.011)	0.067*** (0.019)	0.093*** (0.009)	0.064*** (0.008)
Dummy variable of Year 1	10.379*** (2.758)	8.051*** (1.905)	11.126*** (2.935)	7.735*** (1.703)	0.175 (1.276)	0.424*** (1.476)	0.641*** (1.363)	0.53*** (0.765)
Baseline of Y*Dummy variable of Year 1	-0.142*** (0.031)	-0.112*** (0.024)	-0.138*** (0.033)	-0.109*** (0.021)	-0.091** (0.014)	-0.166*** (0.015)	-0.22*** (0.013)	-0.212*** (0.009)

Significance is indicated as 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1.

Table 10.8 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in U.S. setting

		Treat-to-target hypothetical intensive management				5%-Improvement-based hypothetical intensive management			
		HbA1c		HbA1c+LDL+BMI		HbA1c		HbA1c+LDL+BMI	
		Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU	Max annual price of intervention (\$)	Δ QALYs vs CAU
Base Case	Overall*	160 (111-210)	0.008 (0.006-0.01)	1411 (1144-1705)	0.073 (0.059-0.088)	219 (155-286)	0.011 (0.008-0.014)	549 (451-652)	0.028 (0.023-0.033)
	MARD	69 (32-106)	0.003 (0.001-0.005)	1124 (875-1395)	0.053 (0.042-0.066)	191 (117-270)	0.009 (0.006-0.013)	419 (302-530)	0.02 (0.014-0.025)
	SIRD	114 (59-170)	0.005 (0.002-0.008)	2006 (1532-2565)	0.101 (0.079-0.129)	237 (156-317)	0.011 (0.007-0.015)	759 (593-939)	0.037 (0.029-0.046)
	SIDD	392 (272-511)	0.019 (0.013-0.025)	1500 (1240-1771)	0.077 (0.064-0.091)	275 (189-356)	0.013 (0.009-0.018)	607 (483-736)	0.031 (0.025-0.038)
	MOD	190 (127-253)	0.01 (0.007-0.013)	1283 (992-1641)	0.072 (0.056-0.093)	182 (116-245)	0.01 (0.007-0.013)	543 (440-663)	0.03 (0.024-0.036)
	Range†	323	0.016	882	0.048	93	0.028	340	0.069
Scenario 1	Overall*	314 (223-409)	0.015 (0.011-0.02)	2591 (2103-3117)	0.131 (0.107-0.158)	384 (274-495)	0.019 (0.014-0.024)	1271 (1046-1481)	0.063 (0.052-0.073)
	MARD	161 (93-235)	0.007 (0.004-0.011)	1957 (1572-2382)	0.093 (0.074-0.112)	343 (228-466)	0.016 (0.011-0.022)	935 (718-1161)	0.044 (0.033-0.054)
	SIRD	231 (141-316)	0.01 (0.006-0.014)	3229 (2513-4101)	0.162 (0.126-0.205)	413 (288-532)	0.019 (0.013-0.024)	1479 (1211-1738)	0.071 (0.059-0.084)
	SIDD	698 (517-891)	0.035 (0.026-0.044)	2808 (2254-3342)	0.144 (0.12-0.171)	540 (400-703)	0.027 (0.02-0.034)	1452 (1193-1701)	0.073 (0.061-0.085)
	MOD	349 (253-460)	0.019 (0.014-0.024)	2870 (2262-3561)	0.157 (0.124-0.194)	349 (253-443)	0.018 (0.013-0.023)	1465 (1233-1688)	0.077 (0.065-0.088)
	Range†	537	0.028	1272	0.069	197	0.011	544	0.033

*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 Ahlqvist's data-driven subgroups, 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.

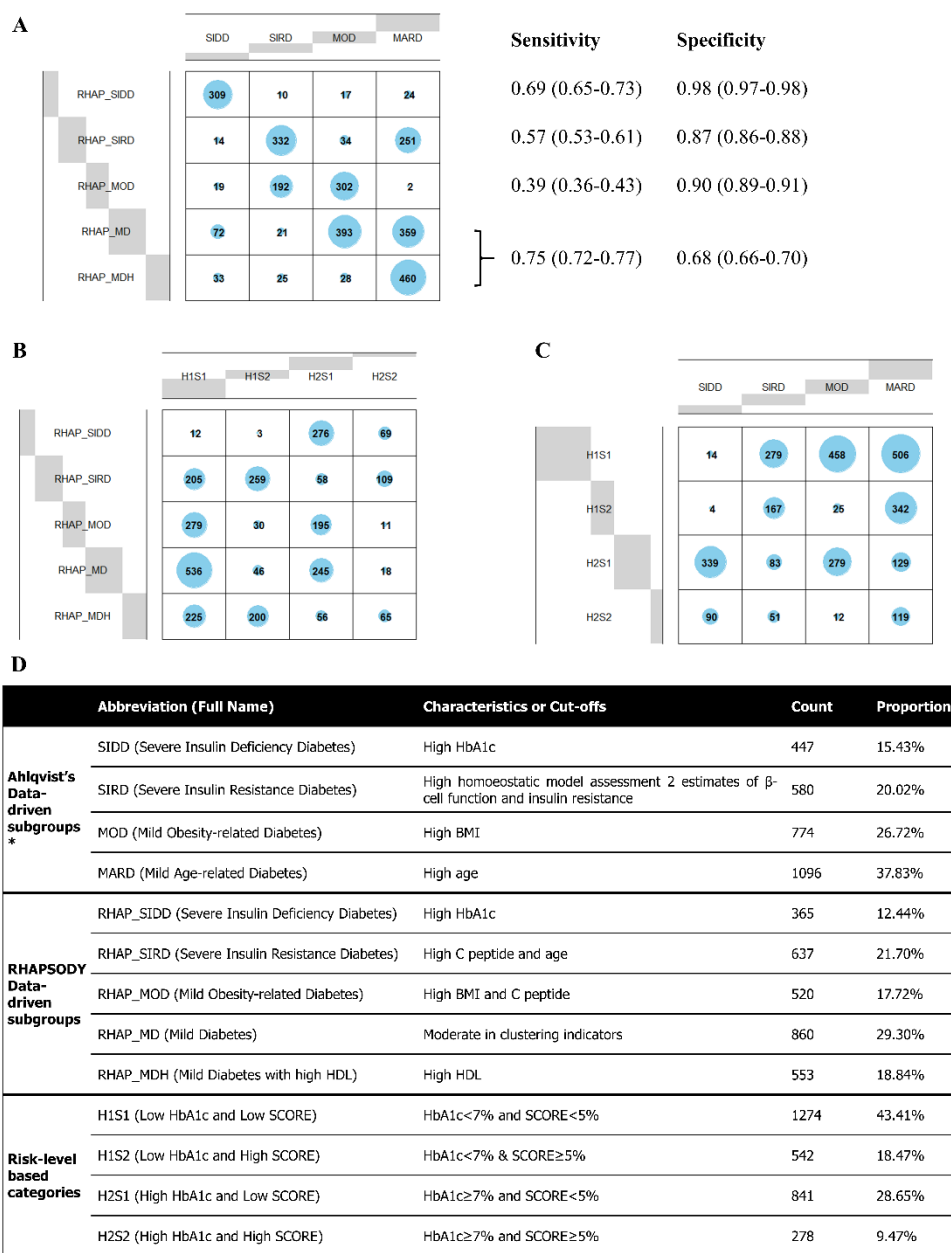
Table 10.9 Outcomes of five-year guideline-based and 5%-improvement based intensive management targeting HbA1c, BMI and LDL, and targeting only HbA1c respectively compared to care-as-usual by subgroups in U.K. setting

		Treat-to-target hypothetical intensive management				5%-Improvement-based hypothetical intensive management			
		HbA1c		HbA1c+LDL+BMI		HbA1c		HbA1c+LDL+BMI	
		Max annual price of intervention (£)	Δ QALYs vs CAU	Max annual price of intervention (£)	Δ QALYs vs CAU	Max annual price of intervention (£)	Δ QALYs vs CAU	Max annual price of intervention (£)	Δ QALYs vs CAU
Base Case	Overall*	38 (123-224)	0.007 (0.005-0.01)	316 (1271-1877)	0.069 (0.056-0.083)	51 (172-304)	0.01 (0.007-0.013)	129 (502-712)	0.026 (0.022-0.031)
	MARD	15 (8-22)	0.003 (0.001-0.005)	250 (197-309)	0.051 (0.04-0.063)	44 (30-59)	0.009 (0.005-0.012)	104 (82-126)	0.019 (0.014-0.024)
	SIRD	26 (16-37)	0.005 (0.002-0.007)	427 (338-542)	0.096 (0.075-0.122)	53 (36-70)	0.01 (0.007-0.014)	162 (131-194)	0.035 (0.028-0.044)
	SIDD	91 (67-116)	0.018 (0.013-0.023)	330 (277-382)	0.073 (0.061-0.086)	64 (47-81)	0.013 (0.009-0.016)	141 (117-163)	0.03 (0.024-0.035)
	MOD	46 (33-59)	0.01 (0.007-0.012)	310 (245-387)	0.068 (0.053-0.086)	48 (34-61)	0.009 (0.006-0.012)	132 (110-156)	0.028 (0.023-0.034)
	Range†	76	0.015	177	0.045	20	0.025	58	0.065
Scenario 1	Overall*	73 (241-428)	0.014 (0.01-0.018)	557 (2304-3329)	0.123 (0.1-0.147)	88 (297-516)	0.018 (0.013-0.022)	281 (1115-1555)	0.058 (0.048-0.067)
	MARD	37 (24-51)	0.007 (0.004-0.01)	422 (337-516)	0.087 (0.07-0.106)	76 (55-99)	0.015 (0.01-0.021)	218 (175-262)	0.041 (0.031-0.051)
	SIRD	51 (35-67)	0.01 (0.006-0.013)	664 (533-829)	0.152 (0.118-0.193)	90 (66-114)	0.018 (0.012-0.023)	304 (250-356)	0.067 (0.055-0.079)
	SIDD	161 (122-203)	0.032 (0.024-0.041)	591 (493-687)	0.135 (0.112-0.159)	122 (91-153)	0.025 (0.019-0.032)	314 (265-361)	0.068 (0.057-0.078)
	MOD	86 (65-107)	0.017 (0.012-0.022)	641 (522-780)	0.144 (0.114-0.179)	85 (62-104)	0.017 (0.012-0.021)	330 (281-376)	0.07 (0.059-0.08)
	Range†	124	0.025	242	0.065	46	0.01	112	0.029

*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 Ahlqvist's data-driven subgroups, 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).

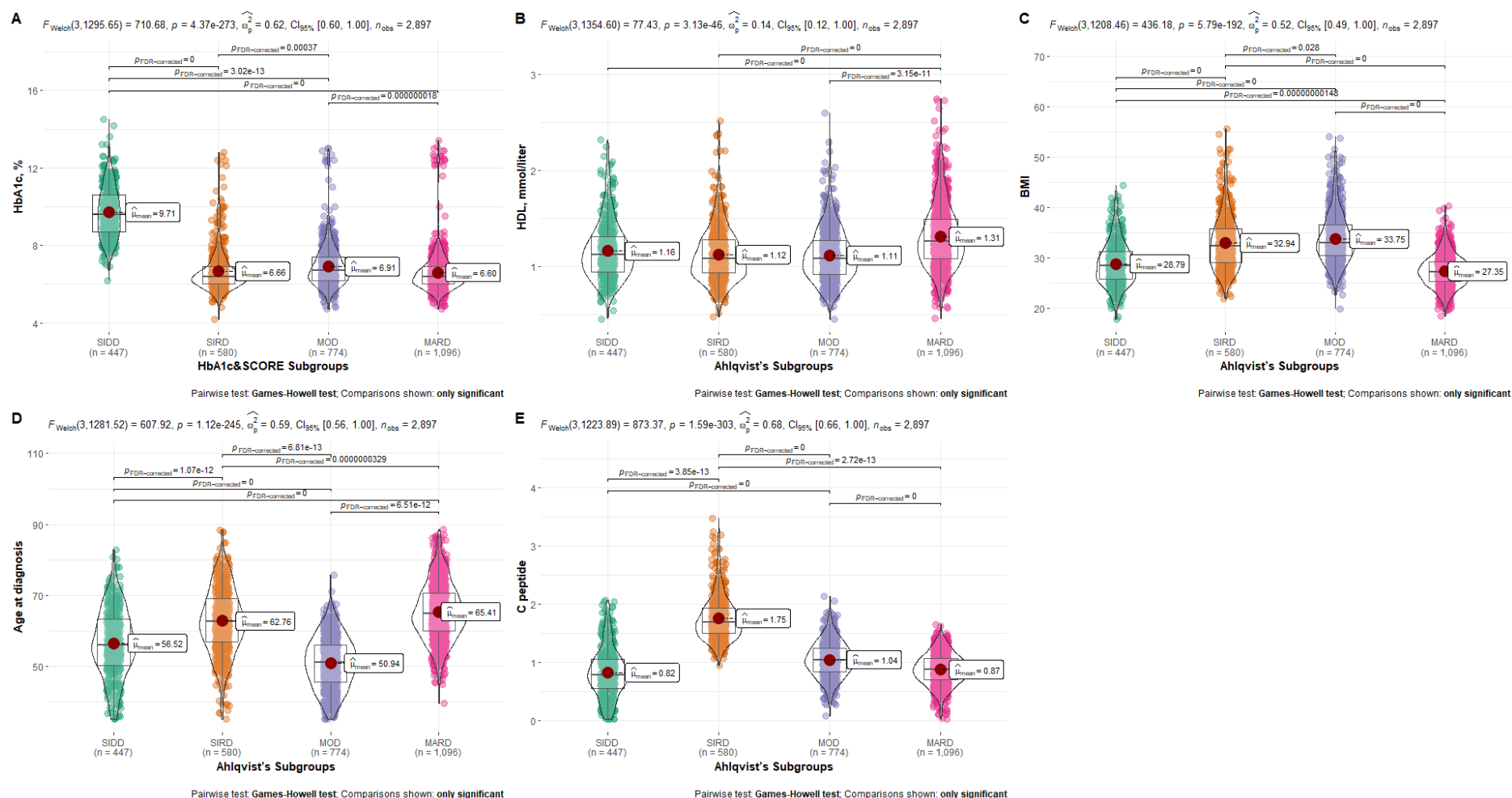
†Range is defined as the maximum – minimum of the mean maximum annual cost-effective price of intervention or incremental QALY.

Figure 10.1 Comparisons among subgroups defined by different methodology



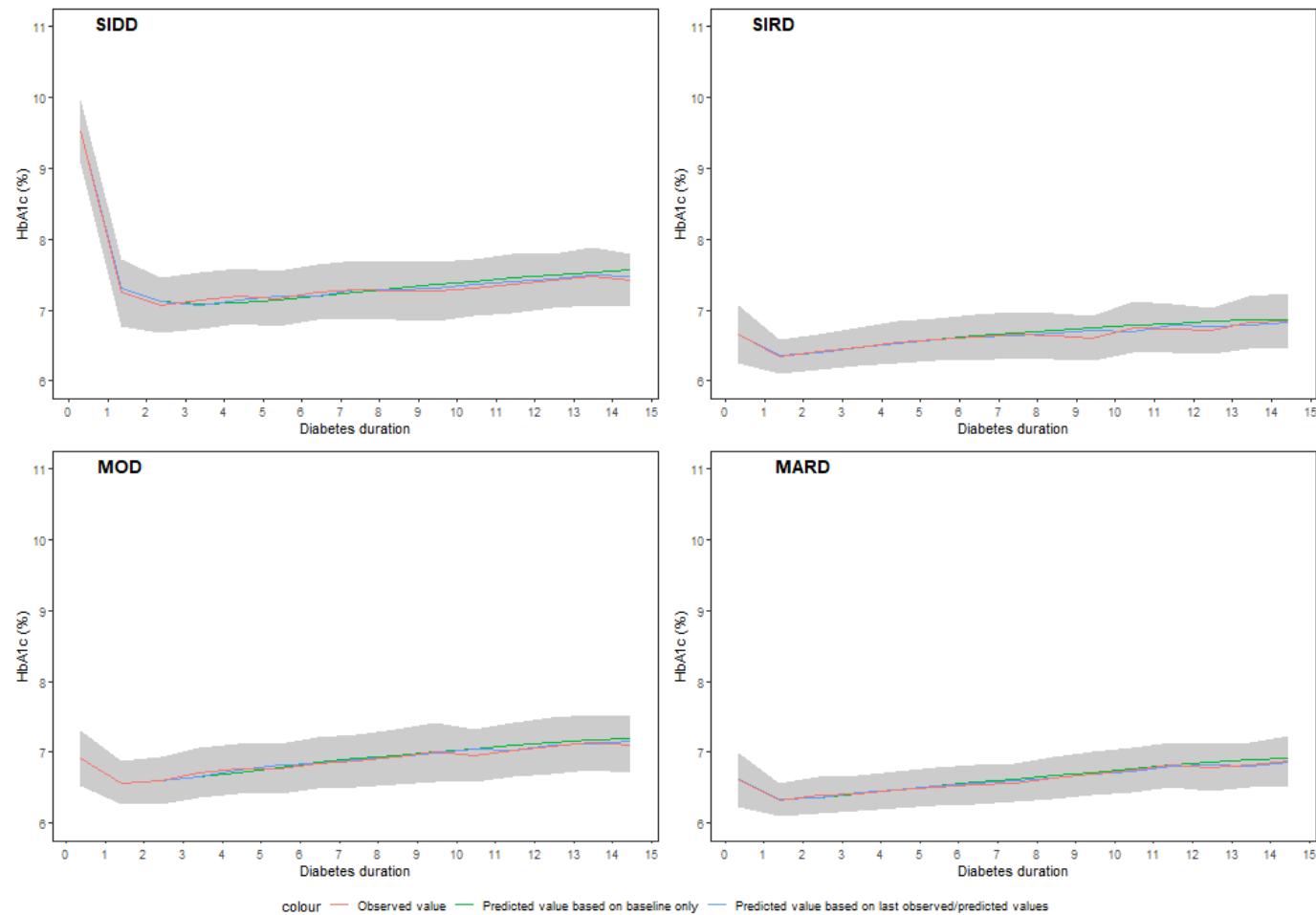
*For assigning Ahlqvist's subgroups Data-driven subgroups, there were 38 individuals less in total compared to RHAPSODY data-driven subgroups and risk-level based categories due to the missing of fasting glucose information which is required in the homeostatic model assessment 2 estimates

Figure 10.2 Violin graphs and pairwise comparisons of baseline clustering indicators of Ahlqvist's data-driven subgroups



The Games-Howell test was applied to enable pairwise comparison and p values were adjusted by using Benjamini-Hochberg correction to account for false discovery rate. P-values below 0.05 were considered statistically significant. These are the baseline characteristics at diagnosis, rather than the baseline simulation characteristics.

Figure 10.3 The graphs for goodness-of-fit over time for HbA1c of each Ahlqvist's data-driven subgroup



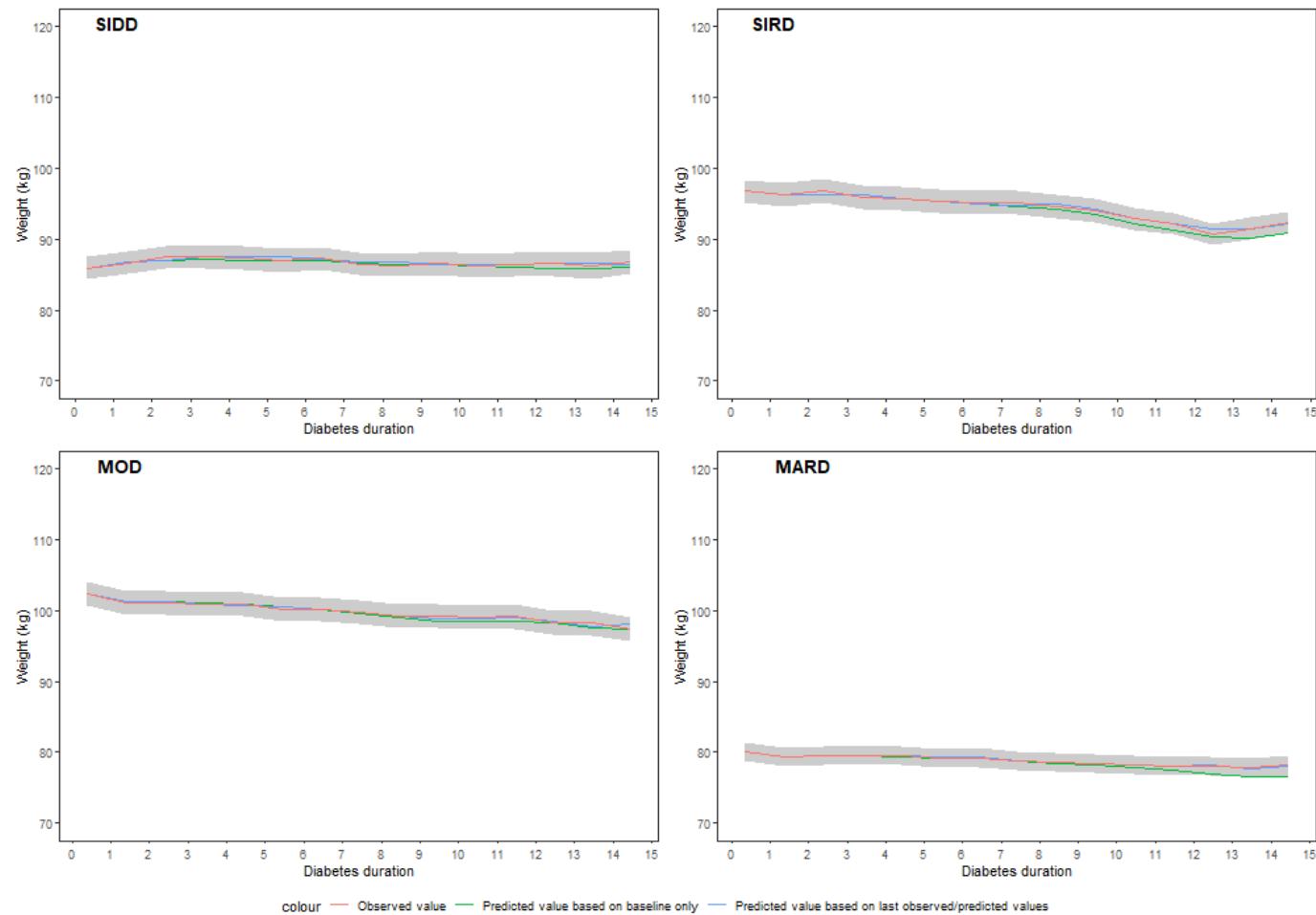
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 10.4 The graphs for goodness-of-fit over time for weight of each Ahlqvist's data-driven subgroup



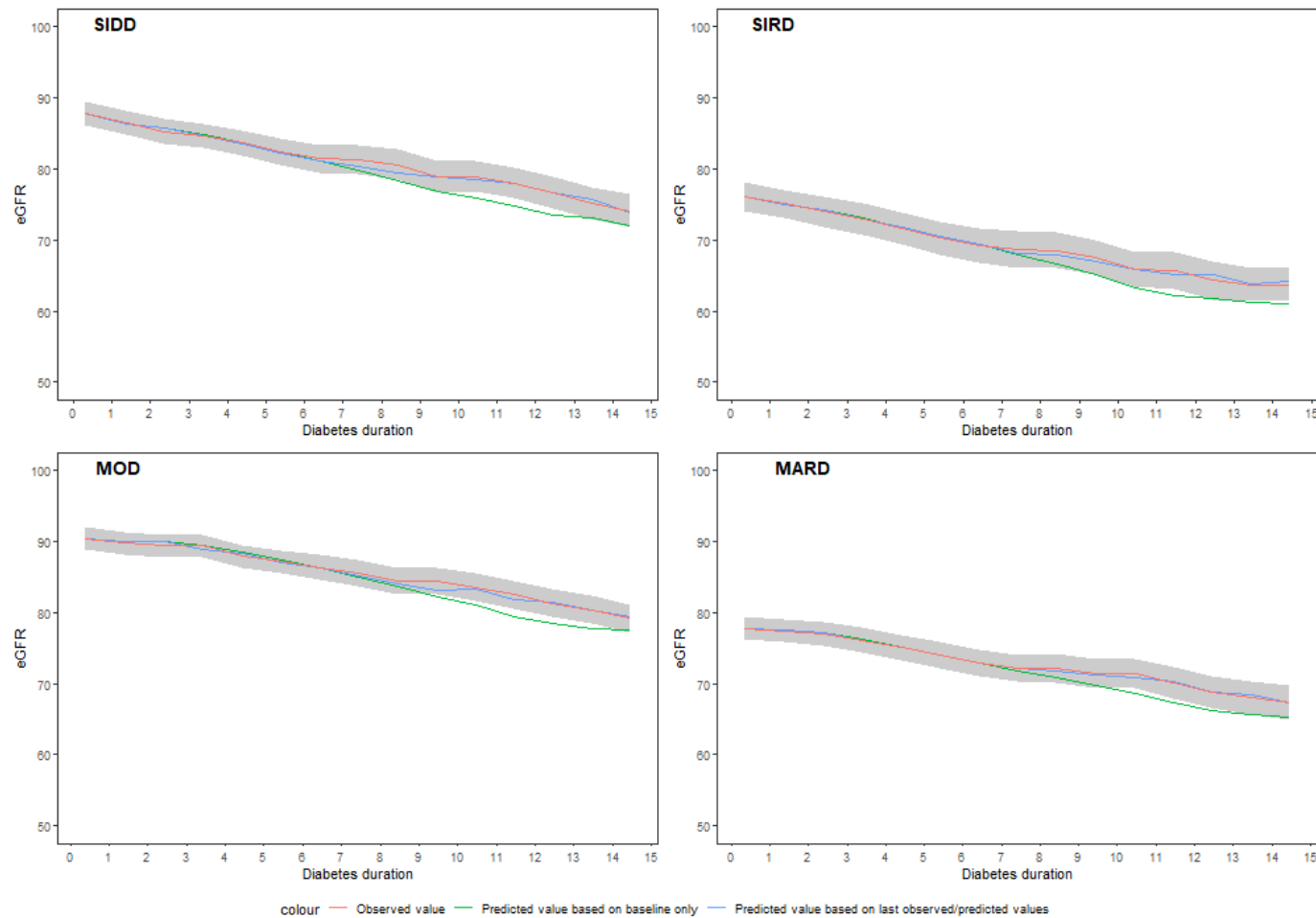
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 10.5 The graphs for goodness-of-fit over time for eGFR of each Ahlqvist's data-driven subgroup



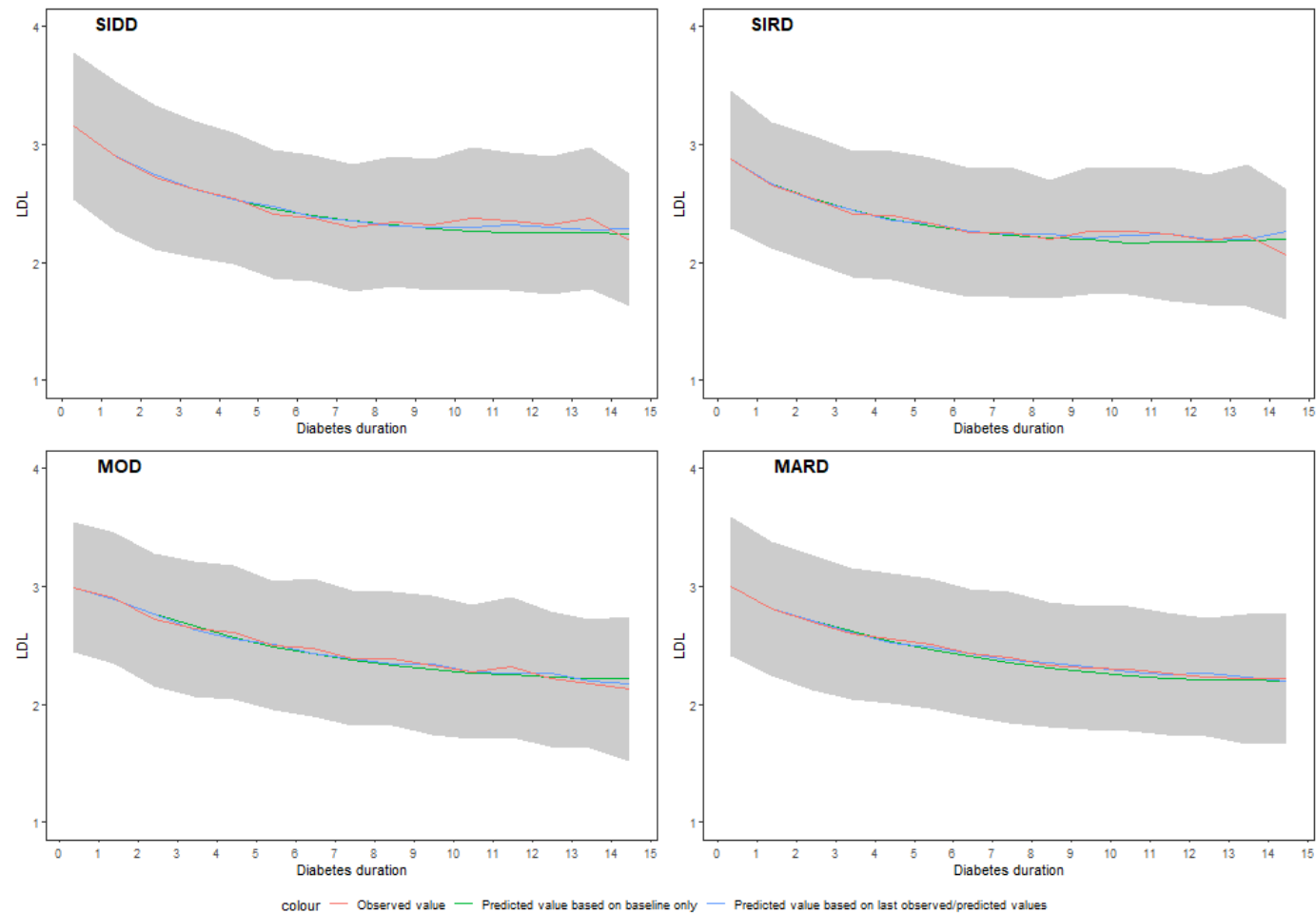
Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 10.6 The graphs for goodness-of-fit over time for LDL of each Ahlqvist's data-driven subgroup



Grey area indicated 95% confidence interval of observed data.

Observed value (red line) denotes the observational data.

Predicted value based on baseline only (green line) denotes the values extrapolated based on only baseline value.

Predicted value based on last observed/predicted (blue line) denotes the values extrapolated based on baseline value and available data observed during follow-up.

Figure 10.7 Non-standardized and standardized mean simulated lifetime QALYs and cost with 95% CI (in U.S. setting) for Ahlqvist's data-driven clustering subgroups

Ahlqvist's Data-Driven Subgroups

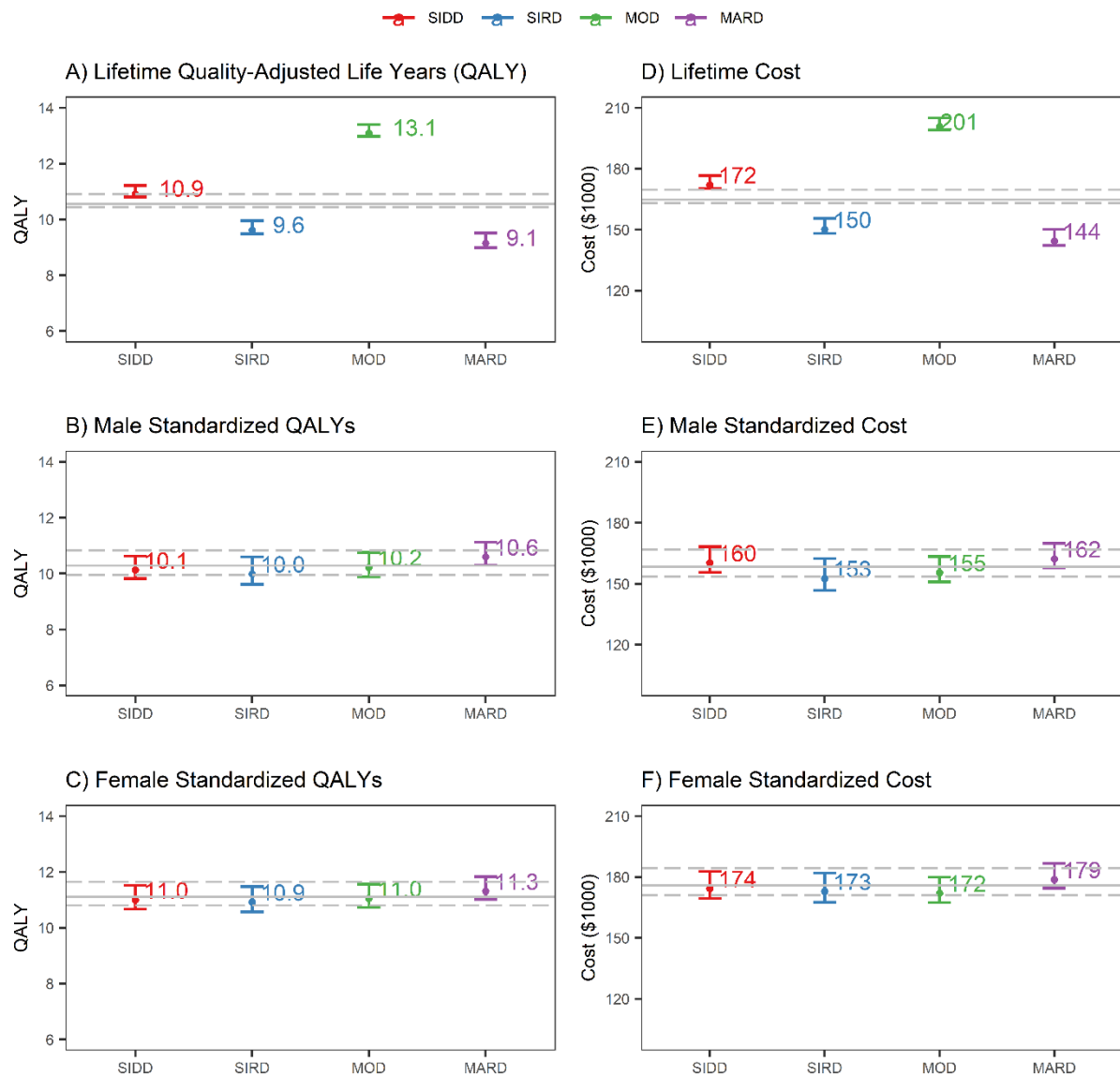


Figure 10.8 Non-standardized and standardized mean simulated lifetime QALYs and cost with 95% CI (in U.K. setting) for Ahlqvist's data-driven clustering subgroups

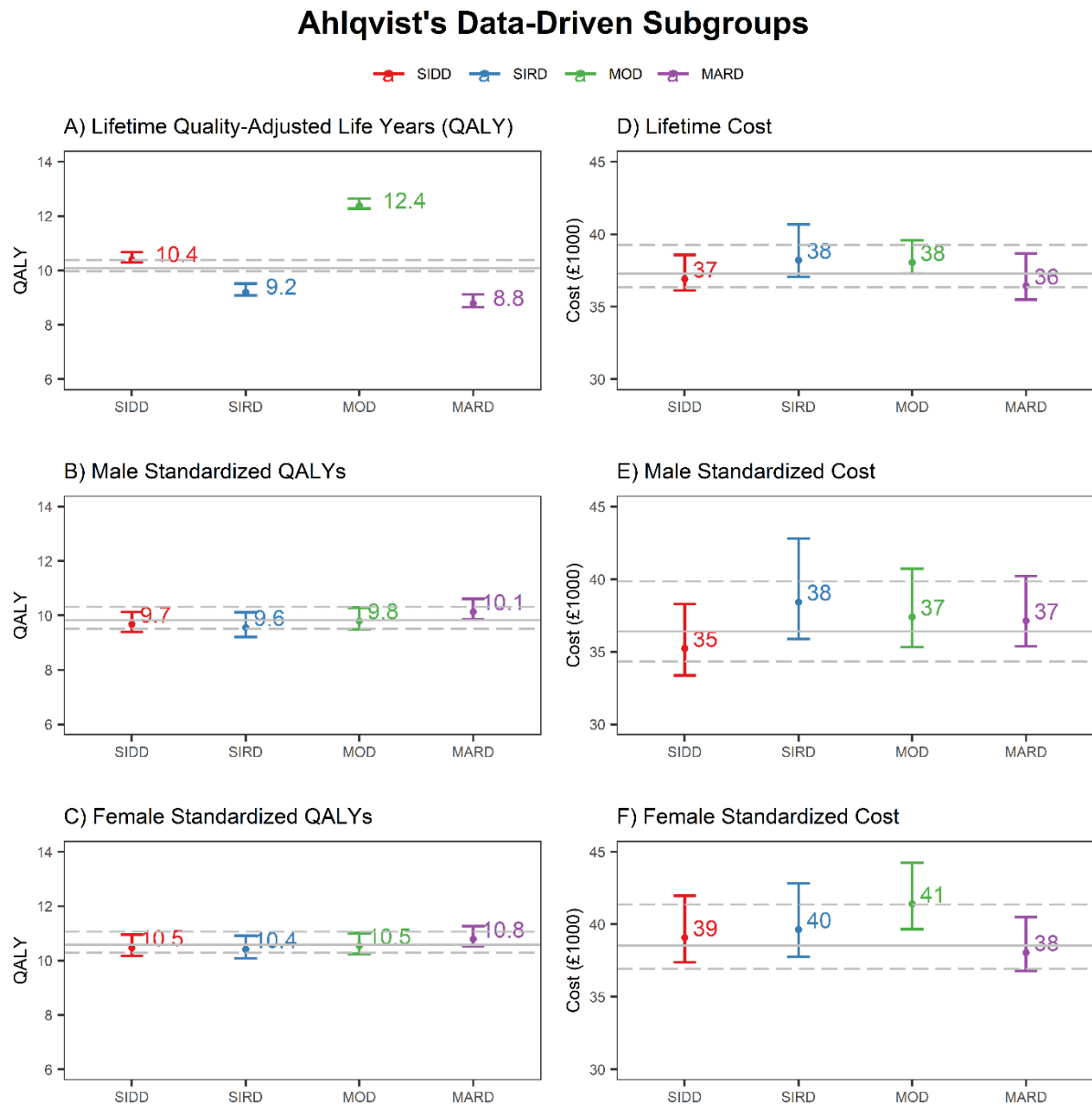
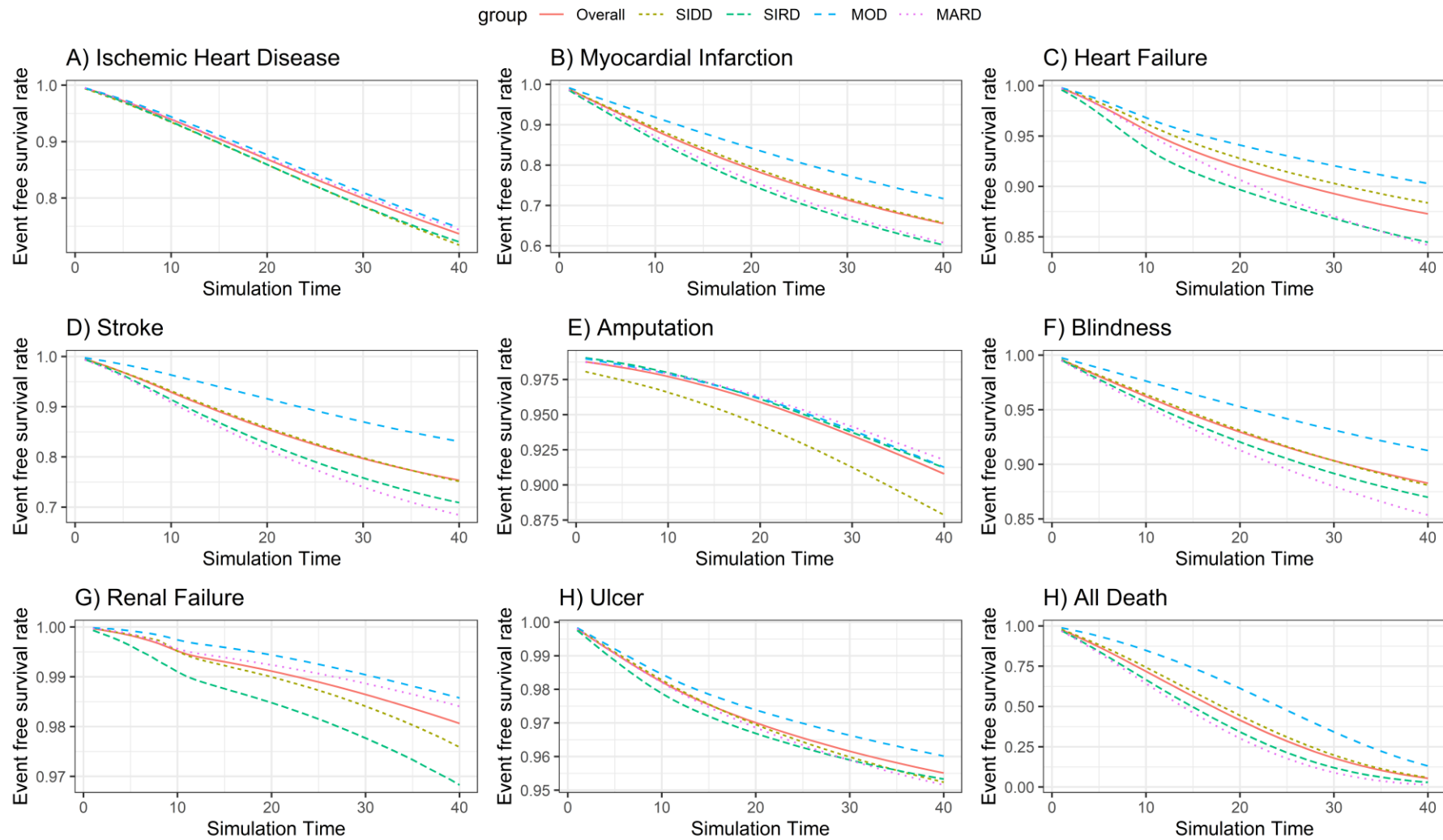
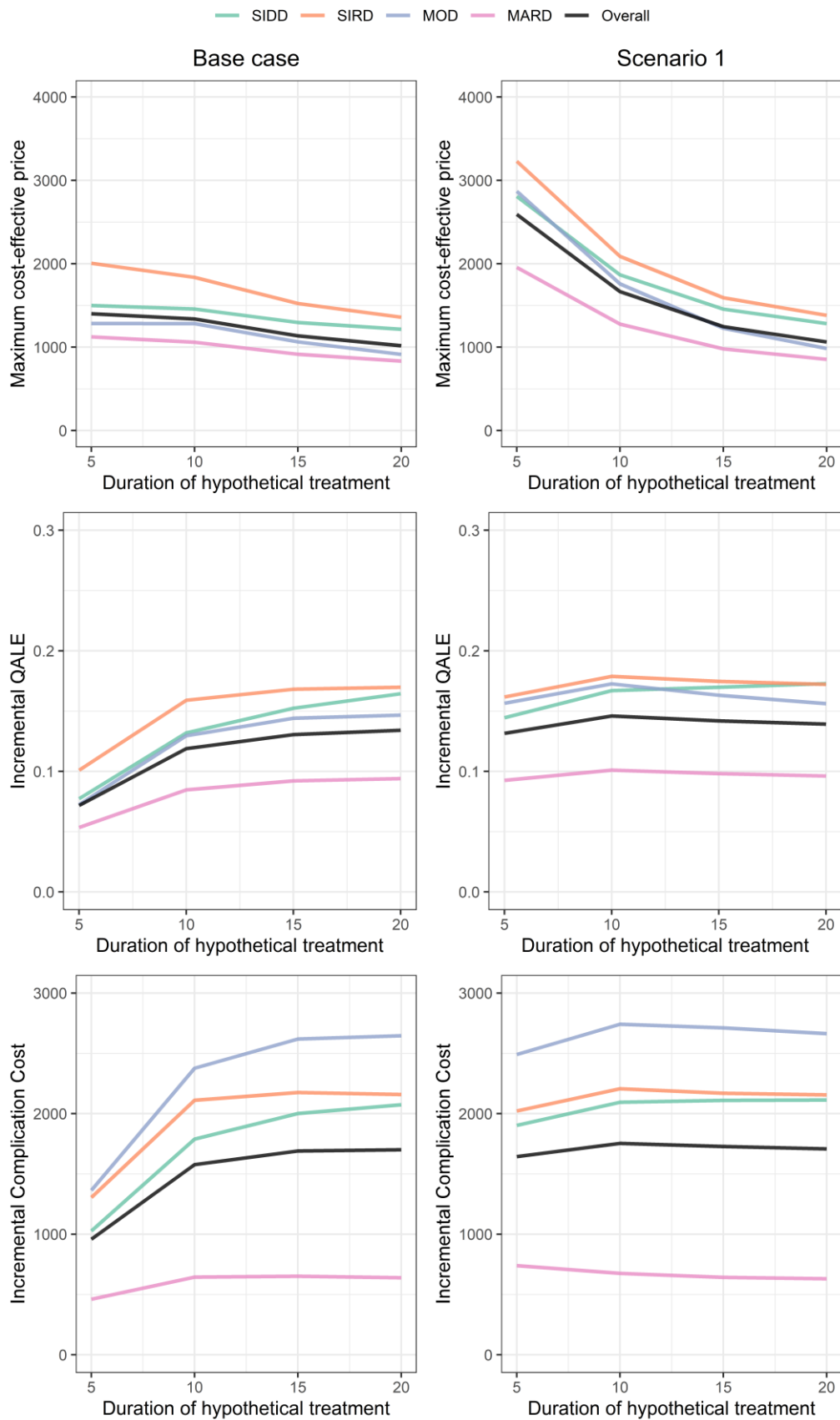


Figure 10.9 Simulated Kaplan-Meier plots of Ahlqvist's data-driven subgroups



Kaplan-Meier survival plots by year of simulation for each event are generated simulated by the model for each subgroup. Higher value indicates less risky. Blindness refers to blindness in one eye.

Figure 10.10 Maximum annual cost-effective price, incremental QALYs and complication costs of multiple-year guideline-based intensive management targeting HbA1c, BMI and LDL by Ahlqvist's data-driven subgroups (in U.S. setting), for a range of treatment durations



References

1. Slieker RC, Donnelly LA, Fitipaldi H, Bouland GA, Giordano GN, Akerlund M, Gerl MJ, Ahlqvist E, Ali A, Dragan I, Festa A, Hansen MK, Aly DM, Kim M, Kuznetsov D, Mehl F, Klose C, Simons K, Pavo I, Pullen TJ, Suvitaival T, Wretling A, Rossing P, Lyssenko V, Legido-Quigley C, Groop L, Thorens B, Franks PW, Ibberson M, Rutter GA, Beulens JWJ, 't Hart LM, Pearson ER. Replication and cross-validation of type 2 diabetes subtypes based on clinical variables: an IMI-RHAPSODY study. *Diabetologia* 2021;64:1982-1989
2. Buse JB, Wexler DJ, Tsapas A, Rossing P. 2019 Update to : Management of Hyperglycemia in Type 2 Diabetes , 2018 . A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes care* 2020;43:487-493
3. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetière P, Jousilahti P, Keil U, Njølstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmsen L, Graham IM, group obotSp. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European Heart Journal* 2003;24:987-1003
4. Alva M, Gray A, Mihaylova B, Clarke P. The effect of diabetes complications on health-related quality of life: the importance of longitudinal data to address patient heterogeneity. *Health economics* 2014;23:487-500
5. Lung TW, Hayes AJ, Hayen A, Farmer A, Clarke PM. A meta-analysis of health state valuations for people with diabetes: explaining the variation across methods and implications for economic evaluation. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2011;20:1669-1678
6. American Diabetes Association. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care* 2018;41:917-928
7. Alva ML, Gray A, Mihaylova B, Leal J, Holman RR. The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84). *Diabetic medicine : a journal of the British Diabetic Association* 2015;32:459-466
8. Ward A, Alvarez P, Vo L, Martin S. Direct medical costs of complications of diabetes in the United States: estimates for event-year and annual state costs (USD 2012). *Journal of medical economics* 2014;17:176-183
9. NHS Blood and Transplant. Factsheet 7: Cost-effectiveness of transplantation. 2009
10. Kerr M, Rayman G, Jeffcoate WJ. Cost of diabetic foot disease to the National Health Service in England. *Diabetic Medicine* 2014;31:1498-1504
11. Khan T, Tsipias S, Wozniak G. Medical Care Expenditures for Individuals with Prediabetes: The Potential Cost Savings in Reducing the Risk of Developing Diabetes. *Population health management* 2017;20:389-396
12. Herman WH, Hoerger TJ, Brandle M, Hicks K, Sorensen S, Zhang P, Hamman RF, Ackermann RT, Engelgau MM, Ratner RE. The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance. *Annals of internal medicine* 2005;142:323-332
13. Pagano E, Konings SRA, Di Cuonzo D, Rosato R, Bruno G, van der Heijden AA, Beulens J, Slieker R, Leal J, Feenstra TL. 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
14. Van Der Heijden AAWA, Rauh SP, Dekker JM, Beulens JW, Elders P, T'Hart LM, Rutters F, Van Leeuwen N, Nijpels G. 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;
15. van der Heijden AA, Rauh SP, Dekker JM, Beulens JW, Elders P, t Hart LM, Rutters F, van Leeuwen N, Nijpels G. 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
 16. M. vP. De Sims classificatie. In: Voeten en diabetes. 2018;
 17. Aldington SJ, Kohner EM, Meuer S, Klein R, Sjolie AK. Methodology for Retinal Photography and Assessment of Diabetic-Retinopathy - the Eurodiab Iddm Complications Study. *Diabetologia* 1995;38:437-444
 18. Leal J, Alva M, Gregory V, Hayes A, Mihaylova B, Gray AM, Holman RR, Clarke P. Estimating risk factor progression equations for the UKPDS Outcomes Model 2 (UKPDS 90). *Diabetic Medicine* 2021;38
 19. Lumley T. mitools: Tools for Multiple Imputation of Missing Data. R package version 2.4. . 2019;
 20. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;45:1-67
 21. Team RC. R: A language and environment for statistical computing. 2021;
 22. Ahlqvist E, Storm P, Karajamaki A, Martinell M, Dorkhan M, Carlsson A, Vikman P, Prasad RB, Aly DM, Almgren P, Wessman Y, Shaat N, Spegel P, Mulder H, Lindholm E, Melander O, Hansson O, Malmqvist U, Lernmark A, Lahti K, Forsen T, Tuomi T, Rosengren AH, Groop L. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes & Endocrinology* 2018;6:361-369