BMJ Open Diabetes Research & Care

Incident dementia risk among patients with type 2 diabetes receiving metformin versus alternative oral glucose-lowering therapy: an observational cohort study using UK primary healthcare records

William Doran,[1] Louis Tunnicliffe,[1] Rutendo Muzambi,[1] Christopher T Rentsch ,[1,2] Krishnan Bhaskaran,[1] Liam Smeeth,[1] Carol Brayne,[3] Dylan M Williams,[4] Nish Chaturvedi,[4] Sophie V Eastwood,[4] Susanna J Dunachie,[5,6] Rohini Mathur,[7] Charlotte Warren-Gash [1]

To cite: Doran W, Tunnicliffe L, Muzambi R, *et al*. Incident dementia risk among patients with type 2 diabetes receiving metformin versus alternative oral glucose-lowering therapy: an observational cohort study using UK primary healthcare records. *BMJ Open Diab Res Care* 2024;**12**:e003548. doi:10.1136/bmjdrc-2023-003548

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjdrc-2023-003548).

Received 23 May 2023
Accepted 21 December 2023

For numbered affiliations see end of article.

Correspondence to
Professor Charlotte Warren-Gash; charlotte.warren-gash1@lshtm.ac.uk

ABSTRACT

Introduction 4.2 million individuals in the UK have type 2 diabetes, a known risk factor for dementia and mild cognitive impairment (MCI). Diabetes treatment may modify this association, but existing evidence is conflicting. We therefore aimed to assess the association between metformin therapy and risk of incident all-cause dementia or MCI compared with other oral glucose-lowering therapies (GLTs).

Research design and methods We conducted an observational cohort study using the Clinical Practice Research Datalink among UK adults diagnosed with diabetes at ≥40 years between 1990 and 2019. We used an active comparator new user design to compare risks of dementia and MCI among individuals initially prescribed metformin versus an alternative oral GLT using Cox proportional hazards regression controlling for sociodemographic, lifestyle and clinical confounders. We assessed for interaction by age and sex. Sensitivity analyses included an as-treated analysis to mitigate potential exposure misclassification.

Results We included 211 396 individuals (median age 63 years; 42.8% female), of whom 179 333 (84.8%) initiated on metformin therapy. Over median follow-up of 5.4 years, metformin use was associated with a lower risk of dementia (adjusted HR (aHR) 0.86 (95% CI 0.79 to 0.94)) and MCI (aHR 0.92 (95% CI 0.86 to 0.99)). Metformin users aged under 80 years had a lower dementia risk (aHR 0.77 (95% CI 0.68 to 0.85)), which was not observed for those aged ≥80 years (aHR 0.95 (95% CI 0.87 to 1.05)). There was no interaction with sex. The as-treated analysis showed a reduced effect size compared with the main analysis (aHR 0.90 (95% CI 0.83 to 0.98)).

Conclusions Metformin use was associated with lower risks of incident dementia and MCI compared with alternative GLT among UK adults with diabetes. While our findings are consistent with a neuroprotective effect of metformin against dementia, further research is needed to reduce risks of confounding by indication and assess causality.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Diabetes is a risk factor for dementia in later life with an estimated population attributable fraction of 1%.
⇒ Evidence that metformin therapy reduces the risk of dementia in diabetes is disputed.

WHAT THIS STUDY ADDS

⇒ We found that metformin users experienced a lower risk of incident all-cause dementia and mild cognitive impairment than users of other oral glucose-lowering therapies (GLTs) in a large, unselected UK primary care population.
⇒ This risk reduction was seen only in individuals aged under 80 years, which is consistent with previous studies. Our findings were robust across a broad range of sensitivity analyses, but risk reduction attenuated when we applied additional control for exposure misclassification.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study highlights a need for further research including randomized controlled trials and instrumental variable analyses to strengthen causal inference.
⇒ The development of dementia prevention strategies is a key concern for current policymakers: optimizing GLT for individuals with diabetes may help to improve brain health.

INTRODUCTION

4.2 million individuals are living in the UK with type 2 diabetes mellitus (hereafter referred to as 'diabetes') and prevalence is increasing due to underlying trends in physical inactivity, obesity and population aging.[1] Diabetes is an established dementia risk factor through

mechanisms including insulin resistance, chronic neuroinflammation and microvascular dysfunction.[2 3] Globally, dementia already costs US$1 trillion annually and the number of individuals living with dementia in the UK is expected to increase from 850 000 individuals to 1.2 million by 2040.[2] Low and middle-income countries are reporting rapid increases in dementia incidence consistent with their demographic transition, but high-income countries such as the UK already have a high burden of dementia and multimorbidity in line with their older population structures.[2] There is an urgent need for dementia prevention strategies, and considering risk factor modification of diabetes is highly relevant to all socioeconomic settings.

Individuals with diabetes are 1.5–2 times more likely to be diagnosed with a cognitive disorder including mild cognitive impairment (MCI) or dementia.[3] Dementia risk appears to increase with duration of diabetes and poor glycemic control, but there is inconsistent evidence on whether this is modifiable by glucose-lowering therapy (GLT) such as metformin. Metformin, a biguanide in widespread global use, has been recommended as first-line treatment for diabetes by the National Institute for Health and Care Excellence (NICE) since 2002.[4 5] Prior to this, alternative GLTs such as sulfonylureas were frequently used as first-line treatment. Traditionally, metformin's principal mechanisms were understood to be suppression of hepatic gluconeogenesis and enhancement of peripheral insulin sensitivity via AMPK activation.[6] However, it also exhibits pleiotropic effects including modifying the gut microbiome, immune function and inflammatory mechanisms, which could plausibly modify the pathogenesis of dementia in diabetes.[7 8]

The association between metformin therapy and incident dementia in patients with diabetes is contested: two recent meta-analyses found evidence that metformin therapy was associated with lower incident dementia risk, but a third reported a pooled null effect.[9–11] Many of the included studies were prone to bias due to the use of cross-sectional designs and lack of active comparators. Subsequent to these meta-analyses, Newby *et al* estimated a 20% lower incident dementia risk associated with metformin use using an active comparator new user design in US health records.[12] Here, we aimed to investigate the association between metformin therapy and incident all-cause dementia and MCI using primary care records from the UK Clinical Practice Research Datalink (CPRD).[13]

RESEARCH DESIGN AND METHODS

CPRD Gold includes longitudinal primary care records from practices using Vision software covering a large, unselected study population, which is demographically representative of the UK.[13] It contains details of symptoms, coded diagnoses, test results, referrals and drug prescriptions. It also includes well-established linkages to Office for National Statistics mortality data, Index of Multiple Deprivation (IMD) and Hospital Episode Statistics (HES) Admitted Patient Care data.[13]

We compared new users of metformin with new users of alternative GLTs in an active comparator new user (ACNU) design. This aims to emulate the advantages of randomized controlled trials in observational settings and mitigates potential confounding by indication and healthy user bias. It excludes individuals without an indication for treatment or with important contraindications including frailty. It also ensures that participants are aligned at a common time point (initiation of treatment).[14]

Source population and cohort identification

The study population included adults who were registered at a CPRD-eligible general practice between January 1, 1990 and December 31, 2019 with a new diagnosis of diabetes at ≥40 years old recorded at least 12 months after registration. Eligible individuals required a record of GLT with first prescription on or at any time after the diabetes diagnosis date and no history of dementia or MCI at first prescription (index date) (online supplemental methods and figure 1).

Follow-up

The index date for follow-up was the first prescription date for GLT. For the dementia outcome, follow-up continued until an instance of death, dementia, CPRD de-registration or December 31, 2019, whichever was first. For MCI, follow-up continued until an instance of death, dementia, MCI, CPRD de-registration or December 31, 2019, whichever was first; this ensured non-sensical diagnoses of MCI made after diagnoses of dementia were discounted. Follow-up was restricted to the end of 2019 to exclude unknown impacts of the SARS-CoV-2 pandemic.

Exposure and outcome

The exposure was defined as the first CPRD-recorded prescription of metformin or alternative GLT among individuals with no prior record of either. Eligible GLT included any formulation of: metformin, alpha-glucosidase inhibitors, sulfonylureas, dipeptidyl peptidase 4 inhibitors, glucagon-like peptide-1 analogs, thiazolidinediones, sodium-glucose cotransporter 2 (SGLT-2) inhibitors and metaglinides. We applied a 'first treatment carried forward' approach, meaning that individuals who first began metformin or alternative GLT remained in their original exposure group regardless of treatment non-adherence, intentional cessation or commencement of any additional GLT (we found no instances of metformin therapy and alternative GLT being initiated simultaneously).

All-cause dementia (primary outcome) and MCI (secondary outcome) were defined using extensive, previously established Read code lists based on clinical diagnoses (dementia) or diagnoses and symptoms (MCI) recorded in primary care. All-cause dementia included all known subtypes such as Alzheimer's disease, vascular

dementia, Lewy body dementia as well as undifferentiated dementia. International Classification of Diseases Tenth Revision (ICD-10) codes from hospital records were additionally used to identify dementia outcomes in a sensitivity analysis restricted to individuals with linked data only. In this analysis, the dementia date was taken from the earliest record in either CPRD or HES.

Covariates

Using year of birth, we assumed a July 2 birth date for all participants. Individual-level and general practitioner (GP)-level IMDs were defined as quintiles obtained from Office for National Statistics linkage. Self-reported ethnicity was classified into five categories: white, South Asian, black, mixed or other and unknown. Smoking status was classified as current, former, and never. Body mass index was taken from the measurement closest to the index date in the 12 months prior or 3 months after and was categorized according to the WHO cut-offs. 'Baseline' hemoglobin A1c (HbA1c) in International Federation of Clinical Chemistry units was measured within 6 months prior to the index date and classified into broad 20 mmol/mol categories.

21 additional covariates were collected that describe participants' health status prior to the index date and reflect known or hypothesized risk factors for dementia. These included: statin use, antihypertensive use, hypertension, asthma, chronic obstructive pulmonary disease, liver disease, coronary heart disease, peripheral vascular disease, stroke, diabetic retinopathy, neuropathy, brain injury, depression, autoimmune disease, chronic kidney disease, heart failure, alcohol excess, skin and soft tissue infection, urinary tract infection, lower respiratory tract infection and sepsis. Code lists are available on London School of Hygiene & Tropical Medicine Data Compass: https://datacompass.lshtm.ac.uk/id/eprint/3402/.

Statistical analysis

Analysis was completed using Stata/SE V.17 (Statacorp). Baseline characteristics were described overall and by exposure status with frequency counts and percentages apart from age, which was described with the median value and IQR. We hypothesised that missing data were likely missing not at random, meaning that multiple imputation was unsuitable.[15] There were 115 092 (54.4%) missing entries for individual-level IMD, which were replaced with GP-level IMD. Three other covariates had high proportions of missingness: ethnicity (50.9%), baseline HbA1c (16.9%) and smoking status (13.8%). Crude rates of incident dementia and MCI were calculated overall and by age and calendar time periods with 95% CIs estimated according to the Poisson distribution.

We conducted a single-failure survival analysis using Cox regression with fixed effects and age as the underlying time scale for dementia and MCI. Age-adjusted, minimally adjusted (adjusted for age as time scale plus sex and calendar time) and fully adjusted models (all covariates) were fitted. The fully adjusted models were fitted using a backwards approach, in which all baseline covariates were considered as potential confounders, informed by our directed acyclic graph (online supplemental figure 2). Covariates with >20% missingness (ethnicity) were excluded. Lexis expansion was used to generate 5-year calendar time bands, which were used to control for secular effects. The three earliest time bands were merged after encountering event sparsity. We calculated root mean-squared error (RMSE) to indicate multicollinearity and bias. If this increased between the minimally adjusted and fully adjusted model for the primary outcome, covariates were individually dropped from repeated regressions to identify which could be excluded to achieve the largest reduction in RMSE. This model reduction process was repeated until the RMSE of the fully adjusted model was lower than that of the minimally adjusted model. The regression models for MCI were specified with the same covariates as for dementia to ensure that estimates were comparable. The Schoenfeld residuals test and log-log Kaplan-Meier survival plots were used to evaluate evidence for non-proportional hazards.

We decided a priori to assess for potential interaction by age and sex as there is plausible evidence to support this.[16 17] We specified stratified models and used the likelihood ratio test (LRT) to assess evidence for interaction. We undertook multiple sensitivity analyses to assess the robustness of our findings. To minimize misclassification of pre-existing dementia as incident dementia, we restricted follow-up to exclude dementia diagnoses up to 2 years after the index date. Then, we performed subgroup analysis for individuals started on treatment ≥2004 and ≥2012. CPRD data quality meaningfully improved after the advent of the NICE Quality of Outcomes Framework (QOF) in 2004[18] and the first SGLT-2 inhibitor (dapagliflozin) was approved for UK use in 2012. We restricted our analysis to individuals with HES linkage as we hypothesized they may have improved data quality for covariates and dementia.[19] After this, we performed a post-hoc 'as-treated' analysis where participants were censored 30 days (a typical medication supply) after their last prescription of metformin or non-metformin GLT. Finally, we applied serial restrictions on calendar period to assess whether spurious calendar effects were biasing effect estimates.

RESULTS

There were 211 396 eligible individuals with a median follow-up of 5.4 years. Of these, 179 333 (84.8%) initiated treatment with metformin and had a median follow-up of 5.2 years, while 32 063 (15.2%) initiated alternative GLT and had a median follow-up of 6.9 years. The proportion of metformin initiators increased steadily over the study period from 30.0% of participants between 1990 and 1994 to 90.4% of participants by 2015–2019. 90.3% of those initiating an alternative GLT were prescribed a sulfonylurea (figure 1). We noted that 51.2% of individuals initiated on metformin were later prescribed an

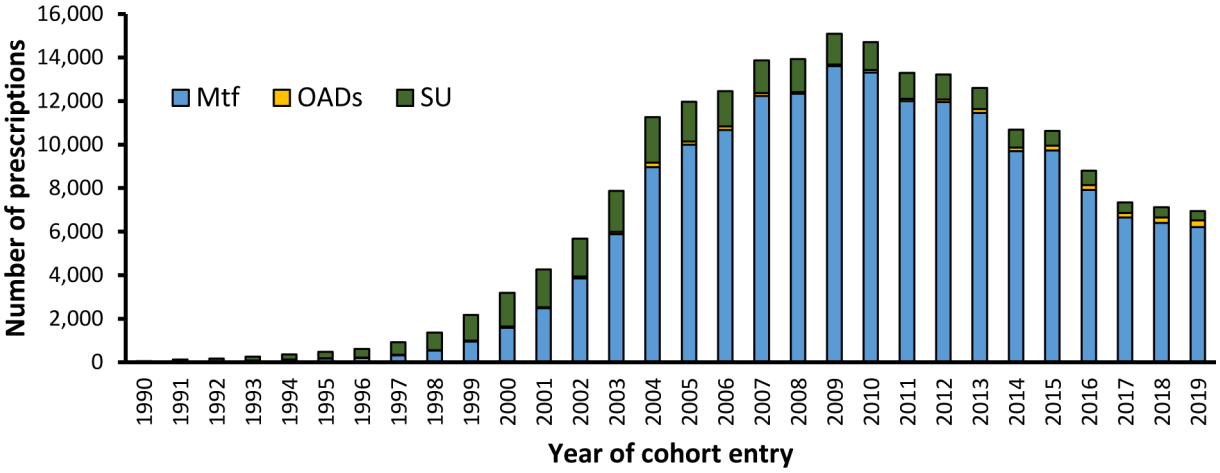


Figure 1 Distribution of first prescriptions of oral GLT over time. NICE guidance changed in 2002 to recommend metformin as first-line treatment for diabetes. GLT, glucose-lowering therapy; Mtf, metformin; NICE, National Institute for Health and Care Excellence; OADs, other anti-diabetic drugs; SU, sulfonylurea.

alternative GLT and 69.0% of individuals initiated on alternative GLT were later prescribed metformin.

Characteristics of baseline population

The cohort was majority male (57.2%), with a median age of 63 years at baseline (IQR: 54–71 years). Metformin users were younger (median age 62 years vs 66 years), more likely to be overweight/obese and to be prescribed statins (65.8% vs 48.4%) or antihypertensives (69.1% vs 65.2%) than those on alternative GLTs. They also had lower baseline HbA1c measurements (median 63.9 mmol/mol, IQR 55.2–80.3) compared with alternative GLT users (median HbA1c 72.7 mmol/mol, IQR 58.5–96.7). Alternative GLT users were more likely to have a history of coronary heart disease, peripheral vascular disease and heart failure (table 1).

Description of the outcomes

We observed 6642 diagnoses of dementia and 10 804 diagnoses of MCI during follow-up. The overall crude incidence rates were 5.0 per 1000 person-years for dementia and 7.6 per 1000 person-years for MCI. Incidence rates of dementia and MCI increased with age: the dementia incidence rate was 0.05 (95% CI 0.02 to 0.11) per 1000 person-years among individuals 40–49 years old, but 34.18 (95% CI 31.36 to 37.25) per 1000 person-years in those ≥90 years old. Dementia and MCI incidence rates also increased across successive calendar periods. Dementia incidence was 1.20 (95% CI 1.01 to 1.43) per 1000 person-years in 1990–2004, but 7.05 (95% CI 6.79 to 7.31) between 2015 and 2019. This pattern appeared preserved when conditioning on age (online supplemental table 1).

Association of new users of metformin versus alternatives with all-cause dementia

The age-adjusted model demonstrated no evidence of association between metformin use and dementia with HR 1.01 (95% CI 0.96 to 1.07). The minimally adjusted model, which also included sex and calendar time, estimated that metformin users experienced a lower risk of incident all-cause dementia compared with alternative GLT users with a best estimate of HR 0.83 (95% CI 0.79 to 0.88). The fully adjusted model, which included all available confounders apart from ethnicity (due to missingness), gave a similar best estimate of HR 0.87 (95% CI 0.79 to 0.94, n=146 883). Covariate missingness meant 64 513 individuals were excluded from the fully adjusted estimate. RMSE estimates were minimized in the fully adjusted model and no model reduction was required (table 2 and online supplemental table 2).

There was evidence of interaction between GLT use and age (LRT p=0.03), but no evidence of interaction with sex (LRT p=0.37). Metformin users under 80 years experienced a lower risk of incident dementia with HR 0.78 (95% CI 0.68 to 0.89), but there was no evidence of a risk reduction in older individuals with HR 0.93 (95% CI 0.83 to 1.03) (table 3). For the primary outcome, model checking with Schoenfeld's residuals showed borderline evidence of non-proportional hazards arising from the exposure with p=0.05 without specifying an age interaction. The log-log Kaplan-Meier survival plot demonstrated converging hazards over time (online supplemental figure 3). After specifying the age interaction as above, repeat calculation of Schoenfeld's residuals yielded a null p=0.70.

Association of new users of metformin versus alternatives with MCI

The age-adjusted model showed no evidence of association between metformin use and incident MCI with HR 1.00 (95% CI 0.95 to 1.05). The minimally adjusted model indicated that metformin users experienced a lower risk of incident MCI (although more modest than for dementia) with a best estimate of HR 0.92 (95% CI 0.87 to 0.96). The fully adjusted model, conditioning on all other available confounders, gave an almost unchanged best estimate of HR 0.92 (95% CI 0.86 to 0.99, n=146 883). This MCI model was specified with the same

Table 1 Baseline characteristics of cohort overall and by GLT type

Variable		Overall N (%$_n$)	Metformin N (%$_n$)	Other GLTs N (%$_n$)
Total		211 396 (100.0)	179 333 (84.8)	32 063 (15.2)
Age at entry	Median (IQR)	63 (54–71)	62 (54–71)	66 (56–75)
Sex	Male	120 902 (57.2)	102 127 (56.9)	18 775 (58.6)
	Female	90 494 (42.8)	77 206 (43.1)	13 288 (41.4)
Ethnicity	White	94 470 (91.0)	81 707 (91.0)	12 763 (91.1)
	South Asian	5352 (5.2)	4677 (5.2)	675 (4.8)
	Black	2253 (2.2)	1921 (2.1)	332 (2.4)
	Mixed or other	1744 (1.7)	1511 (1.7)	233 (1.7)
	Missing	107 577 (50.9*)	89 517 (49.9*)	18 060 (56.3*)
Index of Multiple Deprivation	1 (poorest)	35 512 (16.8)	29 933 (16.7)	5579 (17.4)
	2	37 817 (17.9)	31 624 (17.6)	6193 (19.3)
	3	43 983 (20.8)	37 318 (20.8)	6665 (20.8)
	4	45 679 (21.6)	39 124 (21.8)	6555 (20.4)
	5 (wealthiest)	48 405 (22.9)	41 334 (23.0)	7071 (22.1)
Baseline HbA1c (mmol/mol)	<48	13 120 (7.5)	11 353 (7.3)	1767 (8.3)
	48–67.9	86 038 (49.0)	78 682 (50.9)	7356 (34.7)
	68–87.9	42 042 (23.9)	36 852 (23.9)	5190 (24.5)
	88–107.9	21 153 (12.0)	17 591 (11.4)	3562 (16.8)
	>108	13 374 (7.6)	10 036 (6.5)	3338 (15.7)
	Missing	35 669 (16.9*)	24 819 (13.8*)	10 850 (33.8*)
Body mass index (BMI) (kg/m^2)†	Underweight	566 (0.3)	237 (0.1)	329 (1.2)
	Normal weight	19 226 (10.2)	12 334 (7.6)	6892 (26.1)
	Overweight	58 922 (31.3)	48 770 (30.1)	10 152 (38.5)
	Obesity class I	58 380 (31.0)	52 832 (32.6)	5548 (21.0)
	Obesity class II, III	51 206 (27.2)	47 756 (29.5)	3450 (13.1)
	Missing	23 096 (10.9*)	17 404 (9.7*)	5692 (17.8*)
Smoking status	Never	80 894 (44.3)	70 059 (44.3)	10 835 (44.9)
	Current	33 907 (18.6)	29 223 (18.5)	4684 (19.4)
	Ex	67 517 (37.0)	58 894 (37.2)	8623 (35.7)
	Missing	29 078 (13.8*)	21 157 (11.8*)	7921 (24.7*)
Alcohol misuse		4143 (2.0)	3522 (2.0)	621 (1.9)
Statin		133 510 (63.2)	117 997 (65.8)	15 513 (48.4)
Antihypertensive		144 892 (68.5)	123 982 (69.1)	20 910 (65.2)
Hypertension		107 618 (50.9)	92 783 (51.7)	14 835 (46.3)
Asthma		33 733 (16.0)	29 250 (16.3)	4483 (14.0)
COPD		23 362 (11.1)	19 466 (10.9)	3896 (12.2)
Liver disease		5032 (2.4)	4296 (2.4)	736 (2.3)
Coronary heart disease		34 187 (16.2)	27 841 (15.5)	6346 (19.8)
Peripheral vascular disease		4322 (2.0)	3430 (1.9)	892 (2.8)
Stroke		768 (0.4)	628 (0.4)	140 (0.4)
Diabetic retinopathy		12 797 (6.1)	11 098 (6.2)	1699 (5.3)
Neuropathy		4560 (2.2)	3866 (2.2)	694 (2.2)
Brain injury		2044 (1.0)	1780 (1.0)	264 (0.8)
Depression		56 270 (26.6)	49 195 (27.4)	7075 (22.1)

Continued

Table 1 Continued

Variable	Overall N (%n)	Metformin N (%n)	Other GLTs N (%n)
Autoimmune disease	16 871 (8.0)	13 498 (7.5)	3373 (10.5)
Chronic kidney disease	894 (0.4)	332 (0.2)	562 (1.8)
Heart failure	17 806 (8.4)	13 526 (7.5)	4280 (13.3)
Skin and soft tissue infection	42 387 (20.1)	36 845 (20.5)	5542 (17.3)
Urinary tract infection	36 574 (17.3)	31 081 (17.3)	5493 (17.1)
Lower respiratory tract infection	83 315 (39.4)	71 633 (39.9)	11 682 (36.4)
Sepsis	1964 (0.9)	1550 (0.9)	414 (1.3)

%n refers to the % among those without missing data
*% of total.
†BMI category definitions (kg/m^2): underweight <18.5, normal weight 18.5–24.9, overweight 25–29.9, obesity class I 30–34.9, obesity class II and III ≥35.
COPD, chronic obstructive pulmonary disease; GLT, glucose-lowering therapy; HbA1c, hemoglobin A1c; %n, % of non-missing.

covariates as for dementia (to ensure comparability) and RMSE estimates increased from RMSE 0.022 in the minimally adjusted model to RMSE 0.032 in the fully adjusted model (table 2 and online supplemental table 3).

For MCI, we also found evidence of interaction between GLT use and age, but no evidence of interaction with sex: we estimated a lower incident MCI risk among metformin users under 80 years with HR 0.83 (95% CI

Table 2 Associations of metformin versus alternative GLT with dementia and MCI

Exposure	N persons	N events	HR (95% CI)	SE	RMSE	P value
Dementia						
Age-adjusted						
Other GLTs	32 063	1650	1 (ref)		0.162	
Mtf	179 333	4992	1.01 (0.96 to 1.07)	0.029		0.653
Minimally adjusted (adjusted by age, sex and calendar time)						
Other GLTs	32 063	1650	1 (ref)		0.046	
Mtf	179 333	4992	0.83 (0.79 to 0.88)	0.024		<0.001
Fully adjusted*						
Other GLTs	16 547	713	1 (ref)		0.044	
Mtf	130 336	3282	0.87 (0.79 to 0.94)	0.038		0.001
MCI						
Age-adjusted						
Other GLTs	32 063	2438	1 (ref)		0.086	
Mtf	179 333	8366	1.00 (0.95 to 1.05)	0.023		0.98
Minimally adjusted (adjusted by age, sex and calendar time)						
Other GLTs	32 063	2438	1 (ref)		0.024	
Mtf	179 333	8366	0.92 (0.87 to 0.96)	0.022		<0.001
Fully adjusted*						
Other GLTs	16 547	1105	1 (ref)		0.035	
Mtf	130 336	5814	0.92 (0.86 to 0.99)	0.032		0.017

Excludes ethnicity.
*Adjusted for age, sex, calendar time, IMD, body mass index, smoking status, alcohol excess, statin use, antihypertensive use, hypertension, asthma, COPD, liver disease, coronary heart disease, peripheral vascular disease, stroke, diabetic retinopathy, neuropathy, brain injury, depression, autoimmune disease, CKD, heart failure, skin and soft tissue infection, urinary tract infection, lower respiratory tract infection, sepsis and baseline HbA1c.
CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; GLT, glucose-lowering therapy; HbA1c, hemoglobin A1c; IMD, Index of Multiple Deprivation; MCI, mild cognitive impairment; Mtf, merformin; RMSE, root mean-squared error.

Table 3 Associations of metformin versus alternative GLT with dementia and MCI stratified by age group and sex

	Exposure	N persons	N events	HR (95% CI)	P value
Dementia					
Age					
<80	Other GLTs	14 339	277	1 (ref)	
	Mtf	122 441	1542	0.78 (0.68 to 0.89)	<0.001
80+	Other GLTs	4725	436	1 (ref)	
	Mtf	21 401	1740	0.93 (0.83 to 1.03)	0.173
Sex					
Male	Other GLTs	5791	257	1 (ref)	
	Mtf	44 853	1256	0.87 (0.77 to 0.99)	0.03
Female	Other GLTs	3558	270	1 (base)	
	Mtf	31 499	1334	0.86 (0.76 to 0.97)	0.011
MCI					
Age					
<80	Other GLTs	14 339	623	1 (ref)	
	Mtf	122 441	3863	0.83 (0.76 to 0.91)	<0.001
80+	Other GLTs	4839	482	1 (ref)	
	Mtf	22 060	1951	1.05 (0.95 to 1.17)	0.35
Sex					
Male	Other GLTs	9986	587	1 (ref)	
	Mtf	75 587	2995	0.89 (0.81 to 0.98)	0.018
Female	Other GLTs	6561	518	1 (ref)	
	Mtf	54 749	2819	0.95 (0.86 to 1.05)	0.32

Interaction tests: dementia–age: p=0.03; dementia–sex: p=0.37; MCI–age: 0.0014; MCI–sex: 0.58.
GLT, glucose-lowering therapy; MCI, mild cognitive impairment; Mtf, metformin.

0.76 to 0.91), but no evidence of a risk reduction in older users with HR 1.05 (95% CI 0.95 to 1.17), respectively. LRT yielded strong evidence to support an age interaction with p=0.001 for MCI (table 3).

Sensitivity analyses

Application of serial restrictions in which dementia diagnoses made early during follow-up were excluded showed consistent HR estimates. In the post-2004 subgroup (after the advent of QOF), we found an estimated HR 0.88 (95% CI 0.80 to 0.97; n=138 444). In the post-2012 subgroup (post availability of dapagliflozin), the best estimate was HR 0.88 (95% CI 0.71 to 1.06, n=60 250), which was compatible with the null. In an analysis restricted to 96 308 individuals eligible for HES linkage, we found lower missingness for covariates (online supplemental table 4) and a best estimate that suggested a greater risk reduction among metformin users versus alternative GLT users than was seen in the main analysis (HR 0.80; 95% CI 0.71 to 0.90, n=68 614). The 'as-treated' analysis yielded an effect estimate HR 0.90 (95% CI 0.83 to 0.98, n=146 817) that was closer to the null (table 4). Application of serial restrictions on calendar period did not reveal a clear trend (online supplemental figure 4).

DISCUSSION

In this observational study of >200 000 UK adults treated for diabetes, we found evidence of a reduced risk of incident dementia and MCI among individuals initially prescribed metformin versus alternative GLT. We also found evidence that GLT interacted with age: individuals aged 40–79 years old exposed to metformin appeared to experience a lower risk of incident dementia and MCI than those on other GLTs, but this relationship was not seen in individuals aged ≥80 years. Best estimates suggest 23% lower incident dementia and 17% lower incident MCI associated with metformin use versus alternative GLT among individuals aged 40–79 years old. These findings were consistent across multiple sensitivity analyses. We found evidence of meaningful exposure misclassification in the main analysis as participants initiated additional GLT at later dates: 69% of participants in the alternative GLT group were prescribed metformin during follow-up. Although still compatible with a protective association, the 'as-treated' analysis yielded an effect estimate closer to the null.

Our results are consistent with established observational evidence showing that metformin use in diabetes is associated with a lower risk of neurodegenerative disease.

Table 4 Sensitivity analyses*

Exposure	N persons	N events	HR (95% CI)	P value
(1) (a) Lagged analysis excluding dementia diagnoses in first 3 months				
Other GLTs	16042	697	1 (ref)	
Metformin	127249	3239	0.88 (0.81 to 0.96)	0.004
(b) Excluding dementia diagnoses in first 6 months				
Other GLTs	15498	679	1 (ref)	
Metformin	123699	3182	0.88 (0.81 to 0.97)	0.008
(c) Excluding dementia diagnoses in first year				
Other GLTs	14531	658	1 (ref)	
Metformin	116211	3050	0.88 (0.80 to 0.96)	0.003
(d) Excluding dementia diagnoses in first 2 years				
Other GLTs	12863	595	1 (ref)	
Metformin	102177	2780	0.88 (0.81 to 0.97)	0.01
(2) Subgroup with entry post-2004				
Other GLTs	14070	540	1 (ref)	
Metformin	124374	2982	0.88 (0.80 to 0.97)	0.008
(3) Subgroup with entry post-2012				
Other GLTs	5261	106	1 ref)	
Metformin	54989	681	0.88 (0.71 to 1.10)	0.26
(4) Restricted to those with HES linkage				
Other GLTs	7922	369	1 (ref)	
Metformin	60692	1453	0.80 (0.71 to 0.90)	<0.001
(5) 'As-treated' analysis				
Other GLTs	16539	713	1 (ref)	
Metformin	130278	3282	0.90 (0.83 to 0.98)	0.021

*Confounder adjustment as per fully adjusted models in table 2.
GLTs, glucose-lowering therapies; HES, Hospital Episode Statistics.

Newby *et al*[12] reported comparable estimates from a US retrospective cohort study using an ACNU design with HR 0.80 (95% CI 0.73 to 0.88) for all-cause dementia and HR 0.91 (95% CI 0.79 to 1.04) for MCI. Zhang *et al* reported a pooled relative risk (RR) 0.77 (95% CI 0.67 to 0.88) in a meta-analysis of 12 studies of metformin use and incident neurodegenerative diseases including Parkinson's disease.[11] They included high-quality population-based cohort studies, but many lacked active comparators. Other meta-analyses have, however, been conflicting: Zhang *et al* reported a protective association between metformin use and cognitive dysfunction (HR 0.90 (95% CI 0.88 to 0.92)) from 10 cohort studies (more of which had active comparators), but Ping *et al* found a null effect for the association between metformin use and neurodegenerative diseases including Alzheimer's disease in a meta-analysis of 19 studies, which included cross-sectional studies with a higher risk of bias.[9 20]

Nevertheless, a recent Mendelian randomisation study, which by design eliminates reverse causation and most confounding, showed that genetically proxied metformin use was associated with a small reduction in Alzheimer's disease risk.[21] In that study, mitochondrial function and the *NDUFA2* gene were proposed as dementia protection mechanisms. Metformin has also been shown to have anti-inflammatory effects irrespective of diabetes status and it reduces accumulation of Alzheimer's disease neuropathology in in vitro models.[22 23]

We found that metformin exposure was only associated with lower incident dementia risk for individuals under 80 years old. This is consistent with three previous US retrospective cohort studies and may be because older individuals accumulate multiple other risk factors for dementia, meaning that any potential benefit of metformin becomes increasingly negligible with age.[12 17 24] Similarly, work on the predictive modeling of dementia risk factors has found that models with proven efficacy in younger groups are inaccurate in advanced old age.[25] We observed that dementia rates increased in successive calendar periods (even when conditioning on age). This is consistent with sustained improvements in the routine ascertainment of dementia since 1990 in the UK, but there is still evidence of a dementia diagnosis gap.[26]

As far as we are aware, this is the largest historical cohort study investigating the association between metformin use and incident dementia to date with 211 396 participants. 98% of the UK population is GP-registered and CPRD is known to be representative of the UK general population and is comparable with census data.[13] Findings are likely to be generalizable to the UK population with type 2 diabetes aged 40 years or more in receipt of oral GLTs. We used robust methodology for this study: the ACNU approach helps address some inherent weaknesses of pharmacoepidemiologic studies—namely confounding by indication and healthy user bias—and we used multiple sensitivity analyses to assess the contribution of other plausible sources of bias, which gave similar results.

The ACNU design intends to reduce possible confounding by indication and uses the first recorded drug prescription. This may have increased the potential for exposure misclassification, especially in a study with a long follow-up time, during which time treatment escalation or switching could occur. Around half of individuals classified as initial metformin users later received an alternate GLT, for example, due to treatment escalation, while 69% of individuals classified as alternative GLT users were later prescribed metformin. We note that the effect estimate from the 'as-treated' analysis was closer to the null, but was still consistent with a 10% risk reduction for incident dementia. Furthermore, we cannot be certain of participant adherence to treatment, which has been estimated to be as low as 36% in Western settings for individuals prescribed GLT.[27] While interest is growing in the use of metformin as an 'anti-aging' therapy,[28] it is unlikely that this would have contributed to exposure misclassification in our study. Although exposure misclassification is complex to evaluate, if the alternative GLT group received metformin at a later date, this may have contributed to underestimation of the effect size in the main analysis.

NICE guidelines have favored metformin as a first-line treatment since 2002[5] and prescribing practice for diabetes has changed meaningfully—overall, 95% of individuals were prescribed metformin at least once and new participants were more likely to be prescribed metformin versus alternative GLT in successive calendar periods. Nevertheless, results restricted to later calendar entry were similar to the main analysis. In another sensitivity analysis, there was no evidence of spurious calendar effects, although low numbers of participants from 1990 to 2003 meant that this estimate was relatively underpowered.

Despite using an ACNU approach and multivariate adjustment, residual confounding, especially confounding by indication, remains a considerable risk. This could occur if the choice of GLT prescription was associated with underlying characteristics leading to dementia or MCI. Comorbidities such as renal impairment or heart failure are common among individuals with diabetes[29] and may affect metformin prescribing.[30] In our study, metformin users had lower baseline HbA1c measurements, were more likely to be overweight or obese and to be prescribed statins or antihypertensives than alternative GLT users. While we controlled for these variables in our analyses, it is difficult to rule out confounding by frailty or other factors that are poorly assessed in routine health records and may account for prescribing differences including therapeutic inertia among individuals with diabetes.[31] Confounding may have biased effect estimates away from the null.

UK dementia diagnoses are most frequently made in general practice or in dedicated memory assessment services. CPRD recording of dementia has previously been evaluated and is comparable with other sources,[32] but it is still likely to meaningfully underascertain cases. In England, the observed dementia prevalence from primary care electronic health records (EHRs) is approximately 62% of epidemiological predictions for individuals over 65 years old and is subject to considerable local variation.[33 34] There is also a possibility that frequent users of GP services are likely to have better ascertainment of dementia, MCI and other conditions. Diagnostic codes are complex, overlapping and conditions are not necessarily coded by clinicians. Clinical diagnoses of subtypes are known to correlate poorly with postmortem neuropathological examination and individuals in advanced old age can present with overlapping pathological features.[35] Given these concerns, we chose all-cause dementia for the primary outcome. We used comprehensive code lists to maximise ascertainment of dementia, but note that a substantial proportion of cases will not have a diagnosis in EHRs. Nevertheless, as long as outcome underascertainment is non-differential between metformin and alternative GLT users, it will not have biased our effect estimates.[36]

The substantial missingness for certain covariates, for example, ethnicity is problematic, because it is likely missing not at random, making techniques like multiple imputation unsuitable and results in a smaller population for the complete case analysis, potentially introducing selection bias. Also, missingness for covariates detailing baseline comorbidities could not be assessed as ascertainment relies solely on the presence of relevant Read codes. However, we hypothesize that individuals with diabetes are likely have more frequent GP consultations which would mitigate potential differential information bias. It is also reassuring that the fully adjusted models and minimally adjusted models gave similar estimates for both outcomes. Although the complete case analysis approach reduced the precision of estimates, it does not appear to have introduced bias. We also replicated our main findings in a HES-linked subgroup with improved record completeness.

CONCLUSIONS

This study adds to a growing evidence base that suggests that metformin use in diabetes may be protective against

dementia among individuals aged <80 years. However, although an ACNU approach can help to mitigate confounding by indication and healthy user bias, it may increase the potential for exposure misclassification, especially with long follow-up. Data missingness and likely underascertainment of dementia remain an ongoing concern when using routine health records. Future studies could use additional methods to reduce and explore confounding such as high-dimensional propensity scores and quantitative bias analysis. Triangulating evidence across other study designs that are robust to confounding including randomized controlled trials and Mendelian randomization will help to assess causality.

Author affiliations
[1]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK
[2]Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, USA
[3]Cambridge Public Health, University of Cambridge, Cambridge, UK
[4]MRC Unit for Lifelong Health and Ageing, University College London, London, UK
[5]NDM Centre for Global Health Research, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[6]NIHR Oxford Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[7]Centre for Primary Care, Wolfson Institute of Population Health, Queen Mary University of London, London, UK

Acknowledgements The data are provided by patients and collected by the NHS as part of their care and support.

Contributors CW-G conceived the study. All authors contributed to study design. RM, LT, WD and CW-G managed data. WD analyzed data. WD and CW-G wrote the paper. All authors interpreted findings, revised the manuscript for important intellectual content and approved the final version for publication. CWG is the guarantor.

Funding CW-G was originally supported by a Wellcome Intermediate Clinical Fellowship (201440/Z/16/Z) and now holds a Wellcome Career Development Award (225868/Z/22/Z). SJD is supported by an NIHR Professorship (NIHR300791). KB holds a Wellcome Senior Research Fellowship (220283/Z/20/Z). SVE is supported by a Diabetes UK Sir George Alberti Clinical Training Fellowship (grant number 17/0005588). RM is supported by Barts Charity (MGU0504).

Competing interests NC receives funds from AstraZeneca to serve on data safety and monitoring committees for clinical trials. R. Mathur is part of the Genes & Health Program, which is part-funded (including salary contributions) by a Life Sciences Consortium comprising AstraZeneca, Bristol-Myers Squibb Company, GlaxoSmithKline Research and Development Limited, Maze Therapeutics, Merck Sharp & Dohme, Novo Nordisk, Pfizer, and Takeda Development Center Americas. All other authors report no conflicts.

Patient consent for publication Not required.

Ethics approval This study involves human participants and was approved by LSHTM research ethics committee (27407; 22247). Data are collected as part of routine clinical care. Consent is given at GP practice level to contribute de-identified patient data to the CPRD database for research. Patients are able to opt out of contributing their de-identified data for research.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data used for this study were obtained from the CPRD. Access to CPRD data is subject to protocol approval via CPRD's Research Data Governance Process (see https://www.cprd.com/Data-access). Data acquisition is associated with a fee and data protection requirements. Code lists used to define health conditions in this study have been made openly available on LSHTM Data Compass: (https://datacompass.lshtm.ac.uk/id/eprint/3402/).

ORCID iDs
Christopher T Rentsch http://orcid.org/0000-0002-1408-7907
Charlotte Warren-Gash http://orcid.org/0000-0003-4524-3180

REFERENCES

1 Whicher CA, O'Neill S, Holt RIG. Diabetes in the UK: 2019. *Diabet Med* 2020;37:242–7.
2 Livingston G, Huntley J, Sommerlad A, *et al*. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 2020;396:413–46.
3 Luo A, Xie Z, Wang Y, *et al*. Type 2 diabetes mellitus-associated cognitive dysfunction: advances in potential mechanisms and therapies. *Neurosci Biobehav Rev* 2022;137:104642.
4 National Institute for Health and Care Excellence. Type 2 diabetes in adults: management. NICE guideline 28; 2021.
5 Home P, McKintosh A. NICE sets out strategy for glucose control in diabetes patients. Guidelines in practice; 2002.
6 Rena G, Hardie DG, Pearson ER. The mechanisms of action of metformin. *Diabetologia* 2017;60:1577–85.
7 Sanati M, Aminyavari S, Afshari AR, *et al*. Mechanistic insight into the role of metformin in Alzheimer's disease. *Life Sci* 2022;291:120299.
8 Liao W, Xu J, Li B, *et al*. Deciphering the roles of metformin in Alzheimer's disease: a snapshot. *Front Pharmacol* 2022;12.
9 Ping F, Jiang N, Li Y. Association between metformin and neurodegenerative diseases of observational studies: systematic review and meta-analysis. *BMJ Open Diabetes Res Care* 2020;8:e001370.
10 Zhou J-B, Tang X, Han M, *et al*. Impact of antidiabetic agents on dementia risk: a Bayesian network meta-analysis. *Metabolism* 2020;109:154265.
11 Zhang Y, Zhang Y, Shi X, *et al*. Metformin and the risk of neurodegenerative diseases in patients with diabetes: a meta-analysis of population-based cohort studies. *Diabet Med* 2022;39:e14821.
12 Newby D, Linden AB, Fernandes M, *et al*. Comparative effect of metformin versus sulfonylureas with dementia and Parkinson's disease risk in US patients over 50 with type 2 diabetes mellitus. *BMJ Open Diabetes Res Care* 2022;10:e003036.
13 Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
14 Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep* 2015;2:221–8.
15 White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med* 2010;29:2920–31.
16 Campesi I, Seghieri G, Franconi F. Type 2 diabetic women are not small type 2 diabetic men: sex-and-gender differences in antidiabetic drugs. *Curr Opin Pharmacol* 2021;60:40–5.
17 Orkaby AR, Cho K, Cormack J, *et al*. Metformin vs Sulfonylurea use and risk of dementia in US veterans aged $65 years with diabetes. *Neurology* 2017;89:1877–85.
18 Kontopantelis E, Reeves D, Valderas JM, *et al*. Recorded quality of primary care for patients with diabetes in England before and after the introduction of a financial incentive scheme: a longitudinal observational study. *BMJ Qual Saf* 2013;22:53–64.
19 Wilkinson T, Ly A, Schnier C, *et al*. Identifying dementia cases with routinely collected health data: a systematic review. *Alzheimers Dement* 2018;14:1038–51.

20 Zhang Q-Q, Li W-S, Liu Z, *et al*. Metformin therapy and cognitive dysfunction in patients with type 2 diabetes: a meta-analysis and systematic review. *Medicine (Baltimore)* 2020;99:e19378.
21 Zheng J, Xu M, Walker V, *et al*. Evaluating the efficacy and mechanism of metformin targets on reducing Alzheimer's disease risk in the general population: a mendelian randomisation study. *Diabetologia* 2022;65:1664–75.
22 Gupta A, Bisht B, Dey CS. Peripheral insulin-sensitizer drug metformin ameliorates neuronal insulin resistance and Alzheimer's-like changes. *Neuropharmacology* 2011;60:910–20.
23 Cameron AR, Morrison VL, Levin D, *et al*. Anti-inflammatory effects of metformin irrespective of diabetes status. *Circ Res* 2016;119:652–65.
24 Scherrer JF, Salas J, Floyd JS, *et al*. Metformin and Sulfonylurea use and risk of incident dementia. *Mayo Clin Proc* 2019;94:1444–56.
25 Walters K, Hardoon S, Petersen I, *et al*. Predicting dementia risk in primary care: development and validation of the dementia risk score using routinely collected data. *BMC Med* 2016;14:6.
26 Donegan K, Fox N, Black N, *et al*. Trends in diagnosis and treatment for people with dementia in the UK from 2005 to 2015: a longitudinal retrospective cohort study. *Lancet Public Health* 2017;2:e149–56.
27 Khunti N, Khunti N, Khunti K. Adherence to type 2 diabetes management. *Br J Diabetes* 2019;19:99–104.
28 Soukas AA, Hao H, Wu L. Metformin as anti-aging therapy: is it for everyone? *Trends Endocrinol Metab* 2019;30:745–55.
29 Pearson-Stuttard J, Holloway S, Polya R, *et al*. Variations in comorbidity burden in people with type 2 diabetes over disease duration: a population-based analysis of real world evidence. *EClinicalMedicine* 2022;52:101584.
30 Joint Formulary Committee. *British National Formulary, 84th ed*. London: British Medical Association and Royal Pharmaceutical Society of Great Britain, 2022.
31 Mathur R, Farmer RE, Eastwood SV, *et al*. Ethnic disparities in initiation and intensification of diabetes treatment in adults with type 2 diabetes in the UK, 1990-2017: a cohort study. *PLoS Med* 2020;17:e1003106.
32 McGuinness LA, Warren-Gash C, Moorhouse LR, *et al*. The validity of dementia diagnoses in routinely collected electronic health records in the United Kingdom: a systematic review. *Pharmacoepidemiol Drug Saf* 2019;28:244–55.
33 Walker IF, Lord PA, Farragher TM. Variations in dementia diagnosis in England and association with general practice characteristics. *Prim Health Care Res Dev* 2017;18:235–41.
34 NHS Digital. Primary care dementia data. 2023. Available: https://digital.nhs.uk/data-and-information/publications/statistical/primary-care-dementia-data/january-2023# [Accessed 23 Feb 2023].
35 Selvackadunco S, Langford K, Shah Z, *et al*. Comparison of clinical and neuropathological diagnoses of neurodegenerative diseases in two centres from the brains for dementia research (BDR) cohort. *J Neural Transm (Vienna)* 2019;126:327–37.
36 Alexander LK, Lopes B, Ricchetti-Masterson K, *et al*. Sources of systematic error or bias: information bias. ERIC notebook 2nd edition. Available: https://sph.unc.edu/wp-content/uploads/sites/112/2015/07/nciph_ERIC14.pdf [Accessed 07 Dec 2023].

Supplementary Material

Supplementary methods

Source population & cohort identification – main analysis

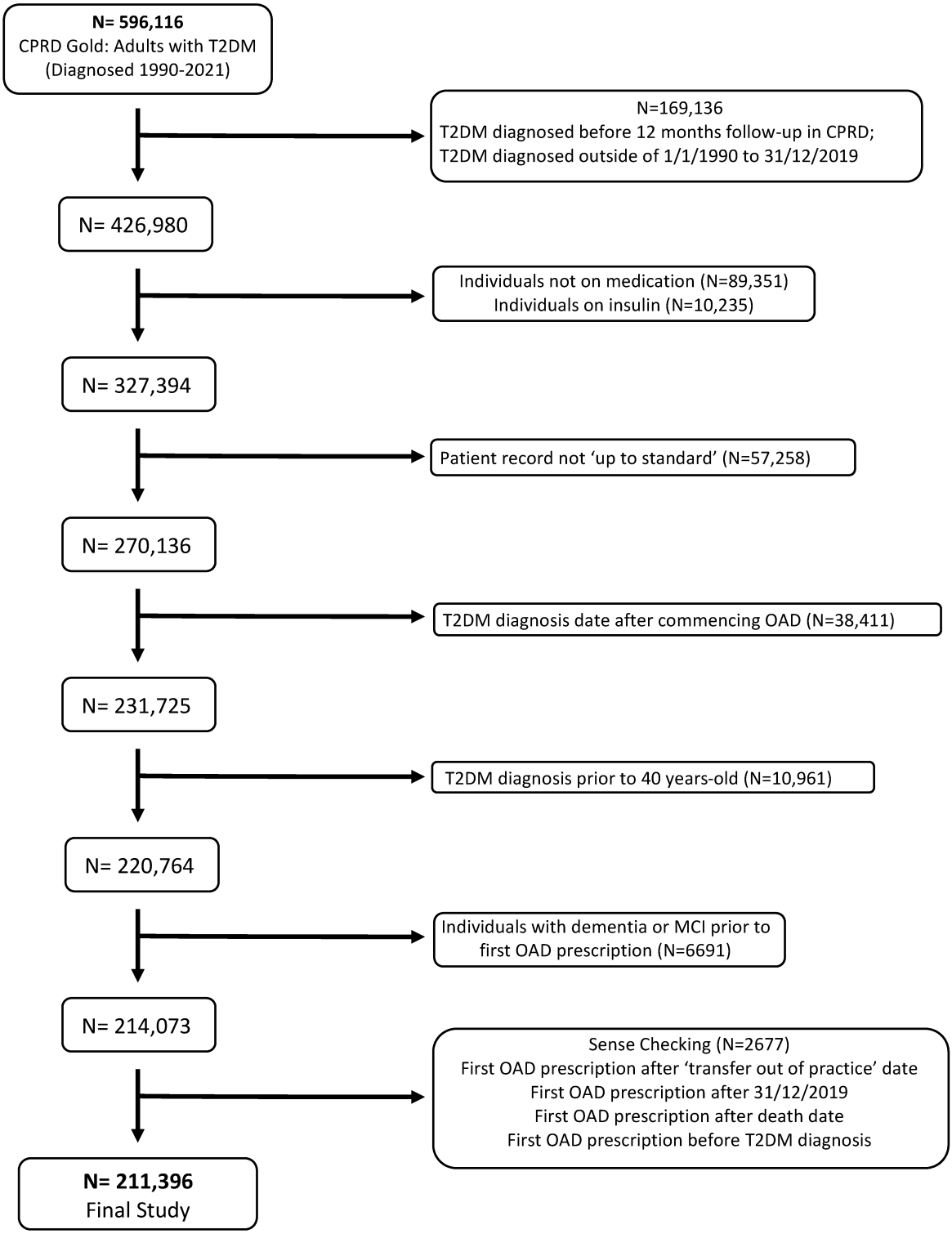
After identification of the initial study cohort (described in p5 in the main methods), we specified further restrictions to ensure data quality, e.g. including only CPRD 'research acceptable' records and including only records from practices that were CPRD 'up-to-standard' prior to the index date. Sense checking was also carried out to exclude individuals with illogical prescription dates e.g. recorded after death or before diabetes diagnosis. The full process of cohort identification is shown in supplementary figure 1

Source population & cohort identification – sensitivity analysis

For the post-hoc "as-treated" sensitivity analysis, LSHTM investigators extracted new data on the study population from CPRD flat files for a later CPRD build (July 2021) as the original raw data had been destroyed. Additional data on dates of all GLT prescriptions from Therapy files was extracted and merged into the existing analysis dataset. This sensitivity analysis dataset included 211,310 individuals – 86 fewer than the original study population.

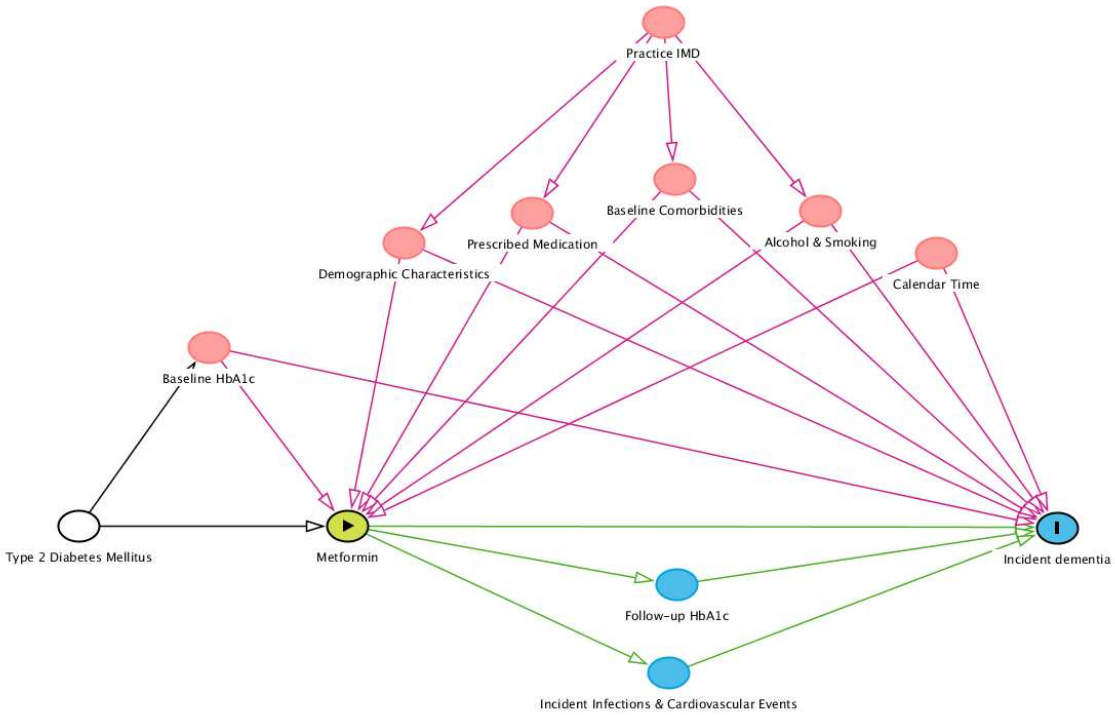
Doran W, *et al. BMJ Open Diab Res Care* 2024; 12:e003548. doi: 10.1136/bmjdrc-2023-003548

Supplementary figure 1: Cohort identification



Footnote supplementary figure 1: 'Up to standard' refers to the practice 'up to standard' date, at which a practice is considered to have continuous high-quality data fit for use in research. It is derived by the data provider using an algorithm that primarily considers practice death recording and gaps in the data.

Supplementary figure 2: Directed acyclic graph



Supplementary table 1: Crude dementia & MCI rates by age band and calendar time

Variable	N Events	P-Y At Risk	Rate Per 1000 (95% CI)
Dementia			
Age bands (years)			
40-49	5	105,309	0.05 (0.02 - 0.11)
50-59	100	298,269	0.34 (0.28 - 0.41)
60-69	511	409,478	1.25 (1.14 - 1.36)
70-79	2,346	353,570	6.64 (6.37 - 6.91)
80-89	3,163	146,808	21.55 (20.81 - 22.31)
≥90	517	15,127	34.18 (31.36 - 37.25)
Calendar Time			
1990-2004	131	108,908	1.20 (1.01 - 1.43)
2005-2009	1,043	318,175	3.28 (3.09 - 3.48)
2010-2014	2,627	498,445	5.27 (5.07 - 5.48)
2015-2019	2,841	403,262	7.05 (6.79 - 7.31)
MCI			
Age bands (years)			
40-49	209	112,218	1.86 (1.63 - 2.13)
50-59	914	319,718	2.86 (2.68 - 3.05)
60-69	1,870	435,772	4.29 (4.10 - 4.49)
70-79	3,844	371,975	10.33 (10.01 - 10.67)
80-89	3,432	158,468	21.66 (20.94 - 22.39)
≥90	535	21,733	24.62 (22.62 - 26.79)
Calendar Time			
1990-2004	383	108,442	3.53 (3.20 - 3.90)
2005-2009	2,157	325,530	6.63 (6.35 - 6.91)
2010-2014	4,561	534,324	8.54 (8.29 - 8.79)
2015-2019	3,703	451,588	8.20 (7.94 - 8.47)

Supplementary table 2: Association between GLT and all-cause dementia. Covariates from fully-adjusted model.

		HR (95% CI)
Calendar Time Band	1990-2004	1 (ref)
	2005-2009	2.25 (1.46 - 3.45)
	2010-2014	4.08 (2.67 - 6.22)
	2015-2019	5.33 (3.5 - 8.14)
Sex	Male	1 (ref)
	Female	1.12 (1.04 - 1.20)
Index of Multiple Deprivation	1 (Poorest)	1 (ref)
	2	1.01 (0.91 - 1.13)
	3	1.05 (0.95 - 1.17)
	4	1.10 (0.99 - 1.21)
	5 (Wealthiest)	1.23 (1.11 - 1.36)
Baseline HbA1c (mmol/mol)	<48	1 (ref)
	48-67.9	0.95 (0.84 - 1.07)
	68-87.9	0.99 (0.87 - 1.12)
	88-107.9	0.97 (0.83 - 1.13)
	>108	0.86 (0.73 - 1.03)
Body Mass Index (kg/m2)‡	Underweight	1 (ref)
	Normal Weight	0.77 (0.49 - 1.20)
	Overweight	0.66 (0.42 - 1.03)
	Obesity Class I	0.58 (0.37 - 0.91)
	Obesity Class II, III	0.60 (0.38 - 0.94)
Smoking Status	Never	1 (ref)
	Current	1.14 (1.02 - 1.26)
	Ex	1.01 (0.94 - 1.08)
Alcohol Misuse		1.39 (1.05 - 1.84)
Statin		1.03 (0.96 - 1.11)
Anti-Hypertensive		1.02 (0.92 - 1.13)
Hypertension		0.94 (0.87 - 1.02)
Asthma		1.02 (0.93 - 1.11)
COPD		0.95 (0.86 - 1.05)
Liver Disease		1.19 (0.92 - 1.53)
Coronary Heart Disease		1.08 (1.00 - 1.17)
Peripheral Vascular Disease		1.31 (1.11 - 1.55)
Stroke		1.11 (0.74 - 1.68)
Diabetic Retinopathy		0.88 (0.78 - 0.99)
Neuropathy		1.18 (0.99 - 1.40)
Brain Injury		1.61 (1.21 - 2.14)
Depression		1.52 (1.41 - 1.63)
Autoimmune Disease		0.99 (0.89 - 1.09)
Chronic Kidney Disease		0.72 (0.47 - 1.11)

Heart Failure	1.14 (1.01 - 1.29)
Skin & Soft Tissue Infection	0.97 (0.89 - 1.05)
Urinary Tract Infection	1.04 (0.96 - 1.13)
Lower Respiratory Tract Infection	0.99 (0.92 - 1.06)
Sepsis	1.24 (0.91 - 1.69)

Supplementary table 3: Association between GLT and MCI. Covariates from fully-adjusted model.

		HR (95% CI)
Calendar Time Band	1990-2004	1 (ref)
	2005-2009	1.55 (1.25 - 1.94)
	2010-2014	1.98 (1.60 - 2.46)
	2015-2019	1.82 (1.46 - 2.26)
Sex	Male	1 (ref)
	Female	1.00 (0.95 - 1.06)
Index of Multiple Deprivation	1 (Poorest)	1 (ref)
	2	1.02 (0.94 - 1.11)
	3	1.02 (0.95 - 1.11)
	4	1.02 (0.95 - 1.11)
	5 (Wealthiest)	1.13 (1.04 - 1.21)
Baseline HbA1c (mmol/mol)	<48	1 (ref)
	48-67.9	0.95 (0.87 - 1.03)
	68-87.9	0.94 (0.85 - 1.03)
	88-107.9	0.85 (0.75 - 0.95)
	>108	0.91 (0.80 - 1.03)
Body Mass Index (kg/m2)‡	Underweight	1 (ref)
	Normal Weight	0.71 (0.50 - 1.02)
	Overweight	0.61 (0.43 - 0.87)
	Obesity Class I	0.61 (0.43 - 0.87)
	Obesity Class II, III	0.62 (0.43 - 0.89)
Smoking Status	Never	1 (ref)
	Current	1.08 (1.00 - 1.16)
	Ex	1.03 (0.98 - 1.09)
Alcohol Misuse		1.14 (0.93 - 1.4)
Statin		1.04 (0.98 - 1.10)
Anti-Hypertensive		1.03 (0.95 - 1.10)
Hypertension		0.93 (0.87 - 0.98)
Asthma		1.09 (1.02 - 1.16)
COPD		1.00 (0.93 - 1.07)
Liver Disease		1.16 (0.97 - 1.38)
Coronary Heart Disease		1.11 (1.04 - 1.17)
Peripheral Vascular Disease		1.29 (1.13 - 1.47)
Stroke		1.20 (0.87 - 1.64)
Diabetic Retinopathy		1.03 (0.94 - 1.12)
Neuropathy		1.16 (1.01 - 1.33)
Brain Injury		1.28 (1.03 - 1.61)
Depression		1.62 (1.54 - 1.71)
Autoimmune Disease		1.16 (1.08 - 1.25)
Chronic Kidney Disease		1.19 (0.89 - 1.59)

Heart Failure	1.21 (1.12 - 1.29)
Skin & Soft Tissue Infection	1.11 (1.05 - 1.18)
Urinary Tract Infection	1.16 (1.09 - 1.23)
Lower Respiratory Tract Infection	1.17 (1.11 - 1.23)
Sepsis	1.09 (0.86 - 1.39)

Supplementary table 4: Descriptive statistics for HES-linked cohort (n=96,308)

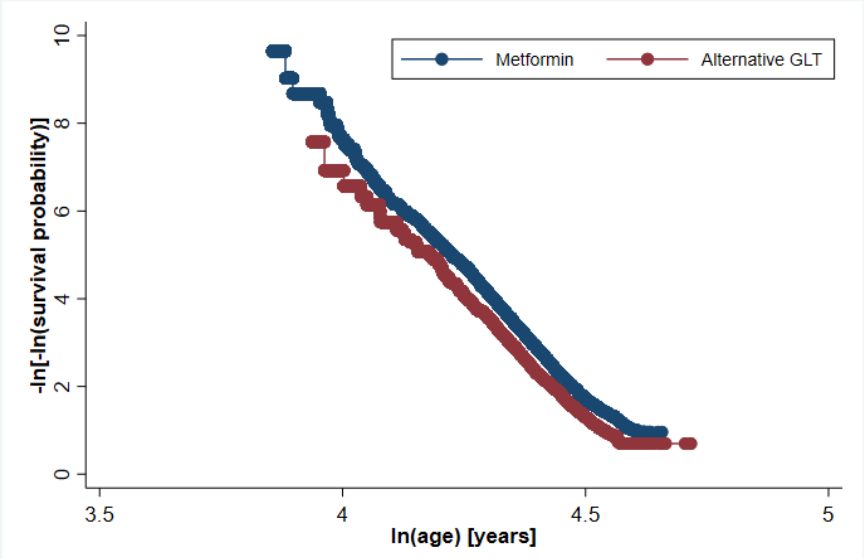
Variable (n missing)		Overall N (%$_n$)†	Metformin N (%$_n$)	Other GLT N (%$_n$)
Totals		96,308 (-)	81,200 (84.3)	15,108 (15.7)
Age at Entry	Median (IQR)	61 (53 - 70)	61 (52 - 69)	64 (55 - 73)
Sex	Male	54,900 (57.0)	45,932 (56.6)	8,968 (59.4)
	Female	41,408 (43.0)	35,268 (43.4)	6,140 (40.6)
Ethnicity	White	48,205 (88.5)	41,510 (88.4)	6,695 (89.0)
	South Asian	3,561 (6.5)	3,123 (6.7)	438 (5.8)
	Black	1,643 (3.0)	1,390 (3.0)	253 (3.4)
	Mixed or Other	1,054 (1.9)	916 (2.0)	138 (1.8)
	Missing	41,845 (43.4*)	34,261 (42.2*)	7,584 (50.2*)
Index of Multiple Deprivation	1 (Poorest)	18,540 (19.3)	15,685 (19.3)	2,855 (18.9)
	2	20,068 (20.8)	16,738 (20.6)	3,330 (22.0)
	3	20,801 (21.6)	17,510 (21.6)	3,291 (21.8)
	4	19,181 (19.9)	16,332 (20.1)	2,849 (18.9)
	5 (Wealthiest)	17,718 (18.4)	14,935 (18.4)	2,783 (18.4)
Baseline HbA1c (mmol/mol)	<48	7,353 (9.0)	6,405 (8.9)	948 (9.3)
	48-67.9	40,532 (49.4)	36,801 (51.2)	3,731 (36.4)
	68-87.9	18,739 (22.8)	16,221 (22.6)	2,518 (24.6)
	88-107.9	9,310 (11.3)	7,751 (10.8)	1,559 (15.2)
	>108	6,130 (7.5)	4,645 (6.5)	1,485 (14.5)
	Missing	14,244 (14.8*)	9,377 (11.5*)	4,867 (32.2*)
Body Mass Index (kg/m^2)‡	Underweight	277 (0.3)	121 (0.2)	156 (1.2)
	Normal Weight	9,712 (11.3)	6,290 (8.5)	3,422 (27.4)
	Overweight	28,372 (32.9)	23,441 (31.8)	4,931 (39.5)
	Obesity Class I	26,188 (30.4)	23,664 (32.1)	2,524 (20.2)
	Obesity Class II, III	21,608 (25.1)	20,148 (27.4)	1,460 (11.7)
	Missing	10,151 (10.5*)	7,536 (9.3*)	2,615 (17.3*)
Smoking Status	Non	36,710 (44.6)	31,714 (44.5)	4,996 (45.2)
	Current	14,582 (17.7)	12,515 (17.6)	2,067 (18.7)
	Ex	30,988 (37.7)	26,991 (37.9)	3,997 (36.1)
	Missing	14,028 (14.6*)	9,980 (12.3*)	4,048 (26.8*)
Alcohol Misuse		947 (1.0)	766 (0.9)	181 (1.2)
Statin		58,926 (61.2)	52,137 (64.2)	67,89 (44.9)
Anti-Hypertensive		65,099 (67.6)	55,426 (68.3)	9,673 (64)
Hypertension		48,843 (50.7)	41,898 (51.6)	6,945 (46)
Asthma		15,416 (16.0)	13,315 (16.4)	2,101 (13.9)
COPD		11,267 (11.7)	9,325 (11.5)	1,942 (12.9)
Liver Disease		2,078 (2.2)	1,795 (2.2)	283 (1.9)
Coronary Heart Disease		14,808 (15.4)	12,009 (14.8)	2,799 (18.5)
Peripheral Vascular Disease		1,798 (1.9)	1,459 (1.8)	339 (2.2)
Stroke		376 (0.4)	309 (0.4)	67 (0.4)
Diabetic Retinopathy		5,133 (5.3)	4,470 (5.5)	663 (4.4)
Neuropathy		2,084 (2.2)	1,756 (2.2)	328 (2.2)

Brain Injury	740 (0.8)	652 (0.8)	88 (0.6)
Depression	24,938 (25.9)	21,666 (26.7)	3,272 (21.7)
Autoimmune Disease	7,484 (7.8)	5,960 (7.3)	1,524 (10.1)
Chronic Kidney Disease	495 (0.5)	219 (0.3)	276 (1.8)
Heart Failure	3,716 (3.9)	2,735 (3.4)	981 (6.5)
Skin & Soft Tissue Infection	21,308 (22.1)	18,442 (22.7)	2,866 (19.0)
Urinary Tract Infection	18,504 (19.2)	15,709 (19.3)	2,795 (18.5)
Lower Respiratory Tract Infection	40,479 (42.0)	34,618 (42.6)	5,861 (38.8)
Sepsis	840 (0.9)	670 (0.8)	170 (1.1)

† $\%_n$: % of non-missing * % of total

‡ BMI Category Definitions (Kg/m2): Underweight <18.5, Normal Weight 18.5-24.9, Overweight 25-29.9, Obesity Class I 30-34.9, Obesity Class II & III ≥ 35

Supplementary figure 3: Log-Log Kaplan Meier Survival Plot



Supplementary figure 4: Serial restrictions on calendar period