Investigating the Relationship Between Type 2 Diabetes and Dementia Using Electronic Medical Records in the GoDARTS Bioresource

Alex S.F. Doney,[1] Wilfred Bonney,[2] Emily Jefferson,[2] Katherine E. Walesby,[3] Rachel Bittern,[2] Emanuele Trucco,[4] Peter Connelly,[5,6,7] Rory J. McCrimmon,[1] and Colin N.A. Palmer[1]

Diabetes Care 2019;42:1973–1980 | https://doi.org/10.2337/dc19-0380

OBJECTIVE

To investigate the impact of type 2 diabetes on incidence of major dementia subtypes, Alzheimer and vascular dementia, using electronic medical records (EMR) in the GoDARTS bioresource.

RESEARCH DESIGN AND METHODS

GoDARTS (Genetics of Diabetes Audit and Research in Tayside Scotland) comprises a large case-control study of type 2 diabetes with longitudinal follow-up in EMR. Dementia case subjects after recruitment were passively identified in the EMR, and using a combination of case note review, an Alzheimer-specific weighted genetic risk score (wGRS), and *APOE4* genotype, we validated major dementia subtypes. We undertook a retrospective matched cohort study to determine the risk of type 2 diabetes status for incident dementia accounting for competing risk of death.

RESULTS

Type 2 diabetes status was associated with a significant risk of any dementia (cause-specific hazard ratio [csHR] 1.46, 95% CI 1.31–1.64), which was attenuated, but still significant, when competing risk of death was accounted for (subdistribution [sd]HR 1.26, 95% CI 1.13–1.41). The accuracy of EMR-defined cases of Alzheimer or vascular dementia was high—positive predictive value (PPV) 86.4% and PPV 72.8%, respectively—and wGRS significantly predicted Alzheimer dementia (HR 1.23, 95% CI 1.12–1.34) but not vascular dementia (HR 1.02, 95% CI 0.91–1.15). Conversely, type 2 diabetes was strongly associated with vascular dementia (csHR 2.47, 95% C 1.92–3.18) but not Alzheimer dementia, particularly after competing risk of death was accounted for (sdHR 1.02, 95% CI 0.87–1.18).

CONCLUSIONS

Our study indicates that type 2 diabetes is associated with an increased risk of vascular dementia but not with an increased risk of Alzheimer dementia and highlights the potential value of bioresources linked to EMR to study dementia.

The increasing global impact of the dementias on health care systems and society has been well-documented and publicized (1,2). While the major driver for this is greater numbers of individuals living to old age, many epidemiological studies have linked type 2 diabetes and related traits with increased dementia risk (3–5). As such traits are

[1]*Division of Population Health and Genomics, Ninewells Hospital & Medical School, University of Dundee, Dundee, Scotland, U.K.*
[2]*Health Informatics Centre, Ninewells Hospital & Medical School, University of Dundee, Dundee, Scotland, U.K.*
[3]*Alzheimer Scotland Dementia Research Centre and Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, Scotland, U.K.*
[4]*VAMPIRE Project, Computing, School of Science and Engineering, University of Dundee, Scotland, U.K.*
[5]*NHS Tayside, Scotland, U.K.*
[6]*Department of Neuroscience, University of Dundee, Dundee, Scotland, U.K.*
[7]*Department of Applied Social Science, University of Stirling, Stirling, Scotland, U.K.*

Corresponding author: Alexander S.F. Doney, a.doney@dundee.ac.uk

Received 22 February 2019 and accepted 17 July 2019

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc19-0380/-/DC1.

also becoming globally more prevalent, this relationship is of major public health interest, as it indicates important opportunities for reducing the future impact of dementia through lifestyle and clinical interventions (6,7). The dementias are pathologically heterogeneous. Alzheimer disease, the most common cause of dementia, is neurodegenerative and caused by accumulation of β-amyloid in plaques and τ-protein in neurofibrillary tangles that ultimately destroy neural cells (8). Vascular dementia, the second most common, is caused by cerebrovascular disease causing cerebral infarction or hemorrhage, which, although also destroying neurons, is not associated with the abnormal protein accumulation of Alzheimer pathology (9). However, it is well established that a significant proportion of individuals with Alzheimer dementia also have, to a greater or lesser extent, cerebrovascular disease. While links between type 2 diabetes and vascular dementia have been largely consistent and compatible with established understanding of micro- and macrovascular complications of type 2 diabetes, for Alzheimer disease, such links are more contentious. Although a number of molecular and cellular processes are shared between type 2 diabetes and Alzheimer pathology (10), the strength of the epidemiological association is overall weaker than it is for vascular dementia and not demonstrated in all studies (11–13). Furthermore, while clinicopathological studies consistently support an association of type 2 diabetes with neurovascular pathologies, this has not been the case for Alzheimer pathology (14–18). Finally, recent Mendelian randomization studies have also indicated that type 2 diabetes may not be directly causal in Alzheimer dementia (19,20). As much of the data indicating the potential to modify dementia risk in general, and Alzheimer dementia in particular, come from observational epidemiological data (7,8) a greater understanding of how type 2 diabetes differentially contributes to the risk of major dementia subtypes is critical in targeting effective preventative strategies for maximum effect.

Bioresources linked to routinely collected electronic medical records (EMR) are powerful, cost-effective, and flexible research resources for investigating the clinical and molecular basis of a wide range of chronic diseases (21,22). In this context, the dementias represent a challenge because individuals with dementia are less likely to engage in population-based research and many years of follow-up are required for the participating, cognitively intact, individuals to develop dementia. Furthermore, the flexibility afforded by using EMR to passively ascertain phenotypes is often at the cost of reduced specificity. Here again the dementias constitute a particular challenge, as underdiagnosis, underrecording, or imprecise recording of dementia diagnoses in the EMRs is widespread (23).

The GoDARTS bioresource is linked to comprehensive and continuously accruing EMRs with a long available follow-up (24) and is potentially an ideal resource to investigate dementia using an EMR approach. As GoDARTS (Genetics of Diabetes Audit and Research in Tayside Scotland) was developed in collaboration with the Wellcome Trust Case Control Consortium (www.wtccc.org.uk) to investigate the genetics of type 2 diabetes, approximately half of recruits had type 2 diabetes and half did not. This experimental case-control design has been validated by the significant contribution by GoDARTS to a number of major genomic discoveries in type 2 diabetes. Here, we report the use of EMR to identify incident dementia cases in GoDARTS, and through a combination of case note validation and the use of an Alzheimer-specific genetic risk score, we classify case subjects into the major clinicopathological subtypes of Alzheimer/mixed dementia or vascular dementia. Finally, we exploit the original experimental case-control design inherent to GoDARTS to investigate the impact of type 2 diabetes on incident dementia and its major subtypes.

RESEARCH DESIGN AND METHODS

Study Population

The GoDARTS type 2 diabetes case-control study population was recruited from the Tayside region of Scotland between the years 1997 and 2007, although recruitment of only type 2 diabetes case subjects continued to 2012. GoDARTS has previously been described fully (24). In brief, type 2 diabetes case subjects were recruited from diabetes outpatient clinics and general practices, and individuals without type 2 diabetes (control subjects) were pragmatically age and sex matched from corresponding general practices to minimize age, sex, and socioeconomic differences but also by a recruitment drive in individuals >30 years of age from the local population. Participants, who were 99.7% Caucasian, provided a sample of blood for genetic studies and permission for this data to be anonymously linked to their EMR. A Community Health Index (CHI) number is provided to all individuals in Scotland when they register with a primary health care provider and, in Tayside in particular, has been used continuously for all National Health Service (NHS) clinical activity over the past 30 years. The Community Health Index facilitates fully deterministic linkage of electronic clinical data sources to assemble comprehensive "cradle to grave" EMRs for all individuals in the region. The most recent availability of follow-up data used for this study was April 2017.

Ethics Approvals for GoDARTS

The GoDARTS bioresource, and its links to the long-term EMR, is approved by the NHS Tayside Caldicott Guardians, the local research ethics committee, and the Tayside Tissue Bank, and access to the resource is regulated by the GoDARTS Access Group. The EMR is fully anonymized and provided to researchers through robust information governance protocols administered by the Health Informatics Centre (HIC), which functions as a secure portal between the NHS Tayside and the University of Dundee research environment.

Identification of Incident Dementia Cases

Supplementary Table 1 lists the main NHS administrative data sets available in GoDARTS that were used in this study to passively identify individuals with dementia. The major data sources in the EMR for identification of dementia cases comprised hospitalization and death data and dispensed prescribing data. From these sources, we categorized any dementia into the following clinicopathological subtypes: Alzheimer dementia (dementia with an Alzheimer component, e.g., Alzheimer or mixed dementia), vascular dementia, and unclassifiable/other dementia (unspecified dementia). For each individual with dementia, the earliest date of occurrence of evidence of dementia in any of the data sources was taken as a surrogate for the incident date of dementia.

Hospitalization and Death Data

This data are coded using ICD codes from both versions ICD-9 and ICD-10. The following codes were used to determine the earliest date of the occurrence of any code for dementia for each dementia category within the ICD classification system: Alzheimer dementia, 331.0, F00, G30; vascular dementia, 290.2, F01; unspecified dementia, 290.0, 290.1 290.2, 290.3, 294.2, 331.2, F03, G31.1; and other dementias, 046.1, 291.2, 292.82, 294.1, 331.1, 331.11, 331.19, 331.82, 797, A81.0, F02, F05.1, F10.27, F10.97, G31.0, G31.10, G31.09., G31.83.

There is no code for mixed dementia in the ICD classification system. Given the longitudinal nature of the EMR data available to GoDARTS, participants may have multiple, independently coded episodes over time and, consequently may receive codes from different dementia categories due to the inherent imprecision in clinical coding. Furthermore, for time-to-event analysis, the first occurrence of a code for dementia may either be uninformative such as a code for unspecified dementia or be a diagnostically inappropriate code. We therefore made use of the entire longitudinal data available for each individual to determine a consensus code category from all dementia coding episodes for that individual and applied that category to the date of the first (incident) episode. This was achieved for each individual by determining the total number of distinct dementia coding episodes and for each coding category, calculating its proportion of the total. This proportion was used as a weighting to adjudicate the likely diagnostic category for an individual. For example, consider a case of an individual having five separate hospital admissions when a code for dementia was recorded: four for Alzheimer dementia and one with vascular dementia and finally death with a code for unspecified dementia recorded in the death certification. Overall, this individual has four of six (0.66) codes for Alzheimer dementia and one of six (0.33) codes for vascular dementia. In a situation where an individual has both codes for Alzheimer dementia and vascular dementia, a greater weighting (or proportion) of Alzheimer codes compared with vascular dementia codes would indicate that the individual had an Alzheimer component to their dementia. Conversely, an individual with a greater proportion of codes for vascular dementia would be classified as having vascular dementia.

Prescribing Data

In the U.K. NHS, medications for dementia (donepezil hydrochloride, galantamine, rivastigmine, and memantine hydrochloride) are largely indicated for individuals with Alzheimer dementia; therefore, in the absence of other information, a history of redeeming a prescription for one of these drugs was considered to indicate a diagnosis of Alzheimer or mixed dementia. We used community-dispensed prescribing data available in the GoDARTS EMR to determine the first date of redeeming a prescription in the community for one of these drugs.

Adjudication of Alzheimer Dementia by EMR

Individuals were adjudicated as having Alzheimer dementia if they ever had a code for Alzheimer dementia or, in situations where there were codes for both Alzheimer and vascular dementia categories, there was a greater proportion of Alzheimer codes. They were also classified as having Alzheimer dementia if they had ever redeemed a prescription for a dementia-specific drug.

Adjudication of Vascular Dementia

Individuals were adjudicated as having vascular dementia if they ever had a code for vascular dementia or, in situations where there were codes for both vascular dementia and Alzheimer dementia categories, there was a greater proportion of vascular category codes.

Validation Data

Through the use of outpatient referral data, all GoDARTS participants who had been referred to outpatient clinics for Psychiatry of Old Age (POA) were identified. A research nurse (R.B.) reviewed the case notes under the supervision of a POA consultant (P.C.) and a consultant cerebrovascular physician (A.S.F.D.). Information relating to the final diagnosis and date was then reincorporated into the anonymized EMR and contributed to the total dementia cases. These data were also used to validate the cases derived from the hospitalization and death data and prescribing data where individuals had both an EMR-derived diagnosis and a validation data–derived diagnosis. Sensitivity, specificity, accuracy (correctly classified), and the positive predictive value (PPV) of the EMR-derived cases were compared with the validation data diagnosis.

Genetic Validation of Dementia Cases

Genome-wide genotyping in GoDARTS has previously been described (24) and is largely confined to the type 2 diabetes cohort. For those individuals with available genome-wide data, 20 confirmed *APOE4*-independent Alzheimer disease susceptibility single nucleotide polymorphisms (SNPs) previously identified by the International Genomics of Alzheimer's Project (IGAP) (25) were used to construct a weighted genetic risk score (wGRS) by weighting each effect allele with the natural log of the published odds ratio and summing the values for each SNP for each individual. The mean available SNPs per individual was 16.4 (SD 6.3), median 19 (interquartile range 18–20). Missing genotypes were imputed with two times effect allele frequency in the study population, providing a probabilistic value and allowing a full score to be determined for all individuals with genotyping. The wGRS was z-transformed such that each unit step in the score comprised 1 SD of the overall score.

While the E4 allele of *APOE* is a major risk factor for Alzheimer disease (26), it is also associated with vascular dementia (27). Therefore, for further validation we also determined the association of the *APOE4* allele with dementia outcomes using the well-established SNP, rs429358, which defines the *APOE4* allele at the *APOE* locus. This had been previously directly genotyped in the majority of GoDARTS participants and was therefore available in similar numbers of case and control subjects. For investigation with dementia outcomes, a codominant model was used, whereas for the investigation of an interaction with type 2 diabetes a dominant model was used for statistical simplicity.

Analysis of Incident All-Cause Dementia and Major Pathological Subtypes

The study comprised a retrospective matched cohort comparing incidence rates of dementia between individuals in GoDARTS recruited with type 2 diabetes

and individuals without. Entry to the study was the latest of either date of becoming 55 years of age or date of recruitment into GoDARTS. Individuals recruited without type 2 diabetes who were subsequently diagnosed with type 2 diabetes during follow-up were classed as having type 2 diabetes at entry. Data from the Scottish Care Information – Diabetes Collaboration (SCI-DC) (www.sci-diabetes.scot.nhs.uk), which records all new diagnoses of diabetes in Scotland, were used for this purpose. Proportional hazards models were used with analysis time being set from date of birth so that time was continuously adjusted for age in years. Individuals were followed up until the first occurrence in the EMR or in the validation data of a diagnosis of dementia. Censoring was the earliest of date of nondementia death or the end of available EMR data. Individuals with evidence of dementia anywhere in their EMR prior to entry date were excluded from the analysis. Overall and age-specific incident rates of dementia with 95% CIs were determined with all incidence rates quoted as per 1,000 person-years (py). Incidence rate ratios were determined using the Mantel-Haenszel method, which because of the age-based time scale incorporated adjustment for age differences between the two groups. Cox proportional hazards was used to compare hazards of dementia of any cause as well as separately for Alzheimer dementia and vascular dementia. Age at recruitment and sex were included in all models. The STATA, version 13, *st* suite of commands was used for all analyses.

Sensitivity Analysis

To further evaluate our validation of dementia subtypes, we undertook a sensitivity analysis by defining a subgroup of Alzheimer-only dementia cases who had codes exclusively for Alzheimer dementia in the hospitalization and death data or a POA diagnosis of Alzheimer dementia. We also excluded case subjects who were defined uniquely on the basis of community prescribing of Alzheimer-specific medication and therefore had no other available diagnostic information. We assessed the association of the wGRS with this Alzheimer-only dementia group as well as the association of type 2 diabetes status.

Competing Risks Analysis

A recent similar investigation of the impact of type 2 diabetes on risk of dementia (13) using EMR highlighted the importance of accounting for competing risk of death. Because of the shorter life span associated with type 2 diabetes, the competing event of nondementia death can influence the measure of the association of type 2 diabetes with dementia. We therefore similarly accounted for the competing risk of death by implementing Fine-Gray competing risks models (28) using the STATA *stcrreg* procedure. In this case, cause-specific hazard ratios (csHRs) provide estimates of the risk of dementia in the absence of competing risk of death, whereas subdistribution hazard ratios (sdHRs) provide estimates of risk in subjects who have not yet experienced either dementia or nondementia death or who have previously died of a nondementia cause.

RESULTS

Identification and Analysis of Incident Dementia

From a total of 18,283 individuals recruited in GoDARTS, the total number identified with any dementia from all data sources was 1,564. Supplementary Fig. 1 provides a breakdown of where the dementia cases came from for each of the three clinical data sources; 127 (8.1%) came from prescribing data alone and 272 (17.4%) from the validation data alone, and 465 (29.6%) came from the hospitalization and death data alone, and from these only, 2.4% had a mixture of both vascular and Alzheimer dementia codes (0.7% of the total cases). A total of 16,461 individuals fulfilled the entry criteria for the study with reasons for exclusion provided in Supplementary Fig. 2. From these, the total number of incident dementia cases used for the analysis was 1,448.

Table 1 provides the population characteristics for the prospective analysis by type 2 diabetes case-control status. Despite the matching process, the mean age at recruitment was significantly higher in type 2 diabetes case compared with control subjects (66.40 vs. 63.34 years old, respectively) and there were significantly fewer females among the participants with type 2 diabetes (44.1% vs. 49.8%). During a median follow-up time of 8.81 years (interquartile range 5.2–10.7), the overall incidence of any dementia was 10.86 (95% CI 10.32–11.43). After adjustment for the age differences, the overall incidence of any dementia among the type 2 diabetes case subjects was significantly higher, 12.66 (95% CI 11.91–13.45), than among the control subjects, 8.06 (95% CI 7.30–8.86). Supplementary Fig. 3 illustrates the overall incidence rates of any dementia by 10-year age-group, and Supplementary Table 2 provides age-specific incident rates by 10-year age-groups broken down by type 2 diabetes case-control status. This demonstrates increased incidence of any dementia across all age-groups in case subjects with type 2 diabetes, with the greatest incident rate ratios in the younger age ranges.

As in the comparison of age-specific incidence rates, in a Cox model that included age at recruitment and sex, type 2 diabetes strongly predicted incident any dementia (csHR 1.46, 95% CI 1.31–1.64, $P = 6.9 \times 10^{-11}$) (Supplementary Table 3). When we accounted for competing risk of nondementia death, this association was attenuated, though still significant (sdHR 1.26, 95% CI 1.13–1.41, $P = 4.9 \times 10^{-5}$). *APOE4* genotype was a strong predictor of any dementia (HR 1.84, 95% CI 1.67–2.04, $P = 9.6 \times 10^{-34}$), and accounting for competing risks did not greatly affect this association. While we found no evidence of a formal interaction of type 2 diabetes status and *APOE4* genotype in predicting any dementia, analysis of combinations of type 2 diabetes status and *APOE4* genotype (Fig. 1 and Table 2) indicated that compared with individuals with neither type 2 diabetes nor *APOE4* genotype, individuals with both type 2 diabetes and *APOE4* genotype had a risk (HR 2.83, 95% CI 2.38–3.36) that was greater than expected based on the combination of risk from either type 2 diabetes alone (HR 1.42, 95% CI 1.22–1.67) or *APOE4* genotype alone (HR 1.8, 95% CI 1.45–2.24).

Validation and Analysis of Alzheimer Dementia and Vascular Dementia Subtypes

Among the total individuals with any dementia in the EMR in GoDARTS, we identified 793 cases of Alzheimer dementia: 417 case subjects who had vascular dementia and 354 who only had a code for unspecified dementia or only had codes for other types of dementia. Of the 884 individuals in the validation data, 574 individuals (~37% of all dementia cases) also had dementia defined purely from the prescribing data or hospitalization

Table 1—GoDARTS study population with crude incidence rates of any dementia by type 2 diabetes case and control status

	Type 2 diabetes case subjects	Type 2 diabetes control subjects	Overall
Age, mean (SD), years	66.40 (10.28)	63.34 (11.3)*	65.23 (10.80)
Female sex, *n* (%)	4,416 (43.6)	3,187 (49.84)†	7,603 (46.19)
Total dementia events, *n* (%)	1,039 (71.8)	409 (28.3)	1,448
Total time (% total), years	82.09 (61.8)	50.84 (38.3)	132.93
Rate (95% CI)	12.66 (11.91–13.45)	8.05 (7.30–8.86)‡	10.89 (10.35–11.47)
Total individuals, *n* (%)	10,129 (61.5)	6,332 (38.5)	16,461

Total time and rate: per 1,000 py. *$P < 0.001$, *t* test. †$P < 0.001$, χ^2 test. ‡$P < 0.001$, Mantel-Haenszel estimate, controlling for age.

and death data (Supplementary Fig. 1). This subset was used to validate these purely EMR-defined dementia subtypes (Supplementary Table 4). The accuracy and PPV of the EMR-defined dementia categories was compared with the validation data for Alzheimer and vascular dementia and was found to be 82% (PPV 86.4%) and 85% (PPV 72%) respectively (Supplementary Table 5).

Supplementary Table 6 provides age-specific incidence rates of Alzheimer and vascular dementia by type 2 diabetes status together with incident rate ratios overall and by 10-year age-groups. Crude incidence rates for Alzheimer disease are similar between type 2 diabetes case and control subjects except in the 65- to 74-year-old age-category, where the incidence rate of Alzheimer dementia is double in the type 2 diabetes case compared with control subjects. The incidence rate of vascular dementia was consistently higher across all age categories in the type 2 diabetes case subjects compared with control subjects, although there were no cases of vascular dementia among the control subjects without type 2 diabetes in the youngest (55–64 years old) age range.

Genetic Validation of Incident Major Pathological Subtypes

In a Cox model with adjustment, as before, for age and sex (Table 3), wGRS significantly predicted incidence of Alzheimer dementia (csHR 1.23 per SD, 95% CI 1.12–1.34, $P = 8.4 \times 10^{-6}$), and as might be expected there was little evidence of an attenuation of this association when competing risks of death were accounted for. In the sensitivity analysis considering the Alzheimer-only dementia cases, the association of the wGRS was slightly stronger (csHR 1.27, 95% CI 1.14–1.42, $P = 2.8 \times 10^{-5}$) compared with the association with overall Alzheimer dementia cases. There was no association of the wGRS with vascular dementia (csHR 1.02, 95% CI 0.91–1.15, $P = 7.0 \times 10^{-1}$). *APOE4* genotype was strongly predictive of both major dementia subtypes, with the association with Alzheimer dementia being larger and more significant (csHR 2.31, 95% CI 2.03–2.63, $P = 4.4 \times 10^{-36}$) than the association with vascular dementia (csHR 1.59, 95% CI 1.31–1.94, $P = 4.0 \times 10^{-6}$). Again, there appeared to be little impact of accounting for competing risk of death with respect to *APOE4* genotype.

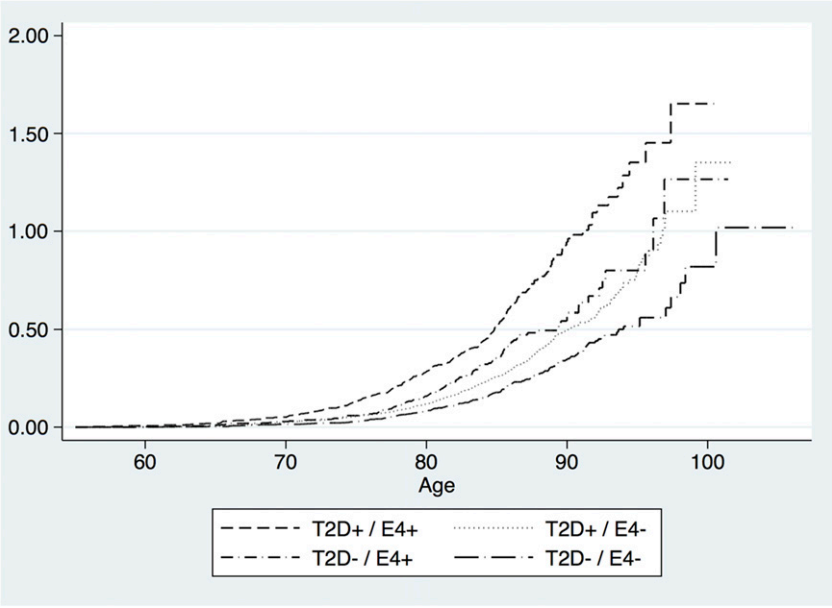

Figure 1—Cumulative incidence of any dementia by combinations of *APOE4* genotype and type 2 diabetes status. T2D+/E4+, type 2 diabetes case subject with *APOE4* genotype; T2D+/E4−, type 2 diabetes case subject without *APOE4* genotype; T2D−/E4+, type 2 diabetes control subject with *APOE4* genotype; T2D−/E4−, type 2 diabetes control subject without *APOE4* genotype. Age is in years.

Association of Type 2 Diabetes Status With Major Dementia Subtypes

Based on the data in Supplementary Table 2 our study had 80% power with $\alpha \leq 0.5\%$ for a minimum detectable HR of 1.29 for vascular dementia by type 2 diabetes status. In fact, we found that type 2 diabetes status was strongly associated with incident vascular dementia (csHR 2.47, 95% CI 1.92–3.18, $P = 2.2 \times 10^{-12}$), with evidence of this association being abrogated by accounting for competing risks (sdHR 2.13, 95% CI 1.67–2.75, $P = 2.6 \times 10^{-9}$). The minimal detectable HR for Alzheimer dementia by type 2 diabetes status was 1.21, and in our study we obtained an estimate that was a slightly lower than this (csHR 1.16, 0.99–1.35, $P = 0.06$). However, accounting for competing risks completely abrogated this association (sdHR 1.02, 95% CI 0.87–1.18, $P = 0.85$). In the sensitivity analysis in consideration of Alzheimer-only dementia cases, as expected, type 2

Table 2—Hazard of any dementia by combinations of *APOE4* genotype and type 2 diabetes status (adjusted for age and sex)

	Cox (csHR)			Competing risks (sdHR)		
	HR	95% CI	*P*	HR	95% CI	*P*
T2D−/E4−	—	—	—	—	—	—
T2D+/E4−	1.42	1.22–1.67	1.3×10^{-5}	1.22	1.05–1.43	0.012
T2D−/E4+	1.80	1.45–2.24	9.6×10^{-8}	1.72	1.39–2.13	6.5×10^{-7}
T2D+/E4+	2.83	2.38–3.36	2.9×10^{-32}	2.38	2.01–2.82	1.4×10^{-23}

T2D+/E4+, type 2 diabetes case subject with *APOE4* genotype; T2D+/E4−, type 2 diabetes case subject without *APOE4* genotype; T2D−/E4+, type 2 diabetes control subject with *APOE4* genotype; T2D−/E4−, type 2 diabetes control subject without *APOE4* genotype.

diabetes status remained similarly an insignificant predictor (csHR 1.15, 95% CI 0.91–1.40, $P = 0.17$) compared with overall Alzheimer cases.

CONCLUSIONS

Our study has demonstrated the feasibility of undertaking clinically relevant long-term prospective studies of dementia incidence in the GoDARTS bioresource making use of EMRs in the Scottish NHS. In particular, it seems possible to successfully differentiate and investigate the major clinicopathological dementia subtypes of Alzheimer dementia and vascular dementia using this approach. Exploiting the experimental type 2 diabetes case-control design inherent in GoDARTS, we have robustly demonstrated that type 2 diabetes is not associated with Alzheimer dementia, the most common pathological dementia subtype, particularly when accounting for competing risk of death, although it is strongly associated with vascular dementia and therefore with all-cause dementia.

The overall effectiveness of passive dementia case finding using EMR depends on the comprehensiveness of health care cover (29). In the U.K. NHS, all health care is free at the point of access and equally available across the entire sociodemographic spectrum with well-developed routine administrative health care data collection. NHS data sets available to GoDARTS are particularly comprehensive and detailed, being available continuously for all participating individuals over a 30-year period. While the precision of defining EMR dementia cases by EMR has been found to be highly variable, the availability of multiple data sources, as we have used, seems to result in the highest PPV in defining dementia clinicopathological subtypes using EMR (30). Despite this, the overall incidence rate of all-cause dementia that we found in GoDARTS is lower compared with other studies. A meta-analysis of studies of dementia incidence from Europe and the U.S. (31) found a pooled estimate of 17.2/1,000 py for individuals >60 years of age (for comparison, in GoDARTS the incidence for the same age cutoff was 12.6/1,000 py). It was acknowledged that there was high heterogeneity of studies included in that analysis. A study from the U.K. (2) from a similar time period and in which there was active and full ascertainment of dementia cases estimated a rate of 17.7/1,000 py in individuals >65 years of age, which was closer to the rate in GoDARTS for the same age cutoff at 15.3/1,000 py. Incidence rates in GoDARTS were again generally lower across all age bands compared with a similar type 2 diabetes case-cohort study from Australia (13), which also made use of EMR. The passive ascertainment approach that we used will have certainly resulted in incomplete ascertainment as a result of individuals who develop dementia not appearing in the EMR. A further important reason for the lower incidence estimation may be the well-known healthy volunteer bias associated with recruitment into bioresources in general, as has been the case with UK Biobank, for example. Typical proportions of Alzheimer dementia and vascular dementia in Western populations are ~55% and ~20%, respectively, and the greater proportion of individuals with vascular dementia (26.7%) and a lower proportion (50%) with Alzheimer dementia in GoDARTS probably reflect the high proportion of individuals with type 2 diabetes in our study.

Case note validation from POA outpatient diagnoses of our EMR-defined dementia subtypes of Alzheimer and vascular dementia indicated a high PPV consistent with other studies with access to multiple data sources (30). Alzheimer disease cases in IGAP studies, from which the wGRS was derived, were rigorously defined (25). Thus, the strong association of the wGRS with Alzheimer disease cases provides robust independent support for the clinicopathological specificity for the Alzheimer dementia cases identified in our study. Furthermore, our point estimate of the HR per SD of the wGRS was very close to the global estimate from the IGAP cohorts (32). Similarly, the lack of an association of the wGRS with vascular dementia in GoDARTS implies the absence of Alzheimer pathology in vascular dementia cases. Conversely, *APOE4* genotype was strongly associated with both Alzheimer dementia and vascular dementia. While the association of *APOE4* genotype with Alzheimer pathology has been consistent and well established (26), its association with vascular dementia has been more variable and generally weaker, although recent meta-analyses have all supported its association with vascular dementia (27,33,34), suggesting potentially a common molecular link between

Table 3—csHRs (Cox) and sdHRs (competing risk) of vascular dementia and Alzheimer dementia by type 2 diabetes, *APOE4* genotype, and wGRS

	Vascular dementia			Alzheimer dementia		
	csHR/sdHR	95% CI	*P*	csHR/sdHR	95% CI	*P*
T2D	2.47	1.92–3.18	2.2×10^{-12}	1.16	0.99–1.35	0.06
T2D CR	2.13	1.67–2.75	2.6×10^{-9}	1.02	0.87–1.18	0.85
APOE4	1.59	1.31–1.94	4.0×10^{-6}	2.31	2.03–2.63	4.4×10^{-36}
APOE4 CR	1.54	1.26–1.89	2.6×10^{-5}	2.24	1.97–2.55	5.0×10^{-35}
wGRS	1.02	0.91–1.15	0.70	1.23	1.12–1.34	8.4×10^{-6}
wGRS CR	1.01	0.90–1.14	0.87	1.21	1.10–1.33	5.1×10^{-5}

Adjusted for age at entry and sex. CR, competing risks analysis.

neurovascular and Alzheimer pathologies (35). Therefore, the combination of the negative association of the wGRS and the positive association of *APOE4* genotype with our vascular dementia phenotype similarly provides additional independent support for the specificity of our vascular dementia phenotype. Such independent support for EMR-defined clinicopathological phenotypes robustly supports our finding of a lack of association of type 2 diabetes with Alzheimer dementia. Furthermore, our sensitivity analysis in which we considered Alzheimer-only dementia cases did not have a notable impact on our findings.

Age-specific incidence rates from our study indicated that in the 65- to 74-year-old age band, the incidence rate of Alzheimer dementia in patients with type 2 diabetes was twice the rate in control subjects without type 2 diabetes, whereas in all other age categories there was no difference. Analysis of this age-group compared with other age-groups did not demonstrate any notable evidence of a difference in the proportion of individuals validated as having Alzheimer dementia, or a difference in wGRS. Furthermore, Cox regression analysis with adjustment for age at recruitment and sex also demonstrate a weak, nonsignificant trend for an increased cause-specific association with Alzheimer dementia. This may indicate an element of age-of-onset dependency of type 2 diabetes and Alzheimer dementia, with younger-onset dementia driven by type 2 diabetes being more likely to have an Alzheimer component. Further studies in larger populations would be required to explore this possibility. However, this weak association between type 2 diabetes and Alzheimer dementia was completely abrogated when we accounted for competing risk of death. This reduction in strength of association in the subdistribution hazard in consideration of competing risks of death in the association of type 2 diabetes and dementia risk has been found in other studies (13). Together these findings underscore the importance of accounting for the increased mortality in patients with type 2 diabetes when considering its links with late-life–onset disease such as Alzheimer dementia. Other factors such as differences between type 2 diabetes case and control subjects in multiple deprivation status, smoking, or indeed epoch of diagnosis, and other potential risk factors, may also be important.

Links between type 2 diabetes and all-cause dementia seem to be consistent and are not in doubt and increasingly appear to be mediated through the relationship of type 2 diabetes with neurovascular pathologies (15). On the other hand, the pathoaetiological relationship between type 2 diabetes and Alzheimer dementia is likely to be complex. Type 2 diabetes is itself emerging as a highly heterogeneous disease (36), and polygenic risk scores, for example, have indicated that insulin resistance might be more important than type 2 diabetes overall in determining the risk of Alzheimer dementia (19). Furthermore, treatments used in the management of type 2 diabetes may also have an influence. For example, in the elderly, treatments aimed at β-cell failure may predispose to hypoglycemia, which may, in turn, predispose to dementia (37), while insulin-sensitizing agents such as metformin and thiazolidinediones may have a protective role (38,39).

There is clearly much more research needed to dissect the relationship between type 2 diabetes and the dementias. The versatility and flexibility of bioresources linked to comprehensive EMR provide powerful opportunities to undertake a wide array of informative investigations into the genomic and clinical determinants of dementia risk. Our study underscores the feasibility of this approach. As a greater number of disease-specific variants are detected through genome-wide association studies, increasingly powerful and clinically useful genomic instruments are rapidly becoming available. This, in turn, leads to greater ability to exploit such instruments to increase precision of phenotype definition in the EMR, as we have done here. It also allows for much needed future adequately powered Mendelian randomization studies using those phenotypes (40) to further understand the potential for reducing the future impact of dementia through risk factor modification.

Acknowledgments. The authors acknowledge all the GoDARTS participants and all personnel at the University of Dundee who have helped with its creation and maintenance. The authors also acknowledge HIC for the provision of anonymized clinical data for GoDARTS from NHS Tayside, the original data owner.

Funding. The GoDARTS bioresource was funded largely by Wellcome Trust (072960/Z/03/Z, 084726/Z/08/Z, 084727/Z/08/Z, 085475/Z/08/Z, and 085475/B/08/Z) and as part of the European Union IMI-SUMMIT program (SUrrogate markers for Micro- and Macro-vascular hard endpoints for Innovative diabetes Tools from the Innovative Medicines Initiative) (www.imi-summit.eu). The case note validation component of the current study was funded by the Engineering and Physical Sciences Research Council (EP/M005976/1).

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

Author Contributions. A.S.F.D. undertook the analysis and drafted the manuscript. W.B. worked with E.J. in HIC to process and provide relevant data sets. R.B. and P.C. were involved in case note validation. A.S.F.D., W.B., E.J., K.E.W., R.B., E.T., P.C., R.J.M., and C.N.A.P. were involved in scientific discussions and cowriting and editing the final manuscript. A.S.F.D. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Prince M, Wimo A, Guerchet M, et al. World Alzheimer report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends [Internet], 2015. London, Alzheimer's Disease International. Available from https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf. Accessed 10 May 2018

2. Matthews FE, Stephan BC, Robinson L, et al.; Cognitive Function and Ageing Studies (CFAS) Collaboration. A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies I and II. Nat Commun 2016;7:11398

3. Schilling MA. Unraveling Alzheimer's: making sense of the relationship between diabetes and Alzheimer's disease1. J Alzheimers Dis 2016;51:961–977

4. Gudala K, Bansal D, Schifano F, Bhansali A. Diabetes mellitus and risk of dementia: a meta-analysis of prospective observational studies. J Diabetes Investig 2013;4:640–650

5. Profenno LA, Porsteinsson AP, Faraone SV. Meta-analysis of Alzheimer's disease risk with obesity, diabetes, and related disorders [article online]. Biol Psychiatry 2010;67:505–512. Available from https://www.sciencedirect.com/science/article/pii/S0006322309002261?via%3Dihub. Accessed 8 January 2019

6. Baumgart M, Snyder HM, Carrillo MC, Fazio S, Kim H, Johns H. Summary of the evidence on modifiable risk factors for cognitive decline and dementia: a population-based perspective [article online]. Alzheimers Dement 2015;11:718–726. Available from https://www.sciencedirect.com/science/article/pii/S1552526015001971. Accessed 13 December 2018

7. Norton S, Matthews FE, Barnes DE, Yaffe K, Brayne C. Potential for primary prevention of Alzheimer's disease: an analysis of population-based data [article online]. Lancet Neurol 2014;13:788–794. Available from https://

www.sciencedirect.com/science/article/pii/S147444221470136X?via%3Dihub. Accessed 13 December 2018
8. Scheltens P, Blennow K, Breteler MM, et al. Alzheimer's disease [article online]. Lancet 2016;388:505–517. Available from https://www.sciencedirect.com/science/article/pii/S0140673615011241. Accessed 13 December 2018
9. O'Brien JT, Thomas A. Vascular dementia [article online]. Lancet 2015;386:1698–1706. Available from https://linkinghub.elsevier.com/retrieve/pii/S0140673615004638. Accessed 29 April 2019
10. Mittal K, Katare DP. Shared links between type 2 diabetes mellitus and Alzheimer's disease: a review [article online]. Diabetes Metab Syndr 2016;10(Suppl. 1):S144–S149. Available from https://www.sciencedirect.com/science/article/pii/S1871402115300709?via%3Dihub. Accessed 7 January 2019
11. MacKnight C, Rockwood K, Awalt E, McDowell I. Diabetes mellitus and the risk of dementia, Alzheimer's disease and vascular cognitive impairment in the Canadian Study of Health and Aging. Dement Geriatr Cogn Disord 2002;14:77–83
12. Akomolafe A, Beiser A, Meigs JB, et al. Diabetes mellitus and risk of developing alzheimer disease: results from the Framingham Study [article online]. Arch Neurol. 2006;63:1551–1555. Available from http://archneur.jamanetwork.com/article.aspx?doi=10.1001/archneur.63.11.1551. Accessed 29 June 2018
13. Davis WA, Zilkens RR, Starkstein SE, Davis TME, Bruce DG. Dementia onset, incidence and risk in type 2 diabetes: a matched cohort study with the Fremantle Diabetes Study Phase I. Diabetologia 2017;60:89–97
14. Dos Santos Matioli MNP, Suemoto CK, Rodriguez RD, et al. Diabetes is not associated with Alzheimer's disease neuropathology. J Alzheimers Dis 2017;60:1035–1043
15. Abner EL, Nelson PT, Kryscio RJ, et al. Diabetes is associated with cerebrovascular but not Alzheimer's disease neuropathology. Alzheimers Dement 2016;12:882–889
16. Sonnen JA, Larson EB, Brickell K, et al. Different patterns of cerebral injury in dementia with or without diabetes. Arch Neurol 2009;66:315–322
17. Beeri MS, Silverman JM, Davis KL, et al. Type 2 diabetes is negatively associated with Alzheimer's disease neuropathology. J Gerontol A Biol Sci Med Sci 2005;60:471–475
18. Javanshiri K, Waldö ML, Friberg N, et al. Atherosclerosis, hypertension, and diabetes in Alzheimer's disease, vascular dementia, and mixed dementia: prevalence and presentation. J Alzheimers Dis 2018;65:1247–1258
19. Walter S, Marden JR, Kubzansky LD, et al. Diabetic phenotypes and late-life dementia risk: a mechanism-specific Mendelian randomization study. Alzheimer Dis Assoc Disord 2016;30:15–20
20. Østergaard SD, Mukherjee S, Sharp SJ, et al.; Alzheimer's Disease Genetics Consortium; GERAD1 Consortium; EPIC-InterAct Consortium. Associations between potentially modifiable risk factors and Alzheimer disease: a Mendelian randomization study. PLoS Med 2015;12:e1001841; discussion e1001841
21. McCarty CA, Chisholm RL, Chute CG, et al. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. BMC Med Genomics 2011;4:13
22. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779
23. Cummings E, Maher R, Showell CM, et al. Hospital coding of dementia: is it accurate? Health Inf Manag 2011;40:5–11
24. Hébert HL, Shepherd B, Milburn K, et al. Cohort profile: Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS) [article online]. Int J Epidemiol 2018;47:380–381j. Available from http://academic.oup.com/ije/article/doi/10.1093/ije/dyx140/4107246/Cohort-Profile-Genetics-of-Diabetes-Audit-and. Accessed 16 January 2018
25. Lambert J-C, Ibrahim-Verbaas CA, Harold D, et al.; European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45:1452–1458
26. Michaelson DM. APOE ε4: the most prevalent yet understudied risk factor for Alzheimer's disease [article online]. Alzheimers Dement 2014;10:861–868. Available from https://www.sciencedirect.com/science/article/pii/S1552526014024996?via%3Dihub. Accessed 7 January 2019
27. Skrobot OA, McKnight AJ, Passmore PA, et al.; Genetic and Environmental Risk for Alzheimer's disease Consortium (GERAD1). A validation study of vascular cognitive impairment genetics meta-analysis findings in an independent collaborative cohort [article online]. J Alzheimers Dis 2016;53:981–989. Available from http://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JAD-150862. Accessed 5 December 2018
28. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509
29. Knopman DS, Petersen RC, Rocca WA, Larson EB, Ganguli M. Passive case-finding for Alzheimer's disease and dementia in two U.S. communities. Alzheimers Dement 2011;7:53–60
30. Wilkinson T, Ly A, Schnier C, et al.; UK Biobank Neurodegenerative Outcomes Group and Dementias Platform UK. Identifying dementia cases with routinely collected health data: a systematic review. Alzheimers Dement 2018;14:1038–1051
31. Fiest KM, Jetté N, Roberts JI, et al. The prevalence and incidence of dementia: a systematic review and meta-analysis [article online]. Can J Neurol Sci 2016;43(Suppl. 1):S3–S50. Available from http://www.journals.cambridge.org/abstract_S0317167116000184. Accessed 7 January 2019
32. Chouraki V, Reitz C, Maury F, et al.; International Genomics of Alzheimer's Project. Evaluation of a genetic risk score to improve risk prediction for Alzheimer's disease. J Alzheimers Dis 2016;53:921–932
33. Rohn TT. Is apolipoprotein E4 an important risk factor for vascular dementia? Int J Clin Exp Pathol 2014;7:3504–3511
34. Yin Y-W, Li J-C, Wang J-Z, et al. Association between apolipoprotein E gene polymorphism and the risk of vascular dementia: a meta-analysis [article online]. Neurosci Lett 2012;514:6–11. Available from https://www.sciencedirect.com/science/article/pii/S0304394012002182?via%3Dihub. Accessed 15 November 2018
35. Lin Y-F, Smith AV, Aspelund T, et al. Genetic overlap between vascular pathologies and Alzheimer's dementia and potential causal mechanisms. Alzheimers Dement 2019;15:65–75
36. McCarthy MI. Painting a new picture of personalised medicine for diabetes [article online]. Diabetologia 2017;60:793–799. Available from http://link.springer.com/10.1007/s00125-017-4210-x. Accessed 16 January 2019
37. Meneilly GS, Tessier DM. Diabetes, dementia and hypoglycemia. Can J Diabetes 2016;40:73–76
38. Orkaby AR, Cho K, Cormack J, Gagnon DR, Driver JA. Metformin vs sulfonylurea use and risk of dementia in US veterans aged ≥65 years with diabetes. Neurology 2017;89:1877–1885
39. Lu C-H, Yang C-Y, Li C-Y, Hsieh C-Y, Ou H-T. Lower risk of dementia with pioglitazone, compared with other second-line treatments, in metformin-based dual therapy: a population-based longitudinal study [article online]. Diabetologia 2018;61:562–573. Available from http://link.springer.com/10.1007/s00125-017-4499-5. Accessed 8 January 2019
40. Kuźma E, Hannon E, Zhou A, et al. Which risk factors causally influence dementia? A systematic review of Mendelian randomization studies [article online]. J Alzheimer's Dis 2018;64:181–193. Available from https://content.iospress.com/download/journal-of-alzheimers-disease/jad180013?id=journal-of-alzheimers-disease%2Fjad180013. Accessed 19 June 2018