Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS   Check for updates

Type 2 diabetes and cognitive performance in middle age: a cross-sectional study

Teppo Sola[a,b], Elina Pimiä[b,c], Elina Lahti[b], Jorma Lahtela[b,d] and Mervi Jehkonen[a,b]

[a]Faculty of Social Sciences, Tampere University, Tampere, Finland; [b]Tays Research Services, Wellbeing Services County of Pirkanmaa, Tampere University Hospital, Tampere, Finland; [c]Diabetes Outpatient Clinic, Wellbeing Services County of Pirkanmaa, City of Tampere Diabetes Outpatient Clinic, Tampere, Finland; [d]Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

ABSTRACT

Introduction: Type 2 diabetes has been associated with cognitive decrements already in middle-age. However, the sample sizes of the studies have been small and the neuropsychological tests used have been heterogeneous. In addition, only a few studies have matched the groups in terms of age, education and gender. In this cross-sectional matched pairs study, we investigated the cognitive performance of Finnish middle-aged type 2 diabetes patients compared to healthy individuals.

Method: A neuropsychological test battery consisting of 16 tests and 21 outcome measures was applied to 28 patients and 28 age-, education- and gender-matched healthy individuals. Various exclusion criteria were applied to minimize the risk of cognitive dysfunction due to factors other than diabetes.

Results: We did not find between-group differences in any of the neuropsychological tests measuring attention, concept formation and reasoning, construction and motor performance, executive functions, memory, processing speed or working memory. In addition, there were no group differences in the frequency or severity of subjective cognitive symptoms, or in anxiety, depression, burnout, fatigue or alcohol use disorder symptoms. The effect sizes in this study were mostly negligible or small, with the mean effect size being −0.12.

Conclusions: In a carefully matched sample of middle-aged type 2 diabetes patients and healthy individuals, we found no significant effects and no meaningful evidence of cognitive differences between the groups.

Introduction

Type 2 diabetes mellitus (T2DM) is a disease characterized by chronic high blood glucose levels (hyperglycemia) that result from variable degrees of insulin resistance and secretory defect (e.g., Alberti & Zimmet, 1998). In 2021, approximately 537 million people worldwide had diabetes, with T2DM accounting for over 90% of all cases, and the prevalence of the disease is predicted to rise to 643 million by 2030 (Magliano et al., 2021). Globally, diabetes is one of the top 10 causes of death, with over 6.7 million adults (aged 20 to 79) predicted to die from diabetes-related causes in 2021 (Magliano et al., 2021). Long-term complications include retinopathy, nephropathy, neuropathy and autonomic dysfunction, as well as a risk of cardiovascular, peripheral vascular and cerebrovascular disease (e.g., Alberti & Zimmet, 1998). In older people with diabetes, risk ratios for developing vascular dementia or Alzheimer's disease

are approximately 2.3 and 1.6, respectively (Gudala et al., 2013).

Multiple systematic reviews and meta-analyses have shown that people with T2DM perform worse than healthy individuals in neuropsychological tests. Effect sizes have been mostly small or medium, with the most affected domains being processing speed or psychomotor speed (Kálcza-Jánosi et al., 2013; Mansur et al., 2018; Monette et al., 2014; Palta et al., 2014; Pelimanni & Jehkonen, 2019; Sadanand et al., 2016), executive functions (Kálcza-Jánosi et al., 2013; Mansur et al., 2018; Monette et al., 2014; Palta et al., 2014; Pelimanni & Jehkonen, 2019; Sadanand et al., 2016; Vincent & Hall, 2015), attention and concentration (Kálcza-Jánosi et al., 2013; Monette et al., 2014; Palta et al., 2014; Pelimanni & Jehkonen, 2019; Vincent & Hall, 2015) and memory (Mansur et al., 2018; Monette et al., 2014; Palta et al., 2014; Pelimanni & Jehkonen, 2019; Sadanand et al., 2016; Vincent & Hall, 2015). The few meta-analyses that have examined language or visuospatial functions

have found small statistically significant effect sizes (visual: Kálcza-Jánosi et al., 2013; Monette et al., 2014; Pelimanni & Jehkonen, 2019; language: Kálcza-Jánosi et al., 2013; Pelimanni & Jehkonen, 2019).

In most meta-analyses the mean age of patients has been around 70 years, and only few studies have investigated the presence and nature of cognitive deficits in middle age. To our knowledge, one meta-analysis exists where the neuropsychological performance of middle-aged people with T2DM is compared to healthy individuals (Pelimanni & Jehkonen, 2019). Patients performed worse in most of the cognitive domains assessed, with medium effect sizes found for processing speed, attention/concentration, working memory and executive functions; small effect sizes for verbal memory, language and perception/construction; and a negligible and statistically nonsignificant effect size for visual memory. Only 12 original studies on middle-aged patients were identified for the meta-analysis and most of them had small sample sizes ranging from 13 to 50 patients, except for one major study that included 1779 patients (Rawlings et al., 2014). However, that study contributed to only three cognitive domains, with one test used per each domain. Some major cognitive domains in which group differences have been observed in older patients, namely attention and working memory, have been studied in only five original studies each, and these have included a maximum of 50 patients. Many of these studies have not controlled for age, education and gender either by matching or by using statistical adjustments. In addition, few original studies have assessed subjective cognitive or psychosocial symptoms. The small number of studies assessing the cognitive performance of middle-aged T2DM patients using multiple neuropsychological tests, adjusting for confounders and examining subjective cognitive and psychosocial symptoms, serves as a motivation for this study.

In this cross-sectional matched pairs study we examine which cognitive domains, if any, are affected in middle-aged Finnish T2DM patients as compared to age-, education- and gender-matched healthy individuals. We use a neuropsychological test battery, emphasizing the domains where previous studies have found differences between patient and healthy comparison groups. Based on the meta-analysis by Pelimanni and Jehkonen (2019), we hypothesize that middle-aged T2DM patients exhibit poorer cognitive performance than healthy individuals in the following cognitive domains as classified by Lezak et al. (2012): attention, concept formation and reasoning, construction and motor performance, executive functions, memory, processing speed and working memory. Our secondary hypothesis is that the patient group will report more

subjective cognitive symptoms than the healthy comparison group. This hypothesis is based on a study by Faiz et al. (2021), in which middle-aged T2DM patients had more subjective memory complaints than healthy individuals, although a smaller study did not find a group difference in subjective cognitive symptoms (Aberle et al., 2008).

Methods

Participants

T2DM patients were recruited from City of Tampere Diabetes Outpatient Clinic, during a four-year period from 1 June 2018 to 31 May 2022. We included male and female T2DM patients and healthy individuals aged 35–65. The exclusion criteria for patients were as follows: type 1 diabetes, diabetes diagnosed less than a year ago, no diabetes medication, diabetes medication for less than one year, pregnancy, hypothyroidism, neurological or psychiatric disorder, developmental neuropsychiatric disorder, severe obstructive sleep apnea, severe late complication of diabetes (dialysis treatment or severe visual impairment), severe sugar imbalance resulting in HbA1c ≥100 mmol/mol, and substance abuse. Every patient who met the inclusion criteria upon visiting the clinic during the recruitment period received written and oral information about the study. The healthy comparison group was recruited by a newspaper advertisement published in a local free distribution magazine on 25 September 2021. The exclusion criteria for the healthy comparison group were type 1 or type 2 diabetes, pregnancy, hypothyroidism, neurological or psychiatric disorder, developmental neuropsychiatric disorder, severe obstructive sleep apnea and substance abuse. The exclusion criteria were chosen to minimize the risk of cognitive dysfunction due to dementia or factors other than T2DM. Pregnancy was selected as an exclusion criteria because during pregnancy, blood sugar balance is regulated very precisely and blood sugar values can be even better during pregnancy than at other times. Fatigue and nausea can also occur during pregnancy, which can impair performance in cognitive tests. In addition, follow-up during pregnancy is organized in a different unit than the unit recruiting for the study.

To overcome bias caused by potential confounding variables, we applied one-to-one Mahalanobis distance matching without replacement, using age, education and gender as confounders. This method was chosen because propensity score matching methods would have resulted either in worse balance or an unnecessarily small sample size. A caliper of 2.5 years was used for education and a caliper of 5 years for age. The caliper for education was

chosen to ensure that people with no more than compulsory basic education (9 years) could not be matched with people with a secondary level education (12 years) or higher, and to ensure that people with an academic degree could only be matched with each other. The caliper for age was chosen to minimize the age difference between each individual matched pair without leaving an unreasonably small sample. After applying the eligibility criteria and matching, we had a total sample size of 28 matched pairs. Twenty patients and 24 healthy individuals were left without a match and excluded from all analyses. In the matched data, standardized mean differences for the covariates were below 0.1. The average age was approximately 55 years and the average years of education approximately 13 years in both matched groups. Table 1 presents descriptive statistics for the background variables of the matched and unmatched patients and healthy individuals. All participants were of Finnish ethnic origin.

Sampling procedure

All patients and healthy individuals who met the inclusion criteria and who were willing to participate were recruited in this study. A non-probability, purposive, self-selection sampling method was used. The healthy comparison group underwent a medical examination with laboratory tests to rule out diabetes and hypothyroidism. For patients, corresponding medical data were drawn from medical records based on their most recent annual diabetes checkup. Both groups underwent a neuropsychological assessment. The medical examination was conducted at the City of Tampere Diabetes Outpatient Clinic, the laboratory tests were taken at Fimlab Laboratories Oy Ltd and the neuropsychological assessment was carried out at Tampere University. All assessments were conducted between 13 June 2018 and 1 June 2022. Participants were given no compensation or reward for their participation. The research project was approved by the Tampere University Hospital Ethics Committee (ETL: R16126, date: 14.10.2016). The Declaration of Helsinki was followed throughout the study and the participants were recruited voluntarily based on their informed consent.

Power analysis

An a priori power analysis using G*Power software version 3.1.9.6 (Faul et al., 2007) indicated that the required sample size was 51 matched pairs for paired-samples t-test to achieve 80% power for detecting a medium effect size ($d = 0.50$ as classified by Cohen [1988]) at a significance criterion of $\alpha = .01$. Due to financial and time constraints we stopped data collection before this requirement was fulfilled. Post hoc power analyses with G*Power revealed that a power of 46% was achieved ($\alpha = .01$) in the study.

Measures

Diabetes status and medical measures
The primary independent variable in this study was diabetes status (T2DM or no diabetes) as determined by two endocrinologists (E.P & J.L) based on HbA1c levels and medical history. HbA1c levels higher than 48 mmol/mol (6.5%) were considered to indicate the presence of diabetes. The laboratory tests performed were complete blood count, sodium, potassium, creatinine, thyroid-stimulating hormone, thyroxine, glucose, HbA1c and lipids. Patients were asked to measure their blood glucose at the beginning and end of the neuropsychological assessment.

Background variables and confounders
The research team constructed a questionnaire that included closed and open-ended questions regarding education, school success, medical history, cognition, occupation and life situation, family situation, social relationships, exercise, sleep habits, tobacco and substance use, diet and hobbies. Patients were asked additional questions regarding diabetes anamnesis, form of treatment, frequency of blood sugar testing and history of hypoglycemia. The purpose of the questionnaire was to further confirm the eligibility of the participants and to collect background variables relevant to our research questions. Age, education in years and gender were considered confounders because there are well known associations between these variables and cognitive performance.

Questionnaires and neuropsychological outcome measures
All the neuropsychological tests used, their cognitive domains and performance measures are presented in Table 2. The neuropsychological assessment consisted of Beck Anxiety Inventory (A. T. Beck et al., 1988), Beck Depression Inventory-II (A. Beck et al., 1996), Fatigue Impact Scale (Fisk et al., 1994), Bergen Burnout Indicator 15 (Näätänen et al., 2003), Alcohol Use Disorders Identification Test (Babor et al., 1992), the aforementioned research questionnaire and standardized neuropsychological tests that were assigned to cognitive domains and their subdomains according to a widely used classification by Lezak et al. (2012). Three tests, namely d2-R (Brickenkamp et al., 2016), Four Word Short-Term Memory Test (STMT [Butters & Cermack, 1980]) and Vilkki Dual Task (Vilkki et al., 1996), are not explicitly mentioned by Lezak et al. (2012), and they were classified by two specialized

Table 1. Participant characteristics.

Group	n	Age	Gender (% female)	Education in years	Working or studying (%)	Body mass index	Waist circumference (cm)	Blood pressure (SBP/DBP)	Cholesterol (HDL/LDL; mmol/l)	Triglyceride (mmol/l)	Disease duration in years	HbA1c (mmol/mol)
Matched patients	28	54.82	61	13.39	75	32.91	111.40	140.00/84.71	1.20/2.84	2.15	10.54	53.86
Matched healthy individuals	28	54.75	61	13.38	82	26.60	93.28	133.70/79.21	1.60/3.09	1.15	-	36.79
Unmatched patients	20	54.30	5	13.62	75	33.78	120.8	142.70/82.95	1.13/2.21	2.36	9.80	57.40
Unmatched healthy individuals	24	50.54	100	14.98	79	26.60	87.64	130.30/78.21	1.71/2.97	0.96	-	36.83

Note: All numbers are mean values unless stated otherwise. DBP = diastolic blood pressure, HbA1c = a hemoglobin A1c, HDL = high-density lipoprotein, LDL = low-density lipoprotein, SBP = systolic blood pressure.

Table 2. Cognitive domains, subdomains, neuropsychological tests used and outcome measures.

Cognitive domain	Subdomain	Neuropsychological test	Outcome
Attention	Attentional capacity	WAIS-IV Digit Span Forward	Total raw score
	Complex attention	WAIS-IV Coding	Total correct
	Divided attention	Trail Making Test B (TMT B)	Time to complete
	Focused attention	Vilkki Dual Task	Worse percentage
		CPT 3	HRT change
		d2-R	Total correct minus errors
		Stroop: interference	Total correct
Concept formation and reasoning	Verbal concept formation	WAIS-IV Similarities	Total raw score
Construction and motor performance	Assembling and building	WAIS-IV Block Design	Total raw score
	Drawing	Rey-Osterrieth Complex Figure Test (ROCF)	Copying score
Executive functions	Planning and decision making	Tower of London (ToL)	Total move score
	Self regulation	Phonemic Fluency: P-A-S	Total raw score
		Semantic Fluency: animals	Total raw score
Memory	Verbal	Rey Auditory Verbal Learning Test (RAVLT)	Trial 1–5 total
	Visual	Rey-Osterrieth Complex Figure Test (ROCF)	Delayed recall
			Immediate recall
			Delayed recall
Processing speed	Processing speed	Trail Making Test A (TMT A)	Time to complete
	Reaction time	CPT 3	Reaction time
Working memory/mental tracking	Working memory	Four Word Short-Term Memory Test (STMT)	Total recall
		WAIS-IV Digit Span Backward	Total raw score

Note: CPT 3 = The Conner's Continuous Performance Test 3rd edition. WAIS-IV = Wechsler Adult Intelligence Scale 4th Edition.

neuropsychologists (T.S & M.J) as follows based on the descriptions of cognitive domains and their subdomains in Lezak et al. (2012). STMT is a variant of the Brown-Peterson task (Brown, 1958; Peterson & Peterson, 1959), which is considered a working memory task by Lezak et al. (2012); d2-R is a cancellation test with minimum demands on attentional capacity or divided attention, classifying it as a focused attention measure; and Vilkki Dual Task requires a person simultaneously to focus on two tasks, backward counting and dot cancellation, making it a test of divided attention. To avoid excessive multiple comparisons, the outcome measures were chosen a priori by two authors (T.S & M.J) with the requirement that the score chosen is the primary measure of the domain that the test is measuring.

Measures of subjective cognitive symptoms

All participants were asked whether they had experienced cognitive decrements and to rate their severeness on a scale from 0 to 5, where 0 indicated barely noticeable symptoms and 5 indicated symptoms that severely affect daily functioning. If a participant reported cognitive symptoms, they were asked to describe them in detail. A licensed neuropsychologist (T.S.) grouped the symptoms based on the participants' narratives.

Quality of measurements

The neuropsychological assessment was carried out by a specialized neuropsychologist (T.S), by a psychologist, or by a master's student in psychology under the supervision of a specialized neuropsychologist (T.S). All questionnaires and neuropsychological tests were first scored by the person who did the

assessment and then reviewed and coded by two specialized neuropsychologists (T.S & M.J). Medical data were collected and coded by an endocrinologist (E.P).

Data collection

After receiving information sheet and having the possibility to ask questions participants gave informed consent. Thereafter medical examination was performed by endocrinologist (E.P) , laboratory tests were performed and participants filled in questionnaires used in this study. Patients' medical history was reviewed through medical records. If the eligibility criteria were still met based on medical data, participants were contacted to make an appointment for the neuropsychological assessment. They were asked to bring completed questionnaires to the neuropsychological assessment. Participants arrived at the neuropsychological assessment individually and were reminded that it would last two to three hours. Water, juice and snacks were provided throughout the assessment and participants were given the opportunity to take breaks if necessary. The person carrying out the assessment ensured that all the questionnaires had been completed and asked for clarification if needed. Patients were asked to measure their blood glucose levels at the beginning and at the end of the assessment. All participants rated their level of fatigue on a scale from 0 (not at all tired) to 10 (very tired) at the beginning and at the end of the assessment. The neuropsychological tests were presented according to the instructions provided in test manuals or original research articles describing the test, and in the same order for all

participants. The order was as follows: ROCF, RAVLT, WAIS-IV Block Design, WAIS-IV Similarities, WAIS-IV Coding, Stroop, phonemic fluency: P-A-S, semantic fluency: animals, STMT, ToL, d2-R, Vilkki Dual Task, TMT A, TMT B, WAIS-IV Digit Span, and CPT 3. Delayed recall tasks were presented after an one-hour delay. All data were stored in a locked cabinet.

Data processing and diagnostics

Participants were excluded from analyses after data collection if they did not complete the neuropsychological assessment or if they could not be matched with another participant. Missing data was handled by replacing each individual missing score with the score of the closest match in the same group, or with the mean score of closest matches if there were multiple exact matches. Values that were over three standard deviations from the group mean were considered outliers. Outliers were not removed since they were minimal and not considered errors by the research group.

Data analysis strategy

Statistical analyses were performed using R statistical software version 4.1.3 within RStudio software version 2022.02.1 (R Core Team, 2022). The "MatchIt" package was used to match participants (Ho et al., 2011). Differences in outcome measures between patient and healthy comparison groups were assessed using paired-samples t-test. A parametric test was employed because visual examination of Q-Q plots showed evidence of normal or near-normal distribution for all outcome measures. To control for Type 1 error inflation, an alpha level of .01 was used for all statistical tests.

Results

Participants

Patients were recruited from the City of Tampere Diabetes Outpatient Clinic between 1 June 2018, and 31 May 2022. Healthy individuals were recruited by a newspaper advertisement published on 25 September 2021. Out of the 58 patients recruited, seven withdrew their participation and three could no longer be contacted. Out of the 60 healthy individuals recruited, eight did not meet the inclusion criteria due to hypothyroidism or diabetes. We were thus left with 48 patients and 52 healthy individuals out of which 28 matched pairs were formed and included in all analyses.

Missing data

The "Tower of London Test – total move score" was missing for one healthy individual and the "semantic fluency: animals – total raw score" for one healthy individual. Both missing values were due to a scoring error by the person carrying out the assessment. The missing scores were replaced with the corresponding score of the closest match in the healthy comparison group.

Selection bias

Unmatched healthy individuals had more years of education than matched healthy individuals $t(47.86) = 2.45$, $p = .02$, Hedge's $g = 0.81$ and there were more women in the unmatched group than in the matched group $\chi2$ (1, $N = 52$) = 11.96, $p = < .001$. No difference was observed in terms of age. Unmatched healthy individuals performed better than matched healthy individuals in RAVLT Trial 1–5 total $t(49.97) = 3.11$, $p = .003$, Hedge's $g = 0.85$, RAVLT delayed $t(48.79) = 2.05$, $p = .045$, Hedge's $g = 0.57$, WAIS-IV Coding $t(47.45) = 3.31$, $p = .002$, Hedge's $g = 0.93$ and Phonemic Fluency $t(50) = 2.21$, $p = .032$, Hedge's $g = 0.61$. No differences between groups were observed in other neuropsychological tests or subjective cognitive symptoms.

Men were overrepresented in the group of unmatched patients $\chi2$ (1, $N = 48$) = 15.45, $p = < .001$. There were no differences between the groups regarding age and years of education. Unmatched patients performed worse than matched patients in RAVLT Trial 1–5 total $t(42.61) = -2.78$, $p = .008$, Hedge's $g = -0.81$. No between-group differences were found for other tests or subjective symptoms.

Descriptive and inferential statistics

Medical variables and questionnaires

For medical background variables, there were statistically significant group differences in body mass index, $t(27) = 4.29$, $p = .000$, Hedge's $g = 1.12$, diastolic blood pressure, $t(27) = 2.35$, $p = .03$, $g = 0.68$, HbA1c, $t(27) = 5.08$, $p = <.001$, Hedge's $g = 1.34$, high-density lipoprotein $t(27) = -4.43$, $p = <.001$, Hedge's $g = -1.25$, triglyceride, $t(27) = 3.61$, $p = .001$, Hedge's $g = 1.07$ and waist circumference, $t(27) = 5.58$, $p = <.001$, Hedge's $g = 1.28$. All differences were in the expected direction, with patients having clinically worse scores than healthy individuals. Differences were not found for systolic blood pressure, $t(27) = 1.53$, $p = .14$, Hedge's $g = 0.48$ and low-density lipoprotein $t(27) = -0.54$, $p = .59$, Hedge's $g = -0.16$. The patient's blood glucose levels were significantly higher at the beginning of the

neuropsychological assessment ($M = 7.96$ mmol/l, $SD = 2.23$) than at the end of the assessment ($M = 6.56$ mmol/l, $SD = 1.48$), $t(27) = 5.18$, $p = <.001$, Hedge's $g = 0.74$.

Both groups scored similarly on Beck Anxiety Inventory ($M = 5.61$, $SD = 4.96$ and $M = 3.00$, $SD = 2.54$, patients and healthy individuals respectively), $t(27) = 2.11$, $p = .04$, Hedge's $g = 0.66$, Beck Depression Inventory ($M = 6.61$, $SD = 5.38$ and $M = 5.50$, $SD = 4.04$), $t(27) = 0.82$, $p = .42$, Hedge's $g = 0.23$, Fatigue Impact Scale ($M = 14.46$, $SD = 13.89$ and $M = 11.68$, $SD = 9.59$), $t(27) = 0.71$, $p = .49$, Hedge's $g = 0.23$, Bergen Burnout Indicator 15 ($M = 29.68$, $SD = 13.23$ and $M = 32.64$, $SD = 13.44$), $t(27) = -0.99$, $p = .33$, Hedge's $g = -0.22$, and Alcohol Use Disorders Identification Test ($M = 3.89$, $SD = 2.94$ and $M = 4.07$, $SD = 2.39$), $t(27) = -0.28$, $p = .78$, Hedge's $g = -0.07$. Both groups reported a similar level of fatigue at the beginning of the neuropsychological assessment ($M = 3.05$, $SD = 1.73$ and $M = 3.07$, $SD = 1.99$, patients and healthy individuals respectively), $t(27) = -0.04$, $p = .97$, Hedge's $g = -0.01$, and at the end of the assessment ($M = 4.00$, $SD = 2.00$ and $M = 3.82$, $SD = 2.24$), $t(27) = 0.32$, $p = .75$, Hedge's $g = 0.08$.

Subjective cognitive symptoms

There was no group difference in the frequency of subjective cognitive symptoms, which were reported by 17 patients (61%) and 13 healthy individuals (46%), $\chi 2$ (1, $N = 28$) = 1.15, $p = .28$. No difference was found between patients ($M = 1.11$, $SD = 1.26$) and healthy individuals ($M = 0.68$, $SD = 0.98$) in the severity of subjective symptoms either, $t(27) = 1.38$, $p = .18$, Hedge's $g = -0.38$. The most frequently reported symptom in the patient group was difficulty concentrating (reported by 53% of the patients who reported subjective cognitive symptoms), followed by difficulties in recall (29%), remembering names (29%), learning new information (24%), finding words (18%) and processing speed (12%). One patient reported working memory problems (6%) and one reported increase in spelling errors (6%). Forgetfulness was the most frequently reported symptom in the healthy comparison group (38%), followed by difficulties in remembering names (31%), learning new information (15%) and finding words (15%). One healthy comparison reported slowing down of information processing (8%).

Table 3. Neuropsychological test results.

Outcome measure	Patient group (n = 28) Mean (SD)	Healthy group (n = 28) Mean (SD)	t	p (two-sided)	95% CI of the difference [LL, UL]	Hedge's g^{b}
Attention						
CPT 3 (HRTC)[a]	3.78 (6.32)	1.71 (7.08)	1.32	0.20	[−1.15, 5.28]	−0.31
d2-R	124.9 (25.22)	136.9 (32.96)	−1.68	0.11	[−26.53, 2.67]	−0.41
Stroop	40.43 (9.56)	42.82 (8.22)	−1.09	0.29	[−6.92, 2.13]	−0.27
Trail Making Test B[a]	73.89 (25.22)	66.75 (21.79)	1.25	0.22	[−4.63, 18.91]	−0.30
Vilkki Dual Task	54.41 (12.33)	52.51 (9.41)	0.75	0.46	[−3.32, 7.13]	0.17
WAIS-IV Coding	61.96 (13.80)	63.07 (12.20)	−0.37	0.72	[−7.28, 5.07]	−0.09
WAIS-IV Digit Span FW	7.82 (2.07)	9.21 (1.71)	−2.36	0.03	[−2.60, −0.18]	−0.73
Concept formation and reasoning						
WAIS-IV Similarities	28.61 (3.54)	29.29 (2.29)	−0.91	0.37	[−2.21, 0.85]	−0.23
Construction and motor performance						
ROCF: Copying	34.43 (2.15)	34.07 (2.75)	0.50	0.62	[−1.11, 1.82]	0.15
WAIS-IV Block Design	40.96 (11.74)	41.25 (11.09)	−0.10	0.92	[−5.91, 5.34]	−0.03
Executive functions						
Phonemic fluency	40.14 (13.47)	44.57 (11.00)	−1.38	0.18	[−11.00, 2.15]	−0.36
Semantic fluency	24.64 (7.48)	25.64 (6.77)	−0.72	0.48	[−3.86, 1.86]	−0.14
Tower of London[a]	28.82 (17.16)	25.61 (20.11)	0.73	0.47	[−5.78, 12.21]	−0.17
Memory						
RAVLT: learning	49.79 (9.86)	48.39 (9.29)	0.69	0.50	[−2.76, 5.55]	0.15
RAVLT: delayed	8.46 (3.34)	8.21 (3.49)	0.30	0.76	[−1.45, 1.95]	0.07
ROCF: immediate	20.45 (6.82)	17.88 (6.11)	2.05	0.05	[−0.00, 5.14]	0.40
ROCF: delayed	19.11 (6.70)	18.25 (6.00)	0.63	0.53	[−1.93, 3.64]	0.14
Processing speed						
CPT 3 (RT)[a]	454.4 (44.69)	444.6 (47.48)	0.98	0.34	[−10.82, 30.46]	−0.21
Trail Making Test A[a]	32.71 (7.64)	30.82 (11.42)	0.78	0.44	[−3.10, 6.89]	−0.19
Working Memory						
STMT	34.57 (9.03)	34.71 (10.21)	−0.07	0.95	[−4.54, 4.25]	−0.01
WAIS-IV Digit Span BW	8.86 (1.69)	9.11 (2.02)	−0.48	0.63	[−1.31, 0.81]	−0.13

Note: BW = backward, CPT 3 = The Conner's Continuous Performance Test 3rd Edition, FW = forward, HRTC = hit reaction time change, LL = lower limit, RAVLT = Rey Auditory Verbal Learning Test, ROCF = Rey-Osterrieth Complex Figure Test, RT = Reaction time, STMT = Four Word Short-Term Memory Test, UL = upper limit, WAIS-IV = Wechsler Adult Intelligence Scale 4th Edition.
[a] a higher score indicates worse performance
[b] a negative value indicates the patient group performing worse than the healthy group

Cognitive performance

The neuropsychological test results are presented in Table 3. Raw scores were used in all analyses. There were no statistically significant differences between the age-, education- and gender-matched patient and healthy groups in any of the 21 neuropsychological test scores analyzed. The effect sizes ranged from −0.73 to 0.40, with the mean effect size being −0.12 and the median effect size being −0.14. All effect sizes were negligible or small, with the exception of WAIS-IV Digit Span Forward in which a medium negative effect size was obtained.

Discussion

We did not find statistically significant differences between the groups in any of the 21 neuropsychological test scores analyzed. Most of the effect sizes obtained were negligible or small, although mostly in the direction of patients performing more poorly than healthy individuals. Furthermore, there was no between-group difference in the number of individuals reporting subjective cognitive symptoms or in the severity of the symptoms.

Our results are not in line with the meta-analysis in which the middle-aged T2DM patients performed worse than healthy individuals in all cognitive domains except visual memory (Pelimanni & Jehkonen, 2019). However, many of the original articles included in the meta-analysis did not control for age-, gender- and education. The studies that did control for all of these confounders were small, with sample sizes ranging from 13 to 38 patients (Aberle et al., 2008; Biessels et al., 2001; García-Casares et al., 2014; Kálcza-Jánosi & Anett, 2016; Mehrabian et al., 2012; Yau et al., 2009). The largest of these studies, in which crystallized intelligence, fluid intelligence, verbal memory, visual memory, executive functioning and psychomotor speed were assessed, found no difference between the groups with all effect sizes being negligible (Aberle et al., 2008). The participants were on average nine years older and had three years less education than the participants in our study. In the five other studies, patients performed worse than healthy individuals in at least one cognitive domain, with all of the studies reporting poorer memory functioning in the patient group than in the healthy comparison group (Biessels et al., 2001; García-Casares et al., 2014; Kálcza-Jánosi & Anett, 2016; Mehrabian et al., 2012; Yau et al., 2009). The effect sizes were mostly medium or large.

In some of these studies where between-group differences were observed, the participants were roughly the same age, slightly more educated, and the duration of diabetes was a few years shorter than in our study (Mehrabian et al., 2012; Yau et al., 2009), making these factors unlikely to explain the differences in the results. The different findings cannot be explained solely by the neuropsychological tests selected either, since in some of the studies the patient group performed worse than the healthy group in tests that were used in our study as well (García-Casares et al., 2014; Mehrabian et al., 2012). In these tests, the mean scores of the healthy individuals were approximately the same as in our study, but the patients performed surprisingly poorly, with mostly large effect sizes obtained. It is possible that some of the previous studies have included a few particularly poorly performing patients, which could have lowered the overall group mean given the small sample sizes. This may be due, for example, to the fact that hypothyroidism, severe obstructive sleep apnea or diabetes complications were not mentioned as exclusion criteria in most of the other studies. For example, in a study with only 13 patients, one patient was reported to have background retinopathy and two patients a history of ischemic heart disease (Biessels et al., 2001). In a small dataset, these factors can already explain the differences in cognitive performance between the groups. In one study, the exclusion criteria were only related to the age and education level of the participants (Kálcza-Jánosi & Anett, 2016). Considering the small number of published studies and the small sample sizes, it is also possible that publication bias distorts the results published so far.

Subjective cognitive symptoms

Similar numbers of patients and healthy individuals reported subjective cognitive symptoms, and there was no difference in the subjective severity of the symptoms reported. A previous study with a larger sample size found that middle-aged T2DM patients had more memory complaints than healthy individuals (Faiz et al., 2021). Aberle et al. (2008) did not find differences in subjective cognitive symptoms between middle-aged T2DM and healthy comparison groups in a study of similar size to ours. In addition to the statistical power of the studies, the differing results could be due to the heterogeneity of the participants studied or the differences in the methods used to investigate the symptoms.

Strengths and limitations

A general weakness of observational studies is that due to unknown confounding variables, they cannot establish causal relationships. In our study we achieved

excellent matching regarding age, education and gender, and found no group differences in any of the psychosocial symptom or alcohol use measures.

The participants of our study were recruited through self-selection, which is a potential source of sampling bias. The healthy individuals in our study were more highly educated and more often female than the patients. When the groups were matched, this resulted in the unmatched healthy individuals performing better than the matched individuals in some of the measures used. Men were overrepresented in the group of unmatched patients and this group performed worse than the matched patient group in one memory measure. The person conducting the neuropsychological assessment was aware of the participants' diabetes status, which may have introduced performance bias. Although we followed an a priori research plan throughout the study and our hypotheses were based on previous research, we did not pre-register our research protocol, which can be considered a limitation.

To our knowledge, this is the first study in middle-aged T2DM patients that uses a careful matching procedure and a neuropsychological test battery covering most cognitive domains where group differences have been observed in older age groups. In addition to cognitive performance, we studied subjective cognitive symptoms and investigated several medical and psychosocial factors that could potentially confound the results. We believe that this study, despite its limitations, brings valuable information about the cognitive performance of middle-aged people with T2DM.

Conclusions

T2DM is a fast-growing worldwide epidemic that has been associated with cognitive decrements already in middle age (Biessels et al., 2001; García-Casares et al., 2014; Kálcza-Jánosi & Anett, 2016; Mehrabian et al., 2012; Yau et al., 2009). However, all of the studies that have controlled for the effects of age, education and gender have been small and heterogeneous in regards to sampling methods, neuropsychological tests used and the participant characteristics.

In our data, which was carefully matched and applied strict exclusion criteria, there were no differences in cognitive performance between the patient and healthy groups. In addition, the effect sizes obtained were mainly negligible. We consider it unlikely that working-age people with T2DM without other medical conditions, such as hypothyroidism, severe obstructive sleep apnea or late complications of diabetes, would, as a group, perform significantly worse than healthy

controls on the neuropsychological tests used in our study. This does not mean that an individual with type 2 diabetes is not at greater risk of developing cognitive symptoms. We encourage readers to investigate this topic further, preferably using a longitudinal design, to better identify factors associated with poorer cognitive performance or decline in cognitive function with aging in people with T2DM.

References

Aberle, I., Kliegel, M., & Zimprich, D. (2008). Cognitive development in young-old type-2 diabetes patients: A longitudinal analysis from the "interdisciplinary longitudinal study of aging. *Current Psychology*, *27*(1), 6–15. https://doi.org/10.1007/s12144-008-9020-3

Alberti, K. G. M. M., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications Part 1: Diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diabetic Medicine*, *15*(7), 539–553. https://doi.org/10.1002/(SICI)1096-913619980715:7

Babor, T. F., de la Fuente, J. R., Saunders, J., & Grant, M. (1992). *AUDIT. The Alcohol Use Disorders Identification Test. Guidelines for use in primary health care*. World Health Organization.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897. https://doi.org/10.1037/0022-006X.56.6.893

Beck, A., Steer, R., & Brown, G. (1996). *Manual for the beck depression inventory-II*. Psychological Corporation.

Biessels, G. J., ter Braak, E. W. M., Erkelens, D. W., & Hijman, R. (2001). Cognitive function in patients with type 2 diabetes mellitus. *Neuroscience Research Communications*, *28*(1), 11–22. https://doi.org/10.1002/1520-6769200101/0228

Brickenkamp, R., Schmidt-Atzert, L., & Liepmann, D. (2016). d2-R Tarkkaavuuden ja keskittymisen testi. Suomenkielisen version käsikirjaliite Kynä-paperitestille ja nettitestille. Hogrefe.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *The Quarterly Journal of Experimental Psychology*, *10*(1), 12–21. https://doi.org/10.1080/17470215808416249

Butters, N., & Cermack, L. S. (1980). *Alcoholic Korsakoff's syndrome: An information processing approach in amnesia*. Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Faiz, S., Qureshi, F. M., Hussain, A. W., & Nigah-E-Mumtaz, S. (2021). Association of subjective memory complaints amid patients of diabetes mellitus type II and hypertension. *Pakistan Journal of Medical Sciences*, *37*(2), 477–482. https://doi.org/10.12669/pjms.37.2.3426

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., & Schlech, W. F. (1994). Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *18*(Suppl 1), S79–S83. https://doi.org/10.1093/clinids/18.Supplement_1.S79

García-Casares, N., Jorge, R. E., García-Arnés, J. A., Acion, L., Berthier, M. L., Gonzalez-Alegre, P., Nabrozidis, A., Gutiérrez, A., José Ariza, M., Rioja, J., & González-Santos, P. (2014). Cognitive dysfunctions in middle-aged type 2 diabetic patients and neuroimaging correlations: A cross-sectional study. *Journal of Alzheimer's Disease*, *42*(4), 1337–1346. https://doi.org/10.3233/JAD-140702

Gudala, K., Bansal, D., Schifano, F., & Bhansali, A. (2013). Diabetes mellitus and risk of dementia: A meta-analysis of prospective observational studies. *Journal of Diabetes Investigation*, *4*(6), 640–650. https://doi.org/10.1111/jdi.12087

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08

Kálcza-Jánosi, K., & Anett, P. (2016). Glucose metabolism and memory functions in adults with type 2 diabetes mellitus. *Romanian Journal of Experimental Applied Psychology*, *7*(1), 146–149. https://doi.org/10.15303/rjeap.2016.si1.a30

Kálcza-Jánosi, K., Lukács, A., Barkai, L., & Szamosközi, I. (2013). Cognitive functions in type 1 and type 2 diabetes: Meta-analysis. *Orvosi Hetilap*, *154*(18), 694–699. https://doi.org/10.1556/OH.2013.29602

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th) ed.). Oxford University Press.

Magliano, D. J., & Boyko, E. J., IDF Diabetes Atlas 10th edition scientific committee. (2021). *IDF DIABETES ATLAS* (10th ed.). International Diabetes Federation.

Mansur, R. B., Lee, Y., Zhou, A. J., Carmona, N. E., Cha, D. S., Rosenblat, J. D., Bruins, R., Kakar, R., Rasgon, N. L., Lovshin, J. A., Wroolie, T. E., Sim, K., Brietzke, E., Gerstein, H. C., Rong, C., & McIntyre, R. S. (2018). Determinants of cognitive function in individuals with type 2 diabetes mellitus: A meta-analysis. *Annals of Clinical Psychiatry: Official Journal of the American Academy of Clinical Psychiatrists*, *30*(1), 38–50. https://pubmed.ncbi.nlm.nih.gov/29373617/

Mehrabian, S., Raycheva, M., Gateva, A., Todorova, G., Angelova, P., Traykova, M., Stankova, T., Kamenov, Z., & Traykov, L. (2012). Cognitive dysfunction profile and arterial stiffness in type 2 diabetes. *Journal of the Neurological Sciences*, *322*(1–2), 152–156. https://doi.org/10.1016/j.jns.2012.07.046

Monette, M. C., Baird, A., & Jackson, D. L. (2014). A meta-analysis of cognitive functioning in nondemented adults with type 2 diabetes mellitus. *Canadian Journal of Diabetes*, *38*(6), 401–408. https://doi.org/10.1016/j.jcjd.2014.01.014

Näätänen, P., Aro, A., Matthiesen, S., & Salmela-Aro, K. (2003). *Bergen Burnout Indicator-15*. Edita.

Palta, P., Schneider, A. L., Biessels, G. J., Touradji, P., & Hill-Briggs, F. (2014). Magnitude of cognitive dysfunction in adults with type 2 diabetes: A meta-analysis of six cognitive domains and the most frequently reported neuropsychological tests within domains. *Journal of the International Neuropsychological Society*, *20*(3), 278–291. https://doi.org/10.1017/S1355617713001483

Pelimanni, E., & Jehkonen, M. (2019). Type 2 diabetes and cognitive functions in middle age: A meta-analysis. *Journal of the International Neuropsychological Society*, *25*(2), 215–229. https://doi.org/10.1017/S1355617718001042

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*(3), 193–198. https://doi.org/10.1037/h0049234

Rawlings, A. M., Sharrett, A. R., Schneider, A. L., Coresh, J., Albert, M., Couper, D., Griswold, M., Gottesman, R. F., Wagenknecht, L. E., Windham, B. G., & Selvin, E. (2014). Diabetes in midlife and cognitive change over 20 years: A cohort study. *Annals of internal medicine*, *161*(11), 785–793.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL. https://www.R-project.org/

Sadanand, S., Balachandar, R., & Bharath, S. (2016). Memory and executive functions in persons with type 2 diabetes: A meta-analysis. *Diabetes/Metabolism Research and Reviews*, *32*(2), 132–142. https://doi.org/10.1002/dmrr.2664

Vilkki, J., Virtanen, S., Surma-Aho, O., & Servo, A. (1996). Dual task performance after focal cerebral lesions and closed head injuries. *Neuropsychologia*, *34*(11), 1051–1056. https://doi.org/10.1016/0028-3932(96)00028-0

Vincent, C., & Hall, P. A. (2015). Executive function in adults with type 2 diabetes: A meta analytic review. *Psychosomatic Medicine*, *77*(6), 631–642. https://doi.org/10.1097/PSY.0000000000000103

Yau, P. L., Javier, D., Tsui, W., Sweat, V., Bruehl, H., Borod, J. C., & Convit, A. (2009). Emotional and neutral declarative memory impairments and associated white matter microstructural abnormalities in adults with type 2 diabetes. *Psychiatry Research*, *174*(3), 223–230. https://doi.org/10.1016/j.pscychresns.2009.04.016