Association Between Age at Diabetes Onset and Subsequent Dementia: An Age-Attained Analysis

Jiwon Lee

Department of Health Informatics & Biostatistics
Graduate School of Public Health
Yonsei University

Association Between Age at Diabetes Onset and Subsequent Dementia: An Age-Attained Analysis

Directed by Professor Sohee Park

A Master′s Thesis

Submitted to the Department of Health Informatics & Biostatistics

and the Committee on Graduate School of Public Health

Yonsei University

in Partial Fulfillment of the Requirements for the Degree of

Master of Public Health

Jiwon Lee

June 2025

This certifies that the Master's Thesis

Jiwon Lee is approved.

_____

Thesis Supervisor: Sohee Park

_____

Thesis Committee Member #1: Min Jin Ha

_____

Thesis Committee Member #2: Jae Hong Joo

Department of Health Informatics & Biostatistics

Graduate School of Public Health

Yonsei University

June 2025

Acknowledgements

I would like to write my deepest gratitude to all those who have supported me-both emotionally and practically-throughout my two years of graduate studies, helping me to begin a new chapter in my life.

First and foremost, I am profoundly grateful to my supervisor, Professor Sohee Park, for her invaluable guidance, insightful wisdom, and unwavering support. With her deep knowledge and clarity of thought, she has played a pivotal role in shaping my understanding of biostatistics, a field that was once unfamiliar to me. She always provided meticulous and thoughtful supervision, and I am deeply inspired by her dedication and precision. I would also like to extend my sincere thanks to Professor Min Jin Ha, whose expertise and thoroughness greatly contributed to my person and academic growth, especially through my assistantship experience. Her high standards and attention to detail consistently motivated me to improve and learn. Special thanks go as well to Professor Chung Mo Nam, whose thought-provoking questions and lively discussions consistently challenged and expanded my insights. I am particularly thankful for his warmth and generosity, treating me not only as a student but also like family. To each of these professors, I am truly indebted.

I would also like to extend my heartfelt thanks to all the colleagues I had the pleasure of studying with at the Statistical Consulting Center. In particular, I am deeply grateful to Dr. Jae Hong Joo. His consistent support and expertise in biostatistics were invaluable throughout my research. My sincere thanks also go to Jiyeon Chun and Jae Heon Yu, with whom I spent the most time. Their willingness to share both academic insights and practical advice about graduate life enriched my experience tremendously. I am also thankful to In Seong Hwang and Nayeon Kim. Sharing this journey with peers brought countless memorable moments, and I truly appreciated the shared

learning. Finally, I would like to thank the many other individuals at the school who are not mentioned here by name, but whose help made a lasting impact on my graduate experience.

Lastly, I would like to express my deepest gratitude to my beloved family and friends, whose unwavering support and love have carried me through this journey. To my father, Young Chan Lee, thank you for your unwavering belief in me and the love you have always shown. Your constant encouragement and deep trust have been a profound source of motivation. To my mother, Inyae Lee, whose listening ear and emotional strength have helped me through difficult times-your presence means the world to me. To my sister Jiwoo Lee, thank you for always believing in me, cheering me on, and being a steadfast source of motivation. I am also immensely grateful to my dearest friends, Youngsun Min and Sumin Kang, who were always there to lift my spirits and make me smile. Your friendship has been an incredible source of comfort and joy. I love you all, and I am endlessly thankful to have you in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Association Between Age at Diabetes Onset and Subsequent Dementia
: An Age-Attained Analysis

Jiwon Lee

Department of Health Informatics & Biostatistics,
Graduate School of Public Health Yonsei University

(Directed by Professor Sohee Park, Ph.D.)

Background: Traditional Cox Proportional hazards models (Time on study analysis), which use absolute time as the time scale, may not adequately account for age-specific risks. Previous studies have suggested that using attained age as the time scale may be more appropriate in age-specific survival analyses. While association between diabetes and dementia has been well established, limited research has explored how the age at diabetes onset influenced the subsequent risk of dementia in Korea. Given the rising global prevalence of diabetes-particularly among younger individuals-further investigation into this relationship is essential. Therefore, the present study investigated the association

between diabetes onset age and subsequent dementia risk using different time scales, addressing a gap in the existing literature.

Methods: This study utilized the National health Insurance Service-Health Screening Cohort (ver.2.1.), which was constructed through simple random sampling of 514,866 individuals (10% of 515 million) who underwent general health screening in 2002-2003. The study population consisted of individuals aged 40 or more, who were diagnosed with diabetes and maintained eligibility for National Health Insurance coverage. The data spans for 18-year period from 2002 to 2019. A two-year wash-out period was applied; individuals diagnosed with diabetes in 2002 and 2003 were excluded. Following these criteria, a total of 77,092 individuals were included in the final study cohort.

Time on study analysis is commonly employed for the analysis of time-to-event data. However, in studies where age plays a critical role- such as the investigation of the association between age at diabetes onset of developing dementia- it is more appropriate to account for the influence of age on disease incidence by adopting age attained analysis, which uses age rather than absolute time for time scale. Therefore, time on study analysis and age attained analysis were conducted and compared to examine the association between age at diabetes onset and the risk of developing dementia. In addition, the effects of several covariates-including income level, sex, smoking frequency, alcohol consumption, frequency of physical activity, and body mass index (BMI)-on dementia risk were also evaluated. To further elucidate the impact of diabetes onset age on distinct dementia types

and to address the potential overestimation of dementia incidence due to competing mortality, a competing risks analysis was also conducted.

Results: Analyses using two different times scales- time on study analysis and age attained analysis-yielded contrasting results regarding the association between age at diabetes onset and the risk of dementia. Using time on study analysis, later onset of diabetes was associated with a higher risk of developing dementia. In contrast, when using attained age as the time scale, earlier onset of diabetes was associated with an increased risk of dementia. Regarding other factors, the results showed consistent directions of association across both models. Higher income levels were linked to a reduced risk. With respect to sex, dementia was found to be significantly more prevalent among females compared to males. A higher frequency of smoking was associated with a decreased risk of dementia. Conversely, increased alcohol consumption as correlated with a higher risk. Greater frequency of physical activity was associated with a reduced risk of subsequent dementia, and higher body mass index was similarly linked to a lower risk. An important finding of this study was that the association between diabetes onset age and dementia risk yielded opposite results depending on the choice of time scale in the Cox regression model. Given the strong influence of age on both diabetes onset and dementia, using attained age as the time scale provided a more appropriate framework for age-specific risk comparison across individuals.

Conclusions: This study investigated the association between age at diabetes onset and dementia risk using a nationally representative Korean cohort. It also examined the influence of various covariates and explored differences by dementia subtype. The results

indicated that earlier-onset diabetes was associated with a higher risk of dementia, underscoring the importance of timing and duration of diabetes in relation to cognitive decline. Moreover, the comparison of models using different time scales emphasized that using attained age as the time scale may be more appropriate when studying age-dependent outcomes such as dementia. These findings contribute to the growing body of evidence on the link between metabolic conditions and neurodegeneration and offer methodological insight for future research.

Key words: Diabetes Mellitus, dementia, Time Scale, Cox proportional hazards regression model, Age-attained analysis, Competing risk analysis

I. INTRODUCTION

1. Background

The incidence of type 2 diabetes has escalated due to factors such as population aging, lack of exercise, growing obesity trends and calorie-dense eating habits. (Chatterjee, Khunti et al. 2017). Over the past four decades, the prevalence of diabetes in Korea has risen dramatically, increasing from 1.5% to 9.9%, a six- to seven- fold rise. (Kim 2011). Diabetes prevalence is not only rising overall, but is also increasingly observed in younger individuals. National surveys conducted in Korea from 2001 and 2013 showed an age-related increases in diabetes prevalence across all age groups, except for those aged 30-39 years (Ha and Kim 2015). According to the Health Insurance Review and Assessment Service (HIRA) statistics, the number of diabetes patients in their 20s and 30s increased from 139,682 in 2018 to 174,485 in 2022, showing a 24.9% increase. This growth rate exceeds the overall increase in diabetes patients, which was 21% during the same period. As the prevalence of diabetes itself increases and the age of diabetes onset decreases, studies on the impact of diabetes on individuals have become more prevalent.

With the accelerating pace of population aging worldwide, dementia has also become one of the most important public health problems globally. In particular, dementia imposes the highest disease burden among the elderly population in Korea, with rates of 528 per 100,000 in the general population and 5,117 per 100,000 among individuals aged 65 and older in 2008. This burden is expected to rise significantly as the population continues to

age (Park, Eum et al. 2013).

With growing interest in the link between diabetes and dementia, studies have started to investigate how diabetes influences dementia development and its clinical implications. Various studies related to these diseases have been conducted. Significant associations were identified between early diabetes diagnosis and increased mortality from vascular diseases and other non-neoplastic causes, especially those related to neurological, respiratory, and infectious origins, with external causes (Kim 2024). Several large-scale multicohort meta-analyses have been published demonstrating a significant association between diabetes and dementia (Amidei, Fayosse et al. 2021). Furthermore, epidemiological reviews have identified consistent cross-sectional and longitudinal association between type 2 diabetes and moderate cognitive decline, especially in domains such as memory and executive functioning (Pasquier, Boulogne et al. 2006). Another study estimated that diabetes increase the risk of developing dementia by approximately 50% (Yu, Han et al. 2020). As interest in this topic continues to grow, the association between diabetes onset age and subsequent risk of dementia has also been increasingly explored in recent years. A prospective cohort study conducted in the UK demonstrated that an earlier age at diabetes onset is associated with a higher risk of developing dementia later in life (Amidei, C. B., et al. 2021). Another study utilizing data from the UK Biobank similarly reported that individuals with an earlier onset of diabetes exhibited a higher incidence of dementia compared to those diagnosed later in life (Wang, Y., et al. 2023.). Although research on the association between age at diabetes onset and dementia has been emerging in recent years,

studies on this topic remain limited in the Korean Population.

2. Study objectives

The main objective of this study is to discover the association between age at diabetes onset and subsequent dementia in Korea. Furthermore, the study aims to conduct analysis using two different time scales, time-on study and age attained study, to compare the results based on the choice of time scale.

(1) To estimate the hazard risks of dementia in relation to diabetes onset age using traditional Cox proportional hazard regression model using time on study as the time scale.

(2) To estimate the hazard risks of dementia in relation to diabetes onset age using Age attained analysis using attained age as the time scale

(3) To compare the association of diabetes onset age on subsequent dementia using different time scales.

(4) To categorize dementia into 3 groups, Alzheimer's disease/ Vascular dementia/ other dementia and estimate the hazard risks of each group in relation to diabetes onset age using competing risk analysis.

3. Literature review

3.1. Diabetes and dementia

Alzheimer's disease and diabetes mellitus are highly prevalent and pose serious health burdens for the aging population. Several studies have confirmed a relationship between diabetes and various types of dementia (Pasquier, Boulogne et al. 2006). Moreover, the Lancet Commission on dementia prevention, intervention, and care classified diabetes as a confirmed risk factor in the development of dementia (Savelieff, Chen et al. 2022). Among studies investigating the impact of diabetes on dementia, a number of studies have focused specifically on how the age at which diabetes or dementia occurs may influence their association. A study examining risk factors for dementia among individuals with type 2 diabetes mellitus in Korea found that, with the exception of sex, most risk factors had a stronger impact on young-onset dementia than late-onset dementia (Yu, Han et al. 2020). The risk factors included factors such as smoking status, drinking frequency, exercise, income level, depression and more. In another study focusing on the impact of duration of diabetes on the risk of dementia in ischemic stroke patients, the results showed that individuals who had diabetes for five years or longer had a higher risk of developing dementia compared to participants who suffered diabetes for less than five years (Kim, Han et al. 2025). From a study identifying factors that are modifiable on dementia risk using UK Biobank showed that the influence of dementia risk varied by category, with lifestyle contributing the most by 16.6%, followed by medical history of 14.0%, and socioeconomic

status of 13.5% (Zhang, Chen et al. 2023).

3.2. Sociodemographic features and dementia

Socioeconomic status, particularly income level, has repeatedly been shown to play as a determinant of dementia risk. A study based on data from the São Paulo Ageing & Health Study (SPAH) study showed that those who lacked literacy, were employed in unskilled labor, and had lower earnings were more likely to experience dementia (Scazufca, Almeida et al. 2010). In another study, both declining income levels and low income in later life showed association with higher risk of dementia. The results were consistent regarding the association between income levels and dementia.

Most studies on the overall prevalence of dementia have reported that it is higher in women than in men. According to information provided by the Korean Dementia Association, Alzheimer's disease, which is the most common neurodegenerative condition causing dementia, is also known to occur at a significantly higher rate in women compared to men. This gender difference becomes more pronounced with increasing age.

3.3. Health behavior characteristics and dementia

There have been various studies on the association between alcohol consumption and dementia. Focusing on research conducted in South Korea, one study involving a cohort of 3,933,382 Korean individuals found that those who engaged in mild to moderate alcohol consumption had a lower risk of developing dementia compared to non-drinkers. However,

individuals who engaged in heavy drinking had a higher risk of dementia. Moreover, those who reduced their alcohol intake from heavy to moderate levels showed a decreased risk of developing dementia (Jeon, Han et al. 2023). The results were consistent in another study searching for the association between consumption of alcohol and risk of dementia with patients suffering depression (Shin, Jung et al. 2025).

Studies conducted for searching association between smoking levels and dementia in Korea also showed consistent results. A study conducted in Korea involving 789,532 participants showed the results that smoking cessation was associated with a reduced risk of dementia. In addition, individuals who either increased or decreased their smoking levels had a higher risk of developing dementia compared to those who maintained a consistent level of smoking (Jeong, Park et al. 2023). In another study conducted in Korea including 46,140 men who were aged 60 years or older, the findings showed that never-smokers and long term quitters had lower risk of dementia compared to continual smokers (Choi, Choi et al. 2018).

In studies regarding association of exercise level and dementia also showed consistent results. In a large Korean cohort of 62,286 individuals, higher levels of physical activity , even light-intensity activity, were linked to a lower risk of dementia among older adults (Yoon, Yang et al. 2021). In another study held in Korea investigating the association between physical activity and dementia among cancer survivors showed the result of dementia risk increasing in inactive groups compared to exercise groups (Lee, Han et al.).

3.4. Medical characteristics and dementia

BMI also appeared to be a contributing factor influencing the risk of dementia. A study conducted in Korea examining the association between BMI and dementia in newly diagnosed diabetes patients reached the conclusion of a higher baseline BMI being independently associated with a lower risk of dementia regardless of confounding factors (Nam, Park et al. 2019). In another study using a cohort of 1,958,191 individuals from United Kingdom Clinical Practice Research Datalink(CPRD), the results showed that higher BMI showed a lower risk of dementia, showing consistent results to studies in Korea (Qizilbash, Gregson et al. 2015).

II. Methods

1. Data used

National Health Insurance Service-Health Screening Cohort (ver. 2.1) was used in this study. To analyze healthcare utilization and health outcomes among recipients of general health screenings, a cohort-based research database was constructed using a 10% sample the 515 million individuals (514,866 individuals) who underwent general health screenings in 2002–2003. Simple random sampling was used. The cohort includes individuals aged 40 to 79 years as of December 2002 who maintained National Health Insurance qualification. The database spans 18 years (2002–2019) and comprises information on qualification and income (socioeconomic variables), records of medical utilization, health screening results, and medical care. All data are fully anonymized to ensure that personal information is not identified. The dataset includes qualification information that contains socioeconomic status, disability status, and mortality data and health utilization information including clinical and health screening records, and information on the status of healthcare institutions.

2. Study population

The subjects of this study were diabetic patients aged more than 40. Diabetic patients were identified using Korean Standard Classification of Diseases (KCD) code and the usage of hypoglycemic agent. KCD code is established to classify diseases and other health problems recorded in all types of heath and vital records in South Korea based on the International Statistical Classification of Diseases (ICD) recommended by the World Health Organization (WHO). Among the 217,304 subjects who had KCD codes “E11”, “E12”, “E13”, “E14” in medical claims data, those who had received hypoglycemic agent in either inpatient or outpatient prescription records were classified as diabetes patients (116,519). For the Identification of users of hypoglycemic agent, the therapeutic drug classification coded provided by the Hearth Insurance Service Review & Assessment service (HIRA) was used. A two-year wash-out period was applied, and individuals diagnosed with diabetes in 2002 and 2003 were excluded from the study leaving the number of 77,969. Among the 77,969 subjects, those who suffered Dementia prior to Diabetes were excluded and a total of 77,092 individuals were selected as the final study population for analysis.

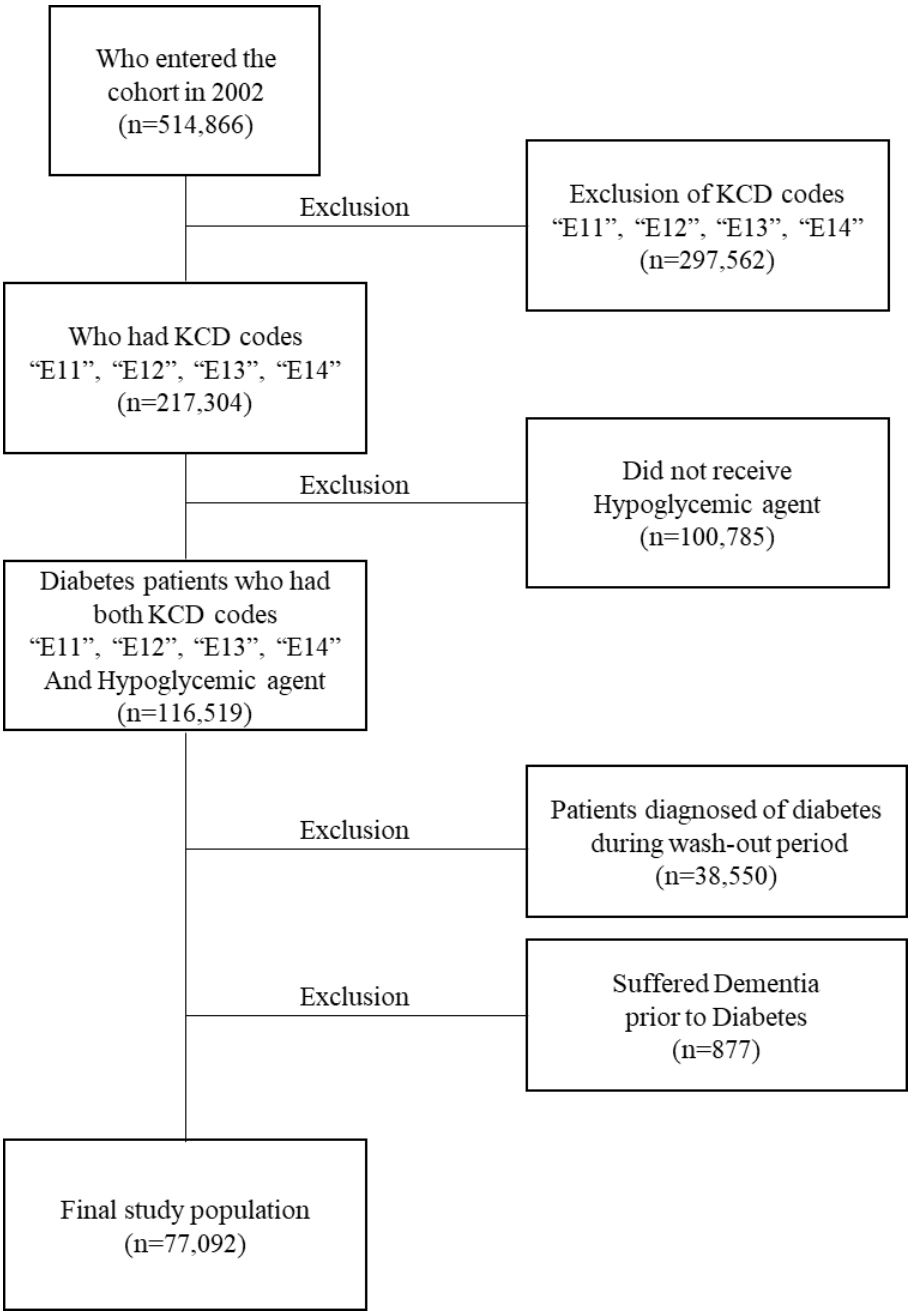

Figure 1. Flow Chart of Study Population

3. Variables

3.1. Outcome variable

In this study, the outcome variable is incidence of dementia. Dementia patients were identified using KCD codes and the usage of dementia medications. Among those who had KCD codes “F00”, “F01”, “F02”, “F03” in medical claims data, individuals who received dementia medications in either inpatient or outpatient prescription records were classified as dementia patients. Dementia medications were selected by the Clinical Practice Guideline for Dementia (Diagnosis and Evaluation): 2021 Revised Edition (Korean Dementia Association, 2021). Donepezil hydrochloride, Rivastigmine, Galantamine, Memantine were used for the identification of Dementia patients. Of the 54,854 individuals identified based on ICD codes, 34,016 had been prescribed dementia medications.

Table 1. KCD codes for different types of Dementia

Types of Dementia	KCD-code
Alzheimer's disease	F00.0 ~ F00.2 F00.9
Vascular dementia	F01.0 ~ F01.3 F01.8 F01.9
Other dementia	F03

KCD: Korean Standard Classification of Diseases, 8th version

Table 2. Dementia Medication Code from HIRA

Name of drug	Code
Donepezil hydrochloride	148630ALQ, 148631ALQ, 148601APD, 148601ATB, 148601ATD, 148602APD, 148602ATB, 148604ATB, 148602ATD, 148603ATB
Rivastigmine	224501ACH, 224503ACH, 224504ACH
Galantamine	385203ACR, 385203ATR, 385204ACR, 385204ATR, 385205ACR, 385205ATR
Memantine	190003ATD, 190004ATB, 190004ATD, 190005ATB, 190006ATD

3.2. Main interest variable

The main interest variable of the study is age of diabetes onset. Diabetes patients were identified using KCD codes and the usage of hypoglycemic agent. Among those who had KCD codes “E11”, “E12”, “E13”, “E14” in medical claims data, individuals who received hypoglycemic agents in either inpatient or outpatient prescription records were classified as Diabetes patients. Hypoglycemic agents were selected by the Clinical Practice Guidelines for Diabetes (Korean Diabetes Association, 2023). Metformin, Voglibose, Repaglinide, Dpp-4, Insulin, Meglitinide and Exenatide were used for the identification of Diabetes patients. The variable containing diabetes onset age was made by the age of the participant on the year of diabetes incidence during the study period. The age of diabetes onset was categorized into 4 groups for the analysis: 40-49, 50-59, 60-69, 70+.

Table 3. Diabetes Medication Code from HIRA

Name of drug	Code
Metformin	191501-191504ATB, 191504ATR, 191505ATR, 443400ATB, 443500ATB, 471900ATB, 474200ATB, 474300ATB, 474300ATR, 497200ATB, 498100ATB, 498600ATB, 502300ATB, 502300ATR, 502900ATB, 507000ATB, 507100ATB, 513700ATB, 513700ATR, 518500ATR, 518600ATR, 518800ATB, 519600ATB, 520500ATB, 520600ATB, 520700ATB, 523600ATB, 523700ATB, 523800ATR, 524700ATR, 631900ATB, 632000ATR, 632100ATB, 635600ATB, 635700ATB, 637200ATB, 639800ATR, 641400ATR, 641800ATR, 641900ATR, 642000ATR, 644900ATB, 645000ATR, 648400ATB, 648500ATB, 648600ATB, 649900ATR, 650000ATR, 650100ATR, 653800ATR, 653900ATR, 654000ATR, 654100ATR, 655700ATR, 191502ATR
Voglibose	249001ATB, 249001ATD, 249002ATB, 249002ATD, 523600ATB, 523700ATB
Repaglinide	379501ATB, 379502ATB, 379503ATB, 632100ATB, 637200ATB
Dpp-4	500801ATB, 501101ATB, 501101ATB, 501101ATB, 501102ATB, 501102ATB, 501102ATB, 501103ATB, 501103ATB, 501103ATB, 501103ATB, 501103ATB, 501103ATB, 501103ATB, 502300ATB, 502300ATR, 502900ATB, 507000ATB, 507100ATB, 513700ATB, 513700ATR, 518500ATR, 518600ATR, 519600ATB, 520500ATB, 520600ATB, 520700ATB, 523800ATR, 524700ATR, 613301ATB, 613302ATB, 616401ATB, 619101ATB, 624201ATB, 624202ATB, 624203ATB, 627301ATB, 630300ATB, 630400ATB, 630500ATB, 630600ATB, 632000ATR, 635600ATB, 635700ATB, 639601ATB, 641900ATR, 642000ATR, 645000ATR, 645301ATB, 648400ATB, 648500ATB, 648600ATB, 649900ATR, 650000ATR, 650100ATR, 654100ATR
Insulin	170130BIJ, 170131BIJ, 170430BIJ, 170431BIJ, 175330BIJ, 175331BIJ, 175332BIJ, 175333BIJ, 441330BIJ, 441331BIJ, 441332BIJ, 441333BIJ, 441334BIJ, 461830BIJ, 461831BIJ, 461832BIJ, 484930BIJ, 484931BIJ, 488730BIJ, 626830BIJ, 626831BIJ
Exenatide	512130BIJ, 512131BIJ

3.3. Covariates

Several factors were accounted for in the analysis when assessing the association of diabetes onset on subsequent dementia. The covariates considered in this study included sociodemographic characteristics (sex and income level), health behavior factors (frequency of alcohol consumption, smoking and physical activity), and medical characteristics (BMI).

The Income level was categorized into three groups based on the health insurance premium percentile. Individuals in the 1st to 3rd quintiles were classified as the low-income group, those in the 4th to 7th quintiles as the middle-income group, and those in the 8th to 10th quintiles as the high-income group. The frequency of alcohol consumption was categorized into four groups: less than three times per month, 4–10 times per month, 10–16 times per month, and 17 or more times per month. Smoking frequency was categorized into four groups: fewer than 10 cigarettes per day, 10–20 cigarettes per day, 20–40 cigarettes per day, and more than 40 cigarettes per day. Physical activity frequency was categorized into four groups: 0–1 time per week, 2–3 times per week, 4–5 times per week, and 6–7 times per week. BMI was categorized into three groups: less than 25, 25 to less than 30, and 30 or higher.

4. Statistical Methods

4.1. Statistical tests for Difference

To examine the general characteristics of the study participants, a Chi-square test was conducted to examine the differences between 4 groups categorized by age of diabetes onset (40-49, 50-59, 60-69, 70+). Kaplan-Meier survival analysis and log-rank test were performed additionally to compare the cumulative incidence rates of the 4 groups categorized by age of diabetes onset.

4.2. Time on study analysis (Traditional Cox proportional hazards model)

Cox Proportional Hazard Regression is a statistical method commonly used in medical research and epidemiology to estimate the effect of covariates on survival time. It is a semi-parametric model used for analyzing time-to-event data. Cox regression is considered a semi-parametric model as it shows both parametric and non-parametric features. There is no assumption about the distribution of survival times; however, the model assumes proportionality hazards assumption, meaning that the effects of different variables on survival are constant over time and additive on a specific scale (Abd ElHafeez, D'Arrigo et al. 2021). The Cox regression model is based on the hazard function. Mathematically, the Cox model is written as follows:

$$h\left(t|X\right) = h_0\left(t\right)\exp\left(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p\right)$$

$H_0(t)$ is the baseline hazard at time t representing the hazard for an individual when all predictor variables are set to zero. The variables x_1 through x_p represent the set of variables that are used to model the relationship between covariates and the outcome of interest. By computing the exponential of the regression coefficient b_1 through b_p, the hazard ratio of a given risk factor or predictor can be calculated in the model. Cox regression does not estimate the baseline hazard directly. Instead, the regression coefficients b_1 through b_p are estimated using partial likelihood. The regression coefficients are estimated using partial likelihood maximization. The equation for calculating the regression coefficients is written as follows:

$$L(\beta) = \prod_{i=1}^{n} \frac{exp(\beta^T X_i)}{\Sigma_{j \in R(t_i)} exp(\beta^T X_j)}$$

For each observed event time t_i, the likelihood contribution is the probability that individual i (who experienced the event) has a higher hazard than others at risk at t_i. $R(t_i)$ is the risk set at time t_i, meaning individuals still at risk just before time t_i. The denominator sums over all individuals still at risk at time t_i. This function is maximized numerically to estimate beta. The followed figure shows the Time-on study scale.

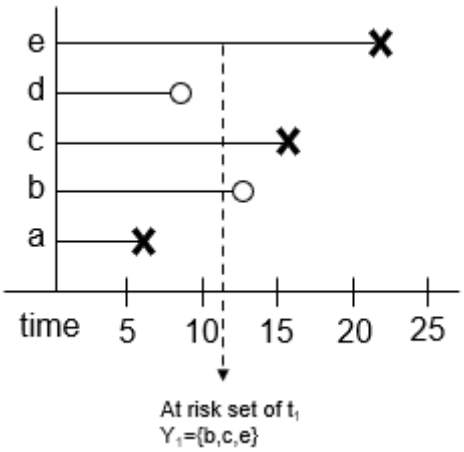

Figure 2. At Risk Set using Time on Study Scale

Time on study analysis was used in this study to estimate the risk of Dementia effected by the several factors of individuals. The factors used in this study were sociodemographic characteristics (sex and income level), health behavior factors (frequency of alcohol consumption, smoking and physical activity), and medical characteristics (BMI). Time to event was defined as the duration from the onset of diabetes to the earliest occurrence of either dementia diagnosis, death, or the end of the study period.

4.3. Age attained analysis

Time on study analysis uses absolute time as the time scale. However, using time-on study as the time scale does not account for the influence of age on disease incidence. For

example, if a 20-year-old and a 60-year-old enters the cohort at the same time, using time-on study as the time scale would mean that they would be given the same time-to-event. Treating as if their risks for a disease were the same would lead to incorrect results as it is a relevant thought that the 60-year-old participant would have a higher chance of developing a disease. Using attained age as the time scale can be a solution to the previously mentioned problem. Employing attained age as the time scale offers maximal flexibility in adjusting for age-related effects (Han, 2018).

Unlike time on study analysis, age attained analysis uses attained age as the time scale for Cox proportional hazards regressions. This means when an individual is observed at a certain age, their risk is compared to others of the same attained age, regardless of when they entered the study. When attained age is used as the time scale of the hazard function, it can be directly interpreted as an age-specific incidence function. Mathematically, the age attained analysis is written as follows:

$$\hbar\left(a \middle| A_{0,i} = a_0, Z_i = z\right) = \hbar_0^I(a) \cdot e^{z'\beta}$$

The equation denotes the hazard function at attained age a, given covariates z. $h_0^I(a)$ is the baseline hazard as a function of attained age. $e^{z'\beta}$ captures the effect of covariates. Using attained age as a time scale gives correct adjustment for the age at entry which is crucial in reducing bias of the estimated coefficients (Pencina, Larson et al. 2007).

In this study, an age attained analysis was chosen, as entry age played a critical role in examining the association between age at diabetes onset and the subsequent risk of dementia. Additional consideration was needed for the proper choice of time scale. As no individuals in the data is followed from birth until the event of interest and as individuals enter cohort at different point of times, left-truncation was taken into account as the most appropriate refinement. Mathematically, the left-truncated age attained analysis is written as follows:

$$\lambda\left(a \middle| A_{0,i} = a_0, Z_i = z\right) = \lambda_0^{II}(a|a_0) \cdot e^{z'\beta}$$

The point of origin is defined at $A=A_{0,i}$ which shows the hazard also depends on a_0. The conditioning on a_0 controls for the truncated individuals, who have developed the event prior to the ages at entry of the cohort. Figures for Age scale and Left truncated Age scale is provided for better understanding.

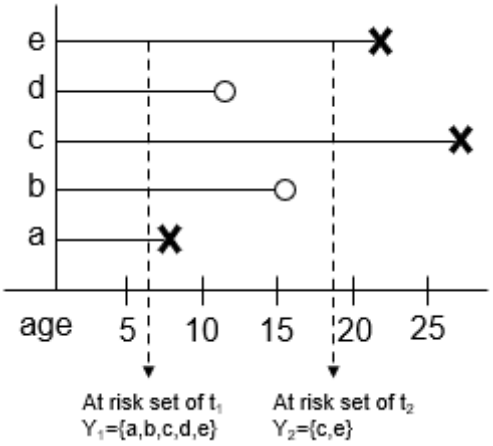

Figure 3. At Risk Set Using Age Scale

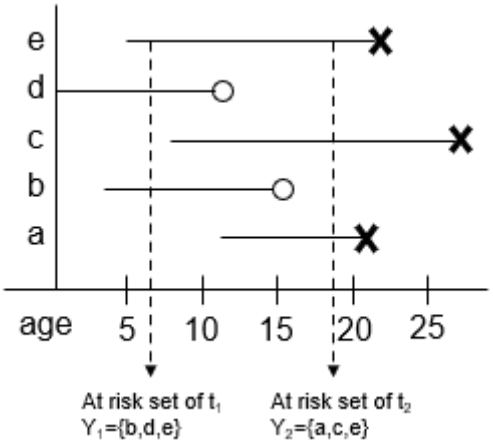

Figure 4. At Risk Set using Age as the Time Scale with Left Truncation

Different from Figure 2, Figure 3 and Figure 4 has the time scale controlled as age. Comparing the two different choice of time scales, individuals included at risk are different. In Figure 2, individuals who are at risk are {a, b, c, d, e} of t1, as in Figure 3, individuals who are at risk are {b, d, e} as it is the time point when participants {a, c} did not enter the

cohort yet. That is, the choice of the time scale determines the composition of the risk set, altering the group of individuals who remain at risk of experiencing the event after time t_i.

4.4. Competing risk analysis

In trials involving survival data, competing risks frequently arise, necessitating the classification of failure events not only by their occurrence of failure and censoring, but also by their specific cause, commonly designated as type I or type II (Lunn and McNeil 1995). Competing risk analysis is needed when multiple types of events occur, and one prevents the occurrence of another while studying time-to-event outcomes. Traditional survival analysis such as Kaplan-Meier and Cox models assumes that all individuals would eventually experience the event of interest if followed long enough. However, this assumption is not true in real data when competing risks exist.

To analyze competing-risk data, cause-specific hazard function was used in this study. The cause-specific hazard function gives the instantaneous risk of experiencing a particular event at a given time, assuming the individual has survived up to that time. Mathematically, the cause-specific hazard function is written as follows:

$$\lambda_k(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, event\ type = k | T \geq t)}{\Delta t}$$

The Cox proportional hazards model can be effectively extended to analyze the cause-specific hazards functions from different event types. For cause k, a separate hazards model can be assumed,

$$\hbar_k(t|Z) = \hbar_{k0}(t)\exp(\beta'_k Z)$$

Where $h_{k0}(t)$ is the baseline of the cause-specific hazard function and the vector β_k represents the covariate effects on the event of interest. Because the K models are mutually exclusive, each can be analyzed independently. The coefficients β_k are estimated by maximizing the modified partial likelihood.

III. Results

1. Characteristics of study population

There were 77,092 individuals selected as the final study population, and the general characteristics of study population were shown in Table 4. Firstly, the association between income level and the groups of diabetes onset age is significant($p<0.0001$). As shown in the table, the income level of the study participants mainly belonged to high-income. Low-income group accounted for 24.24%, middle-income group accounted for 33.79% and high-income group accounted for 41.97% of the study population. Male occupied a higher percentage by 59.40% than Female participants with 40.50%. However, as the age group increases, the proportion of female participants tends to rise. The occupation of Female participants exceeds the male participants from age group of 70-79. Majority of the participants smoke less than 10 cigarettes a day (Less than 10 cigarettes per day: 90.36%, 10-20 cigarettes per day: 2.23%, 20-40 cigarettes per day: 5.03%, more than 40 cigarettes per day: 2.38%). The drink group of the study population mainly belonged to the group who drinks less than 3 times per month with 73.07%, followed by 4-8 times per month (11.59%), 12-16 times per month (7.34%), more than 17 times a month (8.00%). The percentage occupied by participants who drank less than 3 times per month took a bigger portion with the group who were late at diabetes onset. As for frequency of physical activity, the group with 0-1 times per week accounted for more than half of the participants (0-1 times per week: 58.43%, 2-3 times per week: 15.71%, 4-5 times per week: 6.08%, 6-7 times

per week: 19.78%). The groups categorized by BMI rate of less than 25, 25-30 and 30 or higher took 51.30%, 4,25% and 6.45%. The percentage occupied by participants who had a BMI rate less than 25 took a bigger portion with the group who were late of diabetes onset. Participants who had BMI rate that is 30 or higher had a tendency to be in a group with lower age of diabetes onset. Lastly, 7.32 percent of the participants developed dementia during the study period. Late age of diabetes onset seems to take a bigger percentage of Dementia. Since the study period spans 16 years, it appears that the older the entry age, the higher the risk of developing dementia.

For the study participants in the study period, the variables used for the model (Income level, sex, smoking frequency, drinking frequency, frequency of physical activity, BMI, Occurrence of dementia) showed statistically significant differenced by the 4 groups of diabetes onset age (age 41-59, 60-69, 70-79, 80+).

Table 4. General Characteristics of the Study Population

Variables	Total	Age of diabetes onset				*P*-value
		Age (40-49)	Age (50-59)	Age (60-69)	Age (70+)	
Total	77,092	35,136	26,315	13,358	2,283	
Income level						<0.0001
Low	18,684 (24.24)	7,103 (20.22)	7,021 (26.68)	3,887 (29.10)	673 (29.48)	
Middle	26,051 (33.79)	11,835 (33.68)	9,189 (34.92)	4,399 (32.93)	628 (27.51)	
High	32,357 (41.97)	16,198 (46.10)	10,105 (38.40)	5,072 (37.97)	982 (43.01)	
Sex						<0.0001
Male	45,794 (59.40)	23,672 (67.37)	14,657 (55.70)	6,426 (48.11)	1,039 (45.51)	
Female	31,298 (40.50)	11,464 (32.63)	11,658 (44.30)	6,932 (51.89)	1,244 (54.49)	
Smoke						<0.0001
<10 per day	69,657 (90.36)	31,630 (90.02)	23,734 (90.19)	12,202 (91.35)	2,091 (91.59)	
10-20	1,721 (2.23)	986 (2.81)	564 (2.14)	156 (1.17)	15 (0.02)	
20-40	3,878 (5.03)	1,923 (5.47)	1,380 (5.24)	500 (3.74)	75 (3.29)	
≥ 40	1,836 (2.38)	597 (1.70)	637 (2.42)	500 (3.74)	102 (4.47)	
Drink						<0.0001
<3 times per month	56,331 (73.07)	23,387 (66.56)	20,053 (76.20)	10,951 (81.98)	1,940 (84.98)	
4-10 times	8,932 (11.59)	5,362 (15.26)	2,654 (10.09)	822 (6.15)	94 (4.12)	
10-16 times	5,664 (7.34)	3,401 (9.68)	1,613 (6.13)	581 (4.35)	66 (2.89)	
>17 times	6,168 (8.00)	2,986 (8.50)	1,995 (7.58)	1,004 (7.52)	183 (8.02)	

Table 4. General Characteristics of the Study Population (Cont.)

Variables	Total	Age of diabetes onset				*P*-value
		Age (41-59)	Age (60-69)	Age (70-79)	Age (80+)	
Total	77,092	35,136	26,315	13,358	2,283	
Physical activity						<0.0001
0-1 times per week	45,044 (58.43)	19,277 (54.86)	15,767 (59.92)	8,367 (62.64)	1,633 (71.53)	
2-3	12,115 (15.71)	5,271 (15.00)	4,186 (15.91)	2,331 (17.45)	327 (14.32)	
4-5	4,686 (6.08)	2,553 (7.27)	1,431 (5.44)	617 (4.62)	85 (3.72)	
6-7	15,247 (19.78)	8,035 (22.87)	4,931 (18.74)	2,043 (15.29)	238 (10.42)	
BMI						<0.0001
Less than 25	39,549 (51.30)	17,094 (48.65)	13,461 (51.15)	7,547 (56.50)	1,447 (63.38)	
25-30	32,569 (42.25)	15,474 (44.04)	11,215 (42.62)	5,121 (6.64)	759 (33.25)	
30 or higher	4,974 (6.45)	2,568 (7.31)	1,639 (6.23)	690 (5.17)	77 (3.37)	
Dementia						<0.0001
Yes	5,642 (7.32)	515 (1.47)	2,104 (8.00)	2,509 (18.78)	514 (22.51)	
No	71,450 (92.68)	34,621 (98.53)	24,211 (92.00)	10,849 (81.22)	1,769 (77.49)	

2. Cumulative Incidence Curve and Gray's test

Cumulative Incidence functions were used to analyze the probability of developing dementia over time across different age at diabetes onset groups, while accounting for the competing risk of death. In Figure 5, the x-axis represents time in days, and the y-axis shows the cumulative incidence of dementia. The blue line represents the dementia incidence among individuals with diabetes onset at ages 40-49, the red line for ages 50-59, the green line for ages 60-69, and the brown line for those diagnosed at age 70 or older. The cumulative incidence of dementia was higher among those diagnosed with diabetes at younger ages, indicating that earlier onset of diabetes was associated with a greater risk of developing dementia over time. Gray's test was conducted to assess whether the differences among the four age-at-onset groups were statistically significant, and a significant difference was observed ($p < 0.0001$).

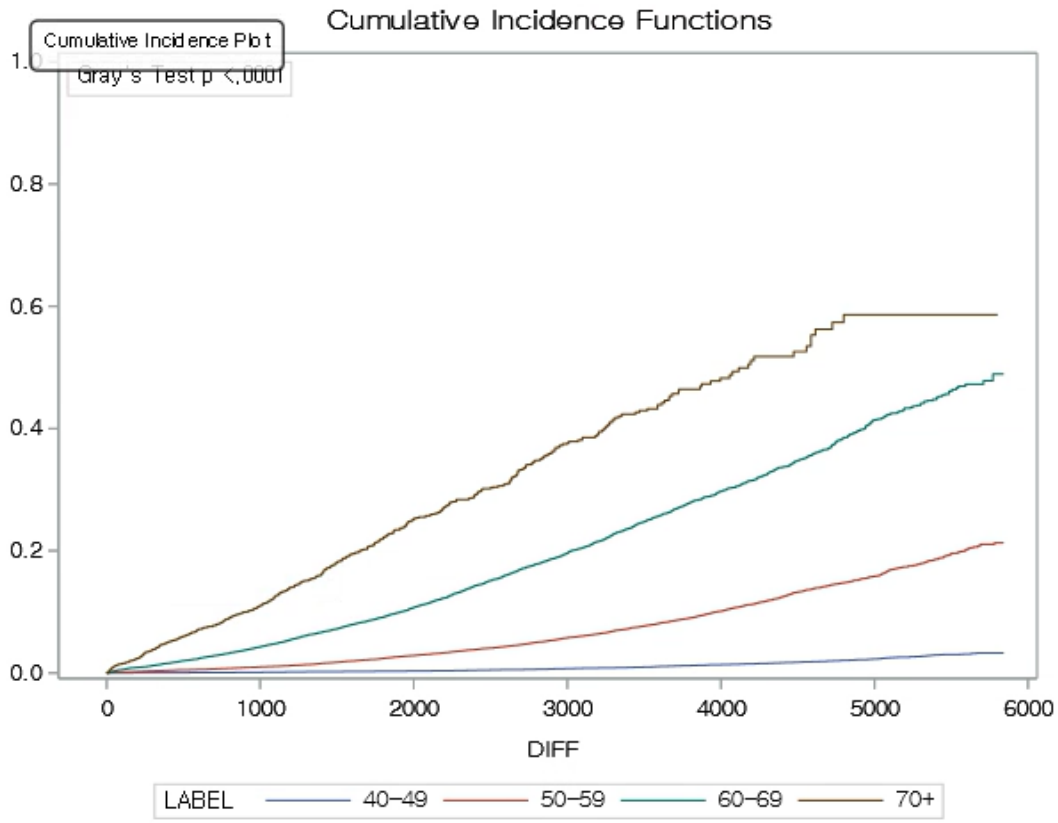

Figure 5. Cumulative Incidence Curves for Dementia Risk by Diabetes Onset Age Group

3. Time on study analysis

Before performing a time on study analysis, a log-log plot was drawn to assess the proportional hazards assumption. This assumption states that the hazard ratio between groups remains constant over time. The parallel lines in Figure 6 suggest that the proportional hazards assumption is satisfied. Table 5 shows the result of the time on study analysis.

Three distinct models were constructed, each employing different covariate adjustment. Model 1 performed time on study analysis using only the diabetes onset age as a variable. In Model 2, sex, income level and age at baseline was added. Model 3 included all the variables which added Frequency of smoking, drinking, physical activity and BMI level to Model 2. From the result of time on study analysis, most of the variables used in model 1,2,3 showed a statistically significant association with the occurrence of dementia. The significant determinants were diabetes onset age, income level, sex, age at baseline, smoking frequency, drinking frequency and frequency of physical activity (Table 5). Especially, after adjusting other risk factors, diabetes onset age showed a statistically significant difference in occurrence of dementia. Compared to diabetes onset age of 40-49, the diabetes onset age group of 50-59 had a 3.02 times higher risk of developing dementia. The diabetes onset age group of 60-69 showed a 7.38 times higher risk of developing dementia than the diabetes onset age group of 40-49. The 70+ diabetes onset age group had a 15.77 times higher risk compared to diabetes onset age group of 40-49 (diabetes onset

age 50-59: HR 3.02, 95% CI 2.66-3.43, diabetes onset age 60-69: HR 7.38, 95% CI 6.38-8.54, diabetes onset age 70+: HR 15.77, 95% CI 13.12-18.95). Regarding the income level, the middle-income level had 1% lower risk of developing dementia compared to the low-income group. The high-income group showed a 22% lower risk of developing dementia compared to low-income group and it was statistically significant (middle-income: HR 0.99, 95% CI 0.92-1.05, high-income: HR 0.78, 95% CI 0.74-0.84). Female had a 1.52 times higher risk of developing dementia and the difference was statistically significant (HR 1.52, 95% CI 1.43-1.61). In addition, compared to individuals who smoked less than 10 cigarettes a day, individuals who smoked from 10 to 20 cigarettes had a 13% lower risk of developing dementia, individuals who smoked from 20-40 cigarettes per day had a 14% lower risk and who smoked more than 40 cigarettes per day had a 28% lower risk of developing dementia and the difference was statistically significant (10-20 per day: HR 0.87, 95% CI 0.72-1.06, 20-40 per day: HR 0.86, 95% CI 0.65-1.13, >=40 per day: HR 0.72, 95% CI 0.63-0.83). With frequency of drinking alcohol, compared to the group that had alcohol less than 3 times per month, the group that consumed alcohol 4-10 times per month had a 1 percent higher risk of developing dementia, the group consuming alcohol 11-16 times per month had a 8% higher risk of developing dementia and the group that consumed alcohol more than 17 times per month had a 1.16 times higher risk of developing dementia (4-10 times per month: HR 1.01, 95% CI 0.90-1.14, 11-16 times per month: HR 1.08, 95% CI 0.94-1.24, >17 times per month: HR 1.16, 95% CI 1.04-1.29). With frequency of physical activity, compared to the group that worked out 0-1 times per week, the group

who worked out 2-3 times per week had the same risk of developing dementia, the group that worked out 4-5 times per week had a 13% lower risk and the group that worked out 6-7 times per week had a 24% lower risk of developing dementia and the difference was statistically significant (2-3 times per week: HR 1.00, 95% CI 0.93-1.07, 3-4 times per week: HR 0.87 95% CI 0.77-0.99, 5-6 times per week: HR 0.76, 95% CI 0.70-0.83). Lastly, compared to the group who had BMI less than 25, the BMI group of 25-30 had a 6% lower risk and the group who had BMI of 30 or higher had a 8% lower risk of developing dementia and the difference was statistically significant (BMI 25-30: HR 0.94, 95% CI 0.89-1.00, BMI 30+: HR 0.92, 95% CI 0.82-1.04).

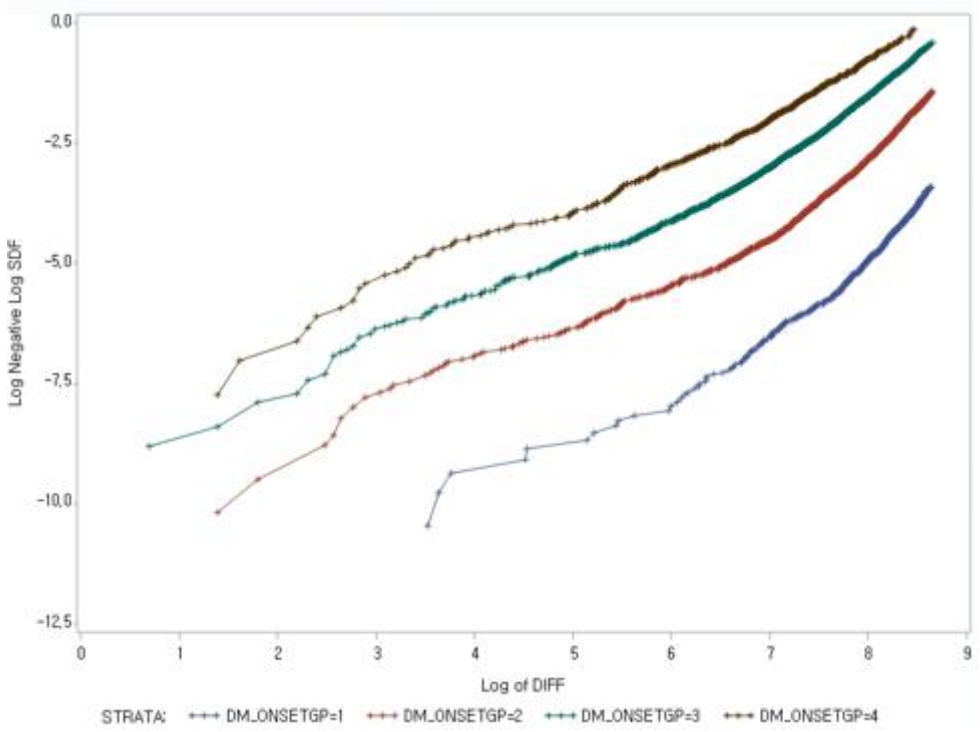

Figure 6. Log-Log Plot of Survival Curves by Diabetes Onset Group

Table 5. Results of time on study analysis on Dementia Associated with Diabetes Onset Age

Variables	Dementia					
	Model 1		Model 2		Model 3	
	HR	95% CI	HR	95% CI	HR	95% CI
Diabetes onset age						
40-49	1.00		1.00		1.00	
50-59	5.88	(5.34 - 6.47)	2.99	(2.64 - 3.40)	3.02	(2.66 - 3.43)
60-69	16.15	(14.69 - 17.76)	7.34	(6.34 - 8.49)	7.38	(6.38 - 8.54)
70+	21.90	(19.38 - 24.75)	15.71	(13.08 - 18.88)	15.77	(13.12 - 18.95)
Income level						
Low			1.00		1.00	
Middle			0.98	(0.91 - 1.04)	0.99	(0.92 - 1.05)
High			0.77	(0.72 - 0.82)	0.78	(0.74 - 0.84)
Sex						
Male			1.00		1.00	
Female			1.43	(1.35 - 1.51)	1.52	(1.43 - 1.61)
Age						
40-49			1.00		1.00	
50-59			2.74	(2.32 - 3.24)	2.67	(2.26 - 3.15)
60-69			4.84	(4.02 - 5.82)	4.69	(3.90 - 5.64)
70+			5.00	(4.08 - 6.12)	4.66	(3.80 - 5.72)

Model 1: Adjusted only for Diabetes onset age
Model 2: Adjusted for Diabetes onset age, sex and income level
Model 3: Adjusted for Diabetes onset age, sex and income level, smoke, drink, physical activity and BMI

Table 5. Results of time on study analysis on Dementia Associated with Diabetes Onset Age (Cont.)

Variables	Dementia					
	Model 1		Model 2		Model 3	
	HR	95% CI	HR	95% CI	HR	95% CI
Smoke						
<10 per day					1.00	
10-20					0.87	(0.72 - 1.06)
20-40					0.86	(0.65 - 1.13)
≥ 40					0.72	(0.63 - 0.83)
Drink						
<3 times					1.00	
4-10 times					1.01	(0.90 - 1.14)
11-16 times					1.08	(0.94 - 1.24)
>17 times					1.16	(1.04 - 1.29)
Physical activity						
0-1 times					1.00	
2-3					1.00	(0.93 - 1.07)
4-5					0.87	(0.77 - 0.99)
6-7					0.76	(0.70 - 0.83)
BMI						
Less than 25					1.00	
25-30					0.94	(0.89 - 1.00)
30 or higher					0.92	(0.82 - 1.04)

Model 1: Adjusted only for Diabetes onset age
Model 2: Adjusted for Diabetes onset age, sex and income level
Model 3: Adjusted for Diabetes onset age, sex and income level, smoke, drink, physical activity and BMI

4. Age attained analysis

Results of time on study analysis indicated that an increase in the age of diabetes onset was associated with a higher risk of developing dementia. However, given that this study focuses on the incidence of dementia, it is crucial to adjust for individual's age at cohort entry. Several studies have demonstrated that simply adjusting for baseline age when using time on study analysis is not sufficient to appropriately account for the influence of age at cohort entry. Therefore, an age attained analysis was conducted. By using this analytical approach, the time scale was set to age, thereby adjusting for entry age into the cohort. Additionally, since this study was conducted on individuals with diabetes, left truncation was applied at the age of diabetes onset.

After adjusting for other risk factors, a higher age at diabetes onset was associated with a reduced risk of developing dementia. Comparing to diabetes onset age of 40-49, diabetes onset age of 50-59 had a 10% lower risk of developing dementia, diabetes onset age of 60-69 showed a 22% lower risk of developing dementia and diabetes onset age over 70 showed a 36% lower risk of developing dementia. The difference was statistically significant except for the age of 50-59 group (diabetes onset age 50-59: HR 0.90, 95% CI 0.78-1.03, diabetes onset age 60-69: HR 0.78, 95% CI 0.67-0.91, diabetes onset age70+: HR 0.64, 95% CI 0.52-0.77). As for income level, higher income showed a tendency to have a lower risk of dementia. Comparing to the low-income group, middle-income group showed a 0.02% decrease in the risk of developing dementia and the high-income group had a 26% lower risk of developing dementia (middle-income: HR 0.998, 95% CI 0.93-1.07, high-income:

HR 0.74, 95% CI 0.74-0.85). Female had a 1.5 times higher risk of developing dementia compared to male individuals (HR: 1.50, 95% CI 1.41-1.60). Regarding to smoking frequency, the group with higher smoking levels had a tendency for a lower risk of dementia. Comparing to the group who smoked less than 10 cigarettes per day, the group who smoked 10-20 cigarettes per day had a 13% lower risk of developing dementia, the group that smoked 20-40 cigarettes per day had a 10% lower risk of developing dementia and the group that smoked more than 40 cigarettes per day had a 26% lower risk of developing dementia (10-20 per day: HR 0.87, 95% CI 0.71-1.07, 20-40 per day: HR 0.90, 95% CI 0.68-1.19, >=40 per day: HR 0.74, 95% CI 0.64-0.85). For frequency of consuming alcohol, higher levels of alcohol consumption showed an increased risk of dementia. Compared to the group that consumed alcohol less than 3 times per month, the group that consumed alcohol 4-10 times a month showed a 1.04 times higher risk of developing dementia, the group that consumed alcohol 11-16 times per month showed a 1.09 times higher risk and the group that consumed alcohol more than 17 times per month showed a 1.13 times higher risk of developing dementia but only the group that consumed alcohol more than 17 times per month were statistically significant (4-10 times per month: HR 1.04, 95% CI 0.93-1.17, 11-16 times per month: HR 1.09, 95% CI 0.94-1.25, >17 times per month: HR 1.13, 95% CI 1.02-1.27). Related to frequency of physical activity, comparing to the group who worked out 0-1 times per week, the group that worked out 2-3 times per week showed a 1.02 times higher risk of developing dementia, the group who worked out 4-5 times per week had a 10% lower risk of developing dementia and the group who worked out 6-7

times per week had a 21% lower risk of developing dementia (2-3 times per week: HR 1.02, 95% CI 0.95-1.09, 3-4 times per week: HR 0.90 95% CI 0.79-1.02, 5-6 times per week: HR 0.79, 95% CI 0.73-0.86). Lastly, the higher BMI showed a tendency for lower dementia. Compared to the group that had lower BMI than 25, the BMI group of 25-30 showed a 4% lower risk of developing dementia and the group with BMI of 30 or higher showed a 7% lower risk of developing dementia but the results were not statistically significant (BMI 25-30: HR 0.96, 95% CI 0.91-1.02, BMI 30+: HR 0.93, 95% CI 0.83-1.05).

Table 6. Results of Age Attained Analysis on Dementia Associated with Diabetes Onset Age

Variables	Dementia					
	Model 1		Model 2		Model 3	
	HR	95% CI	HR	95% CI	HR	95% CI
Diabetes onset age						
40-49	1.00		1.00		1.00	
50-59	0.93	(0.81 - 1.06)	0.90	(0.78 - 1.02)	0.90	(0.78 - 1.03)
60-69	0.82	(0.71 - 0.96)	0.79	(0.68 - 0.92)	0.78	(0.67 - 0.91)
70+	0.69	(0.57 - 0.83)	0.66	(0.54 - 0.80)	0.64	(0.52 - 0.77)
Income level						
Low			1.00		1.00	
Middle			0.99	(0.93 - 1.06)	1.00	(0.93 - 1.07)
High			0.78	(0.73 - 0.84)	0.74	(0.74 - 0.85)
Sex						
Male			1.00		1.00	
Female			1.42	(1.34 - 1.50)	1.50	(1.41 - 1.60)
Smoke						
<10 per day					1.00	
10-20					0.87	(0.71 - 1.07)
20-40					0.90	(0.68 - 1.19)
≥ 40					0.74	(0.64 - 0.85)

Table 6. Results of Age Attained Analysis on Dementia Associated with Diabetes Onset Age (Cont.)

Variables	Dementia					
	Model 1		Model 2		Model 3	
	HR	95% CI	HR	95% CI	HR	95% CI
Drink						
<3 times a month					1.00	
4-10					1.04	(0.93 - 1.17)
11-16					1.09	(0.94 - 1.25)
>17					1.13	(1.02 - 1.27)
Physical activity						
0-1 times a week					1.00	
2-3					1.02	(0.95 - 1.09)
4-5					0.90	(0.79 - 1.02)
6-7					0.79	(0.73 0.86)
BMI						
Less than 25					1.00	
25-30					0.96	(0.91 - 1.02)
30 or higher					0.93	(0.83 - 1.05)

Model 1: Adjusted only for Diabetes onset age

Model 2: Adjusted for Diabetes onset age, sex and income level

Model 3: Adjusted for Diabetes onset age, sex and income level, smoke, drink, physical activity and BMI

5. Comparison of results using different time scales

The findings from time on study analysis and age attained analysis were evaluated to determine the influence of diabetes onset age on the risk of developing dementia. The comparison of results among analysis is summarized in Table 7.

Income level showed an association with risk of dementia occurrence. Higher income level showed a lower risk of subsequent dementia in both models. ([Time on study analysis]: middle-income: HR 0.93 95% CI 0.87-0.99 high-income: HR 0.80, 95% CI 0.75-0.85), [Age attained analysis]: middle-income: HR 0.998, 95% CI 0.93-1.07, high-income: HR 0.74, 95% CI 0.74-0.85). Female had a higher risk of developing dementia in both models ([Time on study analysis]: HR 1.64, 95% CI 1.54-1.74, [Age attained analysis]: HR: 1.50, 95% CI 1.41-1.60). A trend was observed where higher smoking levels were associated with a lower risk of dementia ([Time on study analysis]: 10-20 per day: HR 0.70, 95% CI 0.58-0.85, 20-40 per day: HR 0.60, 95% CI 0.45-0.79, >=40 per day: HR 0.61, 95% CI 0.53-0.70, [Age attained analysis]: 10-20 per day: HR 0.87, 95% CI 0.71-1.07, 20-40 per day: HR 0.90, 95% CI 0.68-1.19, >=40 per day: HR 0.74, 95% CI 0.64-0.85). As for consumption of alcohol, there was a tendency for higher alcohol consumption to be associated with a higher risk of dementia in both models ([Time on study analysis]: 4-12 times per month: HR 0.86, 95% CI 0.77-0.97, 12-16 times per month: HR 0.92, 95% CI 0.80-1.06, >17 times per month: HR 1.13, 95% CI 1.01-1.252, 5-6 times per week: HR 0.62, 95% CI 0.57-0.67, [Age attained analysis]: 4-12 times per month: HR 1.04, 95% CI 0.93-1.17, 12-16 times per month: HR 1.09, 95% CI 0.94-1.25, >17 times per month: HR

1.13, 95% CI 1.02-1.27). Furthermore, there was a tendency for higher exercise frequency to be associated with a lower risk of dementia ([Time on study analysis]: 2-3 times per week: HR 0.89, 95% CI 0.83-0.95, 3-4 times per week: HR 0.72 95% CI 0.63-0.82, 5-6 times per week: HR 0.62, 95% CI 0.57-0.67, [Age attained analysis]: 2-3 times per week: HR 1.02, 95% CI 0.95-1.09, 3-4 times per week: HR 0.90 95% CI 0.79-1.02, 5-6 times per week: HR 0.79, 95% CI 0.73-0.86). Additionally, higher BMI levels showed a lower risk of developing dementia in both models ([Time on study analysis]: BMI 25-30: HR 0.87, 95% CI 0.82-0.92, BMI 30+: HR 0.80, 95% CI 0.71-0.90, [Age attained analysis]: BMI 25-30: HR 0.96, 95% CI 0.91-1.02, BMI 30+: HR 0.93, 95% CI 0.83-1.05).

Although the two models exhibited similar trends for other factors, they showed opposite results regarding diabetes onset age. Using time on study analysis, higher diabetes onset age was associated with an increased risk of dementia. This suggests that a later onset of diabetes is associated with a higher risk of dementia. However, when selecting attained age as a time scale using Age attained analysis and left truncating it by diabetes onset age, higher diabetes onset age showed relations to lower risk of dementia, suggesting a later onset of diabetes having association to a lower risk of dementia ([Time on study analysis]: diabetes onset age 40-49: HR 5.30, 95% CI 4.81-5.84, diabetes onset age 50-59: HR 13.74, 95% CI 12.48-15.12, diabetes onset age 60-69: HR 17.71, 95% CI 15.65-20.05, [Age attained analysis]: diabetes onset age 40-49: HR 0.90, 95% CI 0.78-1.03, diabetes onset age 50-59: HR 0.78, 95% CI 0.67-0.91, diabetes onset age 60-69: HR 0.64, 95% CI 0.52-0.77).

In conclusion, time on study analysis and Age attained analysis identified notable correlations between demographic characteristics, heath behavior and healthcare factors and the risk of developing subsequent dementia. Income level, sex, frequency of smoking, frequency of alcohol consumption, frequency of physical activity and BMI level showed were consistently linked to the risk of subsequent dementia in both of the models. In contrary, the relationship between diabetes onset age and subsequent dementia yielded reverse results depending on the analytical approach employed.

Table 7. **Comparison of results among analyses**

Variables	Time on Study Analysis		Age Attained Analysis	
	HR	**95% CI**	**HR**	**95% CI**
Diabetes onset age				
40-49	1.00		1.00	
50-59	3.02	(2.66 - 3.43)	0.90	(0.78 - 1.03)
60-69	7.38	(6.38 - 8.54)	0.78	(0.67 - 0.91)
70+	15.77	(13.12 - 18.95)	0.64	(0.52 - 0.77)
Income level				
Low	1.00		1.00	
Middle	0.99	(0.92 - 1.05)	1.00	(0.93 - 1.07)
High	0.78	(0.74 - 0.84)	0.74	(0.74 - 0.85)
Sex				
Male	1.00		1.00	
Female	1.52	(1.43 - 1.61)	1.50	(1.41 - 1.60)
Age				
40-49	1.00			
50-59	2.67	(2.26 - 3.15)		
60-69	4.69	(3.90 - 5.64)		
70+	4.66	(3.80 - 5.72)		

Table 7. Comparison of results among analyses (Cont.)

Variables	Time on Study Analysis		Age Attained Analysis	
	HR	95% CI	HR	95% CI
Smoke				
<10 per day	1.00		1.00	
10-20	0.87	(0.72 - 1.06)	0.87	(0.71 - 1.07)
20-40	0.86	(0.65 - 1.13)	0.90	(0.68 - 1.19)
≥ 40	0.72	(0.63 - 0.83)	0.74	(0.64 - 0.85)
Drink				
<3 times	1.00		1.00	
4-12 times	1.01	(0.90 - 1.14)	1.04	(0.93 - 1.17)
12-16 times	1.08	(0.94 - 1.24)	1.09	(0.94 - 1.25)
>17 times	1.16	(1.04 - 1.29)	1.13	(1.02 - 1.27)
Physical activity				
0-1 times	1.00		1.00	
2-3	1.00	(0.93 - 1.07)	1.02	(0.95 - 1.09)
4-5	0.87	(0.77 - 0.99)	0.90	(0.79 - 1.02)
6-7	0.76	(0.70 - 0.83)	0.79	(0.73 - 0.86)
BMI				
Less than 25	1.00		1.00	
25-30	0.94	(0.89 - 1.00)	0.96	(0.91 - 1.02)
30 or higher	0.92	(0.82 - 1.04)	0.93	(0.83 - 1.05)

6. Competing risk analysis

To clarify whether diabetes onset age differentially impacts the different types of dementia and to prevent overestimation of dementia risk by accounting for death as a competing risk, competing risk analysis was used for the study. Also, to adjust for the age at the time of cohort entry, the time scale was chosen as age.

Dementia was classified into three groups, Alzheimer's disease, vascular dementia and other types of dementia. For income level, there was a trend of higher income level associated with lower risk of dementia. It was consistent except for vascular dementia where middle income level had a higher risk of dementia compared to low-income level group ([Alzheimer's disease]: middle-income: HR 0.98, 95% CI 0.91-1.06, high income: HR 0.84, 95% CI 0.78-0.91 [Vascular dementia]: middle-income: HR 1.14 95% CI 0.94-1.37 high-income: HR 0.97 95% CI 0.81-1.18 [Other dementia]: middle-income: HR 0.92 95% CI 0.79-1.07, high-income: HR 0.81 95% CI 0.69-0.94). Female showed a higher risk of dementia in all three dementia groups. Alzheimer's disease showed the strongest effect regarding to sex and vascular dementia showed the smallest effect ([Alzheimer's disease]: HR 1.74 95% CI 1.62-1.87 [Vascular dementia]: HR 1.15 95% CI 0.97-1.36 [Other dementia]: HR 1.56 95% CI 1.3-1.80). For smoking frequency, the association exhibited different patterns across the three types of dementia. In all three types of dementia, there was no clear directional difference based on smoking frequency ([Alzheimer's disease]: 10-20 per day: HR 0.85, 95% CI 0.66-1.08, 20-40 per day: HR 0.75, 95% CI 0.52-1.07, >=40 per day: HR 1.09, 95% CI 0.91-1.30 [Vascular dementia]: 10-20 per day: HR 0.80,

95% CI 0.51-1.25, 20-40 per day: HR 0.70, 95% CI 0.37-1.33, >=40 per day: HR 0.72, 95% CI 0.51-1.02 [Other dementia]: 10-20 per day: HR 0.65, 95% CI 0.41-1.01, 20-40 per day: HR 0.87, 95% CI 0.49-1.54, >=40 per day: HR 0.80, 95% CI 0.59-1.10). Frequency of alcohol consumption showed a trend of heavier alcohol consumption with a higher risk of Alzheimer's disease but didn't show a clear trend in vascular dementia and other types of dementia ([Alzheimer's disease]: 4-10 times per month: HR 1.00, 95% CI 0.87-1.15, 11-16 times per month: HR 1.05, 95% CI 0.89-1.25, >17 times per month: HR 1.14, 95% CI 1.00-1.29 [Vascular dementia]: 4-10 times per month: HR 0.98, 95% CI 0.73-1.31, 11-16 times per month: HR 1.15, 95% CI 0.82-1.60, >17 times per month: HR 1.06, 95% CI 0.80-1.40 [Other dementia]: 4-10 times per month: HR 0.96, 95% CI 0.73-1.25, 11-16 times per month: HR 1.00, 95% CI 0.72-1.38, >17 times per month: HR 0.81, 95% CI 0.61-1.07). Also, frequency of physical activity did not show a clear trend in all three types of dementia ([Alzheimer's disease]: 2-3 times per week: HR 1.25, 95% CI 1.16-1.36, 3-4 times per week: HR 1.23 95% CI 1.07-1.42, 5-6 times per week: HR 0.99, 95% CI 0.90-1.09 [Vascular dementia]: 2-3 times per week: HR 0.96, 95% CI 0.78-1.18, 3-4 times per week: HR 0.62 95% CI 0.40-0.97, 5-6 times per week: HR 0.70, 95% CI 0.54-0.89 [Other dementia]: 2-3 times per week: HR 0.95, 95% CI 0.80-1.13, 3-4 times per week: HR 0.76 95% CI 0.54-1.07, 5-6 times per week: HR 0.90, 95% CI 0.75-1.09). BMI level showed different trends in dementia types. For Alzheimer's disease, higher BMI showed a higher risk of dementia. On the contrary, higher BMI showed a lower risk of dementia for vascular dementia and other types of dementia ([Alzheimer's disease]: BMI 25-30: HR 1.04, 95%

CI 0.97-1.11, BMI 30+: HR 1.05, 95% CI 0.92-1.21 [Vascular dementia]: BMI 25-30: HR 0.93, 95% CI 0.80-1.09, BMI 30+: HR 0.79, 95% CI 0.55-1.14 [Other dementia]: BMI 25-30: HR 0.96, 95% CI 0.84-1.09, BMI 30+: HR 0.73, 95% CI 0.53-0.99). Lastly, when examining the relationship between diabetes onset age -the variable of greatest interest- and different types of dementia, a consistent trend was observed across all three types: the later the onset of diabetes, the lower the risk of dementia. Vascular dementia seemed to have the strongest effect of diabetes onset age followed by Alzheimer's disease and other types of dementia ([Alzheimer's disease]: diabetes onset age 50-59: HR 0.92, 95% CI 0.82-1.04, diabetes onset age 60-69: HR 0.73, 95% CI 0.66-0.82, diabetes onset age 70+: HR 0.48, 95% CI 0.42-0.55 [Vascular dementia]: diabetes onset age 50-59: HR 0.91, 95% CI 0.73-1.13, diabetes onset age 60-69: HR 0.71, 95% CI 0.57-0.88, diabetes onset age 70+: HR 0.46, 95% CI 0.33-0.64 [Other dementia]: diabetes onset age 50-59: HR 0.88, 95% CI 0.72-1.08, diabetes onset age 60-69: HR 0.74, 95% CI 0.61-0.91, diabetes onset age 70+: HR 0.61, 95% CI 0.47-0.78).

In summary, the analysis showed the differential impact of these factors across various types of dementia. Income, sex and diabetes onset age exhibited a similar trend in all types of dementia. As for Frequency of smoking and physical activity, there was no clear directional difference in types of dementia. For frequency of alcohol consumption, a clear trend was observed only in Alzheimer's disease, while no distinct pattern was shown in other types of dementia. Similarly, higher BMI showed higher risk only in Alzheimer's disease and lower risk in vascular dementia and other types of dementia.

Table 8. Results of Competing Risk Analysis for Different Types of Dementia

Variables	Types of Dementia								
	Alzheimer's disease			Vascular dementia			Other		
	HR	95% CI		HR	95% CI		HR	95% CI	
		Lower	Upper		Lower	Upper		Lower	Upper
Diabetes onset age									
40-49	1.00			1.00			1.00		
50-59	0.92	(0.82	- 1.04)	0.91	(0.73	- 1.13)	0.88	(0.72	- 1.08)
60-69	0.73	(0.66	- 0.82)	0.71	(0.57	- 0.88)	0.74	(0.61	- 0.91)
70+	0.48	(0.42	- 0.55)	0.46	(0.33	- 0.64)	0.61	(0.47	- 0.78)
Income level									
Low	1.00			1.00			1.00		
Middle	0.98	(0.91	- 1.06)	1.14	(0.94	- 1.37)	0.92	(0.79	- 1.07)
High	0.84	(0.78	- 0.91)	0.97	(0.81	- 1.18)	0.81	(0.69	- 0.94)
Sex									
Male	1.00			1.00			1.00		
Female	1.74	(1.62	- 1.87)	1.15	(0.97	- 1.36)	1.56	(1.35	- 1.80)
Smoke									
<10 per day	1.00			1.00			1.00		
10-20	0.85	(0.66	- 1.08)	0.80	(0.51	- 1.25)	0.65	(0.41	- 1.01)
20-40	0.75	(0.52	- 1.07)	0.70	(0.37	- 1.33)	0.87	(0.49	- 1.54)
≥ 40	1.09	(0.91	- 1.30)	0.72	(0.51	- 1.02)	0.80	(0.59	- 1.10)

Table 8. Results of Competing Risk Analysis for Different Types of Dementia (Cont.)

Variables	Types of Dementia								
	Alzheimer's disease			Vascular dementia			Other		
	HR	95% CI		HR	95% CI		HR	95% CI	
		Lower	Upper		Lower	Upper		Lower	Upper
Drink									
<3 times	1.00			1.00			1.00		
4-10 times	1.00	(0.87	- 1.15)	0.98	(0.73	- 1.31)	0.96	(0.73	- 1.25)
11-16 times	1.05	(0.89	- 1.25)	1.15	(0.82	- 1.60)	1.00	(0.72	- 1.38)
>17 times	1.14	(1.00	- 1.29)	1.06	(0.80	- 1.40)	0.81	(0.61	- 1.07)
Physical activity									
0-1 times	1.00			1.00			1.00		
2-3	1.25	(1.16	- 1.36)	0.96	(0.78	- 1.18)	0.95	(0.80	- 1.13)
4-5	1.23	(1.07	- 1.42)	0.62	(0.40	- 0.97)	0.76	(0.54	- 1.07)
6-7	0.99	(0.90	- 1.09)	0.70	(0.54	- 0.89)	0.90	(0.75	- 1.09)
BMI									
Less than 25	1.00			1.00			1.00		
25-30	1.04	(0.97	- 1.11)	0.93	(0.80	- 1.09)	0.96	(0.84	- 1.09)
30 or higher	1.05	(0.92	- 1.21)	0.79	(0.55	- 1.14)	0.73	(0.53	- 0.99)

IV. DISCUSSIONS

The findings of this study are consistent with previous research on the relationship between the age of diabetes onset and subsequent dementia. The results indicate that the later the onset of diabetes, the higher the risk of developing dementia. Similar findings have been reported in previous studies using UK Biobank data, further supporting this association. Also, after adjusting for different factors, the strength of association between diabetes and all-cause dementia increased with decreasing onset age of diabetes(Wang, Li et al. 2023). In the 4 groups that were categorized on diabetes onset age (40-49, 50-59, 60-69, 70+), there was a trend of higher diabetes onset age associated with lower risk of dementia in every group. However, in the 50-59 age group, the difference was not statistically significant, whereas in the 60-69 and 70+ age groups, the results were statistically significant. This result aligns with the general characteristics of the study population, where there is a significant difference in dementia incidence across age groups. Given that dementia prevalence is significantly higher in older adults aged 65 and above, these findings appear to be appropriate.

Apart from the main interest factor, the result of other factors was also consistent with findings from previous studies. For instance, higher income levels showed lower risk of dementia. This finding is consistent with a study using data from the São Paulo Ageing & Health Study, which reported that dementia was more prevalent among participants who were illiterate, had non-skilled occupations and had lower income levels (Scazufca,

Almeida et al. 2010). Regarding gender, the study found Dementia was significantly more prevalent in female than men(Kim, Han et al. 2014). This result is also consistent with previous research findings in South Korea. In this study results, a higher frequency of smoking was associated with a lower risk of dementia, which contradicts previous research findings. However, this discrepancy may be due to an imbalanced group, as the majority of participants in the study were in the group that smoked fewer than 10 cigarettes per day. To address this issue, a competing risk analysis was conducted, which yielded results consistent with previous studies. Specifically, the analysis demonstrated an increased risk of Alzheimer's disease with current smoking, as well as not significantly increased risk of vascular dementia, dementia unspecified and cognitive decline(Peters, Poulter et al. 2008). Furthermore, higher alcohol consumption showed higher risk of dementia which was consistent with previous research in South Korea where decreased risk of dementia was associated with maintaining mild to moderate alcohol consumption(Jeon, Han et al. 2023). Moreover, higher frequency of physical activity showed a lower risk of subsequent dementia. This was consistent with the established studies in Korea where increased physical activity level was associated with a reduced risk of dementia in older adults(Yoon, Yang et al. 2021). Lastly, higher BMI was associated with lower risk of dementia. This was consistent with the previous studies where HRs for all-cause dementia tended to increase as BMI decreased(Nam, Park et al. 2019).

An interesting finding of this study was that the results of time on study analysis using time as the time scale and the result of age attained analysis using age as the time scale

showed opposite results on relation of dementia onset age and risk of subsequent dementia. There were studies with the choice of time scale for proportional hazard models in cohort data and the studies showed that unless the baseline hazard is an exponential function of age, using time as a time scale can yield different estimates of relative hazard than using age as the time scale, even when age is adjusted for (Thiébaut and Bénichou 2004). Since the main factor of interest in this study was diabetes onset age, and dementia are strongly influenced by age, using age as the time scale allowed for a direct assessment of the age-specific incidence function, allowing for the risk of participants to be compared to others of the same attained age, regardless of what age they entered the cohort. This can be considered a key strength of the study. Another strength of this study was that as clinical mechanisms differ in different types of dementia, competing risk analysis was used to determine whether the age of diabetes onset has varying effects on different types of dementia and to avoid overestimating dementia risk by considering death as a competing factor.

This study focused on examining the impact of diabetes onset age on dementia risk, which may have limited its ability to fully explore the underlying clinical mechanisms of dementia development. Furthermore, the analysis was based on National Health Insurance Service-Health Screening Cohort, which restricted access to comprehensive health information about participants. Also, the information on medications used for diabetes and dementia was obtained from health insurance claims. Therefore, with cases where prescriptions were not covered by insurance was not included. This means if a patient buys

this medication from a pharmacy without a doctor's prescription, the purchase may not be documented in the data. Additionally, the potential influence of unidentified confounding factors in the relationship between diabetes onset age and dementia should be carefully considered.

Despite these limitations, this study successfully identified the association between diabetes onset age and dementia risk using a dataset that is highly representative of the Korean population by analyzing nationwide health data. Given the scarcity of research on this topic in Korea, investigating this relationship seems particularly significant. As the age of diabetes onset continues to decrease and both diabetes and dementia prevalence rise among Koreans, this study provides valuable insights into an increasingly important public health issue. From the outcome of the study, early cognitive monitoring is suggested for younger-onset diabetes patients. Furthermore, the findings of this study may serve as evidence supporting the importance of midlife and early diabetes prevention.

V. CONCLUSION

This study aimed to investigate the impact of diabetes onset age on dementia using a dataset representative of the Korean population. Also, it showed impact of income level, sex, smoking frequency, frequency of alcohol consumption, frequency of physical activity and BMI level on the risk of dementia. Additionally, it sought to examine whether the effect of diabetes onset age on dementia varies by dementia subtype. In result, significant relationship between age of diabetes onset and subsequent dementia risk was found, emphasizing that later onset diabetes is associated with a lower risk of dementia compared to earlier-onset diabetes. These findings contribute to the growing of evidence on the interaction between metabolic disorders and neurodegeneration.

The outcome of the study suggests that duration and timing of diabetes onset are crucial determinants of dementia risk, highlighting the importance of considering age at diagnosis for future studies. Also, by comparing the two models using each time and age as a time scale, this study suggests that using age as the time scale can be a recommendation for research on diseases that are strongly age-related.

Given the decreasing age of diabetes onset, the rising prevalence of both diabetes and dementia among Koreans, there seems to be a lack of studies on this topic. Future research should consider several key areas. First, studies exploring the mechanisms linking diabetes and dementia are needed to allow for further adjustments for more precise results.

Additionally, since guidelines for diabetes prevention are already well-established, the result of this study can provide evidence for lifestyle interventions aimed for dementia prevention. While numerous studies have examined the impact of diabetes itself on dementia and other diseases, research on how the timing of diabetes onset affects dementia and other conditions remain limited. It is hoped that this study serves as a foundation for future investigations in this field.

VI. REFERENCES

Abd ElHafeez, S., et al. (2021). "Methods to analyze time-to-event data: the Cox regression analysis." Oxidative Medicine and Cellular Longevity 2021(1): 1302811.

Amidei, C. B., et al. (2021). "Association between age at diabetes onset and subsequent risk of dementia." JAMA 325(16): 1640-1649.

Chatterjee, S., et al. (2017). "Type 2 diabetes." The Lancet 389(10085): 2239-2251.

Choi, D., et al. (2018). "Effect of smoking cessation on the risk of dementia: a longitudinal study." Annals of Clinical and Translational Neurology 5(10): 1192-1199.

Ha, K. H. and D. J. Kim (2015). "Trends in the diabetes epidemic in Korea." Endocrinology and Metabolism 30(2): 142.

Jeon, K. H., et al. (2023). "Changes in alcohol consumption and risk of dementia in a nationwide cohort in South Korea." JAMA Network Open 6(2): e2254771-e2254771.

Jeong, S.-M., et al. (2023). "Association of changes in smoking intensity with risk of dementia in Korea." JAMA Network Open 6(1): e2251506-e2251506.

Kim, D. J. (2011). "The epidemiology of diabetes in Korea." Diabetes & Metabolism Journal 35(4): 303.

Kim, J., et al. (2025). "Diabetes status, duration, and risk of dementia among ischemic stroke patients." Alzheimer's Research & Therapy 17(1): 58.

Kim, M. K. (2024). "Younger-Onset Diabetes: Is the Age of Onset More Important than the Duration of Diabetes?" Endocrinology and Metabolism 39(1): 90-91.

Kim, Y. J., et al. (2014). "Prevalence and trends of dementia in Korea: a systematic review and meta-analysis." Journal of Korean Medical Science 29(7): 903.

Lee, S. K., et al. "Association of Physical Activity with Dementia Risk in Cancer Survivors: A Korean Nationwide Cohort Study." Cancer Research and Treatment.

Lunn, M. and D. McNeil (1995). "Applying Cox regression to competing risks." Biometrics: 524-532.

Nam, G. E., et al. (2019). "BMI, weight change, and dementia risk in patients with new-onset type 2 diabetes: a nationwide cohort study." Diabetes Care 42(7): 1217-1224.

Park, J.-H., et al. (2013). "Burden of disease due to dementia in the elderly population of Korea: present and future." BMC Public Health 13: 1-9.

Pasquier, F., et al. (2006). "Diabetes mellitus and dementia." Diabetes & Metabolism 32(5): 403-414.

Pencina, M. J., et al. (2007). "Choice of time scale and its effect on significance of predictors in longitudinal studies." Statistics in Medicine 26(6): 1343-1359.

Peters, R., et al. (2008). "Smoking, dementia and cognitive decline in the elderly, a systematic review." BMC Geriatrics 8: 1-7.

Qizilbash, N., et al. (2015). "BMI and risk of dementia in two million people over two decades: a retrospective cohort study." The Lancet Diabetes & Endocrinology 3(6): 431-436.

Savelieff, M. G., et al. (2022). "Diabetes and dementia: Clinical perspective, innovation, knowledge gaps." Journal of Diabetes and Its Complications 36(11): 108333.

Scazufca, M., et al. (2010). "The role of literacy, occupation and income in dementia prevention: the São Paulo Ageing & Health Study (SPAH)." International Psychogeriatrics 22(8): 1209-1215.

Shin, J., et al. (2025). "Alcohol Consumption and the Risk of Dementia in Patients with Depression: A Nationwide Cohort Study." International Journal of Mental Health and Addiction: 1-19.

Thiébaut, A. C. and J. Bénichou (2004). "Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study." Statistics in Medicine 23(24): 3803-3820.

Wang, Y., et al. (2023). "Onset age of diabetes and incident dementia: A prospective cohort study." Journal of Affective Disorders 329: 493-499.

Yoon, M., et al. (2021). "Association of physical activity level with risk of dementia in a nationwide cohort in Korea." JAMA Network Open 4(12): e2138526-e2138526.

Yu, J. H., et al. (2020). "Incidence and risk factors for dementia in type 2 diabetes mellitus: a nationwide population-based study in Korea." Diabetes Metab J 44(1): 113-124.

Zhang, Y., et al. (2023). "Identifying modifiable factors and their joint effect on dementia risk in the UK Biobank." Nature Human Behaviour 7(7): 1185-1195.

ABSTRACT IN KOREAN (국문 초록)

당뇨병 발병 시기와 이후 치매 위험 간의 연관성
: 연령 도달 분석을 이용하여

연세대학교 보건대학원 보건정보통계학 전공

이지원

연구배경: 콕스 비례위험모형은 생존분석에서 가장 일반적으로 사용되는 통계 기법 중 하나이다. 이런 전통적인 콕스 비례위험모형은 절대적 시간을 시간 척도로 사용하는데 이는 연령 특이적 위험을 충분히 반영하지 못한다는 제한점이 있다. 이를 위한 해결 방안으로 연령을 시간 척도로 사용하는 연령 도달 분석을 제안하는 일부 연구들에 따라 당뇨병 발병 연령과 치매 간의 연관성을 연구하였다. 당뇨병과 치매 간의 연관성은 잘 알려져 있으나, 당뇨병 발병 연령이 이후 치매 위험에 미치는 영향에 대한 연구는 부족하다. 특히 젊은 연령층에서 당뇨병 유병률이 증가하고 있는 점을 고려할 때, 이에 대한 추가 연구가 필요하다.

연구방법: 본 연구는 2002-2003 년 일반건강검진을 받은 514,866 명(약 515 만 명의 10%)을 단순 무작위 추출하여 구축된 국민건강보험 건강검진 코호트 데이터(ver. 2.1.)를 활용하였다. 연구 대상자는 당뇨병 진단을 받은 40 세

이상 성인 중 국민건강보험 자격을 유지한 자로 선정하였다. 본 데이터는 2002 년부터 2019 년까지 총 18 년간의 추적 관찰 자료를 포함하고 있다. 당뇨병 신규 진단자를 명확히 하기 위해 2002-2003 년 당뇨병 진단자를 제외하는 2 년의 세척기간(wash-out period)을 적용하였다. 최종적으로 77,092 명이 연구 코호트에 포함되었다. 당뇨병 발병 연령과 치매 발생 위험 간의 연관성을 확인하기 위해 전통적 Cox 비례위험 회귀분석과 연령 도달 분석(Age-attained analysis)을 수행하였다. 또한, 소득 수준, 성별, 흡연 빈도, 음주 습관, 신체활동 빈도, 체질량지수(BMI) 등 여러 공변량이 치매 발생 위험에 미치는 영향도 함께 평가하였다. 아울러, 치매 유형별로 당뇨병 발병 연령의 영향을 추가로 분석하고, 경쟁위험(Competing risk)으로 인한 치매 발생 과대 추정을 보정하기 위해 경쟁위험 분석을 실시하였다.

연구결과: 당뇨병 발병 연령이 늦을수록 치매 발생 위험이 증가하는 것으로 나타났다. 반면, 소득 수준이 높을수록 치매 위험은 유의하게 낮았다. 성별에 따라서는 여성에서 남성보다 치매 발생률이 유의하게 높은 것으로 확인되었다. 흡연 빈도가 높을수록 치매 위험은 감소하였으며, 음주 빈도가 높을수록 치매 위험은 증가하였다. 또한, 신체활동 빈도가 많을수록 치매 발생 위험은 낮아졌으며, 체질량지수(BMI)가 높을수록 치매 위험이 감소하는 경향을 보였다. 본 연구의 중요한 결과 중 하나는 Cox 회귀분석에서 시간 척도(time scale)의 선택에 따라 당뇨병 발병 연령과 치매 위험 간의 연관성 결과가 상반되게 나타났다는 점이다. 당뇨병 발병과 치매 모두 연령의 영향을 크게 받는 특성을 고려할 때, 도달 연령(attained age)을 시간 척도로 사용하는 것이 개인 간 연령별 위험을 보다 적절하게 비교할 수 있는 분석틀을 제공하는 것으로 확인되었다.

결론: 본 연구는 전국 대표성을 갖춘 한국 코호트 자료를 활용하여 당뇨병 발병 연령과 치매 발생 위험 간의 연관성을 분석하였다. 또한, 여러 공변량의 영향과 치매 유형에 따른 차이도 함께 살펴보았다. 분석 결과, 당뇨병 발병 연령이 어릴수록 치매 발생 위험이 높아지는 경향을 보여, 당뇨병의 발병 시점과 이환 기간이 인지기능 저하에 중요한 영향을 미침을 확인하였다. 아울러, 시간 척도(time scale)에 따른 분석 결과의 차이를 비교한 결과, 도달 연령(attained age)을 시간 척도로 사용하는 것이 치매와 같이 연령에 의존적인 질환의 위험도를 평가할 때 보다 적절한 방법임을 제시하였다. 이러한 결과는 대사성 질환과 신경퇴행성 질환 간의 연관성에 대한 근거를 추가로 제공하며, 향후 관련 연구의 분석 방법론 선택에 있어 유용한 시사점을 제시한다.

핵심어: 당뇨, 치매, 시간척도, 콕스 비례위험모형, 연령 도달 분석,
경쟁위험분석