AGE-PERIOD-COHORT EFFECTS ON THE DEVELOPMENT OF COGNITIVE IMPAIRMENT AMONG THE ELDERLY

by

Xinmei Zhu

BMed, Binzhou Medical College, China, 1995

PhD, Fudan University Medical Center, China, 2004

Submitted to the Graduate Faculty of

the Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Xinmei Zhu

It was defended on

June 9, 2017

and approved by

Committee Members

Hiroko Dodge, PhD, Professor
Department of Neurology
Oregon Health & Science University
Portland, Oregon, USA

Brenda Kurland, PhD, Associate Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Yan Lin, PhD, Research Associate Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Thesis Advisor

Chung-Chou H. Chang, PhD, Professor
Department of Biostatistics, Medicine
Graduate School of Public Health, School of Medicine
University of Pittsburgh

Chung-Chou H. Chang, PhD

AGE-PERIOD-COHORT EFFECTS ON THE DEVELOPMENT OF COGNITIVE IMPAIRMENT AMONG THE ELDERLY

Xinmei Zhu, MS

University of Pittsburgh, 2017

ABSTRACT

The objectives of the study were to investigate the effects of age, calendar period, and birth cohort on the development of cognitive impairment among the elderly, and to identify factors possibly moderating these effects. Harmonized data were drawn from two community-based cohort studies. A total of 3,021 participants, born after 1895, age 65 years or older with normal cognitive capacities were recruited during 1987-2008 and followed for more than 10 years. Cognitive capability was evaluated periodically using the Clinical Dementia Rating (CDR) scale. Incident mild cognitive impairment (MCI) was defined as the CDR value reaching 0.5. Age-period-cohort (APC) modelling approach was used to evaluate three time-varying effects on the development of MCI. Confounding and moderating effects of gender, education, and ApoE4 allele were also examined. Our analysis results showed that age was the most significant time-dependent factor affecting the MCI incident rates. Within the same calendar period and birth cohort, the MCI rate in the older elderly was significantly higher compared with the younger elderly population. A significant period effect was observed in which the MCI incidence rates were decreasing from the period of 1990-1994 through 2015 after controlling for age and birth cohort. No significant cohort effect was found. Gender showed no significant confounding or moderating effects. The age effects on MCI

incidence rate was not moderated or confounded by education, while the period effects were significantly confounded by education. The cohort effect was significantly moderated by education. The cohort effects on MCI incidence rates for individuals who received HS education or higher education were different depending on the levels of education. ApoE4 allele did not show a significant moderating effect.

Public Health Significance

The APC model shows advantages over the traditional modelling approaches as it dissects the independent effects of age, period, and cohort. For public health, chronic disease prevalence often reflects a combination of processes that vary by these three factors. Better understanding the impacts of these time-dependent factors on disease rates help to guide hypotheses about etiologic mechanisms, and more importantly, guides researchers in conducting and presenting surveillance with the best practices.

TABLE OF CONTENTS

PREFACE XII

1.0 INTRODUCTION 1

2.0 MATERIALS AND METHODS 4

2.1 STUDY POPULATION 4

2.2 STATISTICAL ANALYSIS 6

2.2.1 APC model 6

2.2.2 Data analysis 7

2.2.3 Statistical software 8

2.3 ETHICAL CONSIDERATION 9

3.0 RESULTS 10

3.1 DESCRIPTIVE ANALYSIS RESULTS 10

3.2 ESTIMATING MCI INCIDENCE RATES USING THE APC MODELLING APPROACH 17

3.2.1 Single-factor APC models 17

3.2.2 Two-factor APC models 19

3.2.3 Constraint-based full APC model 20

3.3 MODERATOR AND CONFOUNDER EFFECTS OF GENDER AND EDUCATION 24

3.4 MODERATOR EFFECT OF APOE4 ALLELE 31
3.5 MODERATOR EFFECTS OF STUDY 34
3.6 EVALUATION OF MODEL FITNESS 35
4.0 DISCUSSION 37
APPENDIX A: R CODES FOR ANALYSIS 41
APPENDIX B: SUPPLEMENTARY TABLES 50
BIBLIOGRAPHY 58

LIST OF TABLES

Table 1. Data summary by study cohort 12

Table 2. Age-specific MCI incident cases, follow-up person-years and rates by period 14

Table 3. Age-specific MCI incident cases, follow-up person-years and rates by cohort 15

Table 4. Poisson log-linear model with age effects only 17

Table 5. Poisson log-linear model with calendar period effects only 18

Table 6. Poisson log-linear model with birth cohort effects only 18

Table 7. Poisson log-linear model with age and period effects 19

Table 8. Poisson log-linear model with age and cohort effects 20

Table 9. Poisson log-linear model with age, period, and cohort effects* 22

Table 10. Poisson log-linear model with age, period, and cohort effects* 23

Table 11. Poisson log-linear model with age and education effects 27

Table 12. Estimate of Age effects before and after adjusting for education. 27

Table 13. Poisson log-linear model with period and education effects 28

Table 14. Estimates of period effects before and after adjusting for education 28

Table 15. Estimate of cohort effects before and after adjusting for education 30

Table 16. Poisson log-linear model with age and APOE4 effects 32

Table 17. Poisson log-linear model with calendar period and APOE4 effects 32

Table 18. Poisson log-linear model with birth cohort and APOE4 effects 32

Table 19. Poisson log-linear model with age and study effects 34
Table 20. AIC of Different models 36
Table 21. Supplementary Table 1. Data Information 50
Table 22. Supplementary Table 2. Information of MoVIES and MYHAT Study Population 51
Table 23. Supplementary Table 3. Poisson log-linear model with age effects by gender 51
Table 24. Supplementary Table 3.1 Poisson log-linear model with age and gender effects 52
Table 25. Supplementary Table 4. Poisson log-linear model with period effects by gender 52
Table 26. Supplementary Table 4.1. Poisson log-linear model with period and gender effects 52
Table 27. Supplementary Table 5. Poisson log-linear model with cohort effects by gender* 53
Table 28. Supplementary Table 5.1. Poisson log-linear model with cohort and gender effects 53
Table 29. Supplementary Table 6. Poisson log-linear model with age effects by education 54
Table 30. Supplementary Table 7. Poisson log-linear model with period effects by education 54
Table 31. Supplementary Table 8. Poisson log-linear model with cohort effects by education 55

Table 32. Supplementary Table 9. Poisson log-linear model with age effects by ApoE4 56
Table 33. Supplementary Table 10. Poisson log-linear model with period effects by ApoE4 56
Table 34. Supplementary Table 11. Poisson log-linear model with cohort effects by ApoE4 57
Table 35. Supplementary Table 12. Poisson log-linear model with age effects by study 57

LIST OF FIGURES

Figure 1. Consort diagram for data collection information 5

Figure 2. Lexis-diagram 11

Figure 3. Age-specific MCI incidence rates by calendar period and birth cohorts 16

Figure 4. Period-specific and Cohort-specific MCI incidence rate by age 16

Figure 5. Age-specific (A), Period-specific (B) and Birth-cohort-specific (C) MCI incidence rate by gender 25

Figure 6. Age-specific (A), Period-specific (B) and Birth-cohort-specific (C) MCI incidence rate by education 29

Figure 7. Age-specific (A), Period-specific (B) and Birth-cohort-specific (C) MCI incidence rate by ApoE4 allele 33

Figure 8. Age-specific MCI incidence rates by study 35

PREFACE

First, I sincerely thank my thesis advisor, Dr. Chung-Chou H. Chang for all her instructions and time. Dr. Chang has been spending so much time guiding me on the analysis plan, helping me resolve the issues that I had during the analysis, and revising my thesis. Finishing this thesis would not be possible without her insights and guidance. My sincere gratitude would also go to Dr. Hiroko Dodge, Dr. Yan Lin and Dr. Brenda Kurland. They, as my committee members shared their valuable insights and inspiring comments on this thesis.

I am so grateful that Ms. Joanne Pegher had patiently spent so much time on editing my thesis and Ms. Renee Valenti has been always so helpful and supportive throughout my study in the department of Biostatistics of GSPH.

I would like to thank my family, my husband Xingbin Wang and my two little girls for their support. I appreciate that my two little girls have been giving me that much quiet time when I was working on my thesis even they are only 8 and 5 years old only.

Acknowledgement is also given to Drs. Mary Ganguli and Dr. Beth Snitz (Research PIs of projects MoVIES and MYHAT) for the use of data. This work was supported by National Institute on Aging (NIA) and research GRANTS #AG07562 and #AG023651.

NOMENCLATURE

CI, Cognitive Impairment

AD, Alzheimer's Disease

MCI, Mild Cognitive Impairment

CDR, Clinical Dementia Rating

APC, Age-Period-Cohort

MoVIES, Monongahela Valley Independent Elders Survey

MYHAT, Mon-Yough Health Aging Team

ApoE4 allele, Apolipoprotein E ε4 allele

Edugp, Education Group

RR, Risk Ratio

AIC, Akaike Information Criterion

1.0 INTRODUCTION

Dementia or cognitive impairment is an age-related disorder associated with the senior population. The prevalence of dementia increases exponentially with age and it has been a leading cause of disability and dependence among the elderly [1]. It had been predicted that a worldwide figure of 81.1 million people will be affected by dementia by 2040, if there is no curative treatment developed by then [2]. Alzheimer's disease (AD) is the most common form of dementia. Studies showed that mild cognitive impairment (MCI) without dementia is even more prevalent in the United States than dementia [3]. It was observed that people with MCI are likely to progress to dementia at a rate of approximately 12% per year, compared with 1 to 2% for cognitively normal people at the same age [4-6]. Considering declined quality of life and economic burdens among patients and their families, cognitive impairment and AD are becoming a major public health problem [7].

It has been well known that cognitive decline could be a normal process of aging, and the age effects on the development of cognitive impairment (CI) are well documented in the scientific literatures. Many published studies employed a cross-sectional design to compare subjects from different age groups [8]. These studies were subject to confounding due to cohort differences, as people born in different cohorts may have a very different life experience in terms of culture, lifestyle, education, and

social-economic environment [8]. As a result, subjects from one age cohort could perform poorly on the cognitive test as compared to subjects from a different age cohort [9]. Then it is not surprising that effects of aging in cross-sectional studies could be confounded by cohort differences and potentially could be overestimated [10, 11]. Similarly, at certain periods, people may experience some widespread environmental changes and population-wide exposures that were associated with cognitive decline, then the effects of calendar period, could be another confounding factor of aging effects within the existing studies. Given such a situation, we believe all those time related factors, age, calendar period, and birth cohort, could affect the development of incidence rates of CI in different ways. Our study aimed to identify the critical temporal factors that may influence the development of MCI. For this purpose, two promising prospective community-based cohort studies, Monongahela Valley Independent Elders Survey (MoVIES) and Mon-Yough Health Aging Team (MYHAT) were designed and performed. There were 1,608 individuals aged 65 years or older between 1987 and 1989 recruited for MoVIES project to investigate the incidence, risk factors, and outcome in late-life dementia. MYHAT is an on-going project seeking to describe the distribution of cognitive Impairment and its associated features, its outcomes over time, and the predictors of these outcomes. Approximately 1,413 were recruited and assessed periodically to determine their cognitive capability status. Clinical dementia rating (CDR) is a 5-point scale used to characterize the participants' cognitive performances applicable to Alzheimer disease and the related cognitive impairment. In here we harmonized data from MoVIES and MYHAT studies according to their CDR scores. Incident MCI was defined as the CDR value reaching 0.5.

To characterize the temporal effects of age, calendar period, and birth cohort on the development of CI, we used an age-period-cohort (APC) model to dissect the roles of these three time-related factors in the development of MCI. Each component of the APC model provided different insights into the trends of the disease over time. The APC analysis gave us an overview of the magnitude of rates and the variation of rates by these three time-varying components [12].

Cognitive impairment is a complicated age-related disorder resulting from interaction of multiple risk factors including genetic, vascular, and other unknown risk factors. Apolipoprotein E (ApoE) ε4 allele is a known genetic risk factor for AD [13, 14]. Moreover, it was reported that progress of MCI to AD is accelerated by ApoE4 [15, 16]. Within this study, we also investigated whether ApoE4 allele expression would moderate the temporal effects on the MCI incidence rates. Other potential confounders and possible moderators such as gender and education were also examined.

2.0 MATERIALS AND METHODS

2.1 STUDY POPULATION

Data were drawn from MoVIES and MYHAT which were two independent prospective cohort studies designed to identify factors that are associated with cognitive impairment among the community-dwelling older adults (age 65 years or older). There were 1,608 participants recruited between 1987 and 1989 for MoVIES and followed over 15 years, while 1,413 participants were recruited between 2006 and 2008 for MYHAT and followed for about 10 years (which was summarized in **Figure 1** and **supplementary Table 1**). The outcome variable of interest was the incidence rates of MCI where MCI was defined by reaching 0.5 or greater on the CDR scale. The main exposure variables of interest included age, calendar year, and birth year at the time of study entry. Gender and education information, which are potential confounders and may possibly moderate the effects of temporal factors on the development of MCI were collected. Known genetic risk factor for dementia, ApoE4 allele genotyping of participants was tested during the study period. Periodical follow-up assessments were performed to track change in cognitive capabilities. For the MoVIES project, the home interviews were repeated approximately every two years for a total of 15 years, while for the MYHAT

project an annual follow-up assessment was made (which was summarized in **Supplementary Table 2**).

As shown in **Figure 1**, a total of 3,021 participants (1,608 from MoVIES and 1,413 from MYHAT) with normal cognitive capacities (CDR = 0) were included for this study. After excluding subjects whose first interview date was the same as the last follow-up date, our final analytic data set includes 2,772 subjects.

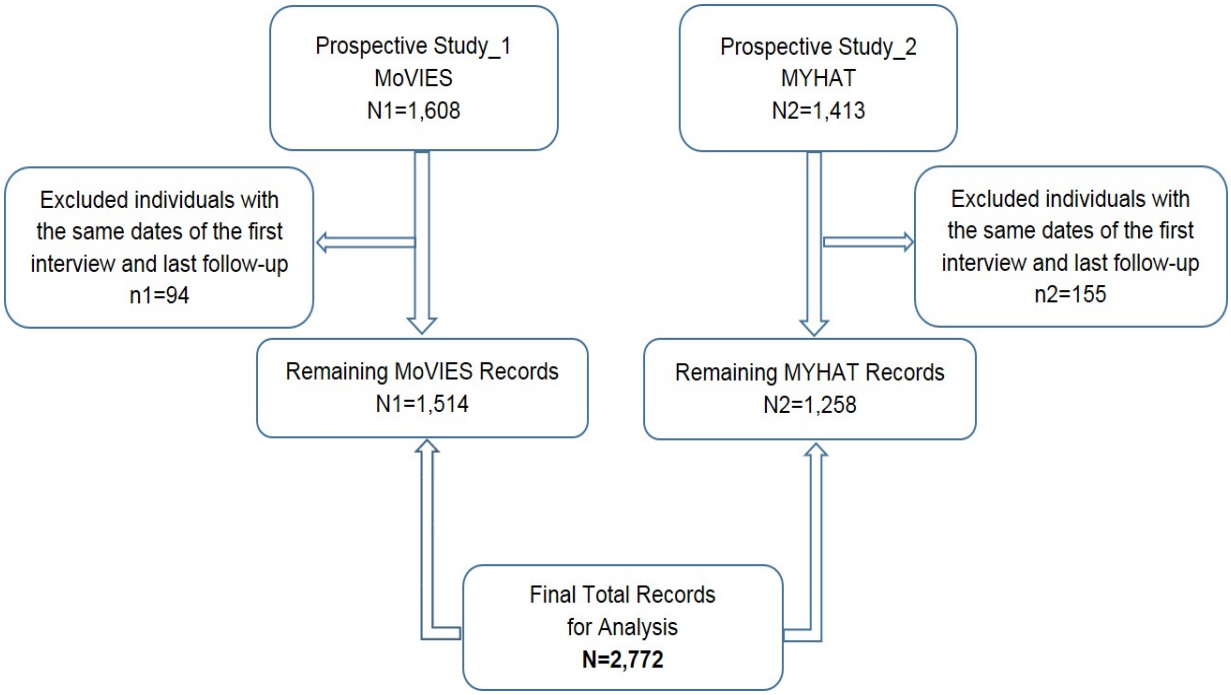

Figure 1. Consort diagram for data collection information

2.2 STATISTICAL ANALYSIS

2.2.1 APC model

We examined the temporal effects of age, calendar period, and birth cohort on the risk of incident MCI using the APC model approach [17-19]. APC model is a popular tool for dissecting the independent effects of all three time-dependent variables (age, period, and cohort) and has received considerable attention in statistical field [20-22]. The model is based on a Poisson log-linear model for the expected rates with additive effects of age, period, and cohorts. The model has the form

$$\mathrm{Ln}(M_{ij}) = \mathrm{Ln}(\theta_{ij}/N_{ij}) = \mu + \alpha_i + \beta_j + \gamma_k, \quad (1)$$

where M_{ij} is the MCI incidence rate for individuals in age group *i* at calendar time *j*; θ_{ij} is the number of incident cases among individuals in age group *i* at calendar time *j*; N_{ij} is the number of follow-up person-years for individuals in age group *i* at calendar time *j*; μ is the overall expected mean log incidence rate; α_i, β_j, and γ_k are the differential rates from the overall mean log incidence rate for individuals in age group *i,* calendar time *j*, and birth cohort *k*, respectively**.**

When we apply an APC model, it is important to recognize the non-identifiability issue resulted from collinearity among these three variables (Cohort = Period − Age), which makes simultaneous mathematical modeling of the linear functions of three effects impossible without additional restrictions [11]. To resolve this identifiability issue, existing literatures have reported a variety of methodological approaches [11, 23-25]. One method suggested is to use a reduced set of covariates (reduced APC model) in the model by including age and period without including birth cohort. In this two-factor

model, cohort effects can be considered as a special form of interaction effects between age and period [26]. Another popular approach to resolve the identification problem was the constraint-based regression analysis. In this strategy, at least one category of age, period, and cohort is constrained in some manner by assuming that some categories of age groups, cohorts, or time periods have identical effects on the dependent variable [22], thus it became possible to estimate independent effect of age period and cohort [19]. It is worth noting that the results from the constraint-based on full APC model might be sensitive to the constraints that investigators chosen and sometimes the validity of the chosen constraints cannot be confirmed [27-29]. To select the appropriate and reasonable constraints, investigators often use graphical data that describes the disease trends by period, age, and cohort to establish the choice of constraints for this model [30].

2.2.2 Data analysis

We analyzed the data in the following steps. First, we classified age into 7 five-year groups (65-69, 70-74, 75-79, 80-84, 85-89, 90-94, and 95-99 years old), calendar time into 7 five-year periods (1985-1989, 1990-1994, 1995-1999, 2000-2004, 2005-2009, 2010-2014, and 2015 after), and birth year into 10 five-year birth cohorts (1895-1899,1900-1904, 1905-1909, 1910-1914, 1915-1919, 1920-1924, 1925-1929, 1930-1934, 1935-1939, and 1940 or after). Second, age-specific MCI incidence rates per 1,000 person-years were calculated for each of the 7 five-year age groups, each of the 7 five-year calendar periods, and each of the 10 five-year birth cohorts. We employed the lexis diagram to depict the MCI incidence rates among different birth cohorts. The

diagram is a two-dimensional figure where horizontal-axis and the vertical axis represent the calendar time and age, respectively. The straight line in the diagram is called an individual's life line, which begins at the time/age of study entry and continues diagonally upwards and ends at the time/age of the individual's last follow-up (either events occurred or last time/age followed). Third, we fit the data into a reduced APC model (one or two temporal factors instead of all 3 factors together) which provided us with unique (but not independent) information of the temporal effects on MCI incidence rates [26]. And in here we would expect to see a possible trend that certain groups would not show significant effects on the MCI incidence rates. We further examine if there are two period groups (or birth cohort) having similar effects, which can be chosen as the constraints in the full APC Model. We then fit the data into the full APC model with constraints to simultaneously estimate the independent effects of all 3 temporal factors. We confirmed the validity of our constraint choice by selecting different birth cohorts and period groups as constraints and regenerated the analysis results. We finally used the Akaike information criterion (AIC) to choose the best APC model among restricted one-factor, restricted two-factor, and full model with different constrains.

2.2.3 Statistical software

All analyses were performed using R version 3.3.3. We included the R codes of our data analyses to the Appendix.

2.3 ETHICAL CONSIDERATION

The community-wide recruitment, and assessment procedures were approved by the Institutional Review Board at University of Pittsburgh.

3.0 RESULTS

3.1 DESCRIPTIVE ANALYSIS RESULTS

1. Descriptive Analysis Results

The harmonized MoVIES and MYHAT data contain 3,021 subjects and 13 variables, as shown in **Figure 1** and **Supplementary Table 1**. Among these 3,021 observations, 1,608 records were from MoVIES and the remaining 1,413 records were from MYHAT. We excluded 94 records from MoVIES and 155 records from MYHAT for the final analysis, because the first interview date and the last follow-up date were the same for each of the excluded subjects. Therefore, the final analytic sample contains 2,772 subjects. The lexis diagram in **Figure 2** depicts the ages and the calendar dates of entry and exit for all 2,772 subjects. Each line on the diagram represents the follow-up time span. If the end of the line has a solid red dot, the corresponding participant experienced MCI incidence at that time point. The lexis diagram shows that there was a gap between the two studies (Year 2002 through 2006) where there was no participant involved in the studies.

There were 655 MCI cases (total incidence rate 35.7 per 1,000 person-years) for the harmonized data, 305 MCI cases from the MoVIES project (incidence rate 24.9 per 1,000 person-years), and 350 MCI cases from the MYHAT project (incidence rate 57.4

per 1,000 person-years), which was summarized in **Table 1**. The average (SD) age in years upon entering was 74.9 (6.6) for the entire harmonized data, 72.8 (5.9) for the MoVIES study, and 77.3 (7.3) for the MYHAT study. The average (SD) age in years at the last follow-up was 81.4 (6.2) for the harmonized data, 80.9 (5.4) for the MoVIES study, and 82.1 (6.9) for the MYHAT study. The average (SD) follow-up time in years was 6.6 (4.1) for the harmonized data, 8.1 (4.4) for the MoVIES study, and 4.8 (2.9) for the MYHAT study. Note that the MYHAT study is on-going for a total of 9-10 years of data collection.

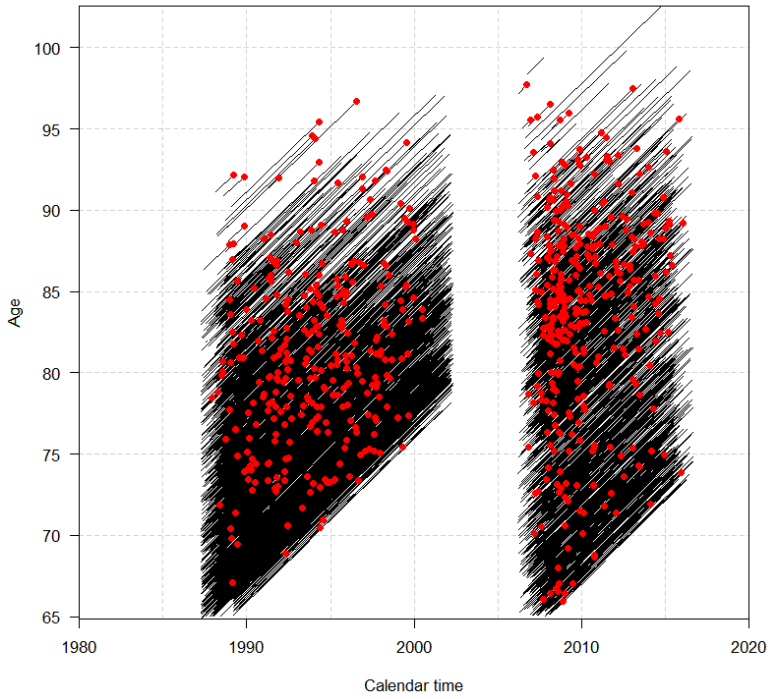

Figure 2. Lexis-diagram

There was a higher proportion of participants with less than high school (HS) education in MoVIES compared with that in MYHAT (42.4% vs. 11.1%). Moreover, participants in MoVIES had a higher average depression score compared with that of MYHAT (mean±SD mCESD score: 1.39±2.56 vs. 0.71±1.74). There were total 2,053

participants who had their ApoE4 allele expression checked. The percentage of ApoE4 allele being positive in MoVIES and in MYHAT were very close (21.1% vs. 20.8%) and 20.9% of all participants expressed ApoE4 allele.

Table 1. Data summary by study cohort

Variable	MoVIES (N=1,514)	MYHAT (N=1,258)	Total (N=2,772)
Incident cases of MCI[1]	305	350	655
Incident rate of MCI Per 1,000 person-years	24.9	57.4	35.7
Age at baseline, in years			
N	1,514	1,258	2,722
Mean (SD)	72.8 (5.9)	77.3 (7.3)	74.9 (6.6)
Age at last follow-up, in years			
N	1,514	1,258	2,722
Mean (SD)	80.9 (5.4)	82.1 (6.9)	81.4 (6.2)
Follow time, years			
N	1,514	1,258	2,772
Mean (SD)	8.1 (4.4)	4.8 (2.9)	6.6 (4.1)
Female	902 (59.6%)	808 (64.2%)	1,710 (61.7%)
APOE*4 allele, n (%)			
N	891	1,162	2,053
Negative	703 (78.9%)	920 (79.2%)	1,623 (79.1%)
Positive	188 (21.1%)	242 (20.8%)	430 (20.9%)
Education n (%)			
< HS education	642 (42.4%)	140 (11.1%)	782 (28.2%)
HS education	499 (33.0%)	575 (45.7%)	1,074 (38.7%)
> HS education	373 (24.6%)	543 (43.2%)	916 (33.0%)
mCESD score			
N	1,272	1,256	2,528
Mean (SD)	1.39 (2.56)	0.71 (1.74)	1.05 (2.22)

[1]Incident MCI = CDR = 0.5

*Abbreviation: HS = high school

[2]mCESD represents a modified version of the original Center for Epidemiologic Studies Depression Scale (mCESD).

The tabulated data in **Table 2** and **Table 3** show the age-specific MCI incidence rates for each calendar period and each birth cohort respectively. As we observed earlier from the lexis-diagram of **Figure 2**, there was a gap where there was no participant between the MoVIES study and the MYHAT study during the calendar period of 2002-2006. There were few three cases in the calendar period of 2000-2004. Considering the low incidence rate of MCI in period of 2000-2004 an artefact effect, we omitted this period in our following data analysis. We performed sensitivity analyses for data including this period and that excluding this period and the results were similar. In this thesis, we presented the analysis results that from data excluding this period. The charts in **Figures 3 and 4** and the summarized age-specific MCI incidence rates in **Tables 2** and **3** showed a trend within the same calendar period or the same birth cohort that the MCI incidence rates increase with age. For people within the same age group, the MCI incidence rates were higher in the calendar period 2005-2009 than that in other calendar periods (**Figures 3 and 4, left panel**), whereas no significant change in the MCI incidence rate was found among different birth cohorts (**Figures 3 and 4, right panel**).

Table 2. Age-specific MCI incident cases, follow-up person-years and rates by period

	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014	2015+
				Incident Cases			
65-69	1	NA	NA	NA	9	NA	NA
70-74	6	12	NA	NA	12	4	NA
75-79	4	41	20	NA	23	17	1
80-84	13	55	43	0	60	22	0
85-89	6	31	33	2	64	43	1
90-94	2	9	19	1	34	33	7
95-99	NA	4	2	0	7	10	2
100+	NA	NA	1	0	1	0	0
				Person-Years			
65-69	538.84	NA	NA	NA	310.88	NA	NA
70-74	817.36	1690.39	NA	NA	593.15	471.93	NA
75-79	498.39	2153.88	1262.09	NA	517.54	820.78	50.85
80-84	258.48	1240.42	1403.28	327.36	618.38	658.88	63.41
85-89	83.35	503.03	598.08	317.39	489.56	670.26	52.52
90-94	9.05	150.42	186.55	120.67	167.64	423.46	43.17
95-99	NA	19.21	32.01	25.20	18.95	89.73	22.87
100+	NA	NA	1.55	3.04	1.46	8.99	3.29
				Rate			
65-69	1.86	NA	NA	NA	28.95	NA	NA
70-74	7.34	7.1	NA	NA	20.23	8.48	NA
75-79	8.03	19.04	15.85	NA	44.44	20.71	19.67
80-84	50.29	44.34	30.64	0	97.03	33.39	0
85-89	71.98	61.63	55.18	6.3	130.73	64.15	19.04
90-94	220.96	59.83	101.85	8.29	202.81	77.93	162.14
95-99	NA	208.23	62.49	0	369.37	111.45	87.44
100+	NA	NA	646.75	0	686.56	0	0

Table 3. Age-specific MCI incident cases, follow-up person-years and rates by cohort

	1895-1899	1900-1904	1905-1909	1910-1914	1915-1919	1920-1924	1925-1929	1930-1934	1935-1939	1940+
65-69	NA	NA	NA	NA	NA	1	NA	NA	NA	9
70-74	NA	NA	NA	NA	6	12	NA	NA	12	4
75-79	NA	NA	NA	4	41	20	NA	23	17	1
80-84	NA	NA	13	55	43	NA	60	22	0	NA
85-89	NA	6	31	33	NA	64	43	1	NA	NA
90-94	2	9	19	NA	34	33	7	NA	NA	NA
95-99	4	2	NA	7	10	2	NA	NA	NA	NA
100+	1	NA	1	0	0	NA	NA	NA	NA	NA
					Person years					
69	NA	NA	NA	NA	NA	533.61	NA	NA	NA	310.88
70-74	NA	NA	NA	NA	817.36	1681.26	NA	NA	593.15	471.93
75-79	NA	NA	NA	498.39	2153.88	1257.09	NA	517.54	820.78	50.85
80-84	NA	NA	258.48	1240.42	1403.28	NA	618.38	658.88	63.41	NA
85-89	NA	83.35	503.03	598.08	NA	489.56	670.26	52.52	NA	NA
90-94	9.05	150.42	186.55	NA	167.64	423.46	43.17	NA	NA	NA
95-99	19.21	32.01	NA	18.95	89.73	22.87	NA	NA	NA	NA
100+	1.55	NA	1.46	8.99	3.29	NA	NA	NA	NA	NA
					Rate					
65-69	NA	NA	NA	NA	NA	1.87	NA	NA	NA	28.95
70-74	NA	NA	NA	NA	7.34	7.14	NA	NA	20.23	8.48
7579	NA	NA	NA	8.03	19.04	15.91	NA	44.44	20.71	19.67
80-84	NA	NA	50.29	44.34	30.64	NA	97.03	33.39	0	NA
85-89	NA	71.98	61.63	55.18	NA	130.73	64.15	19.04	NA	NA
90-94	220.96	59.83	101.85	NA	202.81	77.93	162.14	NA	NA	NA
95-99	208.23	62.49	NA	369.37	111.45	87.44	NA	NA	NA	NA
100+	646.75	NA	686.56	0	0	NA	NA	NA	NA	NA

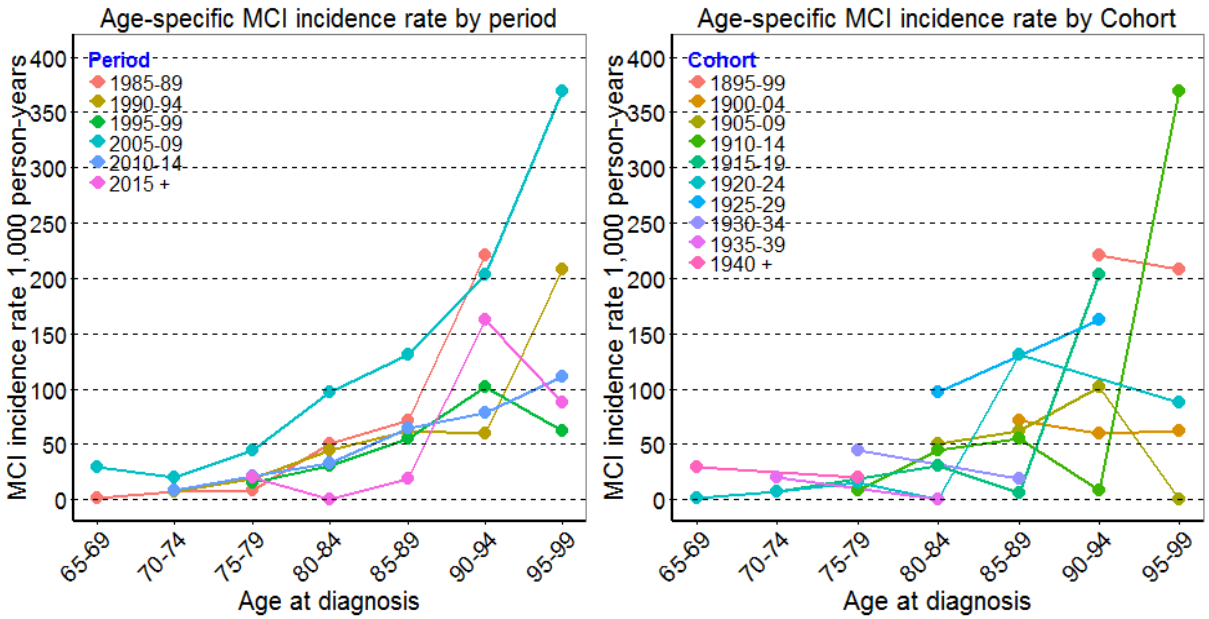

Figure 3. Age-specific MCI incidence rates by calendar period and birth cohorts

Age-specific rates by calendar period (Left panel) and by birth cohort (Right panel).

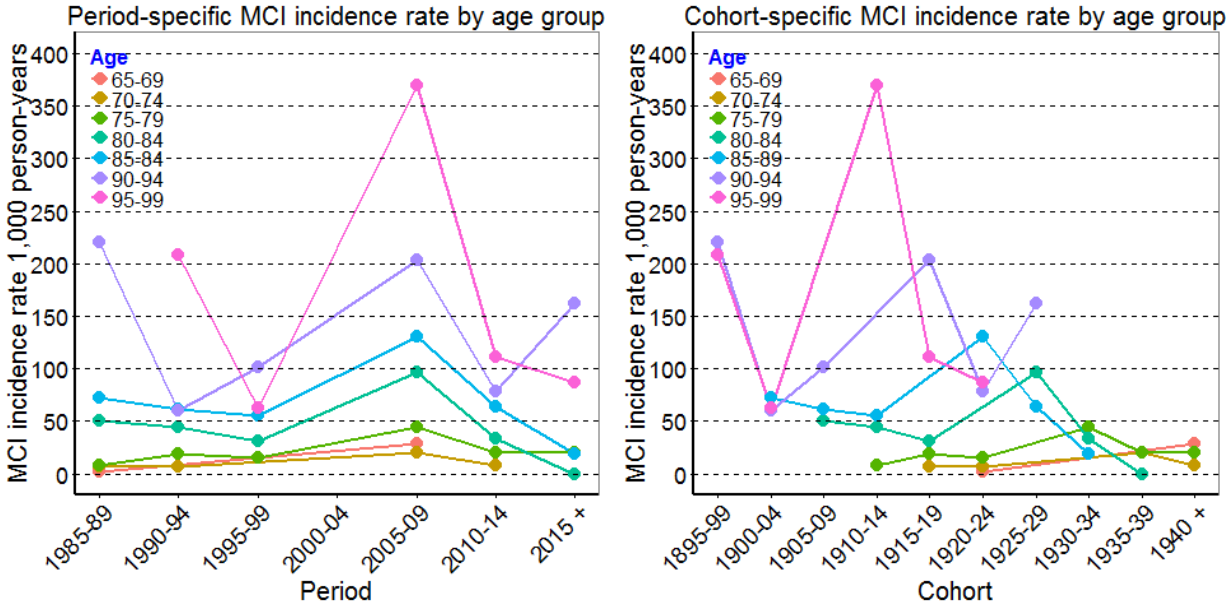

Figure 4. Period-specific and Cohort-specific MCI incidence rate by age

Period-specific (Left panel) and Cohort-specific (Right panel) MCI incidence rates by age group.

3.2 ESTIMATING MCI INCIDENCE RATES USING THE APC MODELLING APPROACH

3.2.1 Single-factor APC models

We first modelled the effects of age, period, and cohort using the traditional method by treating each temporal factor at a single-factor level. We observed that the incidence rates of MCI in the older participants (groups with age >= 80) were significantly higher than that in the younger-old participants. As shown in **Table 4**, the MCI incidence rates for age groups of 80-84 and 85-89 years were 3.8 and 6.3 times of that of the youngest

Table 4. Poisson log-linear model with age effects only

	Estimate	Std. Error	z value	Pr(>|z|)	*RR	p-value[(a)]
(Intercept)	-4.436	0.316	-14.029	0.000	1.0	
A70-74	-0.216	0.360	-0.601	0.548	0.8	
A75-79	0.524	0.331	1.585	0.113	1.7	
A80-84	1.346	0.324	4.150	0.000	3.8	
A85-89	1.836	0.325	5.650	0.000	6.3	
A90-94	2.193	0.331	6.623	0.000	9.0	
A95-99	2.447	0.374	6.539	0.000	11.6	<0.0001

(a) P-value by likelihood ratio test (LRT) for all age groups
*RR: Risk ration relative to the control group

age group (age group of 65-69 which is control group). The MCI incidence rates for age groups of 90-94 and 95-99 years were 9.0 and 11.6 times of that of the youngest age group (age group of 65-69). We found that the MCI incidence rates increased gradually with time when comparing results from the earliest period, 1985-1989, through period 2005-2009, and that the MCI incidence rates decreased starting from period 2010-2014

and remained stable afterwards (**Table 5**). We also found that the first birth cohort of 1895-1899 had the highest MCI incidence rate. However, because participants in the first cohort were all aged at 90 or older, moreover, this earliest cohort only had total seven cases along with relatively short follow-up person-years (which could be seen in **Table 3)**, the unusually high MCI incidence rate of the first cohort may not represent the true MCI incidence rate of that cohort. No clear pattern was observed for the change of MCI incidence rates among all the other cohorts (**Table 6**).

Table 5. Poisson log-linear model with calendar period effects only

	Estimate	Std. Error	z value	Pr(>|z|)	RR*	p-value(a)
(Intercept)	-4.231	0.177	-23.934	0.000	1.0	
P1990-1994	0.598	0.194	3.074	0.002	1.8	
P1995-1999	0.839	0.199	4.205	0.000	2.3	
P2005-2009	1.666	0.190	8.777	0.000	5.3	
P2010-2014	1.040	0.197	5.266	0.000	2.8	
P2015+	1.178	0.349	3.371	0.001	3.2	<0.0001

(a) P-value by likelihood ratio test (LRT) for all period groups
*RR: Risk ration relative to the control group

Table 6. Poisson log-linear model with birth cohort effects only

	Estimate	Std. Error	z value	Pr(>|z|)	RR*	p-value(a)
(Intercept)	-1.550	0.408	-3.796	0.000	1.0	
C1900-1904	-1.200	0.475	-2.527	0.012	0.3	
C1905-1909	-1.162	0.427	-2.719	0.007	0.3	
C1910-1914	-1.620	0.420	-3.853	0.000	0.2	
C1915-1919	-1.993	0.417	-4.776	0.000	0.1	
C1920-1924	-1.959	0.417	-4.692	0.000	0.1	
C1925-1929	-0.944	0.419	-2.252	0.024	0.4	
C1930-1934	-1.736	0.434	-3.998	0.000	0.2	
C1935-1939	-2.381	0.448	-5.309	0.000	0.1	
C1940+	-2.537	0.488	-5.199	0.000	0.1	<0.0001

(a) P-value by likelihood ratio test (LRT) for all birth cohorts
*RR: Risk ration relative to the control group

3.2.2 Two-factor APC models

We fit a two-factor model by including age and period but excluding birth cohort and the results are shown in **Tables 7**. After adjusting the period effect, compared to the single factor model with age only, age showed a more significant effect on the MCI incidence rate. As presented in **Table 7**, the incidence rates for age groups of 80-84 and 85-89 years were 4.8 and 7.3 times of that of the youngest age group (age group of 65-69) respectively, while the incidence rates for age groups of 90-94 and 95-99 years were 10.8 and 15.3 times of that of the youngest age group (age group of 65-69) respectively. The age effects we observed from this two-factor model with age and period are consistent with the findings from the single-factor model. However, after adjusting for age, we found that except for calendar period 2005-2009 all the other calendar periods do not have significantly higher MCI incident rate compared to the baseline period 1985-1999, which is different from the results of the single-factor model.

Table 7. Poisson log-linear model with age and period effects

	Estimate	Std. Error	z value	Pr(>|z|)	RR	p-value[(a)]
(Intercept)	-4.916	0.338	-14.548	0.000	1.0	
A70-74	-0.046	0.364	-0.126	0.899	1.0	
A75-79	0.795	0.339	2.344	0.019	2.7	
A80-84	1.569	0.333	4.713	0.000	4.8	
A85-89	1.987	0.334	5.956	0.000	7.3	
A90-94	2.382	0.341	6.980	0.000	10.8	
A95-99	2.731	0.386	7.080	0.000	15.3	<0.0001
P1990-1994	0.146	0.200	0.730	0.466	1.2	
P1995-1999	-0.027	0.207	-0.131	0.896	1.0	
P2005-2009	0.983	0.196	5.025	0.000	2.7	
P2010-2014	0.077	0.207	0.374	0.708	1.1	
P2015+	-0.076	0.358	-0.212	0.832	0.9	<0.0001

(a) P-value by likelihood ratio test (LRT) for all age groups or all periods

Similarly, after adjusting for age in the analysis using the age-cohort two-factor model, although the entire cohorts still showed an overall significant effect on the MCI incidence rate based on the likelihood ratio test for all cohorts (p=0.01), we found no significant effect on the MCI incidence rate in any single birth cohort; see **Table 8**. These findings indicate that age is the most important temporal factor that affects the development of MCI. A model with period and cohort without including age may not make any sense at all, we did not fit the data into a two-factor model of period and cohort.

Table 8. Poisson log-linear model with age and cohort effects

	Estimate	Std. Error	z value	Pr(>|z|)	RR	p-value[(a)]
(Intercept)	-4.241	0.568	-7.461	0.000	1.0	
A70-74	-0.073	0.374	-0.195	0.846	0.9	
A75-79	0.783	0.372	2.107	0.035	2.2	
A80-84	1.608	0.375	4.289	0.000	5.0	
A85-89	2.049	0.373	5.500	0.000	7.8	
A90-94	2.514	0.372	6.766	0.000	12.4	
A95-99	2.765	0.423	6.541	0.000	15.9	<0.0001
C1900-1994	-0.937	0.492	-1.903	0.057	0.4	
C1905-1999	-0.539	0.460	-1.172	0.241	0.6	
C1910-1914	-0.575	0.454	-1.268	0.205	0.6	
C1915-1919	-0.540	0.444	-1.218	0.223	0.6	
C1920-1924	-0.441	0.447	-0.985	0.325	0.6	
C1925-1929	-0.144	0.455	-0.316	0.752	0.9	
C1930-1934	-0.413	0.469	-0.881	0.378	0.7	
C1935-1939	-0.280	0.485	-0.576	0.564	0.8	
C1940+	0.122	0.545	0.224	0.823	1.1	0.01

(a) P-value by likelihood ratio test (LRT) for all age groups or all cohorts

3.2.3 Constraint-based full APC model

We analyzed the results with a constraint-based full APC model using the constraint cohort1 = cohort2 (i.e., C1895-1899 = C1900-1904) and the results show a clear trend

of significant increase in MCI incidence rates in the older-old population after controlling the effects of period and cohort; see **Table 9**. The finding observed in the constraint-based full APC model confirmed the age effects that we have seen from the reduced one-factor model with age group only (**Table 4**) and two-factor model with age and period (**Table 7**). Moreover, based on the likelihood ratio test for all the calendar periods, the overall entire periods showed a significant effect on MCI incidence rate after adjusting for the effects of age and cohort ($p<0.0001$) although none of the single calendar period show a significant effect on MCI incidence rate after adjusting for the effects of age and cohort. It was observed that the MCI incidence rates was decreasing from period of 1990-1994 through 2015 after compared to the earliest period of 1985-1989. No significant effect from either the entire cohorts or any single cohort was observed.

Table 9. Poisson log-linear model with age, period, and cohort effects*

	Estimate	Std. Error	z value	Pr(>|z|)	p-value[(a)]
(Intercept)	-9.107	2.149	-4.238	0.000	
A70-74	1.128	0.650	1.736	0.083	
A75-79	2.949	1.078	2.736	0.006	
A80-84	4.544	1.544	2.943	0.003	
A85-59	5.891	2.028	2.905	0.004	
A90-95	7.276	2.518	2.890	0.004	
A95-99	8.452	2.947	2.867	0.004	<0.0001
P1990-1994	-0.786	0.532	-1.478	0.139	
P1995-1999	-1.833	1.006	-1.822	0.069	
P2005-2009	-2.649	1.986	-1.334	0.182	
P2010-2014	-4.493	2.478	-1.813	0.070	
P2015+	-5.594	2.981	-1.877	0.061	<0.0001
C1905-1909	1.329	0.659	2.018	0.044	
C1910-1914	2.153	1.136	1.896	0.058	
C1915-1919	2.934	1.620	1.811	0.070	
C1920-1924	3.804	2.106	1.806	0.071	
C1925-1929	4.945	2.602	1.901	0.057	
C1930-1934	5.647	3.099	1.822	0.068	
C1935-1939	6.684	3.594	1.860	0.063	
C1940+	8.010	4.096	1.956	0.051	0.27

*By constraint cohort1 = cohort2 (i.e. C1895-1899 = C1900-1904)

(a) P-value by likelihood ratio test (LRT) for overall all age, overall calendar period, or overall birth cohort effects**.**

We also chose the earliest two calendar periods as constraints to fit another constrain-based full APC model by letting period1=period2 (i.e., P1985-1989 = P1990-1904) and the results are shown in **Table 10.** Note that **Tables 9 and 10** showed similar results for the effects on trend of age, period, and cohort.

Table 10. Poisson log-linear model with age, period, and cohort effects*

	Estimate	Std. Error	z value	Pr(>|z|)	p-value[(a)]
(Intercept)	-5.178	1.228	-4.218	0.000	
A70-74	0.342	0.444	0.772	0.440	
A75-79	1.377	0.537	2.565	0.010	
A80-84	2.187	0.673	3.249	0.001	
A85-89	2.748	0.830	3.309	0.001	
A90-94	3.348	1.013	3.304	0.001	
A95-99	3.737	1.210	3.089	0.002	<0.0001
P1995-1999	-0.261	0.260	-1.004	0.316	
P2005-2009	0.494	0.645	0.766	0.444	
P2010-2014	-0.564	0.842	-0.670	0.503	
P2015+	-0.879	1.077	-0.816	0.414	<0.0001
C1900-1904	-0.786	0.532	-1.478	0.139	
C1905-1909	-0.243	0.606	-0.400	0.689	
C1910-1914	-0.204	0.737	-0.277	0.782	
C1915-1919	-0.209	0.897	-0.233	0.816	
C1920-1924	-0.125	1.083	-0.115	0.908	
C1925-1929	0.230	1.275	0.181	0.857	
C1930-1934	0.147	1.464	0.100	0.920	
C1935-1939	0.398	1.649	0.241	0.809	
C1940+	0.939	1.826	0.514	0.607	0.31

*By constraint Period1 = Period2 (P1895-1899 = P1990-1994)
(a) P-value by likelihood ratio test (LRT) for overall all age, overall calendar period, or overall birth cohort effects

Thus far we observed that in the single-factor and two-factor models, age had the most pronounced effects on MCI incidence, therefore, we decided not to set constrains on the age effect. The earliest two calendar periods and the earliest two birth cohorts did not exhibit much difference and may be deemed to have equal effects. The results showed that the estimates under these two different constraint models were similar. Furthermore, the cohort equality constraints generate remarkably similar results to that from the reduced two-factor models. The results from the constraint-based full APC model on period and cohort effects are highly consistent with that of the two-factor models adjusting for age.

3.3 MODERATOR AND CONFOUNDER EFFECTS OF GENDER AND EDUCATION

We examined the associations between gender and risk of MCI and found that the age-specific MCI incidence rates for men and women were very close to each other; **Figure 5A.** Similar trends were observed from the period-specific and cohort-specific rates; see **Figures 5B and 5C**.

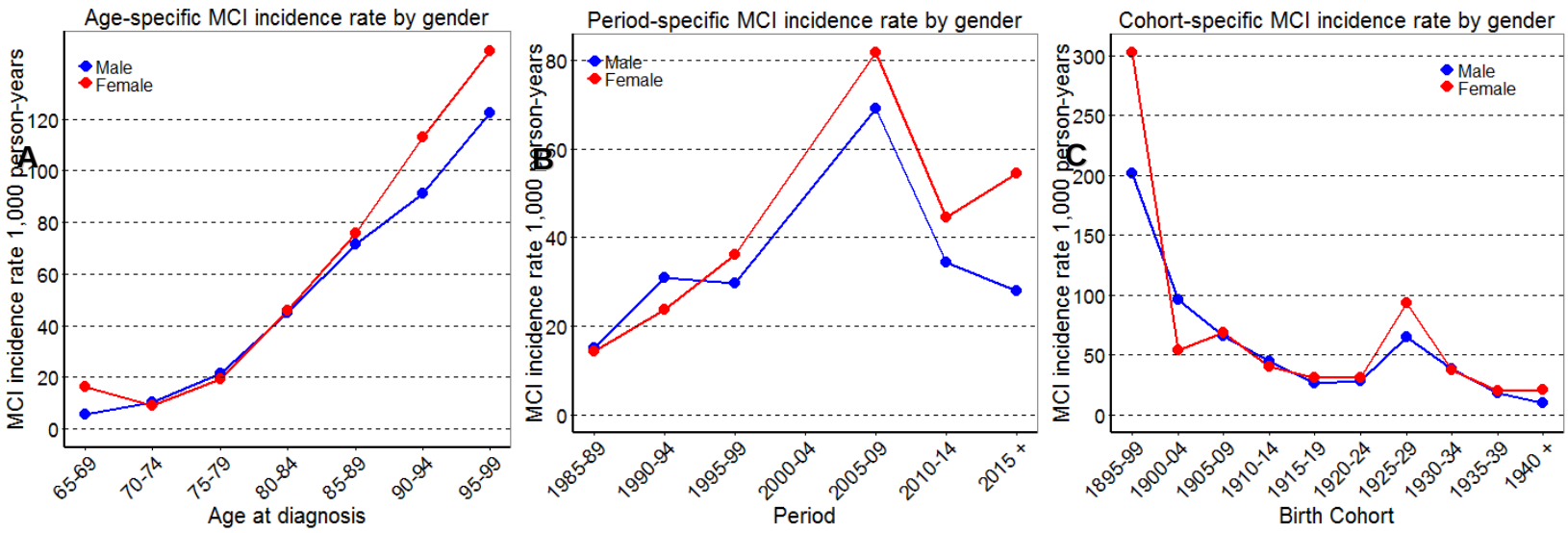

Figure 5. Age-specific (A), Period-specific (B) and Birth-cohort-specific (C) MCI incidence rate by gender

Further analysis with Poisson log-linear model stratified by gender demonstrated that gender did not play any significant role in moderating the effects of age, since none of the interaction terms (age*gender, period*gender, and cohort*gender) is significant; see **Supplementary Tables 3-5**. Gender either did not show any significant effect on the MCI incidence rates after adjusting for age, period, and cohort respectively; see **Supplementary Tables 3.1, 4.1 and 5.1** respectively.

Evidence regarding whether the trajectory of cognitive change in late life could be moderated by education remains conflicting. Early studies suggested that higher levels of education may help slow cognitive decline with aging [31-33]. In contrast, another longitudinal study reported that education does not help attenuate cognitive decline with aging [34]. In here we sought to examine whether education of the participants could have any impact in moderating the temporal effects on the development of MCI. **Figure 6A** showed that participants receiving lower than HS education had significantly higher MCI incidence rates than those who had higher than HS education. After adjusting age, the MCI incidence rates for participants who received HS and higher than HS education were 0.92 (95% CI: 0.77-1.11, p=0.427) and 0.75 (95% CI: 0.61-0.92, p=0.005) times of those who had less than HS education respectively; see **Table 11**. Similar trend was observed in period-specific MCI incident rates of participants who received different level of education; see **Figure 6B**. The ratios of getting MCI for participants who received HS and higher than HS education over those who had less than HS education are 0.54 (95% CI: 0.46-0.67, p=0.000) and 0.42 (95% CI: 0.35-0.53, p=0.000) respectively after adjusting for period; see **Table 13**. Despite the observed effects of education on the MCI incidence rate, education does not moderate the effects of age or

period on MCI incidence rates while the interaction terms age*education and period*education did not show a significant effect in the Poisson log-linear model (**Supplementary Tables 6** and **7**).

Table 11. Poisson log-linear model with age and education effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.306	0.323	-13.341	0.000
A70-74	-0.222	0.360	-0.616	0.538
A75-79	0.508	0.331	1.534	0.125
A80-84	1.324	0.325	4.080	0.000
A85-89	1.821	0.325	5.599	0.000
A90-94	2.177	0.331	6.571	0.000
A95-99	2.410	0.375	6.435	0.000
Edugp1	*-0.074	0.093	-0.794	0.427
Edugp2	#-0.284	0.100	-2.835	0.005

*The ratio of getting MCI for participants received HS education is exp (-0.074) = 0.92 (95% CI: 0.77-1.11) times of those receiving less than HS education after adjusting for age effects.
#The ratio of getting MCI for participants receiving more than HS education is exp (-0.284) = 0.75 (95% CI: 0.61-0.92) times of those receiving less than HS school education after adjusting for age effects.

Table 12. Estimate of Age effects before and after adjusting for education.

	[1]Non-Adj. Est.	[2]Adj. Esti.	*Change (%)
(Intercept)	-4.436	-4.306	2.93
A70-74	-0.216	-0.222	2.78
A75-79	0.524	0.508	3.05
A80-84	1.346	1.324	1.63
A85-89	1.836	1.821	0.82
A90-94	2.193	2.177	0.73
A95-99	2.447	2.410	1.51

[1]Non-Adj. Est: Estimates obtained from the model without adjusting education
[2]Adj. Est: Estimates obtained from the model after adjusting education
*Change: means the change of coefficients before and after adjusting education = [(Adj. Est. – Non-Adj. Est.)/Non-Adj.Est] x 100%

Table 13. Poisson log-linear model with period and education effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.893	0.180	-21.592	0.000
P1990-1994	0.613	0.195	3.146	0.002
P1995-1999	0.909	0.200	4.545	0.000
P2005-2009	1.934	0.194	9.959	0.000
P2010-2014	1.313	0.204	6.449	0.000
P2015+	1.454	0.381	3.819	0.000
edugp1	*-0.619	0.100	-6.218	0.000
edugp2	#-0.871	0.109	-7.959	0.000

*The ratio of getting MCI for participants receiving HS education is exp (-0.619) = 0.54 (95% CI: 0.46-0.67) times of those receiving less than HS education after adjusting period.
#The ratio of getting MCI for participants receiving more than HS is exp (-0.871) = 0.42 (95% CI: 0.35-0.53) times of those receiving less than HS education after adjusting period.

Table 14. Estimates of period effects before and after adjusting for education

	[1]Non-Adj. Est.	[2]Adj. Est.	Change (%)
(Intercept)	-4.231	-3.893	7.75
P1990-1994	0.598	0.613	6.20
P1995-1999	0.839	0.909	8.60
P2005-2009	1.666	1.934	16.80
P2010-2014	1.040	1.313	29.10
P2015+	1.178	1.454	28.90

[1]Non-Adj. Est: Estimates obtained from the model without adjusting education
[2]Adj. Est: Estimates obtained from the model after adjusting education
*Change: means the change of coefficients before and after adjusting education = [(Adj. Est. – Non-Adj. Est.)/Non-Adj.Est] x 100%

The cohort-specific MCI incidence rates among participants with different education levels are depicted in **Figure 6C**. Different from the age-specific and period-specific MCI incidence rates, a significant effect of interaction between cohort and education was found (**Supplementary Table 8**), which means that education level significantly moderates the impacts of birth cohort on the MCI incidence rates. There was not a clear monotone trend in cohort-specific MCI incidence rates among individuals receiving HS or higher than HS education compared to those receiving lower than HS education, in other words, the cohort effect on MCI incidence rates was different depending on the education level.

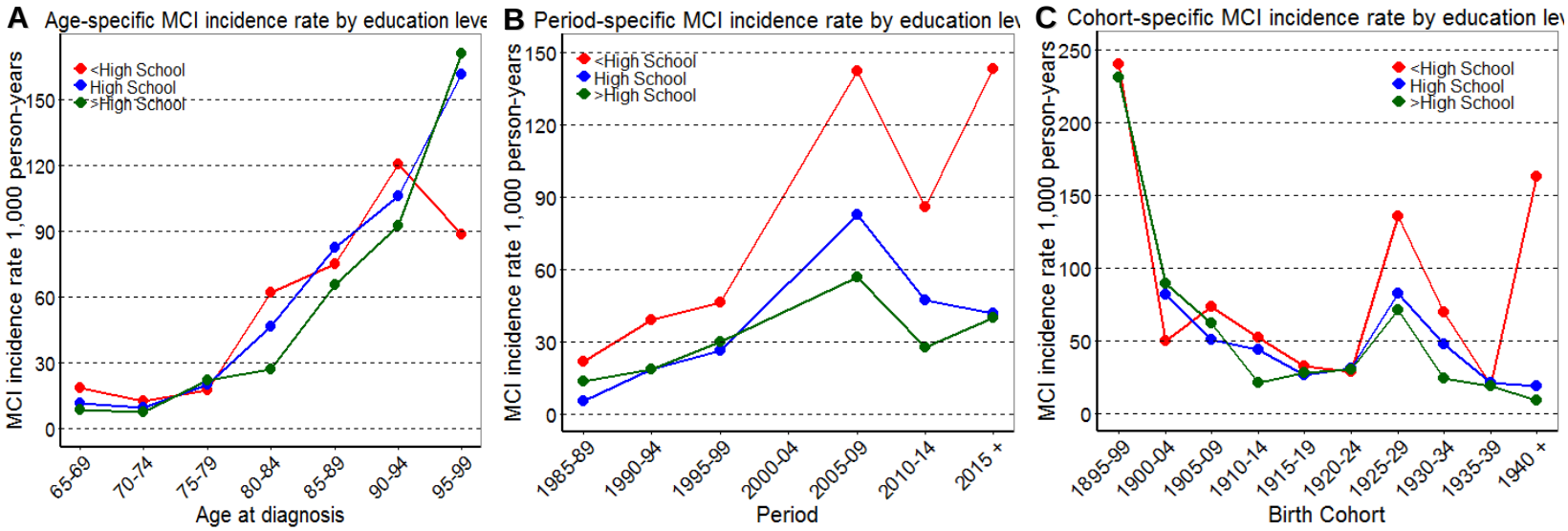

Figure 6. Age-specific (A), Period-specific (B) and Birth-cohort-specific (C) MCI incidence rate by education

We then explored whether effects of age, period, and cohort on the MCI incidence rates were confounded by education. The estimated coefficients for age were compared before and after adjusting for education. If the percentage of the absolute change in the estimated coefficients is greater than 10%, we considered the effect of age as confounded. The percentage of change in coefficients before and after adjusting for education among age groups vary from 0.73% to 3.05 %, as shown in **Table 12**, indicating that the effect of age is not confounded by education. The percentage of change in coefficients before and after adjusting for education during periods 2005-2009, 2010-2014, and 2015-after are 16.80%, 29.10%, and 28.90% respectively, as shown in **Table 14**, indicating that the effects of these three period groups are confounded by education. No confounding effect from education was observed on the cohort effects as the percentage of change in coefficients before and after adjusting education for birth cohorts varies from 0.9% to 8.9 %; see **Table 15**.

Table 15. Estimate of cohort effects before and after adjusting for education

	[1]Non-Adj. Est.	[2]Adj. Est.	Change (%)
(Intercept)	-1.550	-1.344	13.29
C1900-1904	-1.200	--1.307	8.90
C1905-1909	-1.162	-1.264	8.80
C1910-1914	-1.620	-1.701	5.00
C1915-1919	-1.993	-2.050	2.90
C1920-1924	-1.959	-1.981	0.90
C1925-1929	-0.944	-0.921	2.40
C1930-1934	-1.736	-1.692	2.50
C1935-1939	-2.381	-2.335	1.90
C1940+	-2.537	-2.455	3.20

[1]Non-Adj. Est: Estimates obtained from the model without adjusting education
[2]Adj. Est: Estimates obtained from the model after adjusting education
*Change: means the change of coefficients before and after adjusting education = [(Adj. Est. – Non-Adj. Est.)/Non-Adj.Est] x 100%

3.4 MODERATOR EFFECT OF APOE4 ALLELE

We examined the impact of ApoE4 allele on moderating the effects of age, period, and cohort on the development of MCI. We modelled the MCI incidence rates with stratified data of ApoE4 allele expression. There was no significant effect from the interaction between age and ApoE4 allele on the age-specific MCI incidence rate; see **Supplementary Table 9**. No significant effect from the interaction terms between ApoE4 allele and period or cohort was observed when fitting the Poisson log-linear model with the stratified data of ApoE4 allele; see **Supplementary Tables 10 and 11**. These results show that in this study ApoE4 allele did not moderate the temporal effects on the development of MCI.

Nonetheless, as shown in **Figure 7A** and **Table 16**, the age specific MCI rates were significantly higher in the participants expressing ApoE4 allele than those non-ApoE4 allele expressing participants (p=0.012). The average rate of ApoE4 allele positive patients getting MCI is 1.29 (95% CI: 1.05-1.57) times of that for ApoE4 allele negative participants after adjusting age. Marginal effects of ApoE4 allele on the period specific MCI incident rate (**Figure 7B and Table 17**) and the cohort specific MCI incident rate (**Figure 7C and Table 18**) were observed.

Table 16. Poisson log-linear model with age and APOE4 effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.278	0.318	-13.473	0.000
A70-74	-0.575	0.376	-1.529	0.126
A75-79	0.275	0.334	0.823	0.411
A80-84	1.108	0.326	3.395	0.001
A85-89	1.594	0.327	4.875	0.000
A90-94	2.002	0.333	6.003	0.000
A95-99	2.232	0.385	5.804	0.000
Allele	*0.256	0.103	2.498	0.012

*The ratio of getting MCI for APOE positive participants is exp (0.256) = 1.29 (95% CI: 1.05-1.57, p=0.012) times of those APOE negative after adjusted for age.

Table 17. Poisson log-linear model with calendar period and APOE4 effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.903	0.317	-15.452	0.000
P1990-1994	1.026	0.333	3.080	0.002
P1995-1999	1.445	0.331	4.371	0.000
P2005-2009	2.269	0.324	6.992	0.000
P2010-2014	1.686	0.329	5.125	0.000
P2015+	1.896	0.437	4.339	0.000
Allele	*0.184	0.102	1.798	0.072

*The ratio of getting MCI for APOE4 carriers is exp (0.184) =1.20 (95% CI: 0.98-1.46) times higher than those non-APOE*4 carriers after adjusting for period.

Table 18. Poisson log-linear model with birth cohort and APOE4 effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.432	1.000	-2.432	0.015
C1900-1904	-0.324	1.070	-0.303	0.762
C1905-1909	-0.315	1.013	-0.311	0.756
C1910-1914	-0.766	1.007	-0.761	0.447
C1915-1919	-1.130	1.005	-1.125	0.261
C1920-1924	-1.116	1.005	-1.111	0.267
C1925-1929	-0.079	1.005	-0.078	0.938
C1930-1934	-0.837	1.012	-0.828	0.408
C1935-1939	-1.613	1.020	-1.581	0.114
C1940+	-1.726	1.038	-1.662	0.096
Allele	*0.185	0.102	1.805	0.071

*The ratio of getting MCI for APOE*4 carriers is exp (0.185) =1.20 (95% CI: 0.98-1.47) times higher than that of non-APOE*4 carriers after adjusting for cohort.

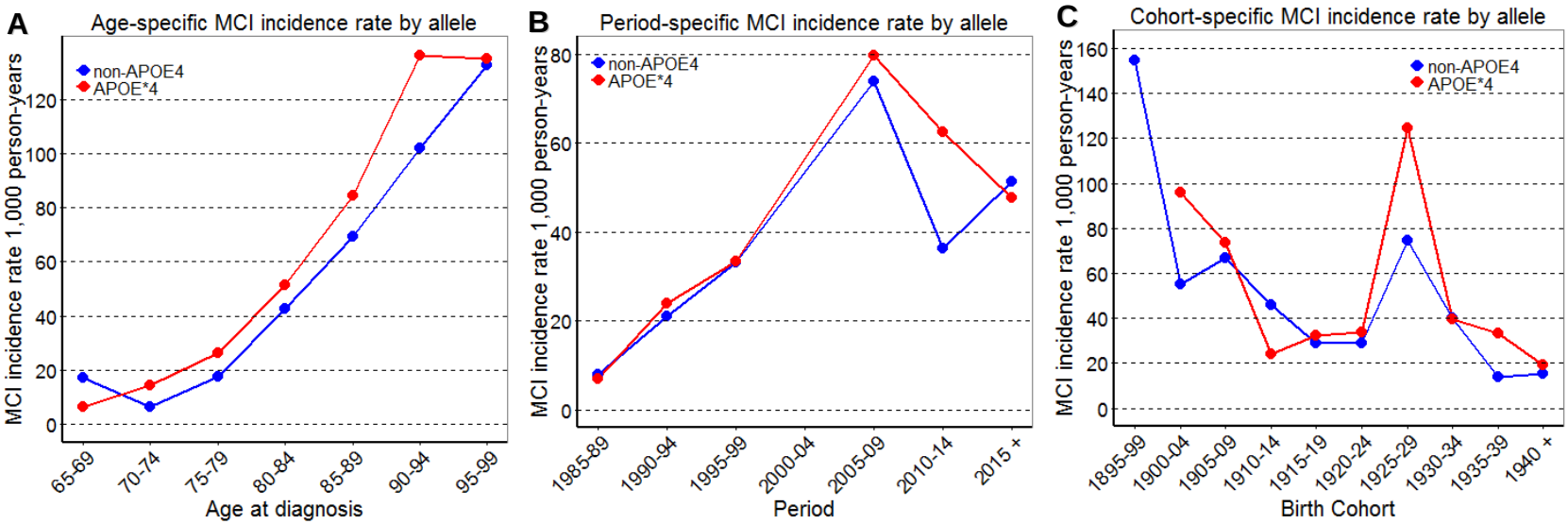

Figure 7. Age-specific (A), Period-specific (B) and Birth-cohort-specific (C) MCI incidence rate by ApoE4 allele

3.5 MODERATOR EFFECTS OF STUDY

We examined whether there was a difference in the MCI incidence rates between the MYHAT and the MoVIES studies. The MCI incidence rate in the MYHAT study cohort, as shown in **Figure 8,** were consistently higher than in the MoVIES study for all age groups. We checked the interaction between age and study with Poisson log linear model to see if the age effect on MCI incidence rates was moderated and found that the study did not moderate the age effect since the interaction of age*study was not significant (p=0.13); see **Supplementary Table 12**. We did find a significant difference in age-specific MCI incidence rates between MoVIES and MYHAT (p <0.001) after dropping the interaction term from our model: the ratio of getting MCI for the MYHAT participants was 1.62 (95% CI: 1.38-1.90) times of that for the MoVIES participants after adjusting for age; see **Table 19**.

Table 19. Poisson log-linear model with age and study effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.641	0.319	-14.562	0.000
A70-74	-0.180	0.360	-0.502	0.616
A75-79	0.579	0.331	1.750	0.080
A80-84	1.372	0.324	4.232	0.000
A85-89	1.769	0.325	5.440	0.000
A90-94	2.061	0.332	6.212	0.000
A95-99	2.284	0.375	6.088	0.000
Study	*0.482	0.082	5.859	0.000

*The ratio of getting MCI for patients in MYHAT study is exp (0.48) = 1.62 (95% CI: 1.38-1.90, p=0.000) times of patients in MOVIES study after adjusting for age group.

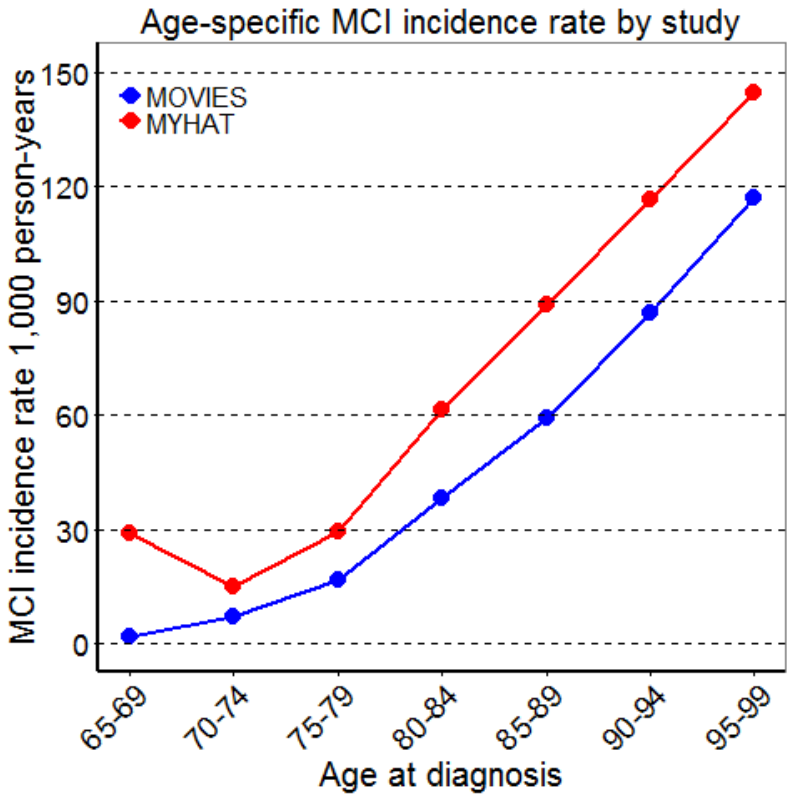

Figure 8. Age-specific MCI incidence rates by study

3.6 EVALUATION OF MODEL FITNESS

For each of the model used in the study, the Akaike information criterion (AIC) is listed in **Table 20**. The model with age and period has the smallest AIC value. The constraint-base full APC model has the second smallest AIC value that is very close to the AIC value of model with age and period. To select the appropriate model, we compared the coefficients of age, period and cohort from different models, including the reduced two-factor and full APC models, it is of note that all the coefficients for these 3 time-related factors age, period and cohort changed quite a lot among different models that means

the effects of these 3 factors could be confounded by each other. Given such a situation, all 3 time related factors should be included into the model. Education was a confounder of period effect on the MCI incidence rate and thus was considered being included in the analysis model. However, when fitting the data into the Poisson log-linear model with age, period, cohort and education, we found a much larger AIC value (444.0) compared to the full APC models without including education. As the education is not a confounder of age effect but period effect only, and age is the most important influence factor of MCI incidence rate. Moreover, education moderated the cohort effects, but we cannot put that many interaction terms into the model with our current sample size. Taken all together, we chose the constraint-based full APC model with period1=period2 or cohort1=cohort2 as constraints for our current data to investigate the independent effects of temporal factors on the MCI incidence rates.

Table 20. AIC of Different models

Model	AIC
Age	292.3
Period	502.5
Cohort	529.6
Age+Period	196.1
Age+Cohort	289.1
Age+period+cohort (P1=P2)	201.2
Age+period+cohort (C1=C2)	202.2
Age+Period+Cohort+Education	444.0

4.0 DISCUSSION

In this study, by using the APC modelling approach, we examined the temporal changes in MCI incidence rates in the senior population (>= 65 years old) recruited to the community-based cohort studies from Year 1985 to 2008. We observed that much of the rise in MCI incidence rates was driven by an age effect. Overall, within the same calendar period and the same birth cohort, the MCI incidence rate was monotonically increasing in age. This finding is consistent with those reported in the recent literatures [35-37].

We also found that among those who are within the same age group and within the same birth cohort, calendar period of 2005-2009 showed a significantly increased MCI incidence rate compared to other calendar periods. A possible reason leading to an increased MCI rate in 2005-2009 might be that the recruitment criteria and the frequency of the outcome assessments for the two study cohorts are different. For the MoVIES study, participants needed to have at least Grade 6 education to be eligible for recruitment. For the MYHAT study, individuals having Mini Mental State Examination (MMSE) score <21 were not eligible. At study baseline, 43.3% of MoVIES and 13.8% of MYHAT participants had less than HS education and their averaged ages were 72.8 ± 5.9 and 77.3 ± 7.3 years old, respectively. The assessments for participants recruited into MoVIES were done approximately once every two years while the participants in

MYHAT were done annually. We suspected that with a higher education level and 5 years older in averaged age, MYHAT participants may contain higher percentage of MCI prevalent cases (false negative for MCI at baseline). Therefore, the incidence rate of MCI in 2005-2009 was much higher compared to other calendar time periods. For the effect of birth cohort, none of the birth cohorts showed a significant effect on the MCI incident rates after adjusting for both age and period. Our study was the first study that examined the effects of age, calendar period, and birth cohort simultaneously on the MCI incidence rates using the APC modelling strategy.

Besides exploring the time trends in MCI incidence rates, we studied the underlying factors that could moderate or confound the effects of age, period, and cohort on the development of MCI. No apparent moderating effect of gender was observed on the development of MCI incident rates. The factors we investigated included gender, education, and ApoE4 allele.

Education played a complicated role in the development of MCI. For the same age group, participants receiving higher than HS education had significantly lower MCI incidence rate. However, the age effect on MCI incidence rate was not moderated or confounded by education. Our study found that, on the other hand, the period effects and cohort effects were significantly confounded and moderated by education, respectively. The effects of the latest 3 periods of 2005-2009, 2010-2014 and 2015 after, especially periods 2010-2014 and 2015 after were significantly confounded by education. It indicates that higher education is both associated with later time period and with lower MCI incidence rate. This also explains the overall trend of MCI incidence rate is decreasing throughout the observation periods of this study after adjusting the

effects of age and cohort. Education moderated the cohort effects on MCI incidence rates, the cohorts receiving HS education showed an increasing trend in MCI incidence rates compared to control, whereas the birth cohorts receiving higher than HS education did not show a significant change compared to the control.

Our study confirmed that genetic factor, ApoE4 allele, is a risk factor in the development of MCI, which is consistent with literatures [4, 5]. We found that, except the youngest and the oldest age groups, the MCI incident rates were significantly and consistently higher in the ApoE4 allele positive population than the ApoE4 allele negative population within the same age group. As mentioned earlier, participants in the youngest and oldest age groups are fewer, therefore, the data from these participants might not predict the true MCI incidence rates in these subpopulations. We concluded that ApoE4 could mediate the decline of cognitive functions, which agreed with the previous reports [38, 39].

Besides the aforementioned findings, however, our study also has some limitations. There was a 5-year gap, approximately, between MoVIES and MYHAT data, which made the follow-up time not continuous. There was a low number of participants and few MCI cases in the period of 2000-2004, which resulted in an unusually low MCI incidence rate in this period. , We performed a sensitivity analysis by fitting the data into Poisson log-linear model with or without period of 2000-2004 and yield the same results, therefore, we removed the period of 2000-2004 in our analysis. Moreover, there were some other groups or periods (the youngest age-group, the earliest period group, and birth cohort) with low number of participants, which may have prevented us to find additional associations. Thus far it is not well documented about which statistical

package(s) could be used to efficiently handle the APC model with unequal age, period, or cohort intervals.

Despite these limitations, this study has several strengths. First, our study was the first time to apply the APC model analysis to study MCI incidence rates. By using this model, we estimated the contributions of age, period, and birth cohort to MCI incidence rates simultaneously. Second, the data used in our study was from a large population, with a long follow-up time span, which allow us to estimate the calendar period effects well. Using the APC model, our studies confirmed the effects of multiple factors, e.g., demographic (e.g. age, education) and genetic (e.g. ApoE4 allele) factors, on the development of MCI.

APPENDIX A: R CODES FOR ANALYSIS

```
#-------Thesis Data Part (I) Primary analysis----------#

        library(sas7bdat)
        library(Epi)
        library(ggplot2)
        library(plyr)
        library(gridExtra)
        library(lmtest)

#-------------------------------------------------------------------------#
# Data pre-processing
#-------------------------------------------------------------------------#
        mydata<- read.sas7bdat("plot3.sas7bdat")
        dim(mydata)#3021
        mydata<-mydata[which(!is.na(mydata$CASE)),]
        dim(mydata)#3021

        Lex.raw<-structure(
        list(id=mydata$researchid,
             birth=as.Date(as.character(format(as.Date(mydata$DTBIR,origin="1960-01-01"),'%d/%m/%Y')),'%d/%m/%Y'),
        INTDT1= as.Date(as.character(format(as.Date(mydata$intdt1,origin="1960-01-01"),'%d/%m/%Y')),'%d/%m/%Y'),
        LTFUDT= as.Date(as.character(format(as.Date(mydata$LTFUDT,origin="1960-01-01"),'%d/%m/%Y')),'%d/%m/%Y'),
        fail=mydata$CASE,
        ageIn=mydata$AGE1,
        allele=mydata$allele,
        study=mydata$STUDY),
        .Names=c("id","DOB","INTDT1","LTFUDT","fail","ageIn","allele","study"),
        row.names=as.character(mydata$researchid),
        class='data.frame')
        Lex.raw$en <- cal.yr( Lex.raw$INTDT1, format="%Y-%m-%d" )
        Lex.raw$ex <- cal.yr( Lex.raw$LTFUDT , format="%Y-%m-%d" )
        Lex.raw$bt <- cal.yr( Lex.raw$DOB , format="%Y-%m-%d" )

#-------------------------------------------------------------------------#
# Convert data.frame to a Lexis project
#-------------------------------------------------------------------------#
        Lex.data<-Lexis(entry=list(per=en,age=en-bt,dob=bt),
                exit=list(per=ex),
                exit.status=fail,
                data=Lex.raw)

#-------------------------------------------------------------------------#
# Figure 2: Lexis-diagram (each line represents one person, each blob      #      # represents an event/case, beginning of the line
represents # the entry date # while the end represents the last follow-up date).
#-------------------------------------------------------------------------#
        png("Figure1 Lexsis Diagram.png",width=800,height=800,res=100)
            plot(Lex.data,time.scale=c("per","age"),grid=0:20*5,
        col="black",xaxs="i",yaxs="i",xlim=c(1980,2020),las=1,
        ylab="Age at entry",xlab="Year of entry")
                points( Lex.data, time.scale=c("per","age"),pch=c(NA,16)[Lex.data$lex.Xst+1], col="red", cex=1 )

        dev.off()
#-------------------------------------------------------------------------# # Function to split data
#-------------------------------------------------------------------------#
```

```
acpSplit<-function(indata,age.breaks,per.breaks,birth.breaks){

split.data1 <- splitLexis( indata, breaks = age.breaks, time.scale="age" )
split.data <- splitLexis( split.data1, breaks =per.breaks, time.scale="per")
split.data$Per_F<-timeBand(split.data,"per","left")
split.data$Age_F<-ncut(split.data$age,age.breaks,type="left")
split.data$Birth_F<-split.data$Per_F-split.data$Age_F
split.data$Birth_F0<-ncut(split.data$bt,birth.breaks,type="left")
index.l<-which(split.data$Birth_F>split.data$Birth_F0)
split.data$Age_F[index.l]<-split.data$Age_F[index.l]+5
split.data$Birth_F<-split.data$Per_F-split.data$Age_F
temp.data<-dply(split.data,
        .(Age_F,Per_F,Birth_F),
        summarize,D=sum(lex.Xst),
        Y=sum(lex.dur),
        Rate=sum(lex.Xst)/sum(lex.dur))
return(list(split.data=split.data,Data=temp.data))

}
#-----------------------------------------------------------------------------#
# Splitting the data
#-----------------------------------------------------------------------------#
Lex.data<-Lex.data[which(Lex.data$age>=65),]
res1<-
acpSplit(indata=Lex.data,age.breaks=seq(65,105,by=5),per.breaks=seq(1985,2020,by=5),birth.breaks=seq(1895,1945,by
=5))

#-----------------------------------------------------------------------------#
# Figure.3 Age-specific MCI incidence rate (per 1,000 person-years) by #      # calendar period and by birth cohort.
# Left panel: Age-specific rates by calendar period. Right panel:          # # Age-specific rates by birth cohorts. (left panel)
#-----------------------------------------------------------------------------#
temp.data<-na.omit(ddply(res1$split.data,.(Age_F,Per_F,Birth_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
temp.data<-temp.data[temp.data$Age_F<100,]
Figure2<-ggplot(data = temp.data,aes(x = Age_F, y = Rate*10^3,color=factor(Per_F))) +
geom_line(lwd=1.5)+
geom_point(size=3)+
xlab("Age") + ylab("MCI Rate Per 1000") +
ggtitle("Age-Specific MCI Incidence Rate by Time Period") +
theme_bw() +
geom_hline(yintercept=seq(0,400,by=50),lty="dashed",colour="black") +
theme(legend.justification=c(0,1),
        legend.position=c(0,1),
        legend.background=element_blank(),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black",size=1),
        legend.title=element_text(colour="blue",size=10,face="bold"),
        legend.text=element_text(colour="black",size=10,face="bold"),
        legend.key=element_blank(),
        axis.title = element_text(colour="black",size=15,face="bold"),
        axis.text = element_text(colour="black", size=12,face="bold"),
        plot.title=element_text(colour="black",  size=16,face="bold")) +
        scale_colour_discrete(name = "Period",labels=c("1985-89","1990-94","1995-99","2000-04","2005-
09","2010-14","2015 +")) +
scale_x_continuous(breaks=c(65,70,75,80,85,90,95),labels=c("65-69", "70-74","75-79","80-84","85-89","90-94","95-99")) +
scale_y_continuous(breaks=seq(0,400,by=50))

png("Figure2_5_all Age-Specific MCI Incidence Rate by Time Period.png",width=800,height=800,res=100)
Figure2
dev.off()

#-----------------------------------------------------------------------------#
# Figure.3 Age-specific MCI incidence rate (per 1,000 person-years)          # bycalendar period and by birth cohort.
# Left panel: Age-specific rates by calendar period. Right panel:
# Age-specific rates by birth cohorts. (right panel)
#-----------------------------------------------------------------------------#
temp.data<-na.omit(ddply(res1$split.data,.(Age_F,Birth_F,Per_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
temp.data<-temp.data[temp.data$Age<100,]
Figure3<-ggplot(data = temp.data,aes(x = Age_F, y = Rate*10^3,color=factor(Birth_F))) +
```

```
        geom_line(lwd=1.5)+
        geom_point(size=3)+
        xlab("Age") + ylab("MCI Rate Per 1000") +
        ggtitle("Age-Specific MCI Incidence Rate by Birth Cohort") +
        theme_bw() +
        geom_hline(yintercept=seq(0,400,by=50),lty="dashed",colour="black") +
        theme(legend.justification=c(0,1),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
                plot.title=element_text(colour="black",  size=16,face="bold")) +
        scale_colour_discrete(name = "Birth",labels=c("1895-99","1900-04","1905-09","1910-14","1915-19","1920-24","1925-29",
                                                        "1930-34","1935-39","1940 +")) +
        scale_x_continuous(breaks=c(65,70,75,80,85,90,95),
        labels=c("65-69", "70-74","75-79","80-84","85-89","90-94","95-99")) +
        scale_y_continuous(breaks=seq(0,400,by=50))

        png("Figure3_5_all Age-Specific MCI Incidence Rate by Birth Cohort.png",width=800,height=800,res=100)
        Figure3
        dev.off()

#-------------------------------------------------------------------------#
# Figure 4: Period-specific MCI incidence rate (per 1,000 person-years) by #age and Cohort-specific incidence rate by age. Left  # #
panel: Period-# # # specific rates by age groups. Right panel: Cohort-specific rates by age # # # groups. (left panel)
#-------------------------------------------------------------------------#
        temp.data<-na.omit(ddply(res1$split.data,.(Per_F,Age_F,Birth_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
        temp.data<-temp.data[temp.data$Age_F<100,]
        Figure4<-ggplot(data = temp.data,aes(x = Per_F, y = Rate*10^3,color=factor(Age_F))) +
        geom_line(lwd=1.5)+
        geom_point(size=3)+
        xlab("Period") + ylab("MCI Rate Per 1000") +
        ggtitle("Period-Specific MCI Incidence Rate by Age Group") +
        theme_bw() +
        geom_hline(yintercept=seq(0,400,by=50),lty="dashed",colour="black") +
        theme(legend.justification=c(0,1),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
                plot.title=element_text(colour="black",  size=16,face="bold"),
                axis.text.x = element_text(angle=45,hjust=1)) +
        scale_colour_discrete(name = "Age",labels=c("65-69", "70-74","75-79","80-84","85-84","90-94","95-99")) +
                scale_x_continuous(breaks=c(1985,1990,1995,2000,2005,2010,2015),labels=c("1985-89","1990-94","1995-
        99","2000-04","2005-09","2010-14","2015 +")) +
        scale_y_continuous(breaks=seq(0,400,by=50))

        png("Figure4_5_all Period-Specific MCI Incidence Rate by Age Group.png",width=800,height=800,res=100)
        Figure4
        dev.off()

#-------------------------------------------------------------------------#
# Figure 4: Period-specific MCI incidence rate (per 1,000 person-years) by #age and Cohort-specific incidence rate by age. Left
# # panel: Period-specific rates by age groups. Right panel: Cohort-specific # rates by age groups. (right panel)
#-------------------------------------------------------------------------#
        temp.data<-na.omit(ddply(res1$split.data,.(Birth_F,Age_F,Per_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
        temp.data<-temp.data[temp.data$Age_F<100,]
```

```
Figure5<-ggplot(data = temp.data,aes(x = Birth_F, y = Rate*10^3,color=factor(Age_F))) +
geom_line(lwd=1.5)+
geom_point(size=3)+
xlab("Birth") + ylab("MCI Rate Per 1000") +
ggtitle("Birth Cohort-Specific MCI Incidence Rate by Age groups") +
theme_bw() +
geom_hline(yintercept=seq(0,400,by=50),lty="dashed",colour="black") +
theme(legend.justification=c(0,1),
                        legend.position=c(0,1),
                        legend.background=element_blank(),
                        panel.grid.major = element_blank(),
                        panel.grid.minor = element_blank(),
                        axis.line = element_line(colour = "black",size=1),
                        legend.title=element_text(colour="blue",size=10,face="bold"),
                        legend.text=element_text(colour="black",size=10,face="bold"),
                        legend.key=element_blank(),
                        axis.title = element_text(colour="black",size=15,face="bold"),
                        axis.text = element_text(colour="black", size=12,face="bold"),
                        plot.title=element_text(colour="black",  size=16,face="bold"),
                        axis.text.x = element_text(angle=45,hjust=1)

                        ) +
scale_colour_discrete(name = "Age",labels=c("65-69", "70-74","75-79","80-84","85-89","90-94","95-99")) +
scale_x_continuous(breaks=c(1895,1900,1905,1910,1915,1920,1925,1930,1935,1940),
labels=c("1895-99","1900-04","1905-09","1910-14","1915-19","1920-24","1925-29",
"1930-34","1935-39","1940 +")) +
scale_y_continuous(breaks=seq(0,400,by=50))

png("Figure5_5_all Birth-Specific MCI Incidence Rate by Age Group.png",width=800,height=800,res=100)
Figure5
dev.off()

#Table 11 Poisson log-linear model with age, calendar period, and birth #cohort effects#
n.a<-length(unique(as.character(temp.data2$A)))#4 number of age groups
n.p<-length(unique(as.character(temp.data2$P)))#10 number of period groups
n.c<-length(unique(as.character(temp.data2$C)))#4 number of cohort groups

#------Contrast matrix for CGLIM 3 (C1=C2)
con.t<-contr.treatment(n.c,base=1)
con.c<-con.t
con.c[2,1]<-0
con.a<-contr.treatment(n.a,base=1)
con.p<-contr.treatment(n.p,base=1)
fit.f.apc <- glm(D ~ A+ P+ C+ offset(log(Y)),data=temp.data2,
            family = quasi(link = log,var="mu"),
            control = glm.control(epsilon = 1e-010, maxit = 10),
            contrasts= list(A=con.a,P=con.p,C=con.c),intercept = intercept)

write.csv(coef(summary(fit.f.apc)),"model_5_APC_all_C1=C2.csv")

anova(fit.f.apc,test="Chisq")


#Table 12 Poisson log-linear model with age, calendar period, and birth #cohort effects
n.a<-length(unique(as.character(temp.data2$A)))#4 number of age groups
n.p<-length(unique(as.character(temp.data2$P)))#10 number of period groups
n.c<-length(unique(as.character(temp.data2$C)))#4 number of cohort groups

#------Contrast matrix for CGLIM 3 (P1=P2)
con.t<-contr.treatment(n.p,base=1)
con.p<-con.t
con.p[2,1]<-0
con.a<-contr.treatment(n.a,base=1)
con.c<-contr.treatment(n.c,base=1)
fit.f.apc <- glm(D ~ A+ P+ C+ offset(log(Y)),data=temp.data2,
            family = quasi(link = log,var="mu"),
            control = glm.control(epsilon = 1e-010, maxit = 10),
            contrasts= list(A=con.a,P=con.p,C=con.c),intercept = intercept)

write.csv(coef(summary(fit.f.apc)),"model_5_APC_all_P1=P2.csv")
```

```
anova(fit.f.apc,test="Chisq")


#---Thesis Data   Part (II) interaction with gender----#
#  Figure 5
Lex.data<-Lex.data[which(Lex.data$age>=65),]
res<-
acpSplit(indata=Lex.data,age.breaks=seq(65,105,by=5),per.breaks=seq(1985,2020,by=5),birth.breaks=seq(1895,1945,by
=5))


temp.data<-na.omit(ddply(res$split.data,.(sexf,Age_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
temp.data<-temp.data[temp.data$Age_F<95,]
P_age<-ggplot(data = temp.data,aes(x = Age_F, y = Rate*10^3,color=factor(sexf))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Age") + ylab("MCI Rate Per 1000") +
        ggtitle("Age-Specific MCI Incidence Rate by Gender") +
        theme_bw() +
        geom_hline(yintercept=seq(0,120,by=20),lty="dashed",colour="black") +
        theme(legend.justification=c(0,1),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
                plot.title=element_text(colour="black",  size=16,face="bold")) +

  scale_colour_discrete(name = "",labels=c("Male","Female")) +
  scale_x_continuous(breaks=c(65,70,75,80,85,90,95),
            labels=c("65-69", "70-74","75-79","80-84","85-89","90-94","95-99")) +
scale_y_continuous(breaks=seq(0,120,by=20))

png("Figure1_5_Age-Specific MCI Incidence Rate by Gender.png",width=800,height=800,res=100)
P_age
dev.off()

# Figure 6
temp.data<-na.omit(ddply(res$split.data,.(sexf,Per_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))

P_period<-ggplot(data = temp.data,aes(x = Per_F, y = Rate*10^3,color=factor(sexf))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Period") + ylab("MCI Rate Per 1000") +
        ggtitle("Period-Specific MCI Incidence Rate by Gender") +
        theme_bw() +
        geom_hline(yintercept=seq(0,80,by=20),lty="dashed",colour="black") +
        theme(legend.justification=c(0,0.9),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
                plot.title=element_text(colour="black",  size=16,face="bold"),
                axis.text.x = element_text(angle=45,hjust=1)

                ) +

  scale_colour_discrete(name = "",labels=c("Male","Female")) +
```

```
  scale_x_continuous(breaks=c(1985,1990,1995,2000,2005,2010,2015),
                     labels=c("1985-89","1990-94","1995-99","2000-04","2005-09","2010-14","2015 +")) +
  scale_y_continuous(breaks=seq(0,80,by=20))
png("Figure1_5_Period-Specific MCI Incidence Rate by gender.png",width=800,height=800,res=100)
P_period
dev.off()

#-----Figure 7----#

temp.data<-na.omit(ddply(res$split.data,.(sexf,Birth_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))

P_cohort<-ggplot(data = temp.data,aes(x = Birth_F, y = Rate*10^3,color=factor(sexf))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Birth Cohort") + ylab("MCI Rate Per 1000") +
        ggtitle("Birth Cohort-Specific MCI Incidence Rate by Gender") +
        theme_bw() +
        geom_hline(yintercept=seq(0,300,by=50),lty="dashed",colour="black") +
        theme(legend.justification=c(0.9,1),
                legend.position=c(0.9,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=12,face="bold"),
                legend.text=element_text(colour="black",size=12,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
                plot.title=element_text(colour="black",  size=16,face="bold"),
                axis.text.x = element_text(angle=45,hjust=1)

              ) +

  scale_colour_discrete(name = "",labels=c("Male","Female")) +
  scale_x_continuous(breaks=c(1895,1900,1905,1910,1915,1920,1925,1930,1935,1940),
                     labels=c("1895-99","1900-04","1905-09","1910-14","1915-19","1920-24","1925-29",
                "1930-34","1935-39","1940 +")) +
  scale_y_continuous(breaks=seq(0,300,by=50))

png("Figure1_5_Cohort-Specific MCI Incidence Rate by gender.png",width=800,height=800,res=100)
P_cohort
dev.off()

#-----Thesis Data   Part (III) interaction with education----------#
#-----Figure 9. Age-specific MCI incidence rate by education level----#

Lex.data<-Lex.data[which(Lex.data$age>=65),]
res<-
acpSplit(indata=Lex.data,age.breaks=seq(65,105,by=5),per.breaks=seq(1985,2020,by=5),birth.breaks=seq(1895,1945,by
=5))

temp.data<-na.omit(ddply(res$split.data,.(edugp,Age_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
temp.data<-temp.data[temp.data$Age_F<95,]
P_age<-ggplot(data = temp.data,aes(x = Age_F, y = Rate*10^3,color=factor(edugp))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Age") + ylab("MCI Rate Per 1000") +
        ggtitle("Age-Specific MCI Incidence Rate by Education Level") +
        theme_bw() +
        geom_hline(yintercept=seq(0,120,by=20),lty="dashed",colour="black") +
        theme(legend.justification=c(0,1),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
```

```
                                axis.title = element_text(colour="black",size=15,face="bold"),
                                axis.text = element_text(colour="black", size=12,face="bold"),
                                plot.title=element_text(colour="black",  size=16,face="bold")
                            ) +

  scale_colour_discrete(name = "",labels=c("<High School","High School",">High School")) +
  scale_x_continuous(breaks=c(65,70,75,80,85,90,95),
                     labels=c("65-69", "70-74","75-79","80-84","85-89","90-94","95-99")) +
  scale_y_continuous(breaks=seq(0,120,by=20))

png("Figure1_5_Age-Specific MCI Incidence Rate by education.png",width=800,height=800,res=100)
P_age
dev.off()

#----Figure 10. Period-specific MCI incidence rate by education levels---#

temp.data<-na.omit(ddply(res$split.data,.(edugp,Per_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))

P_period<-ggplot(data = temp.data,aes(x = Per_F, y = Rate*10^3,color=factor(edugp))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Period") + ylab("MCI Rate Per 1000") +
        ggtitle("Period-Specific MCI Incidence Rate by Education Level") + theme_bw() +
        geom_hline(yintercept=seq(0,150,by=30),lty="dashed",colour="black") +
        theme(legend.justification=c(0,0.9),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
                plot.title=element_text(colour="black",  size=16,face="bold"),
                axis.text.x = element_text(angle=45,hjust=1)

                ) +

  scale_colour_discrete(name = "",labels=c("<High School","High School",">High School")) +
  scale_x_continuous(breaks=c(1985,1990,1995,2000,2005,2010,2015),
                     labels=c("1985-89","1990-94","1995-99","2000-04","2005-09","2010-14","2015 +")) +
  scale_y_continuous(breaks=seq(0,150,by=30))
png("Figure1_5_Period-Specific MCI Incidence Rate by education.png",width=800,height=800,res=100)
P_period
dev.off()

#---Figure 11. Cohort-specific MCI incidence rate by education levels---#

temp.data<-na.omit(ddply(res$split.data,.(edugp,Birth_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))

P_cohort<-ggplot(data = temp.data,aes(x = Birth_F, y = Rate*10^3,color=factor(edugp))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Birth Cohort") + ylab("MCI Rate Per 1000") +
        ggtitle("Birth Cohort-Specific MCI Incidence Rate by Education Level") +
        theme_bw() +
        geom_hline(yintercept=seq(0,250,by=50),lty="dashed",colour="black") +
        theme(legend.justification=c(0.9,1),
            legend.position=c(0.9,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=12,face="bold"),
                legend.text=element_text(colour="black",size=12,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                axis.text = element_text(colour="black", size=12,face="bold"),
```

```
                                plot.title=element_text(colour="black",  size=16,face="bold"),
                                axis.text.x = element_text(angle=45,hjust=1)

                            ) +

  scale_colour_discrete(name = "",labels=c("<High School","High School",">High School")) +
  scale_x_continuous(breaks=c(1895,1900,1905,1910,1915,1920,1925,1930,1935,1940),
                                    labels=c("1895-99","1900-04","1905-09","1910-14","1915-19","1920-24","1925-29",
                        "1930-34","1935-39","1940 +")) +
  scale_y_continuous(breaks=seq(0,250,by=50))

png("Figure1_5_Cohort-Specific MCI Incidence Rate by education.png",width=800,height=800,res=100)
P_cohort
dev.off()

#----Thesis Data   Part (IV) Interaction with Allele----#
#-----Figure 12. Age-specific MCI incidence rate by APOE4------#

Lex.data<-Lex.data[which(Lex.data$age>=65),]
res<-
acpSplit(indata=Lex.data,age.breaks=seq(65,105,by=5),per.breaks=seq(1985,2020,by=5),birth.breaks=seq(1895,1945,by
=5))

temp.data<-na.omit(ddply(res$split.data,.(allele,Age_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))
temp.data<-temp.data[temp.data$Age_F<100,]
P_age<-ggplot(data = temp.data,aes(x = Age_F, y = Rate*10^3,color=factor(allele))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Age") + ylab("MCI Rate Per 1000") +
        ggtitle("Age-Specific MCI Incidence Rate by Allele") +
        theme_bw() +
        geom_hline(yintercept=seq(0,120,by=20),lty="dashed",colour="black") +
        theme(legend.justification=c(0,1),
                legend.position=c(0,1),
                legend.background=element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.line = element_line(colour = "black",size=1),
                legend.title=element_text(colour="blue",size=10,face="bold"),
                legend.text=element_text(colour="black",size=10,face="bold"),
                legend.key=element_blank(),
                axis.title = element_text(colour="black",size=15,face="bold"),
                        axis.text = element_text(colour="black", size=12,face="bold"),
                        plot.title=element_text(colour="black",  size=16,face="bold")) +
        scale_colour_discrete(name = "",labels=c("non-APOE4","APOE*4")) +
        scale_x_continuous(breaks=c(65,70,75,80,85,90,95),
                                labels=c("65-69", "70-74","75-79","80-84","85-89","90-94","95-99")) +
        scale_y_continuous(breaks=seq(0,120,by=20))

png("Figure1_5_Age-Specific MCI Incidence Rate by allele.png",width=800,height=800,res=100)
P_age
dev.off()

#-----Figure 13. Period-specific MCI incidence rate by APOE4------#

temp.data<-na.omit(ddply(res$split.data,.(allele,Per_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))

P_period<-ggplot(data = temp.data,aes(x = Per_F, y = Rate*10^3,color=factor(allele))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Period") + ylab("MCI Rate Per 1000") +
        ggtitle("Period-Specific MCI Incidence Rate by Allele") +
        theme_bw() +
        geom_hline(yintercept=seq(0,80,by=20),lty="dashed",colour="black") +
        theme(legend.justification=c(0,0.9),
                        legend.position=c(0,1),
                        legend.background=element_blank(),
                        panel.grid.major = element_blank(),
                        panel.grid.minor = element_blank(),
                        axis.line = element_line(colour = "black",size=1),
```

```
                                    legend.title=element_text(colour="blue",size=10,face="bold"),
                                    legend.text=element_text(colour="black",size=10,face="bold"),
                                    legend.key=element_blank(),
                                    axis.title = element_text(colour="black",size=15,face="bold"),
                                    axis.text = element_text(colour="black", size=12,face="bold"),
                                    plot.title=element_text(colour="black",  size=16,face="bold")) +

  scale_colour_discrete(name = "",labels=c("non-APOE4","APOE*4")) +
  scale_x_continuous(breaks=c(1985,1990,1995,2000,2005,2010,2015),
                     labels=c("1985-89","1990-94","1995-99","2000-04","2005-09","2010-14","2015 +")) +
  scale_y_continuous(breaks=seq(0,80,by=20))

png("Figure1_5_Period-Specific MCI Incidence Rate by allele.png",width=800,height=800,res=100)
P_period
dev.off()

#-----Figure 14. Cohort-specific MCI incidence rate by APOE4------#

temp.data<-na.omit(ddply(res$split.data,.(allele,Birth_F),summarize,Rate=sum(lex.Xst)/sum(lex.dur)))

P_cohort<-ggplot(data = temp.data,aes(x = Birth_F, y = Rate*10^3,color=factor(allele))) +
        geom_line(lwd=1.5) +
        geom_point(size=3) +
        xlab("Birth Cohort") + ylab("MCI Rate Per 1000") +
        ggtitle("Birth Cohort-Specific MCI Incidence Rate by Allele") +
        theme_bw() +
        geom_hline(yintercept=seq(0,160,by=20),lty="dashed",colour="black") +
        theme(legend.justification=c(0.9,1),
                    legend.position=c(0.9,1),
                    legend.background=element_blank(),
                    panel.grid.major = element_blank(),
                    panel.grid.minor = element_blank(),
                    axis.line = element_line(colour = "black",size=1),
                    legend.title=element_text(colour="blue",size=12,face="bold"),
                    legend.text=element_text(colour="black",size=12,face="bold"),
                    legend.key=element_blank(),
                    axis.title = element_text(colour="black",size=15,face="bold"),
                    axis.text = element_text(colour="black", size=12,face="bold"),
                    plot.title=element_text(colour="black",  size=16,face="bold")) +

  scale_colour_discrete(name = "",labels=c("non-APOE4","APOE*4")) +
  scale_x_continuous(breaks=c(1895,1900,1905,1910,1915,1920,1925,1930,1935,1940),
                                    labels=c("1895-99","1900-04","1905-09","1910-14","1915-19","1920-24","1925-29",
                     "1930-34","1935-39","1940 +")) +
  scale_y_continuous(breaks=seq(0,160,by=20))


png("Figure1_5_Cohort-Specific MCI Incidence Rate by allele.png",width=800,height=800,res=100)
P_cohort
dev.off()
```

APPENDIX B: SUPPLEMENTARY TABLES

Table 21. Supplementary Table 1. Data Information

Variable name in the dataset	interpretation	Unit or level
Research ID	ID of participants	Integer, 1, 2…
Age1	Age at baseline	year
Study	Data collected from which study	1 = MoVIES 2 = MYHAT
Mcesdscore	Depression scale	0-??
Smokeyr (at baseline)	The smoking history at baseline	1 = yes 0 = no
Birthyr	Year of birth	Calendar year
DTBIR	Date of birth	date
APOE*4 allele		1 = ApoE4 Positive 0 = ApoE4 Negative
Sexf	gender	1 = female 0 = male
Edugp	Education Level	0 means less than HS 1 = HS 2 > HS
Case	Reach CDR=0.5	1 = case 0 = non-case
Intdt1	First interview date	Calendar date
LTFUDT	last follow-up date	Calendar date

Table 22. Supplementary Table 2. Information of MoVIES and MYHAT Study Population

	MoVIES	MYHAT
Age at study entry in years	>=65	>=65
Time of entering the study in calendar year	1987-1989	2006-2008
Follow-up period in years	Over 15 years	Over 10 years
Follow-up frequency	Every 2 years	Annually
Recruited participants	1,608	1,413

Table 23. Supplementary Table 3. Poisson log-linear model with age effects by gender

	Estimate	Std. Error	z value	Pr(>|z|)	*P-value*
(Intercept)	-5.166	0.707	-7.308	0.000	
A70-74	0.582	0.756	0.770	0.442	
A75-79	1.322	0.724	1.828	0.068	
A80-84	2.061	0.717	2.873	0.004	
A85-89	2.527	0.719	3.517	0.000	
A90-94	2.773	0.731	3.794	0.000	
A95-99	3.065	0.782	3.921	0.000	
sexf	1.043	0.790	1.320	0.187	
A70:sexf	-1.155	0.864	-1.338	0.181	
A75:sexf	-1.153	0.815	-1.415	0.157	
A80:sexf	-1.020	0.805	-1.267	0.205	
A85:sexf	-0.984	0.806	-1.221	0.222	
A90:sexf	-0.829	0.820	-1.011	0.312	
A95:sexf	-0.862	0.893	-0.965	0.335	0.75

*P-value by LRT for overall all the interaction terms

Table 24. Supplementary Table 3.1 Poisson log-linear model with age and gender effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.467	0.320	-13.957	0.000
A70-74	-0.218	0.360	-0.605	0.545
A75-79	0.522	0.331	1.578	0.115
A80-84	1.343	0.324	4.140	0.000
A85-89	1.833	0.325	5.639	0.000
A90-94	2.188	0.331	6.607	0.000
A95-99	2.446	0.374	6.538	0.000
sexf	0.052	0.083	0.627	0.531

Table 25. Supplementary Table 4. Poisson log-linear model with period effects by gender

	Estimate	Std. Error	z value	Pr(>|z|)	*P-value
(Intercept)	-4.200	0.267	-15.717	0.000	
P1990-1994	0.721	0.294	2.450	0.014	
P1995-1999	0.685	0.316	2.166	0.030	
P2005-2009	1.530	0.293	5.213	0.000	
P2010-2014	0.829	0.313	2.653	0.008	
P2015+	0.657	0.756	0.869	0.385	
sexf	-0.053	0.356	-0.149	0.882	
P1990:sexf	-0.204	0.392	-0.519	0.603	
P1995:sexf	0.234	0.410	0.572	0.567	
P2005:sexf	0.214	0.386	0.556	0.578	
P2010:sexf	0.320	0.405	0.789	0.430	
P2015:sexf	0.696	0.859	0.810	0.418	0.270

*P-value by LRT for overall all the interaction terms

Table 26. Supplementary Table 4.1. Poisson log-linear model with period and gender effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.275	0.183	-23.300	0.000
P1990-1994	0.594	0.195	3.053	0.002
P1995-1999	0.832	0.200	4.170	0.000
P2005-2009	1.661	0.190	8.748	0.000
P2010-2014	1.035	0.198	5.237	0.000
P2015+	1.169	0.350	3.342	0.001
sexf	0.076	0.083	0.916	0.360

Table 27. Supplementary Table 5. Poisson log-linear model with cohort effects by gender*

	Estimate	Std. Error	z value	Pr(>|z|)	p-value
(Intercept)	-16.039	773.784	-0.021	0.983	
C1900-1904	13.854	773.784	0.018	0.986	
C1905-1909	13.317	773.784	0.017	0.986	
C1910-1914	12.884	773.784	0.017	0.987	
C1915-1919	12.370	773.784	0.016	0.987	
C1920-1924	12.466	773.784	0.016	0.987	
C1925-1929	13.308	773.784	0.017	0.986	
C1930-1934	12.767	773.784	0.016	0.987	
C1935-1939	12.034	773.784	0.016	0.988	
C1940+	11.411	773.784	0.015	0.988	
sexf	15.517	773.784	0.020	0.984	
C1900:sexf	-16.330	773.784	-0.021	0.983	
C1905:sexf	-15.499	773.784	-0.020	0.984	
C1910:sexf	-15.641	773.784	-0.020	0.984	
C1915:sexf	-15.406	773.784	-0.020	0.984	
C1920:sexf	-15.431	773.784	-0.020	0.984	
C1925:sexf	-15.156	773.784	-0.020	0.984	
C1930:sexf	-15.536	773.784	-0.020	0.984	
C1935:sexf	-15.407	773.784	-0.020	0.984	
C1940:sexf	-14.758	773.784	-0.019	0.985	0.31

*Model fitting does not converge

Table 28. Supplementary Table 5.1. Poisson log-linear model with cohort and gender effects

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.547	0.708	-2.185	0.029
C1900-1904	-1.275	0.754	-1.692	0.091
C1905-1909	-1.228	0.719	-1.709	0.088
C1910-1914	-1.752	0.715	-2.450	0.014
C1915-1919	-2.113	0.713	-2.964	0.003
C1920-1924	-2.033	0.713	-2.852	0.004
C1925-1929	-1.008	0.714	-1.412	0.158
C1930-1934	-1.805	0.723	-2.497	0.013
C1935-1939	-2.449	0.731	-3.348	0.001
C1940+	-2.602	0.756	-3.441	0.001
sexf	0.096	0.085	1.137	0.255

Table 29. Supplementary Table 6. Poisson log-linear model with age effects by education

	Estimate	Std. Error	z value	Pr(>|z|)	*p-value
(Intercept)	-3.996	0.577	-6.921	0.000	
A70-74	-0.375	0.658	-0.570	0.569	
A75-79	-0.055	0.611	-0.091	0.928	
A80-84	1.215	0.588	2.068	0.039	
A85-89	1.404	0.592	2.373	0.018	
A90-94	1.882	0.600	3.138	0.002	
A95-99	1.570	0.707	2.220	0.026	
Edugp1	-0.467	0.764	-0.611	0.541	
Edugp2	-0.733	0.816	-0.898	0.369	
A70:edugp1	0.180	0.869	0.207	0.836	
A75:edugp1	0.601	0.804	0.748	0.454	
A80:edugp1	0.179	0.780	0.230	0.818	
A85:edugp1	0.567	0.785	0.723	0.470	
A90:edugp1	0.335	0.799	0.420	0.675	
A95:edugp1	1.069	0.917	1.166	0.244	
A70:edugp2	0.233	0.931	0.251	0.802	
A75:edugp2	0.980	0.856	1.145	0.252	
A80:edugp2	-0.086	0.840	-0.102	0.919	
A85:edugp2	0.604	0.838	0.721	0.471	
A90:edugp2	0.467	0.852	0.548	0.584	
A95:edugp2	1.396	0.979	1.426	0.154	0.130

*P-value by LRT for overall all the interaction terms
Edugp1: Participants receiving HS education; Edugp2: Participants receiving higher than HS education

Table 30. Supplementary Table 7. Poisson log-linear model with period effects by education

	Estimate	Std. Error	z value	Pr(>|z|)	*p-value
(Intercept)	-3.823	0.218	-17.519	0.000	
P1990-1904	0.565	0.244	2.319	0.020	
P1995-1999	0.727	0.259	2.813	0.005	
P2005-2009	1.825	0.273	6.681	0.000	
P2010-2014	1.442	0.302	4.778	0.000	
P2015+	1.683	1.024	1.644	0.100	
edugp1	-1.388	0.546	-2.544	0.011	
edugp2	-0.469	0.436	-1.075	0.282	
P1990:edugp1	0.664	0.579	1.146	0.252	
P1995:edugp1	0.847	0.588	1.441	0.150	
P2005:edugp1	0.881	0.579	1.522	0.128	
P2010:edugp1	0.620	0.600	1.034	0.301	
P2015:edugp1	0.168	1.277	0.131	0.896	
P1990:edugp2	-0.361	0.492	-0.734	0.463	
P1995:edugp2	0.018	0.495	0.036	0.971	
P2005:edugp2	-0.444	0.482	-0.920	0.357	
P2010:edugp2	-0.751	0.508	-1.479	0.139	
P2015:edugp2	-0.527	1.179	-0.447	0.655	0.360

*P-value by LRT for overall all the interaction terms

Table 31. Supplementary Table 8. Poisson log-linear model with cohort effects by education

	Estimate	Std. Error	z value	Pr(>|z|)	*p-value
(Intercept)	-1.395	1.000	-1.395	0.163	
C1900-1904	-1.586	1.069	-1.484	0.138	
C1905-1909	-1.240	1.012	-1.225	0.221	
C1910-1914	-1.580	1.008	-1.567	0.117	
C1915-1919	-2.031	1.010	-2.011	0.044	
C1920-1924	-2.193	1.018	-2.155	0.031	
C1925-1929	-0.606	1.031	-0.588	0.557	
C1930-1934	-1.270	1.069	-1.188	0.235	
C1935-1939	-2.540	1.225	-2.074	0.038	
C1940+	-0.419	1.118	-0.375	0.708	
edugp1	-2.148	0.671	-3.202	0.001	
edugp2	-0.217	1.414	-0.154	0.878	
C1900:edugp1	2.713	0.918	2.955	0.003	
C1905:edugp1	1.794	0.800	2.241	0.025	
C1910:edugp1	1.920	0.717	2.678	0.007	
C1915:edugp1	1.807	0.701	2.577	0.010	
C1920:edugp1	2.235	0.709	3.153	0.002	
C1925:edugp1	1.654	0.728	2.272	0.023	
C1930:edugp1	1.766	0.797	2.217	0.027	
C1935:edugp1	2.208	1.013	2.179	0.029	
C1900:edugp2	0.655	1.547	0.424	0.672	
C1905:edugp2	0.066	1.446	0.046	0.964	
C1910:edugp2	-0.944	1.455	-0.649	0.516	
C1915:edugp2	0.041	1.434	0.029	0.977	
C1920:edugp2	0.297	1.435	0.207	0.836	
C1925:edugp2	-0.426	1.445	-0.295	0.768	
C1930:edugp2	-0.844	1.486	-0.568	0.570	
C1935:edugp2	0.174	1.604	0.109	0.913	
C1940:edugp2	-2.662	1.565	-1.701	0.089	0.048

*P-value by LRT for overall all the interaction terms

Table 32. Supplementary Table 9. Poisson log-linear model with age effects by ApoE4

	Estimate	Std. Error	z value	Pr(>|z|)	*P-value
(Intercept)	-4.058	0.333	-12.173	0.000	
A70-74	-0.957	0.422	-2.269	0.023	
A75-79	0.012	0.357	0.034	0.973	
A80-84	0.905	0.346	2.618	0.009	
A85-89	1.387	0.346	4.002	0.000	
A90-94	1.774	0.354	5.014	0.000	
A95-99	2.037	0.405	5.034	0.000	
Allele	-0.986	1.054	-0.936	0.350	
A70:Allele	1.748	1.135	1.540	0.124	
A75:Allele	1.396	1.081	1.292	0.196	
A80:Allele	1.169	1.071	1.091	0.275	
A85:Allele	1.187	1.073	1.107	0.268	
A90:Allele	1.277	1.085	1.177	0.239	
A95:Allele	1.004	1.290	0.779	0.436	0.660

*P-value by LRT for overall all the interaction terms

Table 33. Supplementary Table 10. Poisson log-linear model with period effects by ApoE4

	Estimate	Std. Error	z value	Pr(>|z|)	*p-value
(Intercept)	-4.832	0.354	-13.668	0.000	
P1990-1994	0.968	0.373	2.594	0.009	
P1995-1999	1.412	0.370	3.815	0.000	
P2005-2009	2.222	0.363	6.120	0.000	
P2010-2014	1.522	0.370	4.119	0.000	
P2015+	1.881	0.486	3.871	0.000	
Allele	-0.128	0.790	-0.161	0.872	
P1990:Allele	0.257	0.829	0.311	0.756	
P1995:Allele	0.151	0.825	0.183	0.855	
P2005:Allele	0.208	0.810	0.257	0.797	
P2010:Allele	0.669	0.815	0.821	0.412	
P2015:Allele	0.037	1.112	0.034	0.973	0.510

*P-value by LRT for overall all the interaction terms

Table 34. Supplementary Table 11. Poisson log-linear model with cohort effects by ApoE4

	Estimate	Std. Error	z value	Pr(>|z|)	*P-value
(Intercept)	-2.432	1.000	-2.432	0.015	
C1900-1904	-0.465	1.118	-0.416	0.677	
C1905-1909	-0.303	1.017	-0.298	0.766	
C1910-1914	-0.649	1.008	-0.644	0.520	
C1915-1919	-1.113	1.006	-1.106	0.269	
C1920-1924	-1.110	1.006	-1.104	0.270	
C1925-1929	-0.164	1.007	-0.163	0.871	
C1930-1934	-0.789	1.015	-0.777	0.437	
C1935-1939	-1.833	1.033	-1.775	0.076	
C1940+	-1.732	1.049	-1.651	0.099	
Allele	0.211	0.658	0.320	0.749	
C1900:Allele	0.342	1.008	0.339	0.735	
C1905:Allele	-0.084	0.769	-0.109	0.913	
C1910:Allele	-0.844	0.769	-1.097	0.273	
C1915:Allele	-0.096	0.695	-0.138	0.890	
C1920:Allele	-0.051	0.694	-0.074	0.941	
C1925:Allele	0.302	0.692	0.436	0.663	
C1930:Allele	-0.223	0.750	-0.298	0.766	
C1935:Allele	0.654	0.775	0.844	0.398	0.220

*P-value by LRT for overall all the interaction terms

Table 35. Supplementary Table 12. Poisson log-linear model with age effects by study

	Estimate	Std. Error	z value	Pr(>|z|)	*P-value
(Intercept)	-6.280	1.000	-6.280	0.000	
A70-74	1.347	1.027	1.311	0.190	
A75-79	2.183	1.008	2.166	0.030	
A80-84	3.016	1.004	3.002	0.003	
A85-89	3.451	1.007	3.427	0.001	
A90-94	3.834	1.017	3.772	0.000	
A95-99	4.135	1.080	3.829	0.000	
Study	2.737	1.054	2.597	0.009	
A70:Study	-2.003	1.109	-1.806	0.071	
A75:Study	-2.164	1.073	-2.017	0.044	
A80:Study	-2.268	1.064	-2.131	0.033	
A85:Study	-2.327	1.065	-2.185	0.029	
A90:Study	-2.441	1.076	-2.268	0.023	
A95:Study	-2.528	1.153	-2.192	0.028	0.130

*P-value by LRT for overall all the interaction terms

BIBLIOGRAPHY

[1] Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimer's & dementia : the journal of the Alzheimer's Association 2013;9:63-75 e2.
[2] Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Global prevalence of dementia: a Delphi consensus study. Lancet 2005;366:2112-7.
[3] Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, et al. Prevalence of cognitive impairment without dementia in the United States. Annals of internal medicine 2008;148:427-34.
[4] Petersen RC. Challenges of epidemiological studies of mild cognitive impairment. Alzheimer disease and associated disorders 2004;18:1-2.
[5] Petersen RC. Mild cognitive impairment as a diagnostic entity. Journal of internal medicine 2004;256:183-94.
[6] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Archives of neurology 1999;56:303-8.
[7] Perquin M, Schuller AM, Vaillant M, Diederich N, Bisdorff A, Leners JC, et al. The epidemiology of mild cognitive impairment (MCI) and Alzheimer's disease (AD) in community-living seniors: protocol of the MemoVie cohort study, Luxembourg. BMC public health 2012;12:519.
[8] Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clinics in geriatric medicine 2013;29:737-52.
[9] Williams JD, Klug MG. Aging and cognition: methodological differences in outcome. Experimental aging research 1996;22:219-44.
[10] Hedden T, Gabrieli JD. Insights into the ageing mind: a view from cognitive neuroscience. Nature reviews Neuroscience 2004;5:87-96.
[11] Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971-2006. Social science & medicine 2010;70:1100-8.
[12] Robertson C, Boyle P. Modeling temporal aspects of an exposure. Epidemiology 1998;9:361-2; author reply 2-3.
[13] Agarwal R, Tripathi CB. Association of apolipoprotein E genetic variation in Alzheimer's disease in Indian population: a meta-analysis. American journal of Alzheimer's disease and other dementias 2014;29:575-82.
[14] Fan YY, Cai QL, Gao ZY, Lin X, Huang Q, Tang W, et al. APOE epsilon4 allele elevates the expressions of inflammatory factors and promotes Alzheimer's disease progression: a comparative study based on Han and She populations in the Wenzhou area. Brain research bulletin 2017.

[15] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. Jama 1997;278:1349-56.
[16] Scarabino D, Broggio E, Gambina G, Maida C, Gaudio MR, Corbo RM. Apolipoprotein E genotypes and plasma levels in mild cognitive impairment conversion to Alzheimer's disease: A follow-up study. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics 2016;171:1131-8.
[17] Robertson C, Boyle P. Age-period-cohort models of chronic disease rates. II: Graphical approaches. Statistics in medicine 1998;17:1325-39.
[18] Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: Age-period-cohort models. Statistics in medicine 1987;6:469-81.
[19] Mason KO, Mason WM, Winsborough HH, Poole WK. Some methodological issues in cohort analysis of archival data. American sociological review 1973:242-58.
[20] Rosenbauer J, Strassburger K. Comments on 'age-period-cohort models for the Lexis diagram' by Carstensen B. Statistics in Medicine 2007; 26:3018-3045. Statistics in medicine 2008;27:1557-61; author reply 61-4.
[21] Carstensen B. Age-period-cohort models for the Lexis diagram. Statistics in medicine 2007;26:3018-45.
[22] Mason W, Fienberg S. Cohort Analysis in Social Research: Beyond the Identification Problem. New York: Springer-Verlag 1985.
[23] Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: Age-period and age-cohort models. Statistics in medicine 1987;6:449-67.
[24] Yang Y, Schulhofer-Wohl S, Fu WJJ, Land KC. The intrinsic estimator for age-period-cohort analysis: What it is and how to use it. Am J Sociol 2008;113:1697-736.
[25] Holford TR. The estimation of age, period and cohort effects for vital rates. Biometrics 1983;39:311-24.
[26] Fienberg SE, Mason WM. Specifi cation and Implementation of Age, Period, and Cohort Models. New York: Springer-Verlag 1985:45-88.
[27] Kupper LL, Janis JM, Karmous A, Greenberg BG. Statistical age-period-cohort analysis: a review and critique. Journal of chronic diseases 1985;38:811-30.
[28] Glenn, ND. Cohort analysis. 2. Thousand Oaks, CA: Sage Publications Inc 2005.
[29] Holford, TR. Understanding the Effects of Age, Period, and Cohort on Incidence and Mortality Rates. Annual Reviews in Public Health 1991;12:425-57.
[30] Utz R. Obesity in America, 1960–2000: Is it an Age, Period, or Cohort Phenomenon? Population Association of America; Philadelphia, PA 2005.
[31] Stern Y. Cognitive reserve. Neuropsychologia 2009;47:2015-28.
[32] Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. Journal of the International Neuropsychological Society : JINS 2002;8:448-60.
[33] Bosma H, van Boxtel MP, Ponds RW, Houx PJ, Burdorf A, Jolles J. Mental work demands protect against cognitive impairment: MAAS prospective cohort study. Experimental aging research 2003;29:33-45.
[34] Zahodne LB, Glymour MM, Sparks C, Bontempo D, Dixon RA, MacDonald SW, et al. Education does not slow cognitive decline with aging: 12-year evidence from the victoria longitudinal study. Journal of the International Neuropsychological Society : JINS 2011;17:1039-46.
[35] Reitz C, Brayne C, Mayeux R. Epidemiology of Alzheimer disease. Nature reviews Neurology 2011;7:137-52.

[36] Wecker NS, Kramer JH, Hallam BJ, Delis DC. Mental flexibility: age effects on switching. Neuropsychology 2005;19:345-52.
[37] Wecker NS, Kramer JH, Wisniewski A, Delis DC, Kaplan E. Age effects on executive ability. Neuropsychology 2000;14:409-14.
[38] Sakurai R, Montero-Odasso M. Apolipoprotein E4 Allele and Gait Performance in Mild Cognitive Impairment: Results From the Gait and Brain Study. The journals of gerontology Series A, Biological sciences and medical sciences 2017.
[39] Doi T, Shimada H, Makizako H, Tsutsumimoto K, Uemura K, Suzuki T. Apolipoprotein E genotype and physical function among older people with mild cognitive impairment. Geriatrics & gerontology international 2015;15:422-7.