Experimental Gerontology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Experimental Gerontology

journal homepage: www.elsevier.com/locate/expgero

Causal effects of cardiovascular risk factors on onset of major age-related diseases: A time-to-event Mendelian randomization study

Liang He*, Irina Culminskaya, Yury Loika, Konstantin G. Arbeev, Olivia Bagley, Matt Duan, Anatoliy I. Yashin, Alexander M. Kulminski*

Biodemography of Aging Research Unit, Social Science Research Institute, Duke University, Durham, NC 27708, USA

ARTICLE INFO

Keywords:
Mendelian randomization
Body mass index
Causal effect
Age-related disease
Time-to-event analysis

ABSTRACT

Backgrounds: Elucidating the causal effects of common intermediate risk factors on the onset of age-related diseases is indispensable for developing prevention and intervention procedures.

Methods: We conducted two-stage time-to-event Mendelian randomization meta-analyses combining five large-scale longitudinal cohorts to investigate dynamic causal effects of cardiovascular disease risk factors including body mass index (BMI), systolic blood pressure (SBP), and lipids on the age-at-onset of age-related diseases. We constructed weighted polygenic scores based on genetic markers from previously reported genome-wide association studies as instrumental variables to estimate the causal effects. To avoid false positive due to potential pleiotropic effects of the genetic markers, we performed a leave-one-out sensitivity analysis and an MR-Egger sensitivity analysis that we expanded in the survival context.

Results: Our results show that elevated BMI increases the absolute risk of type 2 diabetes (T2D) ($p = 7.68e-04$), heart failure ($p = 9.03e-03$), and cardiovascular diseases (CVD) ($p = 1.69e-03$) and the causal effects start at different ages. A significant association between BMI and the risk of stroke is observed; however, the sensitivity analyses suggest that the association is attributed to the potential pleiotropic effects of rs2867125 and rs1558902. Raised SBP levels are significantly associated with the development of atrial fibrillation ($p = 6.42e-03$). Low-density lipoprotein cholesterol (LDL-C) levels are inversely associated with the age-at-onset of T2D ($p = 1.05e-02$). In addition, LDL-C and triglycerides are inversely associated with the risks of cancer and T2D, respectively. Nevertheless, the sensitivity analyses suggest that these associations are probably due to pleiotropic effects of several single-nucleotide polymorphisms including rs4970834 and rs1260326.

Conclusions: Our results highlight the involvement of BMI in the development of multiple age-related diseases. Some observed causal associations can attribute to pleiotropic effects of some genetic variations. These findings have important implications in unraveling causal effects of common risk factors on age-related diseases and guiding effective intervention strategies to reduce the incidence of these diseases.

1. Introduction

Common aging-associated diseases including cancer, various cardiovascular diseases (CVDs), type 2 diabetes (T2D) and neurodegenerative diseases (NDs) such as Alzheimer's disease (AD) are the leading causes of death and major contributors to morbidity, disability, and mortality at old age (Sahyoun et al., 2001). Evidence has shown that most of these diseases are correlated with certain common intermediate risk factors such as body mass index (BMI), blood pressure (BP) and lipids. A fundamental hypothesis in gerontology is that the biological aging process that leads to physical dysfunction and deviation from normal physiological indices and levels of some crucial biomarkers would be implicated in the onset and progression of multiple diseases (Arbeev et al., 2011; Kaeberlein et al., 2015). Nevertheless, it is still elusive whether the deviation from the optimal values of the biomarkers directly results in onset of these diseases or they are concurrent ramifications of certain more fundamental molecular mechanisms underlying the biological aging process, i.e., pleiotropy. Thus, elucidating the potential causal relationship between the intermediate risk factors and the onset of the age-related diseases can contribute to the development of effective intervention procedure and management policy that may lead to remarkable improvement of human healthspan and lifespan. It can also provide more insights into the underlying biological implications in the aging process and pinpointing potential

* Corresponding authors at: Erwin Mill Building, 2024 W. Main St., Durham, NC 27705, USA.
E-mail addresses: lianghe@mit.edu (L. He), kulminsk@duke.edu (A.M. Kulminski).

http://dx.doi.org/10.1016/j.exger.2017.09.019
Received 15 March 2017; Received in revised form 1 September 2017; Accepted 26 September 2017

physiological determinants of these diseases.

Despite the evidence of association, it is still unclear whether some of these biomarkers play a causal role in the etiology of these diseases and, more interestingly, how the causal effects vary with ageing if they exist. This is because inferring causation from general association analyses is often not straightforward as it is complicated by potentially unmeasured confounders and reverse causality (Pearl, 2000). One of the feasible approaches for causal inference is to construct an instrumental variable (IV) under a parametric model. With the advent of sequencing and microarray techniques in molecular biology, causal effects can be evaluated with the Mendelian randomization (MR) analyses that leverage a large number of identified genetic variants from genome-wide association analyses (GWAS) as IVs (Davey Smith and Hemani, 2014; Didelez and Sheehan, 2007). Previous MR studies based on cross-sectional or longitudinal datasets have reported that BMI, BP and lipids may have causal effects on the incidence of multiple age-related diseases (Hägg et al., 2015; Holmes et al., 2014, 2015; Østergaard et al., 2015; Proitsi et al., 2014). However, to our knowledge, few MR studies have focused on the dynamic causal effects on the age-at-onset of these diseases.

Most recently, several methods using IVs based on the additive hazards model (Aalen, 1978) have been proposed (Li et al., 2015; Tchetgen Tchetgen et al., 2015), which are characterized by straightforward estimation of the dynamic effects over age and more intuitive interpretation of the estimated parameters. MR using the additive hazards model enjoys robustness due to collapsibility compared to proportional hazards model (Martinussen and Vansteelandt, 2013). Compared to the MR analyses in cross-sectional studies, MR analyses with time-to-event outcomes leverage more information from the longitudinal data, so it can provide an estimate of time-varying causal effects on the risk of diseases. The time-to-event MR analysis gives more information about the age at which the effects start and how long they persist. This is of importance because in many cases intermediate risk factors are measured at different ages across the sample at the baseline. In the survival context using the additive hazards model, the assumption of a constant causal effect at different ages, which is not always realistic, can be relaxed (Tchetgen Tchetgen et al., 2015). For example, large BMI at early stage of life may have different effects on T2D compared to middle age. Therefore, it is of enormous interest to investigate how the causal effects are changed and modulated by other environmental factors during the life course.

In this work, we perform time-to-event MR analyses to investigate the causal effects of five cardiovascular disease risk factors including BMI, systolic blood pressure (SBP), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C) and triglycerides on eight age-related diseases including T2D, coronary heart disease (CHD), heart failure (HF), myocardial infarction (MI), stroke, cancer, atrial fibrillation (AF) and AD. These diseases are major causes of death in developed societies. We include > 30,000 individuals from five NIH-funded large-scale longitudinal cohorts (Atherosclerosis Risk in Communities study (ARIC), the Framingham Heart Study (FHS), the Multi-Ethnic Study of Atherosclerosis (MESA), the Cardiovascular Health Study (CHS), and the Women's Health Initiative (WHI)). We perform a meta-analysis to combine the causal effects from each cohort, which we estimate based on an additive hazards model. We construct polygenic scores as IVs based on well recognized genetic loci with known mechanisms. The MR method based on polygenic scores has been shown to avoid weak instrument bias (Burgess and Thompson, 2013). As some factors such as pleiotropy and population stratification can violate assumptions in the MR analysis (Burgess et al., 2015, 2016; Taylor et al., 2014), we conduct various sensitivity analyses to minimize the possibility of reporting false positives.

2. Material and methods

2.1. Study cohorts

We collected datasets from five large-scale longitudinal cohorts including ARIC, FHS, MESA, CHS, and WHI (total 30,505 individuals included) (Table 1). In each cohort, we only included non-Hispanic Caucasian subjects. For FHS, we only included the datasets of the FHS Original cohort (FHS cohort 1) and the FHS Offspring cohort (FHS cohort 2) because the third generation was too young to observe sufficient informative outcomes. Details of the study design, collection of samples, measurement of risk factors and diagnosis of diseases have been described in previous publications (ARIC: (Sharrett, 1992; The ARIC investigators, 1989), FHS: (Cupples et al., 2009; Govindaraju et al., 2008; Splansky et al., 2007), MESA: (Bild et al., 2002), CHS: (Gottdiener et al., 2000), WHI: (The Women's Health Initiative Study Group, 1998)). Table 1 summarizes the basic characteristics of the samples, risk factors, covariates and diseases we investigated.

2.2. Phenotypes and covariates

In each cohort, we used the measurements of BMI, SBP, HDL-C, LDL-C and triglycerides at the entry of enrollment, so that the precedence of the factors to the diagnosis of diseases was guaranteed to minimize the possibility of reverse causation. For the disease outcomes, we included T2D, CHD, HF, MI, stroke and CVD in all five cohorts. The CVD variables were constructed separately in each cohort from CHD, HF, MI and stroke to harmonize the definition across the cohorts (The details of the construction of the CVD variable are provided in Table S9). The age at onset of T2D in WHI was reckoned from self-reported treatment history or the age at diagnosis that was categorized into age groups. We included cancer in ARIC, FHS, CHS and WHI, which was not included in MESA due to missing information on age at onset in the original data. In FHS and WHI, we used the cancer variable that excluded skin cancer because the information was either not available or inaccurate. Additionally, we included AF in ARIC and WHI, and AD in FHS and CHS in which we selected AD or mild dementia. In the sensitivity analyses adjusted for potential confounders (described later), we included sex, birth cohort, which were non-heritable variables, and further education

Table 1
The basic characteristics of the cohorts included in the MR analyses. Sample size: the total number including non-Caucasian whites. Diseases: the age-related diseases examined in each cohort. AF is only available in ARIC and WHI. AD is only available in FHS and CHS. Cancer is not included in MESA because its age-at-onset information is missing in the original data. Covariate: the covariates adjusted in the main MR analysis. 'Cohort' in FHS is the indicator for generation, i.e., FHS or FHSO. Other confounders: the covariates adjusted in the follow-up sensitivity analyses. BC: birth cohort. SM: smoking history (current smoker, ex-smoker, non-smoker). DR: alcohol consumption history (the concrete definition varies among the cohorts). EDU, education level (in FHS and CHS, the binary variable indicating low education level is used due to large missing data in education level).

Study	Sample size	Number of families	Male (%)	Average age at baseline (± sd)	Diseases	Covariate	Other confounder
ARIC	9810	9105	47.13	54.33 (5.67)	T2D, CHD, HF, MI, AF, cancer, stroke, CVD	Site	BC, Sex, SM, DR, EDU
FHS	4700	1482	45.28	35.91 (9.25)	T2D, CHD, HF, MI, cancer, stroke, AD, CVD	Cohort	BC, Sex, SM, DR, EDU
CHS	3310	3310	39.76	72.41 (5.41)	T2D, CHD, HF, MI, cancer, stroke, AD, CVD	–	BC, Sex, SM, DR, EDU
MESA	2685	2649	47.86	62.74 (10.16)	T2D, CHD, HF, MI, stroke, CVD	Site	BC, Sex, SM, EDU
WHI	10,000	9990	0	67.05 (6.45)	T2D, CHD, HF, MI, AF, cancer, stroke, CVD	Region	BC, SM, DR, EDU