Somatic and occupational outcomes in adult attention-deficit/hyperactivity disorder: epidemiology studies based on real-world data

To my beloved family

献给我最爱的爷爷奶奶

Örebro Studies in Medicine 257

LIN LI

李　琳

Somatic and occupational outcomes in adult attention-deficit/hyperactivity disorder: epidemiology studies based on real-world data

Title: Somatic and occupational outcomes in adult ADHD: epidemiology studies based on real-world data
Publisher: Örebro University 2022
www.oru.se/publikationer

Print: Örebro University, Repro 03/2022

ISSN 1652-4063
ISBN 978-91-7529-429-2

Abstract

Lin Li (2022): Somatic and occupational outcomes in adult ADHD: epidemiology studies based on real-world data. Örebro Studies in Medicine 257.

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common neurodevelopmental disorders, characterized by inattention or hyperactivity–impulsivity, or both. ADHD is a multifactorial disorder influenced by the complex interplay between genetic and environmental risk factors, but a detailed understanding of the causal status of these factors is lacking. ADHD is associated with many psychiatric disorders, but somatic comorbidity in ADHD has received less attention in the research literature. Pharmacological treatment is effective in reducing the core symptoms of ADHD, but the effects on occupational outcomes remain unclear. The overarching aim of this thesis is to extend previous knowledge on the early risk factors of ADHD, and to increase the awareness and the understanding on somatic and occupational outcomes of ADHD in adults.

In Study I, we combined a systematic review and mate-analysis with a population based cohort of 971,501 individuals born between 1992 and 2004 in Sweden. The meta-analysis revealed a positive association between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring. However, these associations gradually attenuated toward the null when adjusted for measured confounders, unmeasured factors shared by cousins and unmeasured factors shared by siblings. In Study II, by using a Swedish population-based twin study with 17,999 individuals aged 20–47 years, we found both inattention and hyperactivity/impulsivity was associated with higher consumption of high-sugar food and unhealthy dietary habits, although these associations were generally weak. Further, the observed associations was explained by both genetic and non-shared environmental factors. In Study III, we explored the prospective associations between ADHD and a broad range of cardiovascular diseases in 5,389,519 adults from Sweden, and found that ADHD may be a novel and independent risk factor for cardiovascular diseases. In Study IV, based on the longitudinal cohort of 12,875 middle-aged adults with ADHD, we found the use of ADHD medications during the previous two years was associated with a 10% reduction in the risk of long-term unemployment in the following year.

Taken together, findings from the thesis highlight the need of future studies with various study designs, to fully understand the aetiology and long-term health outcomes of ADHD across the lifespan.

Keywords: ADHD, Comorbidities, Obesity, Dietary habits, ADHD medication, Cardiovascular disease, Epidemiology, Casual inference, Adults

Lin Li, School of Medical Sciences, Medicine
Örebro University, SE-701 82 Örebro, Sweden, lin.li@oru.se.

LIST OF SCIENTIFIC PAPERS

I. **Li L,** Lagerberg T, Chang Z, Cortese S, Rosenqvist MA, Almqvist C, D'Onofrio BM, Hegvik TA, Hartman C, Chen Q, Larsson H. Maternal pre-pregnancy overweight/obesity and the risk of attention-deficit/hyperactivity disorder in offspring: a systematic review, meta-analysis and quasi-experimental family-based study. International journal of epidemiology. 2020 Jun 1;49(3):857-75.

II. **Li L,** Taylor MJ, Bälter K, Kuja-Halkola R, Chen Q, Hegvik TA, Tate AE, Chang Z, Arias-Vásquez A, Hartman CA, Larsson H. Attention-deficit/hyperactivity disorder symptoms and dietary habits in adulthood: A large population-based twin study in Sweden. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2020 Dec;183(8):475-85.

III. **Li L,** Chang Z, Sun J, Garcia-Argibay M, Rietz ED, Dobrosavljevic M, Brikell I, Jernberg T, Solmi M, Cortese S, Larsson H. Attention-deficit/hyperactivity disorder as a novel risk factor for cardiovascular diseases: a nationwide population-based cohort study. (Submitted)

IV. **Li L,** Chang Z, Sun J, Jangmo A, Zhang L, Andersson LM, Werner-Kiechle T, Ahnemark E, D'Onofrio BM, Larsson H. Association between Medication Use and Long-term Unemployment in Working-aged Individuals with Attention-Deficit/Hyperactivity Disorder. Accepted by JAMA Network Open.

RELATED PUBLICATIONS

(Not included in the thesis)

I. **Li L**, Sujan AC, Butwicka A, Chang Z, Cortese S, Quinn P, Viktorin A, Öberg AS, D'Onofrio BM, Larsson H. Associations of prescribed ADHD medication in pregnancy with pregnancy-related and offspring outcomes: a systematic review. CNS drugs. 2020 Jul;34(7):731-47.

II. **Li L**, Taylor M.J, Bälter K, Xie T, Solberg B.S, Haavik J, Arias Vásquez A, Hartman C.A, Larsson H. Gene-Environment Interactions in Attention-Deficit/Hyperactivity Disorder Symptom Dimensions: The Role of Unhealthy Food Habits. Genes. 2022; 13(1):47.

III. Schweren LJ, Larsson H, Vinke PC, **Li L**, Kvalvik LG, Arias-Vasquez A, Haavik J, Hartman CA. Diet quality, stress and common mental health problems: A cohort study of 121,008 adults. Clinical Nutrition. 2021 Mar 1;40(3):901-6.

IV. Kase BE, Rommelse N, Chen Q, **Li L**, Andersson A, Du Rietz E, Vos M, Cortese S, Larsson H, Hartman CA. Longitudinal Associations Between Symptoms of ADHD and BMI From Late Childhood to Early Adulthood. Pediatrics. 2021 Jun 1;147(6).

V. Schweren LJS, van Rooij D, Shi H, Larsson H, Arias-Vasquez A, **Li L**, Grimstvedt Kvalvik L, Haavik J, Buitelaar J, Hartman C. Diet, Physical Activity, and Disinhibition in Middle-Aged and Older Adults: A UK Biobank Study. Nutrients. 2021; 13(5):1607.

VI. Kvalvik LG, Klungsøyr K, Igland J, Caspersen GH, Brantsæter AL, Solberg BS, Hartman C, Schweren LJS, Larsson H, **Li L**, Forthun I, Johansson S, Arias-Vasquez A, Haavik J. Association of sweetened carbonated beverage consumption during pregnancy and ADHD symptoms in the offspring. A study from the Norwegian, Mother, Father and Child Cohort Study (MoBa). European Journal of Nutrition. 2022; 1-14.

LIST OF ABBREVIATION

ADHD	Attention-deficit/hyperactivity disorder
AIC	Akaike information criterion
ASD	Autism spectrum disorder
ATC	Anatomical Therapeutic Chemical classification system
BMI	Body mass index
CDR	The Causes of Death Register
CI	Confidence interval
CTCT	Cross-twin cross-trait correlation
CVD	Cardiovascular disease
DSM	The Diagnostic and Statistical Manual of Mental Disorders
DZ	Dizygotic twins
EHR	Electronic health record
FFQ	Food frequency questionnaire
GEE	Generalized estimating equations
HI	Hyperactivity/impulsivity
HKD	Hyperkinetic disorder
HR	Hazard ratio
IA	Inattention
ICC	Intra-class correlation
ICD	The International Classification of Diseases and Related Health Problems
LISA	Longitudinal integration database for health insurance and labor market studies register
MGBA	Microbiome–gut–brain axis
MGR	Multi-generation register
MR	Mendelian randomization
MZ	Monozygotic twins
NCR	The National Crime Register
NPR	National Patient Register
OR	Odds ratio
PDR	Prescribed Drug Register
PIN	Personal identity number
PRISMA	Preferred reporting items for systematic reviews and Meta-Analyses
RCT	Randomized controlled clinical trials
RR	Relative risk
RWD	Real-world data
SEM	Structural equation modelling
SES	Socioeconomic status
SMI	Serious mental illness
STAGE	Study of Twin Adults: Genes and Environment
STR	The Swedish Twin Registry
T2DM	Type 2 diabetes mellitus
TPR	Total Population Register

Table of Contents

1. INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common neurodevelopmental disorder, characterized by inattention or hyperactivity–impulsivity, or both.[1] With onset in early childhood, ADHD commonly persists into adulthood.[2, 3] The disorder is estimated to affect about 5% of school-age children and 2.5% of adults across the world.[4] The etiology is considered multifactorial with the complex interplay between many genetic and environmental factors. To date, no single factor or biomarker has been identified for screening the risk of the disorder or facilitating the diagnostic procedure.

Substantial evidence shows comorbidity between ADHD and psychiatric disorders such as conduct disorders,[5] substance use disorders,[6] mood and personality disorders,[7, 8] as well as other neurodevelopmental disorders such as autism spectrum disorders.[9] Throughout the lifetime, ADHD impairs psychosocial functioning, leading to academic and occupational difficulties, social disability and risky behaviors.[1, 10] Furthermore, increasing evidence suggests that people with ADHD have higher risk of developing a broad range of somatic diseases, including obesity, epilepsy, asthma, sleep disorder, infections and autoimmune diseases,[4] as well as unhealthy lifestyles,[11] such as unhealthy dietary habits.[12] However, few studies have explored how ADHD is associated with cardiovascular diseases (CVDs), the leading cause of death across the world.[13]

Given its high prevalence and associated impairments and adverse outcomes, ADHD is a major public health concern. Randomized controlled clinical trials (RCTs) have demonstrated efficacy of pharmacologic treatments in reducing the core symptoms of ADHD in adults. Observational pharmacoepidemiological studies have suggested beneficial effects of ADHD medications on behavioral and neuropsychiatric outcomes, including injuries and traumas, criminality, motor vehicle accidents, education and substance use disorder.[14] However, little is known about effects of pharmacological treatment for ADHD on long-term unemployment, which is associated with economic difficulties, worse mental and physical health, and higher mortality rates.[15]

Based on large-scale real-world data, this thesis seeks to extend previous knowledge about the risk factors and long-term consequences of ADHD in adults, by focusing on somatic and occupational outcomes.

2. BACKGROUND

2.1 Diagnosis and symptoms of ADHD

Attention-deficit/hyperactivity disorder (ADHD) 'syndromes' were first described in a German textbook in 1775 and was later recognized as a disorder in 1902 by George Frederic Still, a physician in the United Kingdom.[4] ADHD is diagnosed according to two parallel diagnostic systems: The Diagnostic and Statistical Manual of Mental Disorders (DSM),[16] which is predominantly used in the United States and The International Classification of Diseases and Related Health Problems (ICD), [17] which is predominantly used in Europe and other parts of the world. The DSM-5, published in 2013, allows for an ADHD diagnosis based on the presence of either inattentive or hyperactive/impulsive symptoms, together with impairments in social, academic, and/or occupational functioning caused by the symptoms. The DSM defines ADHD as three different subtypes: predominantly inattentive type, predominantly hyperactive/impulsive type and ADHD combined type.[16] It is important to note that DSM-V now uses the term "presentation", to acknowledge the heterogeneity of the behavioral manifestations of ADHD in different developmental stages.[18] In the ICD, ADHD is described as Hyperkinetic disorder (HKD). HKD requires the presence of both inattentive and hyperactive/impulsive symptoms across two or more settings (e.g. school and home) and tends to capture more severe cases.[17] In the ICD-11 (new version published in 2018), the name of the disorder has changed to "Attention deficit hyperactivity disorder", listing in the "Neurodevelopmental disorders" category. The description of ADHD in ICD-11 resembles DSM-V criteria and describes the same presentations as DSM-V [19].

Rating scales are often used as screening tools and supplementary support for the clinical interview. Although not accurate enough to estimate the onset age of ADHD, studies have reported that teacher-rated questionnaires are valuable in early identification and treatment decision making among children.[20] Rating scales for adults such as Conners' Adult ADHD Diagnostic Interview [21] and the Adult Self-Report Scale[22] have also been shown reliable in structured diagnostic interviews for both the patients and informants.

2.2 Prevalence

The prevalence of ADHD among children has been estimated to be around 2–7% worldwide, with an average prevalence of 5%.[23] Variation in the prevalence across studies could primarily be explained by methodological differences, including different diagnostic criteria (DSM vs. ICD), information sources (patient vs. informants), and requirement of functional impairments.[1, 24] In childhood, boys are significantly more likely to have an ADHD diagnosis than girls, with a male-to-female sex ratio of 4:1 in clinical samples and 2.4:1 in population-based samples.[25] These differences decrease with age.[26]

ADHD affects around 2.5% of adults in the world.[27-29] The prevalence for strictly applied operational definitions of ADHD declines with age. Therefore, the prevalence of ADHD among adults substantially depends on how ADHD persistence into adulthood is defined, following the childhood onset of the disorder. [30] Evidence from a systematic review using a thorough assessment of ADHD indicated that 60% of the ADHD group demonstrated symptom persistence and 41% met both symptom and impairment criteria in adulthood.[3] Despite the increased interest in ADHD in adults, more research is needed to understand better its manifestations and associated impairment in adults.[31, 32]

2.3 Management and treatment of ADHD

The European clinical guidelines for management of ADHD propose a multimodal approach, where psychological interventions, educational change, medication and dietary measurements are integrated.[33]

2.3.1 Pharmacological treatment of ADHD

It is well established that both non-stimulant and central stimulant medication effectively reduce the core symptoms of ADHD.[34] Stimulants are the front-line medications for ADHD, which include methylphenidate, dexmethylphenidate, mixed amphetamine salts, and lisdexamfetamine.[33] Non-stimulant medications (e.g., atomoxetine) have also been used to treat ADHD, although their efficacy seems to be slightly lower than that of stimulants.[35] Choice of medication depends on clinical severity and presentation of symptoms, as well as comorbid somatic or psychiatric conditions. Specific treatment intentions concerning different situations are also important, as medications with transient vs. relatively long-lasting effects are

now available.[24, 36, 37] Findings from RCTs indicate that ADHD medications have beneficial short-term effects on the core symptoms of ADHD, cohort studies on long-term effects suggest that ADHD medication may contribute to improvements in multiple functional domains,[35, 38] including educationally relevant outcomes[39] and occupational performance.[40, 41]

2.3.2 Non-pharmacological treatment

Non-pharmacological treatments include patient education and behavioral interventions toward both parents and children and adolescents with ADHD.[42] Efficacy of behavioral approaches is relatively good, but such interventions may only act as a complement to ADHD medication treatment.[43, 44] Apart from the multimodal standard of ADHD therapy with pharmacological treatments and/or behavioral intervention,[45] diet intervention (e.g. artificial food color elimination, a few-foods diet and poly-unsaturated fatty acid supplementation) were proposed as another possible therapy for ADHD management. However, based on evidence from systematic reviews, none of these dietary interventions are recommended as part of ADHD treatment, due to a lack of well-designed and large-sample-size RCT/cohort studies, the inconclusive findings across studies, and the unclear underlying mechanisms of the effects of diet intervention on ADHD development.[45-47] Given the generally low adherence to medications after adolescence,[10, 48] a better understanding of how diets associated with ADHD may shed new insight into treatment planning and management of the disorder in adults.

2.4 Etiology of ADHD

ADHD is a complex disorder and its pathophysiology is not very clear. Growing evidence indicates that genetic and environmental risk factors, and their complicated interactions play important roles in the etiology of the disorder.

2.4.1 Genetic factors in the etiology of ADHD

ADHD runs in families, with first-degree relatives having a five to tenfold increased risk of developing the disorder compared to the general population.[49-51] Twin studies show that clinical diagnosed ADHD has a heritability (i.e. the proportion of variation in a trait that can be attributed to genetic variation) of 70%-80% in childhood and adolescence,[52] suggesting a substantial role of genetic factors in the development of ADHD. Similar heritability estimates have been found across sex and the two ADHD

symptom dimensions: inattention (IA) and hyperactivity/impulsivity (HI).[53-57] Although heritability estimates of ADHD in adults tend to be lower, this appears to be largely explained by rater effects. Self-ratings, commonly used in studies of adult samples, tend to lead to lower heritability estimates. When estimated from multiple raters or clinical diagnoses, heritability estimates for ADHD do not differ substantially with age.[58] A recent twin study has reported a genetic correlation of 0.56 between ICD-based ADHD diagnoses and traits of ADHD, suggesting that genetic factors associated with clinically relevant ADHD are also associated with milder variation in ADHD traits throughout the general population.[59]

Single nucleotide polymorphisms (SNPs)-based heritability for ADHD symptom scores was estimated to be between 5 and 34%, in a large study including several population-based samples.[60] The GWAS meta-analysis, including over 20,000 individuals diagnosed with ADHD and over 35,000 controls, has identified the first 12 genome-wide significant risk loci associated with ADHD and calculated a SNP-based heritability for ADHD around 22%.[61] Findings from this study also indicated an important role for common variants in the polygenic architecture of ADHD.

2.4.2 Environmental factors in the etiology of ADHD

Several environmental factors may increase the risk for ADHD, including prenatal and perinatal conditions, toxins, dietary factors, and psychosocial adversities.[62-64] Effects of each environmental factor are relatively small and usually accumulate to increase the overall risk.[24, 65] However, the causal nature of these factors is largely unknown. Randomized experiments are preferred for testing causal environmental hypotheses. However, the assumptions of randomized experiments cannot always be met and they are not always feasible.[66] For such cases, several quasi-experimental designs to test causal hypotheses by ruling out plausible alternative explanations have been devised, including co-twin control, sibling- and cousin-comparison, and adoption studies.[67, 68] According to the findings from genetically informative studies, several observed associations between prenatal and perinatal factors (e.g. smoking during pregnancy[69, 70] and stress during pregnancy[71]) and ADHD are better explained by unmeasured familial confounding. On the contrary, low birth weight,[72, 73] advanced paternal age at childbearing[74] and low family income[75] might be causal risk factors for ADHD.

Maternal overweight and obesity prior to pregnancy are increasingly being recognized as potential modifiable risk factors for ADHD in offspring.[76] Systematic reviews have suggested that maternal pre-pregnancy overweight/obesity may be associated with suboptimal neurodevelopment in offspring, including an increased risk for ADHD.[76-78] To date, the precise mechanisms underlying the association between maternal pre-pregnancy overweight/obesity and ADHD in offspring remain unclear. Some biological mechanisms have been proposed as mediators for a causal association, including fetal programming hypotheses,[79] placental and intrauterine environment alterations and inflammatory mechanisms [80]. Alternatively, the associations might be explained by unmeasured confounders, such as genetic and shared environmental influences. Indeed, recent register-based within-family studies [81, 82] have suggested that the associations of ADHD with high body mass index (BMI), including clinically diagnosed obesity, could be attributed to genetic factors shared by the two conditions. Additionally, a large genome-wide association study [61] of clinically diagnosed ADHD reported a modest genetic correlation (r_g) between ADHD and obesity-related phenotypes, including BMI ($r_g = 0.26$), waist-to-hip ratio ($r_g = 0.30$), and childhood obesity ($r_g = 0.22$). Unmeasured environmental confounders, such as lifestyle factors (e.g. dietary habits, physical activities), might also influence maternal overweight/obesity [83], as well as the risk of ADHD in offspring.[84] Although two previous sibling-comparison studies have examined the association between maternal pre-pregnancy BMI and risk of ADHD in offspring, [85, 86] women who change pre-pregnancy weight between pregnancies may be systematically different from women whose pre-pregnancy weight remain stable.[87-89] Therefore, complementary designs, such as cousin comparisons, are needed to address these limitations.

2.5 Consequences of adult ADHD

2.5.1 Lifestyle and dietary habits

As mentioned above, dietary intervention could be a potential non-pharmacological treatment for ADHD, while abnormal dietary habits may also be a consequence of ADHD. Meta-analytic evidence on ADHD and obesity has suggested that the two core symptom domains of ADHD may contribute to the development of unhealthy or less-balanced dietary habits via in related yet separated means.[90-92] The core symptoms of inattention, poor planning and self-regulation deficits, may cause difficulties in adhering to a

regular eating pattern, favoring abnormal eating behaviors.[90, 91] In contrast, deficient inhibitory control and delay aversion, which are expressions of the hyperactivity/impulsivity component of ADHD may translate into impulsive eating of highly palatable foods or having no patience for eating vegetables, which are less rewarding than high-caloric foods.[90, 91, 93] Further, dietary habits may also play a key role in pathways linking ADHD to health-related outcomes such as metabolic syndromes (e.g., obesity). An advanced understanding of the association between ADHD symptoms and dietary habits could harbour new interventions and non-pharmacological treatments for ADHD, and may also contribute to the important knowledge gap of long-term consequences of ADHD in adults.

2.5.2 Cardiovascular diseases

In addition to the core clinical symptoms of ADHD, psychiatric and non-psychiatric coexisting problems and clinical conditions have been described in individuals with ADHD.[94] In particular, psychiatric comorbid conditions are recognized in both children and adults, and have posed considerable clinical and public health challenges [94-96]. Compared with the extensive descriptions of psychiatric comorbidity, somatic comorbidity in ADHD has received less attention in the research literature, particularly among adults.[97] Evidence from a systematic review suggested that obesity, sleep related problems and asthma were well documented somatic comorbidities in adults with ADHD.[97] However, the results were less robust or controversial for other disorders such as cardiovascular diseases (CVDs), enuresis, irritable bowel syndrome, restless legs, epilepsy, chronic fatigue syndrome, fibromyalgia syndrome, systemic lupus erythematosus and atopic dermatitis.[97] Further, important methodological limitations (e.g. small sample size, cross-sectional study design, or self-report questionnaires on ADHD diagnosis and health problems) exists across studies, and the role of potential confounders between ADHD and diseases have not been fully addressed.

Previous researches have established consistent associations of several psychiatric disorders, such as schizophrenia,[98] bipolar disorder,[99] depressive disorders[100] and obsessive-compulsive disorder[101] with metabolic and cardiovascular diseases, which in turn are thought to contribute to the increased mortality associated with these conditions.[102, 103] Unhealthy lifestyles factors (e.g. smoking, absence of physical activities) and prolonged use of psychiatric medications may contribute to the observed risk of CVDs in these psychiatric disorders.[101, 104, 105] However, in

ADHD, CVDs have been studied mainly as potential adverse effect of ADHD pharmacologic treatment,[106] and only a few previous studies have focused more specifically on the associations between ADHD per se and CVDs. A Dutch study [107] found elevated levels of ADHD symptoms were associated with increased risk of CVDs. Recently, a Swedish register-based cohort study [13] showed that adults with ADHD are at increased risk of a wide range of physical health conditions, including CVDs, compared with adults without ADHD. However, there are still several knowledge gaps limiting our understanding of how ADHD is associated with CVDs. For example, no studies have explored the role of other psychiatric comorbidities and well-established risk factors for cardiovascular diseases (e.g. smoking,[108] sleep disorder,[109] and metabolic conditions (e.g. obesity [110], type 2 diabetes mellitus (T2DM),[111] and dyslipidemia[112]).Furthermore, broad measures of CVDs, encompassing a wide range of circulatory system diseases, has been used in previous studies.[13, 107] Thus, little is known about the risks of specific groups of CVDs in ADHD. This is important to inform prevention and treatment strategies, which may vary substantially depending on which specific CVDs are most strongly associated with ADHD.

2.5.3 Functional impairments

Despite a decline of some ADHD symptoms over time, functional impairments often remain.[40, 113] Previous studies have reported that adult ADHD was associated with poor functional outcomes in multiple life domains such as psychosocial function and role functioning in education and work life,[41, 114-116] of which, work-related impairments (e.g. unemployment, having trouble keeping work or financial problems) represent high economic costs.[116] In adults with ADHD, the severity of ADHD symptoms might play an important role in occupational functioning.[41, 115] Being in work seems to include the ability to perform skills and strategies relevant to most functional domains. Executive function deficits such as problems with self-management of time, self-motivation, and self-discipline are found to contribute to occupational problems in adults. [117] ADHD comorbid with other psychiatric disorders has also been associated with lower occupational functioning in terms of higher risk of unemployment.[118]

Pharmacologic treatment is effective in reducing the core symptoms and functional impairments of ADHD, [14] but it is unclear whether it helps

improve occupational outcomes. The effects of ADHD medication on occupational outcomes have only been explored in two observational studies [40, 116] with results suggesting a positive correlation between employment status and treatment with stimulants, both past and present. However, these cross-sectional/retrospective studies need replication with larger samples and longitudinal methods.

2.6 Epidemiology studies based on real-world data

2.6.1 Large-scale real-world data (RWD)

Real-world data (RWD) refers to data obtained from daily clinical practice. Specifically, RWD could be defined as data collected in a non-RCT, non-interventional/non-controlled or non-experimental setting.[119] RCTs provide the ideal study design to demonstrate causality between two or more traits and determine the efficacy and safety of a therapy under ideal conditions. However, in addition to potential ethical issues in some research topics, the highly selective populations examined within the setting of RCTs are often not comparable with the more heterogeneous populations in clinical practice, which limits the transfer of their results.[119] A recent study suggested that RCTs may represent only about 20% of individuals with schizophrenia spectrum disorders in real-world. Therefore, more studies, such as observational cohorts and RCTs with broader inclusion criteria, are need to explore effect among underrepresented patients. [120]

Digital technology has been rapidly accelerating the access to vast amounts of health-related data on real-world conditions from recent years. In epidemiology, the most common sources of large-scale RWD include routinely collected health records (the electronic health record, EHR) and questionnaire data.

Compared with traditional approaches to gathering information from questionnaires (e.g. face-to-face/telephone interview, paper/pencil questionnaires), web-based questionnaires are more important tool for big data collection when considering cost- and time-effectiveness.[121] While questionnaire survey has the advantage of addressing a specific research question straightforwardly, the quality of data collection heavily depends on the respondents (e.g. background knowledge of the studied topic, cognitive ability, and recall bias). Another main disadvantage of web-based questionnaire survey is high nonresponse rate, which may introduce selection bias in epidemiology studies, especially in long-term follow-up studies.[121] Therefore, the use of EHR data may be an attractive alternative for large-scale

big data collections. By documenting the real care that patients receive in the clinic and include a variety of cases (usually comorbid with other diseases), EHR data typically comes with lower cost, larger sample size, longer follow-up periods, greater representativeness and higher external validity than those with other approaches. EHR opens new possibilities for providing clinical evidence regarding the effectiveness and safety of new treatments during routine care.[122] However, EHR-based research relies on secondary data as the original purpose of EHR is to record and manage information from clinical care. Additionally, lifestyle factors (e.g. diets, physical activities, and screening time), important covariates for many studied diseases, are typically not available in EHR. Therefore, the use of EHRs for epidemiology requires throughout consideration on each specific research question regarding target population, exposure/outcome definition, and privacy concerns.[123]

2.6.2 Causal inference in Epidemiology studies

Causal inference from large-scale RWD is a complicated concept in epidemiology, which relies on triangulating evidence from multiple sources and on the application of a variety of methodological approaches.[124] That is, no single study with a specific method can provide a straightforward answer to a causal question, but a thoughtful combination of different approaches, where each approach has its own strengths and limitations, can provide much stronger evidence for causal inference. [125]

Traditionally, one of the most widely used approach for causal inference is using statistical methods to adjust for confounding.[124] However, such methods rely on the assumption that all relevant or potential confounders have been correctly identified and measured, which only hold true in ideal conditions but not in the real world when we are studying complex health and social outcomes. For example, genetic influences on human traits and diseases are often unknown or unmeasured in many traditional epidemiology studies.

A different approach to causal inference is to use design-based approaches, to ensure covariate balance before doing any analyses, or compare results across methods with different sources and directions of potential bias. [124]

Traditional epidemiological studies generally involve unrelated individuals, often using population based sampling, while genetically informative studies focus on individuals varying in their degree of genetic relatedness in setting of families and relatives, seeking to dissect the relative contributions

of genetic and environmental factors in the development of diseases.[126] One important application of genetically informative designs in epidemiology is to assess and adjust for genetic confounding and other unmeasured familial confounding in an exposure-outcome association. For example, comparisons of full siblings raised in the same family could help to rule out confounding effect by all shared environmental and genetic factors (full siblings shared 50% of their genetic makeup).[68] When multiple types of sibling comparisons are included in one study (e.g., identical and fraternal twins, full sand half sibling), researchers can explore whether the observed associations is causal, or the degree to which an association could be explained by unmeasured familial confounding.

3. AIMS

3.1 Overarching aims

The overarching aims of the four studies are to extend previous knowledge on the early risk factors for ADHD, and to better understand and increase the awareness of somatic and occupational outcomes of ADHD in adults.

3.2 Specific aims

Study I To clarify associations between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring, and explore the potential influence of shared familial confounding.

Study II To identify associations between ADHD symptom dimensions and different dietary habits in adults, and then investigate the relative contribution of genetic and environmental factors to the associations.

Study III To investigate the risk of a broad range of cardiovascular diseases in adults with ADHD, and explore the potential influence of psychiatric comorbidities and traditional risk factors for cardiovascular diseases.

Study IV To explore the association between ADHD medication use and long-term unemployment in individuals with ADHD.

4. MATERIALS AND METHODS

4.1 Data Sources and main measurements

Study I, III and IV of this thesis were register-based cohort studies based on data derived from a linkage of several Swedish nationwide registers. Study II used questionnaire data from Study of Twin Adults: Genes and Environment (STAGE).

4.1.1 Swedish health care and population registers

Every resident in Sweden has been assigned a ten-digit unique personal identity number (PIN) since 1947. PIN enables record linkage with all administrative and health registers in Sweden.[127] The data for this thesis was extracted mainly from the following registries:

Total population register (TPR) includes information on demographics (date of birth, sex, country of birth, migration, date of death) on all individuals residing in Sweden who were born after 1932 and were alive in 1968 or later.[128]

Multi-generation register (MGR) provides information on biological (and, when applicable, adoptive) parents of all individuals born after 1932, alive and living in Sweden after 1961, with the exception of those whose parents died or migrated out of the country before 1947.[129] For individuals born since 1950, information on mothers is complete while the coverage of information on fathers is slightly lower.

The Prescribed Drug Register (PDR) includes data on all prescribed drugs dispensed at pharmacies in Sweden since June 1, 2005. It contains information regarding drug identity, according to the Anatomical Therapeutic Chemical classification system (ATC), quantity and dosage of the prescribed drug, date of prescription/dispensing, and specific patient information (sex, age and residential area).[130]

The Swedish National Patient Register (NPR) has national coverage of psychiatric inpatient care since 1987, and information on psychiatric outpatient care since 2001. Every record has a discharge date, a primary discharge diagnosis and up to eight secondary diagnoses assigned by the treating medical doctor according to the International Classification of Diseases (ICD).[131] ICD codes of ADHD includes code 314 (ICD-9) and F90 (ICD-10). Information from the NPR was used to identify ADHD cases, other psychiatric and somatic disease diagnosis as outcomes in different studies.

Longitudinal integration database for health insurance and labor market studies register (LISA) contains information on highest level of education, unemployment, social benefits, and family income for all Swedish residents aged 16 years or older since 1990. The database integrates existing data from the labor market, educational and social sectors and is updated each year with a new annual register.[132] The register was used to retrieve information on work absences and occupational outcomes. For instance, disposable income, social allowance, unemployment, and need for social welfare benefits (i.e., rehabilitation, or unemployment benefits).

The national crime register (NCR) contains records of all criminal convictions in Swedish district courts since 1973. Age 15 is the minimum age of criminal responsibility in Sweden, so NCR only included individuals at aged 15 years or older.[133]

The cause of death register (CDR) provides information on all deaths among individuals registered in Sweden from 1952. The specific causes of death were coded based the ICD coding system.[134]

4.1.2 Main measures in the registers

4.1.2.1 ADHD and ADHD medication

Individuals with ADHD were identified according to ADHD diagnosis (ICD-9 or ICD-10: 314/F90) from NPR at the age of 3 years or older, and ADHD medication prescription (ATC codes: N06BA01/ N06BA02/N06BA12/N06BA09) from PDR, or both. This approach to identify individuals with ADHD has been validated and is widely used in Swedish register-based studies, as ADHD medication can only be prescribed for ADHD treatment by physicians specialized in psychiatry or neurology.[13, 135] Therefore, prescription of ADHD medication is a valid indicator of ADHD diagnoses in Sweden. However, ADHD medication was an important mediator for studied association (ADHD and CVDs) in study III, we only used diagnoses of ADHD from NPR for case identification in a sensitivity analysis to reflect clinically diagnosed cases.

In study III, ADHD was treated as a time-varying variable. That is, individuals were assigned to the unexposed group before the diagnosis of ADHD, and were assigned to the exposed group from the first diagnosis or medication prescription of ADHD to the end of follow-up.

In study IV, the main exposure is pharmacological treatment of ADHD during the previous two years. We used a validated natural language processing algorithm to estimate the duration of ADHD medication from "free-

text" prescription in the PDR.[136] We assessed the exposure in two aspects: 1) Treated status: whether or not treated with ADHD medications during the previous two years. 2) Duration of treatment: using six months as the time interval, we further divided the treated ADHD group into four levels: treated for <6 months, 6-<12 months, 12-<18 months and 18-24 months.

4.1.2.2 **Maternal pre-pregnancy overweight/obesity**

The MBR includes information on self-reported height and measured weight at the first prenatal visit for pregnant women, which was used to calculate body mass index (BMI) in Study I. According to the World Health Organization guidelines, overweight is defined as a BMI greater than or equal to 25, and obesity is a BMI greater than or equal to 30.[137] Obesity is further classified into three levels: obesity class I ($30.0 \leqslant BMI < 35.0$), obesity class II ($35.0 \leqslant BMI < 40.0$) or obesity class III ($BMI \geqslant 40.0$).

4.1.2.3 Cardiovascular diseases

Consistent with previous studies,[138, 139] incident CVD events (including any CVDs and specific diseases subtypes included: ischemic heart disease, cerebrovascular disease, venous thromboembolism, hypertensive diseases, heart failure, arrhythmias, cardiac arrest and peripheral vascular disease/arteriosclerosis) were defined as the first diagnosis of CVD from NPR or death with CVD as the underlying cause from the CDR. A complete list of all specific CVDs, and corresponding ICD-8, ICD-9, and ICD-10 codes, used in study III is presented in Table 4.1.2.3.

Table 4.1.2.3 ICD codes for cardiovascular diseases from the Swedish National Patient Register

	ICD-8	ICD-9	ICD-10
Any cardiovascular disease	390-438, 440, 444, 445, 450-453, 458	390-438, 440,444,445	I00-I70, I73.0, I74-I75
Ischemic heart disease			
Acute coronary syndrome (ACS)	410-412	410-412	I21-I24, I25.2, I20.0
Chronic coronary syndrome (without ACS)	413-414	413-414	I20.1-I20.9, I25.1, I25.5-I25.9
Cerebrovascular disease			
Subarachnoidal bleeding	430	430	I60
Hemorrhagic stroke	431	431, 432	I61-I62
Ischemic stroke	432-434	433, 434	I63-I64

Other cerebrovascular disease	436-438	436-438	I65-I69
Venous thrombo-embolism (VTE)			
Deep vein thrombosis (DVT)	451	451	I80
Pulmonary emboli (PE)	450	415	I26
Hypertensive diseases			
Essential hypertension	400	401	I10
Other hypertensive disease	401-404	402-405	I11- I16, I674
Heart failure			
Heart failure	428	428	I50
Ischemic cardiomyopathy	-	-	I25.5
Other cardiomyopathy	-	-	I42.0-I42.9
Arrhythmias			
Bradyarrhythmias:	-	427W,426A, 426B	I49.5, I44.1-I44.2
Takyarrhythmias:	-	427A/B/D/E	I47.0-I47.2,I48, I49.0, I49.8
Cardiac arrest	-	427F	I46
Peripheral vascular disease/Arteriosclerosis	440-444	440-444	I70-I74

4.1.2.4 Long-term unemployment

Information on yearly accumulated days registered as unemployed was obtained from administrative records reported by the Public Employment Service (registered in the LISA database). The variable equals to zero for those who have never been registered as unemployed in each year. Consistent with previous studies, [140-142] unemployment groups were classified into those with 1–89 days per year (short-term unemployment) and those with 90 or more days per year (long-term unemployment) of accumulated registered unemployment. The division by 90 days was chosen based on the theoretical assumption of duration dependence, that is, the probability of leaving unemployment for work declines with the duration of unemployment,[140] and also as an indicator for identify short breaks in employment related to job change. In study IV, we use long-term unemployment as the primary outcome of interest, as it is a better proxy of reduced work ability that may be caused by ADHD.

4.1.2.5 Covariates

In study I, III and IV, sociodemographic information on year of birth, birth country, sex, and education levels were obtained from MBR and LISA.

In study I, we further extracted information on maternal age at delivery ($\leqslant$19, 20–24, 25–29, 30–34 or $\geqslant$35 years), smoking during pregnancy (0, 1–9 or $\geqslant$10 cigarettes per day), and cohabitation with child's father at childbirth (yes or no) from MBR. In study III, type 2 diabetes mellitus

(T2DM), obesity, hyperlipidemia, sleep disorders, heavy smoking (including tobacco abuse and nicotine dependence) were obtained from NPR. In study IV, crime records (length of incarceration) were obtained from The National Crime Register. In study III and IV, we also identified common psychiatric comorbidities of ADHD from NPR. A complete list of all the ICD-9 and ICD-10 codes used for abovementioned covariates can be found in Table 4.1.2.5.

Table 4.1.2.5 ICD-codes for covariates used in study III and IV from the Swedish National Patient Register

Covariates	ICD-8	ICD-9	ICD-10
Obesity	277	278.A, 278.B	E65,E66
Type 2 diabetes mellitus	250	250	E11
Hyperlipidemia	279	272	E78
Sleep problems	347, 780.60	347,780F	F51,G47.0- G47.4, G47.8, G47.9
Smoking	-	305B	F17, T65.2, Z71.6, Z72.0
Anxiety Disorder	300.0	300.00,300.02	F40-F41
Autism spectrum disorders	-	299	F84
Bipolar Disorder	296.1, 296.3, 296.8	296A/ C/D/E/W	F30-F31
Conduct disorder	-	312	F91
Depressive disorder	296.2, 298.0,300.4	296B,300E	F32-33
Eating disorders	306.5x	307.5	F50
Intellectual disability	310-315	317-319	F7x
Personality disability	301	301	F60, F69
Schizophrenia	295 except 295.7	295A-E/G/W/X	F2x
Substance use disorder	291,303,304	291,292,304,305A, 305X	F10-16,F18-F19

4.1.3 The Swedish Twin Registry

The Swedish Twin Registry (STR) contains more than 170,000 twins born in Sweden from 1886 and onwards. In addition to the PIN that everyone has in the Swedish administrative registers, twins in the STR have a unique twin identification number enabling record linkage with all sub-cohorts of the register.[143] Data have been collected through multiple waves of questionnaires. Specifically, in May 2005, a total of 42,582 Swedish twins born between 1959 and 1985 who have survived their first birthday, were identified from the population based Swedish Twin Register. Of the target population, 25,364 (59.6%) individuals participated in the Study of

Twin Adults: Genes and Environment (STAGE).[144] Participants were given access to the web-based survey, containing a questionnaire with 1,300 questions, in 34 sections, regarding lifestyle, mental and physical health. We identified information on ADHD symptoms and dietary habits from STAGE.

4.1.4 Main measures in STAGE

4.1.4.1 ADHD symptoms

Self-reported ADHD symptoms were obtained via a checklist of 18 DSM-IV symptoms, consisting of nine items for inattention and nine items for hyperactivity/impulsive.[145] Each item had a three-point answer format (0="No", 1="Yes, to some extent", 2="Yes"), the items were then summed up to create a two dimensional scale of inattention and hyperactivity/impulsivity.

4.1.4.2 Dietary habits

Dietary habits were assessed by a semi-quantitative food frequency questionnaire (FFQ), which consisted of 94 food items usually consumed by Swedish adults. For each food item, the participants were requested to indicate their average consumption frequency over the past year in terms of the specified serving size by checking one of the following frequency categories: never, 1-3 times/month, 1-2 times/week, 3-4 times/week, 5-6 times/week, 1 time/day, 2 times/day, 3 times/day. The selected frequency category for each food item was converted to a value in a number of servings per day. As is shown in Table 4.1.4.2, dietary habits were expressed in three ways (a) consumption of food groups (fruits, vegetables, dairy, meat, and seafood), (b) consumption of food items rich in a particular macronutrient (high-fat food, high-carbohydrate food, high-sugar food, and high-protein food), and (c) healthy dietary patterns (fruits, vegetables, fish, and white meat) and unhealthy dietary patterns (food with high proportions of refined sugar and saturated fat). Such definition is in line with previous studies on the associations between ADHD and dietary habits.[146] The scores of each dietary habit were calculated as the total frequency of each food group.

Table 4.1.4.2 The definitions of dietary habits. The intake frequency for each food item was converted into number of servings per day and summarized for each food group, respectively.

Dietary habits	Food items
1. Food Groups	
Fruits	Orange, apple, banana, berries, other fruit
Vegetables	Salad, cabbage, cauliflower, broccoli, tomato, pepper, spinach, green peas, onion, garlic, mixed vegetables, carrot
Dairy	Cheese, low-fat cheese, cottage cheese, low-fat milk, milk 1.5% fat, regular milk, low-fat yoghurt, yoghurt
Meat	Minced meat/meatball, pork, beef, chicken, sausage, liver, liver pâté s, ham, lamb, other meat
Seafood	Herring/mackerel, salmon, fish fingers, tuna, other fish, caviar, prawns
2.Food items rich in macronutrients	
High in fat	Pizza, fried potato, French fries, chips, nuts, dressing, mayonnaise, cream, cheese
High in carbohydrates	bread (white bread, loaf of bread, wholemeal bread, crisp bread), oats porridge, porridge, cereal (sweet cereal, cereal, muesli), pasta, rice, wheat, boiled potato, pancakes, vegetables and fruits
High in sugar	Cookies, crackers, cake, chocolate, candy, ice cream, jam, berry cream, juice, energy drink, honey, ketchup
High in protein	Minced meat/meatball, pork, beef, chicken, sausage, liver, liver pâté, ham, lamb, other meat, beans/lentils, tofu/soy, quorn, cheese, low-fat cheese, cottage cheese, low-fat milk, milk 1.5% fat, regular milk, low-fat yoghurt, yoghurt, Herring/mackerel, salmon, fish fingers, tuna, other fish, caviar, prawns, egg, high calorie drink/ protein drink
3. Dietary patterns	
Unhealthy dietary pattern	Pizza, sausage, fried potato, French fries, minced meat/meatball, pork, beef, Cookies, crackers, cake, chocolate, candy, ice cream, jam, berry cream, juice, energy drink, Hamburg, chips, dressing, mayonnaise, cream, ketchup
Healthy dietary pattern	Herring/mackerel, salmon, fish fingers, tuna, other fish, prawns, orange, apple, banana, berries, other fruit, salad, cabbage, cauliflower, broccoli, tomato, pepper, spinach, green peas, onion, garlic, mixed vegetables, carrot, beans/lentils, nuts, boiled potato, chicken, flax, sweet cereal, cereal, muesli

4.2 Overview of study methods and materials

Table 4.2 describes an overview of study design, data sources, study population, exposures, outcomes, covariates and statistical methods used in the four individual studies included in the thesis.

Table 4.2 The overview of methodologies in Study I-IV

	Study design	Data sources	Study population	Exposures and outcomes	Covariates	Statistical methods
I	Part I: Systematic review, meta-analysis Part II: quasi-experimental family-based study	Part I: Previous published studies Part II: Swedish national registers	Part I: 784,804 mother–child pairs from eight pertinent cohort studies Part II: 971,501 individuals born in Sweden between 1992 and 2004	Exposure: pre-pregnancy overweight/obesity Outcome: ADHD in offspring	Part II: sex, birth order, year of birth, mother's country of birth, maternal education, maternal age at delivery, smoking during pregnancy and cohabitation with child's father at childbirth	Part II: Cox proportional hazard models
II	Twin study	STAGE	Swedish population-based twin study with 17,999 individuals aged 20–47 years	Exposure: dietary habits Outcome: ADHD symptoms	Sex, age and Socioeconomic status (SES)	Structural equation modelling (SEM)
III	Population-based cohort study	Swedish national registers	5,389,519 individuals born in Sweden between 1941 and 1983	Exposure: ADHD Outcome: cardiovascular diseases	Sex, year of birth, education level, birth country, T2DM, obesity, dyslipidemia, sleep problems, smoking, psychiatric comorbidities, family history and psychotropic medication	Cox proportional hazard models
IV	Longitudinal cohort study	Swedish national registers	12,875 middle-aged adults with ADHD born between 1958 and 1978	Exposure: ADHD medication Outcome: Long-term unemployment	Age at baseline, sex, country of birth, highest educational level, crime records, psychiatric comorbidities	Generalized estimating equations (GEE)

4.3 Study designs

4.3.1 Systematic review and meta-analysis

The growing number of publications makes it difficult for a clinician or a researcher to keep their knowledge up to date. Thus, systematic reviews and meta-analyses are being conducted to help better extract high quality evidence from the flood of data being produced.[147] For a given topic, systematic reviews systematically summarize the results from all available studies, evaluate the quality and assess the potential bias of included studies. When possible, a statistical meta-analysis of the individual results is conducted to obtain the pooled estimates for a specific research question. In epidemiology, systematic reviews and meta-analyses have been used as essential tools for summarizing evidence accurately and reliably. To ensure that a systematic review is well-conducted, valuable to users, the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement is developed and widely used by researchers. PRISMA 2020 statement [148] is recently introduced to replaces the PRISMA 2009 statement,[149] with new reporting guidance that reflects advances in methods and many innovations in the conduct of systematic reviews (e.g. natural language processing and machine learning). In study I, we conducted a systematic review and meta-analysis to summarize the available evidence on the association between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring.

4.3.2 Register-based cohort study

Cohort designs are one of the most important study designs for observational studies that aid in evaluating exposure-outcome associations.[150] In a cohort study, the study population is first classified into two or more groups by their exposure status or censoring, and then followed over time until the outcome of interest occurs. Given exposure is identified before the occurrence of outcome, cohort studies have a temporal framework to assess causality from exposure and outcome, and provide the strong scientific evidence.[150]

Data from medical records and registers are widely used in epidemiological studies, especially in Nordic countries, where nationwide health care registries have been established for research since the mid-1990s.[151] Comparing with traditional cohort studies, register-based cohort studies are based on larger samples, or even the entire population, with longer or complete follow-up and have the possibility to examine associations between

rare exposures and outcomes with advanced statistical methods and sufficient power.[151] Register-based cohort studies were used in study I, III and IV to explore risk factors and long-term consequences of ADHD.

4.3.3 Twin studies

Twin studies are a special type of epidemiological study that rely on the different levels of genetic relatedness between monozygotic twins (MZ), who share all of their segregating genes, and dizygotic twins (DZ), who share on average half of the polymorphic genetic variation; MZ and DZ twins are assumed to share 100% of their common environmental factors.[152] Based on this information and assumption, it is possible to test if genetic factors contribute to variation in a given trait. That is, a greater intra-class correlation (ICC)- the correlation of a trait between the members of a twin pair, was expected in MZ twins than in DZ twins, if there are genetic influences on this trait (univariate analysis). Similarly, higher cross-twin cross-trait correlations (CTCT), the correlations between trait 1 of twin 1 and trait 2 of twin 2, in MZ twins than in DZ twins indicate genetic influences on the covariation across traits. In Study II, information on dietary habits and ADHD trait dimensions reported by MZ and DZ twins was used to explore the relative contribution of genetic and environmental factors to the observed variations and covariations.

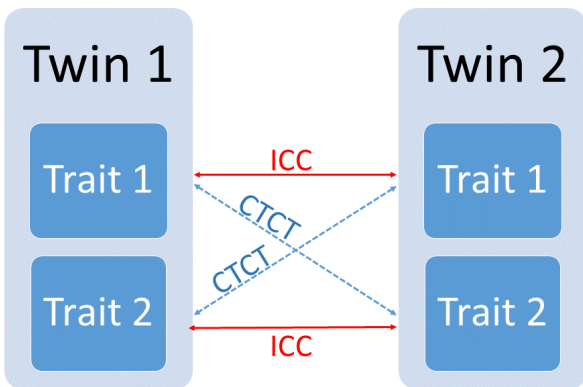

Figure 4.3.3 A schematic representation of the relationships intra-class (ICC) and cross-twin cross-trait correlations (CTCT). Twin 1 and twin 2 are members of a monozygotic or dizygotic twin pair.

4.4 Statistical methods

4.4.1 Meta-analysis

Meta-analysis is a statistical analysis used to systematically assess the results of multiple scientific studies to derive conclusions on a specific research question.[153] Fixed-effect and random-effect models are the most commonly used statistical models for aggregate data. However, fixed-effect models rely on assumptions that there is no heterogeneity across all individual studies (e.g. same population and same definitions for key variables), which is typically unrealistic in real-world scenarios, especially for observational studies. In contrast, random-effect models assume that the observed estimates can vary across studies because of real differences in study conditions across studies.[154] Outcomes from a meta-analysis may include a forest plot presenting pooled estimates of the effect of treatment or risk factor for disease, or other outcomes from included studies, and the results from the examination of heterogeneity and publication bias.

In study I, we aim to clarify the associations between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring based on available evidence. We used random-effects models to calculate the pooled estimates and presented the results in forest plots. Heterogeneity among studies was assessed by the Cochran Q test and I^2 statistic (level of significant $P<0.10$ or $I^2>70\%$). The presence of publication bias was assessed by the Begg's test and Egger's test.

4.4.2 Cox proportional hazard regression

Cox proportional hazard regression is a widely used statistical method in medical research for time-to-event outcomes on one or more predictors. It is a semi-parametric model without any assumption on the baseline hazards, but assuming that the hazards in exposed and unexposed group are proportional across time.[155] In this thesis, the proportionality of hazards across time has been assessed by using a Schoenfeld residuals-based test. Cox proportional hazard regression model was used in study III, to estimate hazard ratios (HRs) with 95% confidence intervals (CI) expressing the risk of CVDs in individuals with ADHD, compared with those without ADHD, taking attained age as the underlying time scale.

Stratified Cox regression can be used to compare the hazards across siblings, cousins or twins in the same family to one another and thus to control for all unmeasured shared family-level variables. In Study I, stratified Cox regression was used in cousin and sibling comparisons, to examine the role

of shared familial confounding on the associations between maternal pre-pregnancy overweight and obesity and ADHD in offspring with each set of maternal full cousins and full siblings as a separate strata.

4.4.3 Structural equation modelling (SEM)

Structural equation modelling (SEM) is a multivariate statistical method that simultaneously unites factor analysis and multiple regression analysis, enables researchers to examine causal relationships among observed variables and latent constructs.[156] SEM includes two basic types of models: 1) The measurement model specifies how a set of observed variables measure the latent constructs. 2) The structural model shows how latent constructs are causally related to each other. It is now widely used in quantitative genetics research, for example twin model-fitting analysis.

In twin analysis, the total phenotypic variance (P) of a trait was decomposed into additive genetic factors (A), dominant genetic factors (D), shared environmental factors (C) and non-shared or unique environmental factors (E), including measurement error.[156]

$$P = A + D + C + E$$

It is important to be aware of the assumptions that are made in the classical twin study, including

a) MZ and DZ twin pairs share their environments to the same extent (i.e. the equal environments assumption);
b) Gene–environment correlations and interactions are minimal for the trait;
c) Twins are no different from the general population in terms of the trait;
d) Matings in the population occur at random (no assortative mating).[156]

In twin data, MZ and DZ twins have different degrees of correlation for the genetic components (A and D), but the same degrees of correlation for the environmental components (C and E). Figure 4.5.3 shows a path diagram for basic univariate twin model.

In study II, SEM was used to conduct univariate and bivariate analysis on associations between dietary habits and ADHD symptoms. In the univariate and bivariate analyses, we first fitted a fully saturated model and several sub-models to test assumptions. Akaike Information Criterion (AIC)

was additionally used to assess the fit of each solution. The estimates from best-fitted models were reported to show the relative contributions of genetic and environmental factors to the potential associations between dietary habits and ADHD symptoms.

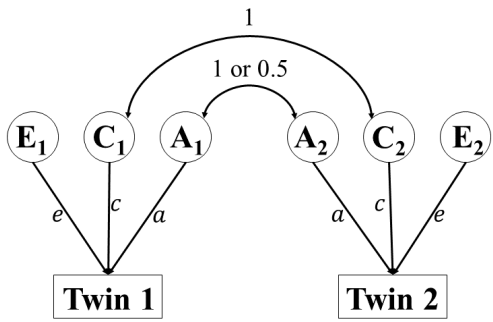

Figure 4.5.3 Path diagram for the basic univariate twin model. The additive (A) genetic factors are correlated 1 between MZ twins and 0.5 between DZ twins, respectively. Shared family environment (C) is correlated 1 for both MZ and DZ twins. Unique environment (E) is uncorrelated between members of MZ and DZ pairs. a, c and e are the path coefficients for the A, C and E effects, respectively.

4.4.4 Generalized estimating equations (GEE)

The generalized estimating equation (GEE) is a general statistical approach facilitates analysis on longitudinal or nested data, and repeated measures designs. Comparing with traditional regression models (e.g. Logistic regression), GEE can take into account the correlation of within-cluster data (longitudinal studies). GEE is a flexible method without any assumptions about the joint distribution of the response vector beyond the marginal means, therefore, it can be used for outcome variables with many different distributions, including normal, binomial, and Poisson.[157]

In Study IV, the dependent variable (long-term unemployment) was measured repeatedly from 2008 to 2013. To take into account the interdependence between repeated within-subject measurements, we performed GEE analysis with a log link to calculate the relative risks (RRs) and 95% confident intervals (CIs), and adjusted for age, sex, country of birth, highest education level and psychiatric disorders.

5. RESULTS

5.1 Maternal pre-pregnancy overweight/obesity and the risk of ADHD in offspring (Study I)

In study I, we first conducted a systematic review and meta-analysis on the associations of maternal pre-pregnancy overweight/obesity with ADHD in offspring. We then tested the results of the meta-analysis and further explored the role of measured and unmeasured familial confounding in a nationwide population-based cohort study.

5.1.1 Systematic review and meta-analysis

As shown in Figure 5.1.1.1, we included eight studies with 784,804 individuals in the meta-analysis, all of them were cohort studies.

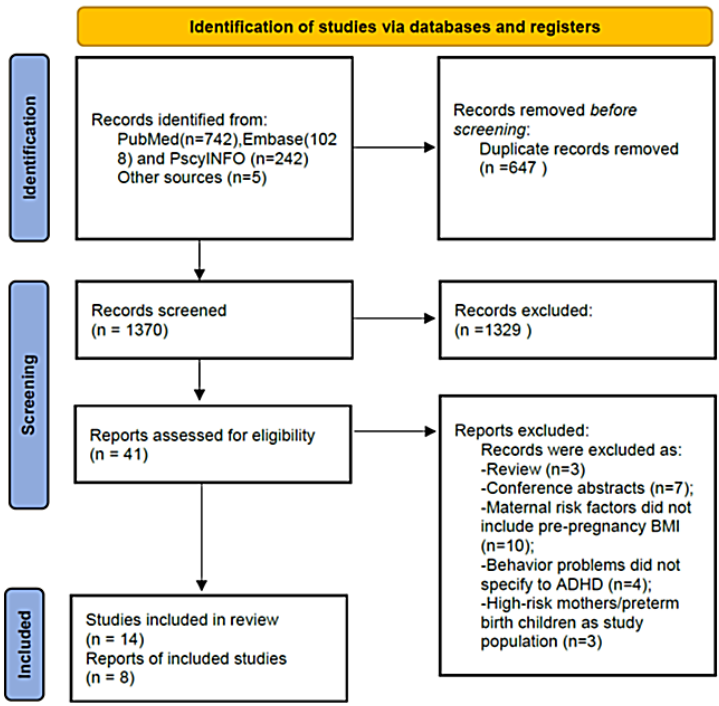

Figure 5.1.1.1 PRISMA flow diagram of selecting studies for systematic review and meta-analysis.

The primary analysis with crude estimates (Figure 5.1.1.2) indicated a significant association of maternal pre-pregnancy overweight (RR=1.31, 95%CI=1.25-1.38) and obesity (RR=1.92, 95%CI=1.84-2.00) with increased risk of ADHD in offspring. When adjusted estimates from individual studies were included in the meta-analysis, the pooled estimates were slightly attenuated but remained significant (HR for maternal pre-pregnancy overweight=1.28, 95% CI=1.17–1.40; HR for maternal pre-pregnancy obesity=1.64, 95% CI=1.57–1.73).

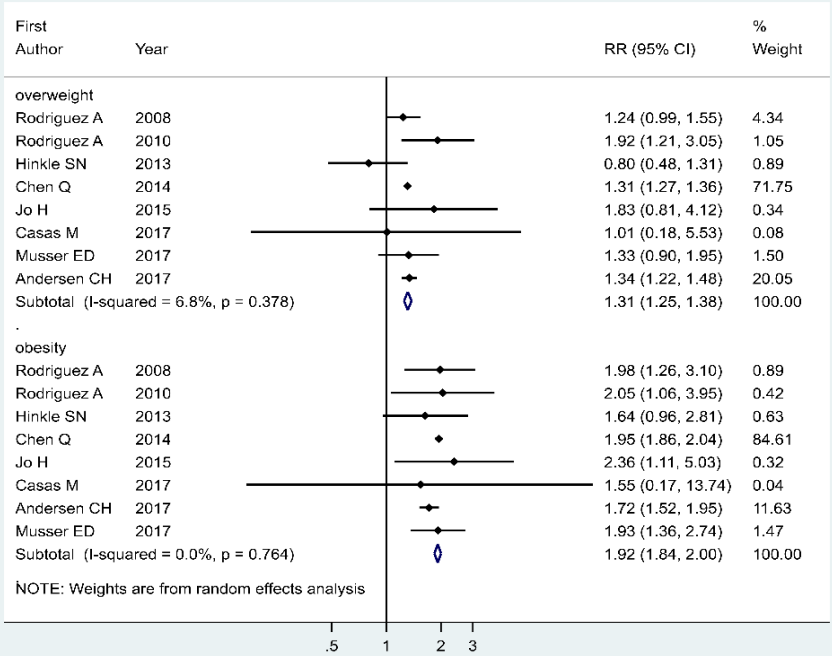

Figure 5.1.1.2 The pooled estimates of the associations between maternal pre-pregnancy overweight and obesity with increased risk of ADHD in offspring.

5.1.2 Nationwide population-based cohort study

We included 971,501 individuals in the nationwide population-based cohort, of whom 43,916 (4.52%) had a diagnosis of ADHD. At the population level, we found maternal pre-pregnancy overweight (HR=1.21, 95%CI=1.19–1.25) and obesity (HR=1.60, 95%CI=1.55–1.65) was associated with increased risk of ADHD in offspring, after adjusted for sex, birth

order, year of birth, mother's country of birth, highest maternal education, maternal age at delivery, smoking during pregnancy and cohabitation with child's father at childbirth.

In full-sibling and full-cousin comparisons, we further identified 463,474 full biological siblings nested within 216,084 families and 155,841 first-born maternal full cousins nested within 74,057 extended families from the entire study population. The observed associations at the population level gradually attenuated towards null, when adjusted for unmeasured factors shared by cousins and siblings (Figure 5.1.2), indicting the important role of unmeasured familial confounding in the association between maternal pre-pregnancy overweight/obesity and ADHD in offspring.

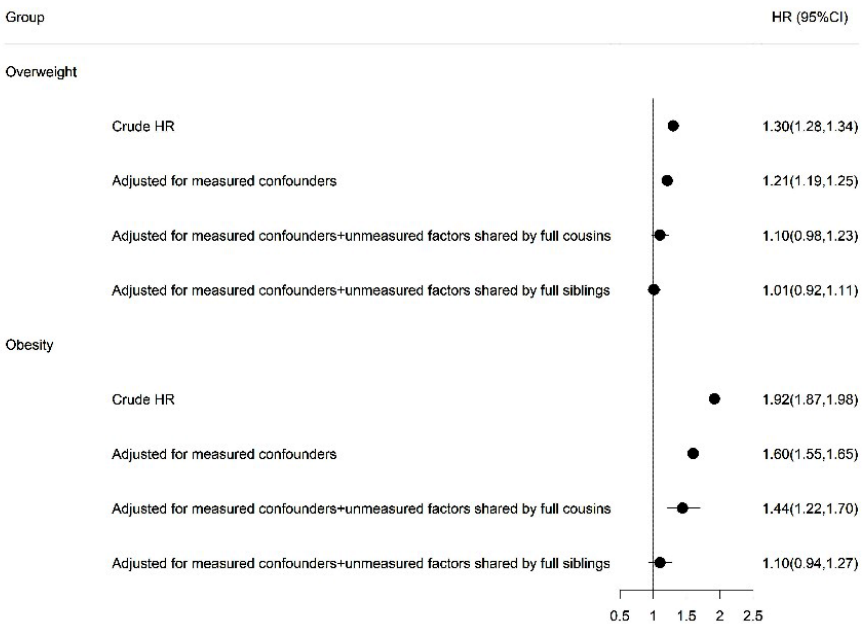

Figure 5.1.2 The estimates of the associations between maternal pre-pregnancy overweight and obesity with increased risk of ADHD in offspring from Swedish population based cohort.

5.2 ADHD symptoms and dietary habits in adults (Study II)

From STAGE, a total of 17,999 individuals with information on either ADHD symptoms and/or FFQ were included in the analyses. At the pheno-

typic level, we found more inattention and hyperactivity/impulsivity symptoms were associated with higher consumption of high-sugar food and unhealthy diets but lower consumption of vegetables, fruits and healthy diets. Generally, all associations were weak but the strongest correlation was found between inattention and high-sugar food intake (0.13, 95%CI=0.11-0.15), see Figure 5.2.

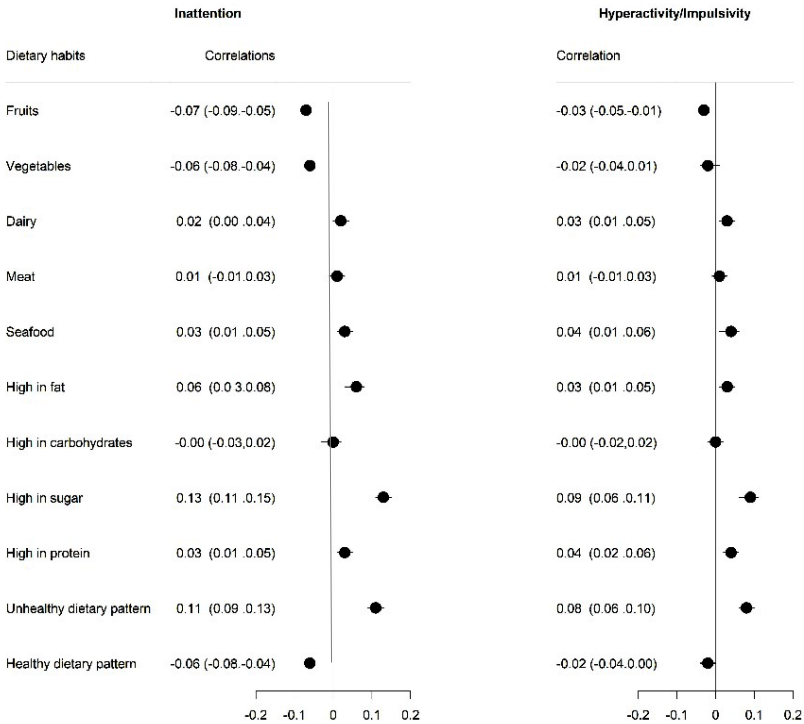

Figure 5.2 The associations between ADHD symptoms and dietary habits in adults.

In twin analyses, higher ICCs and CTCTs were found in MZ twins than in DZ twins, indicting the presence of genetic effects on the associations of ADHD symptoms with high-sugar and unhealthy dietary habits. AE models provided the best fits to the data and the results from bivariate model-fitting analyses are displayed in Table 5.2. Genetic correlations between ADHD symptoms and dietary habits were also weak and the strongest correlation was consistently found between inattention and high-sugar food intake (0.16, 95%CI=0.07-0.25). Genetic factors explained 44% (95% CI=18-

70%), 40% (95% CI=10-69%), and 37% (95% CI=1-71%) of the associations between inattention and high-sugar food, inattention and unhealthy dietary pattern, and hyperactivity/impulsivity and high-sugar food, respectively. Non-shared environmental factors contributed substantially to abovementioned associations.

Table 5.2 Estimates of genetic and environmental effect (95% confidence intervals) from bivariate AE models

	rA	Bivariate A	Bivariate E
IA and High-sugar food	0.16 (0.07,0.25)	0.44 (0.18,0.70)	0.56 (0.30,0.82)
IA and Unhealthy diets	0.13 (0.03,0.22)	0.40 (0.10,0.70)	0.60 (0.30,0.90)
HI and High-sugar food	0.09 (0.002, 0.19)	0.37 (0.01, 0.71)	0.63 (0.30, 0.99)
HI and Unhealthy diets	0.05 (-0.05,0.14)	0.20 (-0.21,0.56)	0.80 (0.44,1.20)

IA: Inattention; HI: Hyperactivity/impulsivity; rA: genetic correlation

5.3 ADHD as a novel risk factor for CVDs (Study III)

By linking several Swedish registers, we created a population-based cohort with 5,389,519 adults born between 1941 and 1983, without pre-existing cardiovascular diseases. The follow-up started from January 1st 2001 or the first diagnosis of ADHD, whichever came later; and ended up with the first diagnosis of any cardiovascular disease, death, emigration, or December 31st, 2013, whichever occurred first. After an average of 11.8 years follow-up, a total of 37,027 (0.68%) individuals had a diagnosis of ADHD and 746,572 individuals was newly diagnosed with CVDs.

We found a more than two-fold increased risk of developing any type of CVDs among adults with ADHD, compared with those without ADHD (HR=2.05, 95% CI=1.98-2.13) after adjusting for sex and age. The associations attenuated but remained significant (HR=1.65, 95%CI=1.59-1.71), when we further adjusted for other sociodemographic characteristics (education level and birth country), well-established risk factors for CVDs (T2DM, obesity, dyslipidemia, sleep problems and smoking), and psychiatric comorbidities (anxiety, ASD, bipolar disorder, conduct disorder, depressive disorder, eating disorder, intellectual disability, personality disorder, schizophrenia, substance use disorder). Among 22 specific CVDs, the strongest associations, after adjusted for covariates, were found for cardiac arrest

(HR=1.88, 95%CI=1.49-2.36) and peripheral vascular disease/arteriosclerosis (HR=1.71, 95%CI=1.47-1.99) (Figure 5.3).

In subgroup and sensitivity analyses, the associations between ADHD and CVDs were stronger in males and younger adults across all levels of adjustments, while similar association patterns were found among individuals with or without psychotropic medications and family history of CVDs.

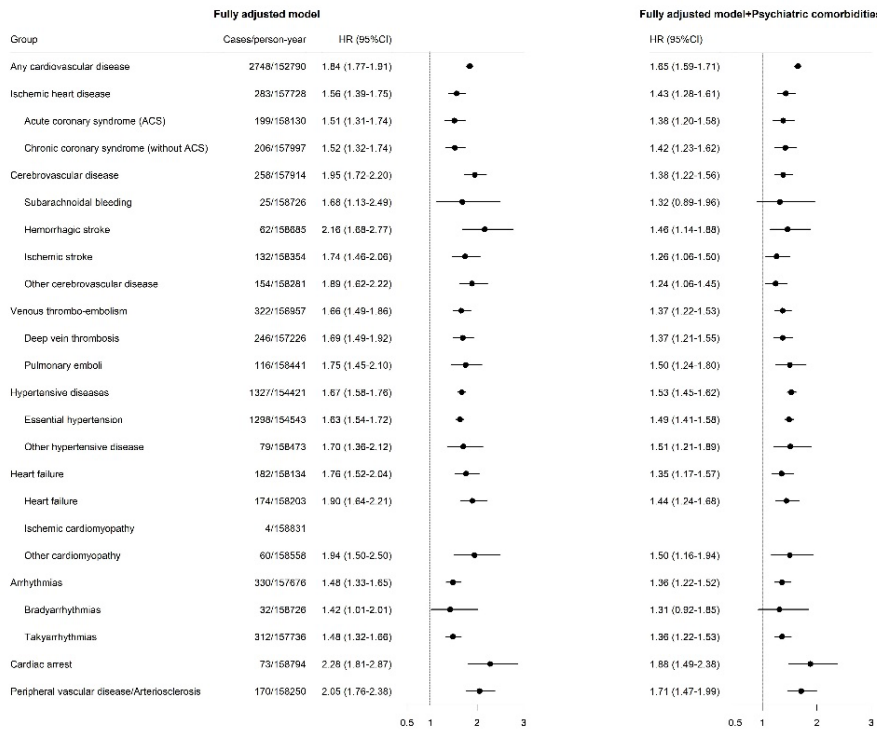

Figure 5.3 Hazard ratio with 95% CIs of developing different types of CVDs among adults with ADHD, compared with those without ADHD. The first part showed results from fully adjusted model, which adjusted for year of birth, sex, educational level, education level, birth country, obesity, T2DM, dyslipidemia, sleep problems, smoking. The second part shown results from Cox regression models adjusted for year of birth, sex, educational level, birth country, obesity, T2DM, dyslipidaemia, sleep problems, smoking and psychiatric comorbidities. Attained age was used as underlying time scale.

5.4 ADHD Medication Use and Long-term Unemployment (Study IV)

Using Swedish national registers, we obtained a cohort of 12,875 middle-aged adults (41.5% females) with ADHD, in the labor force or under risk of unemployment at baseline, and followed them from January 1st, 2008 until death, emigration, or December 31st, 2013, whichever came first. During the follow-up period, 31.29% (31.65% for females) of included individuals with ADHD were never treated with ADHD medication, and 38.85% of individuals were recorded as long-term unemployment at least once across the follow-up period.

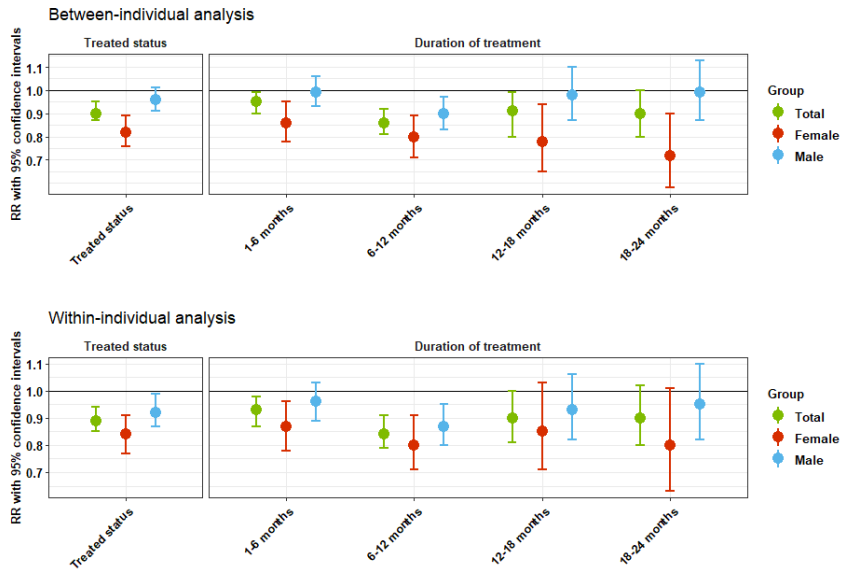

Figure 5.4 Associations between ADHD medication use and long-term unemployment status by using between- and within-individual design. In between-individual models, age, sex, birth country, highest education level and psychiatric comorbidities were adjusted. In with-in individual models, only age was adjusted in the GEE models.

In the between-individual analyses, we found protective associations between ADHD medication use and subsequent long-term unemployment, after adjusting for age, sex, birth country, highest education level and psychiatric comorbidities (RR=0.90, 95% CI=0.87-0.95). The associations were stronger in females (RR= 0.82, 95% CI=0.76-0.89) than males (RR, 0.96;

95% CI, 0.91-1.01). Comparable associations were observed in the within-individual analysis (RR=0.89, 95% CI, 0.85-0.94), where all the time-invariant confounders were implicitly adjusted for, such as sex, genetic background, and the severity of ADHD symptoms. Results were robust across subgroup and sensitivity analyses (Figure 5.4).

6. DISCUSSION

By using large-scale RWD, we investigated the potential risk factor (maternal pre-pregnancy overweight/obesity) and adverse outcomes for ADHD in adults, including dietary habits, CVDs and unemployment. The findings of this thesis indicate that the associations between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring could be largely explained by unmeasured familial confounding; More ADHD symptoms are associated with higher consumption of high-sugar and unhealthy dietary habits; ADHD predicts higher risk of a broad range of CVDs; and ADHD medications appear to be useful in reducing the rate of long-term unemployment. Our findings provide evidence for future researches on risk factors and adverse outcomes of ADHD and highlight the importance of triangulating research findings with a combination of various methods that differ in their inherent assumptions and limitations.

6.1 Interpretation of the main findings

6.1.1 Maternal pre-pregnancy overweight/obesity and the risk of ADHD in offspring

In study I, the pooled estimates from systematic review and meta-analysis revealed a significant positive association between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring. Even though the included studies were overall with high quality regarding study design, most of them failed to fully address all potential confounding. The adjusted effect size from the population-based cohort study was comparable with the results from meta-analysis. However, the substantially attenuated associations obtained from cousin and sibling comparisons, suggesting that the association between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring could be largely ascribed to unmeasured familial confounding. Therefore, the causality of the associations on population-level is not evident.

Such findings suggested maternal pre-pregnancy overweight/obesity probably represents, at least in part, a genetic predisposition to ADHD in offspring, which was consistent with previous findings from familial co-aggregation study[158] and genome-wide association study.[61] Familial environmental factors (e.g. lifestyle and SES) may also play a role on the observed associations, even though the effect may be limited in ADHD according to twin studies.[159]

The study also highlighted the importance of accounting for unmeasured familial confounders when exploring early-life risk factors for ADHD. By combining a systematic review and meta-analysis, the results showed that if unmeasured familial factors have been neglected, the causation may have been claimed erroneously. This reiterates that genetically informative study designs, such as quasi-experimental methods, may play a unique position in testing causal hypotheses in observational studies.

6.1.2 ADHD symptoms and dietary habits in adulthood

In study II, a nationwide population-based twin study with young to middle aged adults, we found both ADHD trait dimensions were associated with higher consumption of high-sugar food and unhealthy diets, but lower consumption of fruits, vegetables and healthy diets. We further found both genetic and non-shared environmental factors contributed substantially to the observed associations.

The results were consistent with previous large-scale studies based on children and adolescents, which indicted that ADHD was positively associated with unhealthy dietary habits (more refined sugar and saturated fat), and negatively associated with healthy dietary habits (more fruits and vegetables).[146] Nemours potential biological pathways, by which ADHD and other psychiatric disorders could associate with dietary intake, have been proposed during the past several years. For example, iron and zinc are cofactors for dopamine and norepinephrine production (essential factors in the etiology of ADHD), so unbalanced diet with lower levels of iron and zinc may further contribute to the development of ADHD.[160] Also, evidence suggested the microbiome–gut–brain axis (MGBA), the bidirectional relationship existent between the gut microbiome and the central nervous system, is involved in the pathophysiological mechanisms of neuro-inflammation and oxidative stress that give rise both to the ADHD core symptoms and to related comorbidities. [161] However, we cannot test the potential causality and the direction of observed associations in the current cross-sectional sample.

In the genetically informative analyses, we found the associations between ADHD symptoms and dietary habits could be explained by common genetic and environmental determinants, that is, these two traits share some etiological factors. Both ADHD and diet composition is strongly or moderately influenced by genetic factors, with the heritability estimates of 70%-80% for ADHD [59, 162, 163] and 20%-70% for diet composition, [164, 165]separately. Recent evidence from GWAS reported significant genetic

correlations of ADHD [61] and diet composition [166] with several metabolic traits (e.g. obesity), which also support for the hypothesis that there was genetic overlap between ADHD and diets. On the other hand, no previous studies have explores the non-shared environmental overlap between ADHD symptoms and dietary habits, but some lifestyle factors (e.g. screen time and physical activity) are reported to be associated with both traits.[93, 167] However, it remains to be investigated whether and how these factors contribute to the co-occurrence of ADHD symptoms and different dietary habits. Future longitudinal studies are needed to explore the causal associations between ADHD symptoms and dietary habits in adults.

6.1.3 ADHD as a novel risk factor for cardiovascular diseases

In study III, we explored the prospective associations between ADHD and a broad range of CVDs in over five million adults from Sweden. We found that adults with ADHD were more than twice as likely to develop at least one CVD, compared with those without ADHD, independently from well-established risk factors for CVDs, i.e. obesity and smoking, psychiatric comorbidities, medication use and cardiovascular family history. The increased risk was present across all types of cardiovascular diseases, but the strength of the associations was strongest for cardiac arrest, haemorrhagic stroke, and peripheral vascular disease/arteriosclerosis.

The potential causal associations between ADHD and CVDs were supported by a recent Mendelian Randomization (MR) study, reporting a direct causal effect of ADHD on coronary artery disease.[168] Further, many biological mechanisms are proposed to explain the observed associations, including immune system abnormalities,[169, 170] neuromodulator dysregulation,[171, 172] and dysregulation of the hypothalamic–pituitary–adrenal (HPA) axis.[173, 174] However, as suggested in a genetically informed study based on sibling pairs, the associations between ADHD and CVDs could alternatively be partly explained by shared etiological components.[13] It is clear that more studies are needed to explore the potential mechanisms linking ADHD and CVDs.

Notably, the observed strength of associations between ADHD and CVDs was largely comparable to estimates of associations between cardiovascular risk and serious mental illness (SMI), for example, schizophrenia,[175] depression[176] and bipolar disorders,[177] but stronger than those with anxiety,[178] obsessive-compulsive disorder[139] and stress-related disorders.[138] However, the associations between ADHD and CVDs are substantially understudied compared to other psychiatric disorders.[179]

These findings call for carefully monitoring cardiovascular health in adults with ADHD, and future studies are needed to confirm our findings.

6.1.4 ADHD medication use and long-term unemployment in working-aged adults with ADHD

In study IV, we explored the association between ADHD medication use and long-term unemployment among 12,875 middle-aged adults with ADHD. The results suggested that the use of ADHD medications associated with a 10% lower risk of the subsequent risk of long-term unemployment, when taking measured and unmeasured confounders into consideration. Therefore, treating patients with ADHD medication might help to improve their occupational outcomes.

Previous studies provided controversial findings regarding the associations between ADHD medication use and related occupational outcomes. For example, Norwegian studies found that early treatment of ADHD might have a beneficial effect on being in work in adulthood,[180, 181] but no differences on work productivity were reported between ADHD patients treated with or without atomoxetine in an US clinical sample.[182] However, our results are consistent with the available evidence on ADHD medication and other functional problems, indicating that ADHD medication is associated with a reduced risk of academic performance, social function.[183-185] It is important to note that the effect size was generally small in magnitude, but a reduction of 10% in the risk of long-term unemployment might translate into a substantial decrement of the economic burden at the societal level. Regardless, the small effect size suggested that other treatment programs, such as psychotherapy, are also needed to help individuals with ADHD in work-related settings.

6.2 Methodological considerations

6.2.1 Measurement error and misclassification

Measurement error refers to difference between the measured value of a variable and its true value, including random error and systematic error.[186] Misclassification is incorrectly assigning an individual, a value or an attribute into a category. The misclassification of exposure or outcome can be considered as either differential (systematic) or non-differential (random).[187]

6.2.1.1 ADHD and ADHD medications

We defined ADHD based on ICD codes from the NPR and ADHD medication prescriptions from the PDR (ATC-codes) in Study I and III. This register-based ADHD definition showed high specificity in Sweden.[13] However, ICD-based ADHD definition mainly capture the most severe cases, as only treatment seeking cases were included in medical registers. Therefore, false negatives cannot be avoided, whereas bias due to false positives is less possible. In Sweden, only physicians specialised in psychiatry or neurology are authorised to prescribe ADHD medication in Sweden, which supports that prescription of ADHD medication is a valid indicator of ADHD diagnoses.[13] It is important to note that information on ADHD medication is available in PDR from 2006, possibly leading to misclassification due to incomplete coverage.

ADHD trait dimensions (inattention and hyperactivity/impulsivity) used in Study II were based on self-rated checklist of DSM-IV. It is well-established that self-reported ADHD-symptoms are more likely to be underestimated, compared to reports from other informants.[58, 188] Therefore, the heritability estimates for self-reported ADHD symptoms in adults, and the reported associations between ADHD symptoms and dietary habits might be biased toward null. In addition, we used a dimensional measure of ADHD symptoms rather than as a discrete category when evaluating the relative genetic and environmental contribution to the observed associations in Study II. We identified a dimension specific overlap between ADHD and different dietary habits, with a stronger correlation of inattention with dietary habits than hyperactivity/impulsivity with dietary habits. Thus, Future genomic studies on the studied association may benefit from including information about ADHD symptom dimensions.

In Study IV, the main exposure is ADHD medication use in previous two years. The two-year limit for defining exposure was chosen in order to examine the accumulative effects of long-term ADHD medication use. Given the repeated nature of measurements and information on PDR started from 2006, the statistical power was reduced substantially when we used more than two years as the time window. A key concern of register-based definition of medication use is that the information might not reflect the actual consumption of medication, either due to non-adherence, or misspecification of the treatment periods. This type of misclassification would lead to an underestimation of the effect of medication.

6.2.1.2 Maternal pre-pregnancy BMI

In Study I, we used maternal early-pregnancy BMI as a proxy for pre-pregnancy BMI, and then defined overweight and obesity, accordingly. Even though pre-pregnancy weight were highly correlated with early gestational weight, evidence from systematic review (the first part of Study I) suggested the magnitude of the association may be overestimated.[189] However, the conclusion from sibling comparisons, no causal associations between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring, was unlikely to be affected.

6.2.1.3 Dietary habits

Dietary habits in Study II was measured by a Swedish FFQ with 94 typical food items. As a voluntary section of a comprehensive questionnaire, the response rate of FFQ had a low response rate of 36.8%. We found that females with lower SES and higher levels of behavior problems were more likely to respond to both the ADHD questionnaire and the FFQ, when comparing with those who only responded to the ADHD questionnaire; but the response rate was quite similar between DZ and MZ twins. Therefore, our findings from Study II may not generalize to the general population, especially for the phenotypic results.

6.2.1.4 CVDs

In Study III, CVDs were defined based on ICD-10 codes from the NPR and CDR, which may only capture the most severe cases, leading to an underestimation of the number of patients with less severe CVDs. Notably, given the median age of the study population at the end of the follow-up was 50.49 (range 31-73) years, we might have mostly captured early onset cases of CVDs. As it is well known that the risk for CVDs increases with age, especially for those who are over 65.[190]

6.2.1.5 Long-term unemployment

Long-term unemployment in Study IV was defined as having 90 or more days of unemployment in a calendar year, obtained from LISA register. This definition could only capture individuals who are actively looking for work through the official employment agency in Sweden. To minimize the misclassification caused by unemployed and non-economically active individuals, including studying, sick leave or in prison, we only included middle-

aged adults who belonging to the workforce. We, therefore, may underestimate the associations if some of the sick leave or disability pension was caused by ADHD.

6.2.2 Causal inference from observational studies

In this thesis, we aimed to test the causal effects in populations and testing causal hypotheses using large-scale RWD, therefore, efforts have been taken to control or reduce biases produced by measured or unmeasured confounding. In order for a variable to be a potential confounder, it needs to have the following three properties: (1) the variable should be a risk factor for the outcome (e.g. disease); (2) it must be associated with the exposure; and (3) it must not be an intermediate variable in the causal pathway from the exposure to the outcome.[191]

In traditional epidemiological studies, there are various ways to modify a study design to actively exclude or control for confounding variables, for example restriction (e.g. selecting participants with same age or gender to control potential confounding effect by sex and age) and matching (e.g. selecting a comparison group according to the distribution of one or more potential confounders).[192] To control for confounding in the analyses, stratification (Study II-III) and multivariate regression analyses (Study I-IV) are the most commonly used methods to control or adjust for the effect of confounding.[192] However, these approaches can only control for measured confounders, and relying heavily on the accuracy of a measurement.

On the other hand, quasi-experimental designs allow for stronger tests of causal inferences by taking into account unmeasured genetic and environmental confounding into account. In this thesis, we employed sibling and cousin comparisons (Study I) to control for unmeasured genetic and environmental factors shared by full siblings and cousins; and within-individual comparisons (Study IV) to adjust for all time-stable confounders during the follow-up period (e.g. genetic predisposition and early environments).

Obviously, none of the above-mentioned approaches is likely to completely eliminate bias from confounding, as they are reliant upon varying assumptions, each has the potential to address differing sources of confounding to differing extents. Therefore, future efforts should attempt to triangulate research findings based on a combination of different designs that differ in their underlying assumptions and limitations.[124, 193]

6.2.3 Generalizability

All studies involved in this thesis used data from the Swedish population. The rates of ADHD and ADHD medication use, and other key variables studied in this thesis may differ from those in other countries. Given the specific economic, political, and cultural factors, the dietary habits (Study II) and unemployment rate (Study IV) could be very specific to people in Sweden and other Nordic countries with similar social and cultural background. Generalization of the findings to other countries should be made with cautions.

6.3 Challenges and opportunities with large-scale real-world data

The availability of large-scale RWD provides important insights and opportunities in epidemiology in real clinical conditions. We could conduct studies with better designs to answer complicated research questions, which are previously thought infeasible, with the development of sophisticated methods and new analytical capabilities.[194, 195] However, there are still challenges need to be considered when deriving evidence from RWD.[122] First, data quality is fundamental in epidemiology. Large-scale RWD should be carefully assessed with five main dimensions of data quality, including availability, usability, reliability, relevance, and presentation quality.[196] Assessing data quality appropriately is essential to draw valid conclusions by RWD. Second, combining data from different countries is more and more common in epidemiology studies, especially when the intervention or the outcome of interest is rare so that very large sample size is required. Additionally, observational data is usually needed to be combined with other data sources, for example biobank and omics data, to answer causal research questions. Therefore, it will be important to create national and international level data-sharing network to improve the coordination between different organizations and further develop means to standardize and harmonize RWD management, in order to translate RWD into evidence.[197] However, the technical, logistical ethical and legal challenges are difficult to overcome due to issues of data access, data security, and potential conflicts of interest.[198] Finally, as the sample size and number of variables of large-scale RWD is growing rapidly to the millions, more advanced studies designs, more complicated analytic methods and more carefully interpretations are required.[199] For example, almost all comparisons of interest in very large samples become statistically significant, but most of them may not have meaningful implications for clinical practice or policy. Therefore, in order to identify significant associations in real world

among hundreds or thousands of comparisons, it is essential to develop more sophisticated variable selection and machine learning methods.[199]

6.4 Ethical considerations

Large-scale RWD was used in this thesis work. Study I (the cohort part), III and IV in this thesis were based on data from a record linkage of Swedish national registers, except for the meta-analysis part of Study I. According to current Swedish regulation, informed consent is not required when register data is used for the purpose of research. Even though register data used in the thesis was anonymized before being used for research, it is still sensitivity personal data, as described in the Personal Data Act in Sweden (Swedish abbreviation: PUL) and the European General Data Protection Regulation (GDPR). Therefore, register-based research should be performed in the corresponding data protection framework to ensure personal integrity of research participants. Study I, III, IV were approved by the Regional Ethical Review board in Stockholm, Sweden (DNR: 2013/862-31/5).

In Study II, data was obtained from large-scale questionnaire data: the Study of Twin Adults: Genes and Environment (STAGE), which included 60% of all participants from Swedish Twin Register. All participants provided informed consent. The project has been reviewed and approved by the Regional Ethics Committee at the Karolinska Institutet, Stockholm, Sweden (DNR: 03–224).

In addition, results from observational studies should be clearly stated and interpreted. It is not uncommon among the public to misinterpret potential risk factors as deterministic cause, as a result, individuals, families and patient groups could potentially be labelled and stigmatized. Therefore, we always emphasize the complexity of ADHD etiology in research papers to minimize social and ethical harms.

7. CONCLUSION

This thesis has demonstrated how epidemiologic research with various designs and analytical approaches could answer causal research questions by using large-scale RWD. The main conclusions gained from the four studies are as follows:

I. The observed association between maternal pre-pregnancy overweight/obesity and ADHD in offspring is largely ascribable to unmeasured familial confounding but not a strong causal relationship, highlighting the importance of accounting for unmeasured familial confounders in risk-factor studies of ADHD in offspring.

II. There are phenotypic and genetic correlations between ADHD symptoms and unhealthy dietary habits in adulthood, but future studies are needed to further explore the direction and causality of the observed associations between the two traits.

III. ADHD is a novel risk factor for a wide range of CVDs, independent from traditional risk cardiovascular risk factors, psychiatric comorbidity, and psychotropic medication treatment. Additional studies are needed to investigate the mechanisms underlying the association between ADHD and cardiovascular diseases.

IV. The use of ADHD medication is associated with lower risk of subsequent long-term unemployment, especially in females. This finding provides evidence for developing treatment plans for working-aged adults with ADHD.

8. FUTURE PERSPECTIVES

This thesis addresses a range of research questions on the etiology and long-term consequences of ADHD. Future studies are warranted to explore the potential causal pathways and underlying mechanisms by which ADHD is associated with maternal overweight/obesity, dietary habits, CVDs and long-term unemployment.

In the systematic review, we found all studies on the associations between maternal pre-pregnancy overweight/obesity and risk of ADHD in offspring were conducted in Europe and the US, therefore future studies are needed to examine the associations using other samples, especially in countries outside Europe and the US. Additional studies with different study designs, e.g. intergenerational Mendelian randomization or children-of-twins design, are also needed to triangulate our findings.

Although ADHD symptoms were associated with unhealthy dietary habits, the direction of the observed associations remains unclear. Future longitudinal studies (e.g., cross-lagged models) and other study designs (e.g., Mendelian randomization) are needed to replicate this finding in children and adolescents, and to explore the directions of the associations in adults.

We found adults with ADHD were more than twice as likely to develop at least one CVD, compared with those without ADHD, independently from traditional risk factors for CVD and psychotropic medications. However, compared to other psychiatric disorders, the associations between ADHD and cardiovascular diseases are substantially understudied. Our finding needs to be tested in different settings and the underlying mechanisms linking ADHD with CVDs are yet to be elucidated in future studies.

In study IV, ADHD medication use in females showed stronger negative associations with long-term unemployment than in males. However, it is largely unknown whether there are true differences in the effectiveness of ADHD medications among males and females, which need to be explored in future studies.

Taken together, findings from the thesis highlight the need for future studies with various study designs, to fully understand the etiology and long-term health outcomes of ADHD across the lifespan.

9. ACKNOWLEDGEMENTS

My PhD journey has been challenging, but also the most shining and rewarding part of my working life so far. I would like to express my deepest sincere gratitude to everyone who have accompanied me through this adventure and made all these things possible.

Henrik Larsson, my main supervisor. Thank you for offering the special and precious opportunity to study with your excellent research groups in both ORU and KI. Staying with my family means a lot to me. Your generous support have made my life much easier than planned. I thank you, from the bottom of my heart, for all the encouragement and trust throughout the journey, for having me to be part of different training programs and research projects, for sharing your vast knowledge on psychiatric epidemiology, for always being there when I was lost, for being such an outstanding professor and group leader! I could not ask for a more perfect and reliable supervisor than you.

Zheng Chang, my co-supervisor. You are a legend and role model to me. I appreciated all our discussions about research and life. Thank you for providing all the support and guidance that a guest doctoral student may need, which made me feel so safe and comfortable in KI.

Qi Chen, my co-supervisor. I thank you deeply for your help to start with my first doctoral project, for guiding me thoroughly in statistical methods, scientific writing and preparing a rebuttal letter. You are so warm and always caring about my personal life. I really appreciate to have a supervisor and a friend like you!

Mark Taylor, my co-supervisor. Thank you for your generous help on the twin studies in my thesis. I have learnt a great dealt about the complex twin model fitting analyses from you. These studies would not be possible without your contribution.

To my coauthors: **Tyra Lagerberg, Samuele Cortese, Mina A Rosenqvist, Catarina Almqvist, Brian M D'Onofrio, Tor-Arne Hegvik, Catharina Hartman, Katarina Bälter, Ralf Kuja-Halkola, Alejandro Arias-Vásquez, Ashley E Tate, Andreas Jangmo, Le Zhang, Lars Magnus Andersson, Tamara Werner-Kiechle, Ewa Ahnemark, Miguel Garcia-Argibay, Ebba Du Rietz, Maja Dobrosavljevic, Isabell Brikell, Tomas Jernberg, and Marco Solmi.** Thank you for the wonderful collaboration and for your invaluable inputs. I have learnt a lot from you.

To my brilliant friends and colleagues in **Larsson-ORU** group: **Anneli Andersson, Maja Dobrosavljevic, Sofi Oskarsson, Carmen Solares Canal,**

Miguel Garcia-Argibay, Rickard Ahlberg, Anna-Karin Ångström, Rebecca Siponen. Thank you for all the inspirational discussions and constructive feedback, and for all the joyful moments we shared.

It is my great honor to work so closely with **MEB-Psychepi** group in KI, including but not limited to: **Paul Lichtenstein, Ebba Du Rietz, Isabell Brikell, Laura Ghirardi, Mina A Rosenqvist, Ralf Kuja-Halkola, Agnieszka Butwicka, Andreas Jangmo, Marica Leone, Sarah Bergen, Vide Ohlsson Gotby, Erik Pettersson, and Gunilla Sonnebring.** I have learned a lot from working with you.

The fantastic **Eat2beNice** team**: Liv Grimstvedt Kvalvik, Lizanne Schweren, Berit Skretting Solberg, Jan Haavik** and **Catharina Hartman,** my mentor. Thank you for the fruitful collaboration from our work package 1, and I also appreciate the great time we shared together in Madrid, Brussel and Lisbon.

To the '**Big Data Analytics**' Book Group: **Jiangwei Sun, Yunzhang Wang, Zheng Ning, Yinxi Wang, Xia Li, Weiwei Bian,** I still cannot believe that we have published a thick book in China! Thank you for sharing your years of experiences and knowledge on each software, the book already helped me a lot in my studies.

The awesome professor couple, **Yudi Pawitan & Marie Reilly.** Thank you for sharing your work and life experiences, and a lot of interesting food, with us! ***Yudi*****,** I will never forget how you encouraged me to make the correct decision in the darkest winter of 2019, your hugs and powerful words brightened up my days during the most difficult time. You made me a stronger person!

Big thanks to all the administrative staff at School of Medical Sciences (MV) of ORU, for making our department such a great and warm place to work and study. **Ina Johansson,** thank you for helping me throughout my half-time review and public defense, and many other important paper work during the past years.

My dearest friends in KI and ORU: **Xiaoying Kang (KK), Can Cui, Jingyan He, Jiangnan Liu, Tong Jiao, Qian Yang, Jie Song, Shuyang Yao, Ji Zhang, Cen Chen, Ge Bai, Bowen Tang, Nanbo Zhu, Guannan Zhou, Le Zhang, Honghui Yao, Chenxi Qin, Dang Wei, Hua Chen, Jet Termorshuizen, Kejia Hu, Linghua Kong, Mailin Zhou, Wenjiang Deng, Chen Wang, Tianyang Zhang, Ruyue Zhang, Yufeng Chen, Jingru Yu, Xinhe Mao, Yun Du, Tor-Arne Hegvik, Xueli Zhang, Kaya, and many other great friends.** Thank you for filling my life with pleasure and amusement and adding so many good memories and times of happiness to my PhD journey.

KK, there are no words that can express my gratitude for having such a thoughtful and kind friend like you! You have been there for me and my family, through so many ups and downs, and you are always so reliable. Love you forever!

A special thanks to our lovely family friends in Sweden: **Shihua Sun & Qian Zhou & Chengan; Fen Yang & Jianjian Gao & Leo; Yiliu Chen & Yibo Wang & Andy, Ruoqing Chen & Yiqiang Zhan & Muning**. Thank you for celebrating so many important days and festivals with us, for sharing your experiences of being a strong parent, for all the delicious food of course! **Shihua,** Thank you for your kind help even before we arrived!

My past and current officemates and my close colleagues: **Hong Xu, Tyra Lagerberg, Ashley E Tate, and Shengxin Liu.** Thank you for sharing the PhD journey and all the interesting discussions about research and life. You made me feel so welcomed in MEB!

Maohua Miao, Wei Yuan and Hong Liang, my excellent supervisors from my master study. Thank you for guiding me in the world of epidemiology and for your love, trust and support over the years.

To my beloved family, thank you for being the best family I could ever ask for. Thank you, **mom and dad**, for being my biggest supporter. Your unconditional love and support has made me all I am today. My **parents-in-law**, thank you for so much love, care, and joy you have given me all these years! **Jiangwei**, my husband and best friend. It is the power of your love and belief that has made my dreams come true! Thank you for being a great husband, father and you! We have shared so many beautiful moments together. I am eagerly looking forward to future adventures with you! **Ethan**, our precious son. My life has never been so beautiful since the day that you came into it. Thank you for being the best part of my life. You are everything to me. I love you all!!

10. REFERENCES

1. Faraone, S.V., et al., *Attention-deficit/hyperactivity disorder*. Nat Rev Dis Primers, 2015. **1**: p. 15020.
2. Faraone, S.V., et al., *Attention-Deficit/Hyperactivity Disorder in adults: an overview*. Biological Psychiatry, 2000. **48**(1): p. 9-20.
3. Sibley, M.H., et al., *Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity*. J Child Psychol Psychiatry, 2017. **58**(6): p. 655-662.
4. Faraone, S.V., et al., *The World Federation of ADHD International Consensus Statement: 208 Evidence-based conclusions about the disorder*. Neuroscience & Biobehavioral Reviews, 2021. **128**: p. 789-818.
5. Dalsgaard, S., et al., *Conduct problems, gender and adult psychiatric outcome of children with attention-deficit hyperactivity disorder*. British Journal of Psychiatry the Journal of Mental Science, 2002. **181**(9): p. 416.
6. Quinn, P.D., et al., *ADHD Medication and Substance-Related Problems*. American Journal of Psychiatry: p. appi.ajp.2017.1.
7. Marco, et al., *Attention deficit-hyperactivity disorder in adult bipolar disorder patients*.
8. Cole, J., et al., *Genetic Overlap Between Measures of Hyperactivity/Inattention and Mood in Children and Adolescents*. Journal of the American Academy of Child & Adolescent Psychiatry. **48**(11): p. 1094-1101.
9. Ghirardi, L., et al., *The familial co-aggregation of ASD and ADHD: a register-based cohort study*. Molecular Psychiatry.
10. Posner, J., G.V. Polanczyk, and E. Sonuga-Barke, *Attention-deficit hyperactivity disorder*. Lancet, 2020.
11. Björk, A., et al., *Health, lifestyle habits, and physical fitness among adults with ADHD compared with a random sample of a Swedish general population*. Society, Health & Vulnerability, 2018. **9**(1): p. 1553916.
12. Li, L., et al., *Attention - deficit/hyperactivity disorder symptoms and dietary habits in adulthood: A large population - based twin study in Sweden*. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 2020. **183**(8): p. 475-485.
13. Du Rietz, E., et al., *Mapping phenotypic and aetiological associations between ADHD and physical conditions in adulthood in Sweden: a genetically informed register study*. Lancet Psychiatry, 2021. **8**(9): p. 774-783.

14. Chang, Z., et al., *Risks and Benefits of Attention-Deficit/Hyperactivity Disorder Medication on Behavioral and Neuropsychiatric Outcomes: A Qualitative Review of Pharmacoepidemiology Studies Using Linked Prescription Databases*. Biol Psychiatry, 2019. **86**(5): p. 335-343.
15. Nichols, A., J. Mitchell, and S. Lindner, *Consequences of long-term unemployment*. Washington, DC: The Urban Institute, 2013.
16. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, 2013.
17. *International Statistical Classification of Diseases and Related Health Problems 10th Revision*. 2016; Available from: https://icd.who.int/browse10/2016/en.
18. Nigg, J.T., R. Tannock, and L.A. Rohde, *What is to be the fate of ADHD subtypes? An introduction to the special section on research on the ADHD subtypes and implications for the DSM-V*. J Clin Child Adolesc Psychol, 2010. **39**(6): p. 723-5.
19. *International statistical classification of diseases and related health problems (11th Revision)*. 2018.
20. Power, T.J., et al., *The Predictive Validity of Parent and Teacher Reports of ADHD Symptoms*. Journal of Psychopathology and Behavioral Assessment, 1998. **20**(1): p. 57-81.
21. Epstein, J.N. and S.H. Kollins, *Psychometric properties of an adult ADHD diagnostic interview*. Journal of Attention Disorders, 2006. **9**(3): p. 504-514.
22. Adler, L.A., et al., *Validity of pilot Adult ADHD Self-Report Scale (ASRS) to rate adult ADHD symptoms*. Annals of Clinical Psychiatry, 2006. **18**(3): p. 145-148.
23. Sayal, K., et al., *ADHD in children and young people: prevalence, care pathways, and service provision*. The Lancet Psychiatry, 2017.
24. Biederman, J. and S.V. Faraone, *Attention-deficit hyperactivity disorder*. Lancet, 2005. **366**(9481): p. 237-248.
25. Polanczyk, G., et al., *The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis*. Am J Psychiatry. **164**(6): p. 942-948.
26. Caye, A., et al., *Attention-Deficit/Hyperactivity Disorder Trajectories From Childhood to Young Adulthood: Evidence From a Birth Cohort Supporting a Late-Onset Syndrome*. 2016. **73**(7): p. 705.
27. Simon, V., et al., *Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis*. The British Journal of Psychiatry, 2009. **194**(3): p. 204-211.

28. Michielsen, M., et al., *Prevalence of attention-deficit hyperactivity disorder in older adults in The Netherlands*. The British Journal of Psychiatry, 2012. **201**(4): p. 298-305.
29. Guldberg-Kjär, T. and B. Johansson, *Old people reporting childhood AD/HD symptoms: Retrospectively self-rated AD/HD symptoms in a population-based Swedish sample aged 65–80*. Nordic journal of psychiatry, 2009. **63**(5): p. 375-382.
30. Faraone, S.V., J. Biederman, and E. Mick, *The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies*. Psychol Med, 2006. **36**(2): p. 159-65.
31. Asherson, P., et al., *Under Diagnosis of Adult ADHD: Cultural Influences and Societal Burden*. Journal of Attention Disorders. **16**(5 Suppl): p. 20S-38S.
32. Ginsberg, Y., et al., *Underdiagnosis of Attention-Deficit/Hyperactivity Disorder in Adult Patients: A Review of the Literature*. Primary Care Companion to the Journal of Clinical Psychiatry, 2014. **16**(3).
33. Taylor, E., et al., *European clinical guidelines for hyperkinetic disorder -- first upgrade*. Eur Child Adolesc Psychiatry, 2004. **13 Suppl 1**: p. I7-30.
34. Fredriksen, M., et al., *Long-term efficacy and safety of treatment with stimulants and atomoxetine in adult ADHD: a review of controlled and naturalistic studies*. Eur Neuropsychopharmacol, 2013. **23**(6): p. 508-27.
35. Faraone, S.V. and S.J. Glatt, *A comparison of the efficacy of medications for adult attention-deficit/hyperactivity disorder using meta-analysis of effect sizes*. J Clin Psychiatry, 2010. **71**(6): p. 754-63.
36. Faraone, S.V., et al., *Attention-deficit/hyperactivity disorder in adults: An overview*. Biological Psychiatry, 2000. **48**(1): p. 9-20.
37. Biederman, J., *Attention-deficit/hyperactivity disorder: a selective overview*. Biol Psychiatry, 2005. **57**(11): p. 1215-20.
38. Faraone, S.V. and J. Buitelaar, *Comparing the efficacy of stimulants for ADHD in children and adolescents using meta-analysis*. Eur Child Adolesc Psychiatry, 2010. **19**(4): p. 353-64.
39. Lu, Y., et al., *Association between medication use and performance on higher education entrance tests in individuals with attention-deficit/hyperactivity disorder*. JAMA psychiatry, 2017. **74**(8): p. 815-822.
40. Halmoy, A., et al., *Occupational outcome in adult ADHD: impact of symptom profile, comorbid psychiatric problems, and treatment: a cross-sectional study of 414 clinically diagnosed adult ADHD patients*. J Atten Disord, 2009. **13**(2): p. 175-87.

41. de Graaf, R., et al., *The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the WHO World Mental Health Survey Initiative*. Occup Environ Med, 2008. **65**(12): p. 835-42.
42. Pfiffner, L.J. and L.M. Haack, *Behavior management for school-aged children with ADHD*. Child and Adolescent Psychiatric Clinics, 2014. **23**(4): p. 731-746.
43. Evans, S.W., et al., *Middle school–based and high school–based interventions for adolescents with ADHD*. Child and Adolescent Psychiatric Clinics, 2014. **23**(4): p. 699-715.
44. Daley, D., et al., *Behavioral interventions in attention-deficit/hyperactivity disorder: a meta-analysis of randomized controlled trials across multiple outcome domains*. Journal of the American Academy of Child & Adolescent Psychiatry, 2014. **53**(8): p. 835-847. e5.
45. Faraone, S.V., et al., *Attention-deficit/hyperactivity disorder*. Nature Reviews Disease Primers, 2015. **1**: p. 15020.
46. Heilskov Rytter, M.J., et al., *Diet in the treatment of ADHD in children - a systematic review of the literature*. Nord J Psychiatry, 2015. **69**(1): p. 1-18.
47. Pelsser, L.M., et al., *Diet and ADHD, Reviewing the Evidence: A Systematic Review of Meta-Analyses of Double-Blind Placebo-Controlled Trials Evaluating the Efficacy of Diet Interventions on the Behavior of Children with ADHD*, in *PLoS One*. 2017.
48. Adler, L.D. and A.A. Nierenberg, *Review of medication adherence in children and adults with ADHD*. Postgrad Med, 2010. **122**(1): p. 184-91.
49. Biederman, J., et al., *Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder: Patterns of comorbidity in probands and relatives in psychiatrically and pediatrically referred samples*. Archives of general psychiatry, 1992. **49**(9): p. 728-738.
50. Biederman, J., et al., *Family-genetic and psychosocial risk factors in DSM-III attention deficit disorder*. Journal of the American Academy of Child & Adolescent Psychiatry, 1990. **29**(4): p. 526-533.
51. Chen, Q., et al., *Familial aggregation of attention-deficit/hyperactivity disorder*. J Child Psychol Psychiatry, 2016.
52. Larsson, H., et al., *The heritability of clinically diagnosed attention deficit hyperactivity disorder across the lifespan*. Psychological Medicine, 2014. **44**(10): p. 2223-2229.
53. Larsson, H., P. LICHTENSTEIN, and J.-O. LARSSON, *Genetic Contributions to the Development of ADHD Subtypes From*

Childhood to Adolescence. Journal of the American Academy of Child & Adolescent Psychiatry, 2006. **45**(8): p. 973-981.
54. Nikolas, M.A. and S.A. Burt, *Genetic and environmental influences on ADHD symptom dimensions of inattention and hyperactivity: A meta-analysis*. Journal of Abnormal Psychology, 2010. **119**(1): p. 1-17.
55. Greven, C.U., et al., *A Longitudinal Twin Study on the Association Between Inattentive and Hyperactive-Impulsive ADHD Symptoms*. 2011. **39**(5): p. 623-632.
56. Greven, C.U., F.V. Rijsdijk, and R. Plomin, *A Twin Study of ADHD Symptoms in Early Adolescence: Hyperactivity-impulsivity and Inattentiveness Show Substantial Genetic Overlap but Also Genetic Specificity*. 2011. **39**(2): p. 265-275.
57. Larsson, H., et al., *Childhood attention-deficit hyperactivity disorder as an extreme of a continuous trait: A quantitative genetic study of 8,500 twin pairs*. Journal of Child Psychology & Psychiatry & Allied Disciplines, 2011. **53**(1): p. 73-80.
58. Brikell, I., R. Kuja-Halkola, and H. Larsson, *Heritability of attention-deficit hyperactivity disorder in adults*. Am J Med Genet B Neuropsychiatr Genet, 2015. **168**(6): p. 406-413.
59. Taylor, M.J., et al., *Association of Genetic Risk Factors for Psychiatric Disorders and Traits of These Disorders in a Swedish Population Twin Sample*. JAMA Psychiatry, 2019. **76**(3): p. 280-289.
60. Middeldorp, C.M., et al., *A Genome-Wide Association Meta-Analysis of Attention-Deficit/Hyperactivity Disorder Symptoms in Population-Based Pediatric Cohorts*. Journal of the American Academy of Child & Adolescent Psychiatry, 2016. **55**(10).
61. Demontis, D., et al., *Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder*. Nat Genet, 2018.
62. Thapar, A., et al., *Practitioner Review: What have we learnt about the causes of ADHD?* Journal of Child Psychology and Psychiatry, 2013.
63. Thapar, A. and M. Rutter, *Do prenatal risk factors cause psychiatric disorder? Be wary of causal claims*. Br J Psychiatry, 2009. **195**(2): p. 100-101.
64. Lahey, B.B., B.M. D'Onofrio, and I.D. Waldman, *Using epidemiologic methods to test hypotheses regarding causal influences on child and adolescent mental disorders*. 2009.
65. Banerjee, T.D., F. Middleton, and S.V. Faraone, *Environmental risk factors for attention - deficit hyperactivity disorder*. Acta paediatrica, 2007. **96**(9): p. 1269-1274.

66. West and S. G., *Alternatives to Randomized Experiments*. Current Directions in Psychological Science, 2009. **18**(5): p. 299-304.
67. Benjamin B. Lahey, B.M.D.O., *All in the Family: Comparing Siblings to Test Causal Hypotheses Regarding Environmental Influences on Behavior*. 2010. **19**(5): p. 319-323.
68. D'Onofrio, B.M., et al., *Critical need for family-based, quasi-experimental designs in integrating genetic and social science research*. Am J Public Health, 2013. **103 Suppl 1**: p. S46-55.
69. Skoglund, C., et al., *Familial Confounding of the Association between Maternal Smoking During Pregnancy and ADHD in Offspring*. Journal of Child Psychology & Psychiatry & Allied Disciplines, 2014. **55**(1): p. 61-68.
70. Langley, K., et al., *Maternal and Paternal Smoking During Pregnancy and Risk of ADHD Symptoms in Offspring: Testing for Intrauterine Effects*. American Journal of Epidemiology, 2012. **176**(3): p. 261-268.
71. Rice, F., et al., *The links between prenatal stress and offspring development and psychopathology: disentangling environmental and inherited influences*. Psychological Medicine, 2010. **40**.
72. Hultman, C.M., et al., *Birth weight and attention-deficit/hyperactivity symptoms in childhood and early adolescence: a prospective Swedish twin study*. J Am Acad Child Adolesc Psychiatry, 2007. **46**(3): p. 370-377.
73. Pettersson, E., et al., *Birth weight as an independent predictor of ADHD symptoms: a within-twin pair analysis*. J Child Psychol Psychiatry, 2015. **56**(4): p. 453-9.
74. D'Onofrio, B.M., et al., *Paternal Age at Childbearing and Offspring Psychiatric and Academic Morbidity*. Jama Psychiatry, 2014. **71**(4): p. 432.
75. Larsson, H., et al., *Family income in early childhood and subsequent attention deficit/hyperactivity disorder: a quasi-experimental study*. J Child Psychol Psychiatry, 2014. **55**(5): p. 428-35.
76. Adane, A.A., G.D. Mishra, and L.R. Tooth, *Maternal pre-pregnancy obesity and childhood physical and cognitive development of children: a systematic review*. Int J Obes (Lond), 2016. **40**(11): p. 1608-1618.
77. Sanchez, C.E., et al., *Maternal pre-pregnancy obesity and child neurodevelopmental outcomes: a meta-analysis*. Obes Rev, 2018. **19**(4): p. 464-484.
78. Van Lieshout, R.J., V.H. Taylor, and M.H. Boyle, *Pre-pregnancy and pregnancy obesity and neurodevelopmental outcomes in offspring: a systematic review*. Obes Rev, 2011. **12**(5): p. e548-59.

79. O'Donnell, K.J. and M.J. Meaney, *Fetal Origins of Mental Health: The Developmental Origins of Health and Disease Hypothesis*. Am J Psychiatry, 2017. **174**(4): p. 319-328.
80. Edlow, A.G., *Maternal obesity and neurodevelopmental and psychiatric disorders in offspring*. Prenat Diagn, 2017. **37**(1): p. 95-110.
81. Chen, Q., et al., *Shared familial risk factors between attention-deficit/hyperactivity disorder and overweight/obesity - a population-based familial coaggregation study in Sweden*. J Child Psychol Psychiatry, 2017. **58**(6): p. 711-718.
82. Chen, Q., et al., *Attention-deficit/hyperactivity disorder and clinically diagnosed obesity in adolescence and young adulthood: a register-based study in Sweden*. Psychol Med, 2018: p. 1-9.
83. Malik, V.S., W.C. Willett, and F.B. Hu, *Global obesity: trends, risk factors and policy implications*. Nature Reviews Endocrinology, 2012. **9**: p. 13.
84. Sullivan, E.L., E.K. Nousen, and K.A. Chamlou, *Maternal high fat diet consumption during the perinatal period programs offspring behavior*. Physiol Behav, 2014. **123**: p. 236-42.
85. Musser, E.D., et al., *Maternal prepregnancy body mass index and offspring attention-deficit/hyperactivity disorder: a quasi-experimental sibling-comparison, population-based design*. J Child Psychol Psychiatry, 2017. **58**(3): p. 240-247.
86. Chen, Q., et al., *Maternal pre-pregnancy body mass index and offspring attention deficit hyperactivity disorder: a population-based cohort study using a sibling-comparison design*. Int J Epidemiol, 2014. **43**(1): p. 83-90.
87. Frisell, T., et al., *Sibling comparison designs: bias from non-shared confounders and measurement error*. Epidemiology, 2012. **23**(5): p. 713-20.
88. Sudan, M., et al., *Complexities of sibling analysis when exposures and outcomes change with time and birth order*. J Expo Sci Environ Epidemiol, 2014. **24**(5): p. 482-8.
89. Sjolander, A. and J. Zetterqvist, *Confounders, Mediators, or Colliders: What Types of Shared Covariates Does a Sibling Comparison Design Control For?* Epidemiology, 2017. **28**(4): p. 540-547.
90. Nigg, J.T., et al., *Attention-deficit/hyperactivity disorder (ADHD) and being overweight/obesity: New data and meta-analysis*. Clinical psychology review, 2016. **43**: p. 67-79.
91. Cortese, S., et al., *Attention-deficit/hyperactivity disorder (ADHD) and obesity: a systematic review of the literature*. Critical reviews in food science and nutrition, 2008. **48**(6): p. 524-537.

92. Cortese, S., et al., *Association between ADHD and obesity: a systematic review and meta-analysis*. American Journal of Psychiatry, 2015. **173**(1): p. 34-43.
93. Mian, A., et al., *Children's Attention-Deficit/Hyperactivity Disorder Symptoms Predict Lower Diet Quality but Not Vice Versa: Results from Bidirectional Analyses in a Population-Based Cohort*. J Nutr, 2019. **149**(4): p. 642-648.
94. Angold, A., E.J. Costello, and A. Erkanli, *Comorbidity*. J Child Psychol Psychiatry, 1999. **40**.
95. Singh and Ilina, *Beyond polemics: science and ethics of ADHD*. 2008. **9**(12): p. 957-964.
96. Friedrichs, B., et al., *Coexisting Psychiatric Problems and Stressful Life Events in Adults With Symptoms of ADHD--A Large Swedish Population-Based Study of Twins*. Journal of Attention Disorders, 2012. **16**(1): p. 13-22.
97. Instanes, J.T., et al., *Adult ADHD and Comorbid Somatic Disease: A Systematic Literature Review*. J Atten Disord, 2018. **22**(3): p. 203-228.
98. Mitchell, A.J., et al., *Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders--a systematic review and meta-analysis*. Schizophr Bull, 2013. **39**(2): p. 306-18.
99. Czepielewski, L., et al., *Bipolar disorder and metabolic syndrome: a systematic review*. Braz J Psychiatry, 2013. **35**(1): p. 88-93.
100. Vancampfort, D., et al., *Metabolic syndrome and metabolic abnormalities in patients with major depressive disorder: a meta-analysis of prevalences and moderating variables*. Psychol Med, 2014. **44**(10): p. 2017-28.
101. Isomura, K., et al., *Metabolic and Cardiovascular Complications in Obsessive-Compulsive Disorder: A Total Population, Sibling Comparison Study With Long-Term Follow-up*. Biol Psychiatry, 2018. **84**(5): p. 324-331.
102. Wang, J., et al., *The metabolic syndrome predicts cardiovascular mortality: a 13-year follow-up study in elderly non-diabetic Finns*. Eur Heart J, 2007. **28**(7): p. 857-64.
103. Li, Z., et al., *The Cohort Study on Prediction of Incidence of All-Cause Mortality by Metabolic Syndrome*. PLoS One, 2016. **11**(5): p. e0154990.
104. van Dijk, G. and B. Buwalda, *Neurobiology of the metabolic syndrome: an allostatic perspective*. Eur J Pharmacol, 2008. **585**(1): p. 137-46.
105. Vancampfort, D., et al., *Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic*

disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. World Psychiatry, 2015. **14**(3): p. 339-47.
106. Mick, E., D.D. McManus, and R.J. Goldberg, *Meta-analysis of increased heart rate and blood pressure associated with CNS stimulant treatment of ADHD in adults*. European Neuropsychopharmacology, 2013. **23**(6): p. 534-541.
107. Semeijn, E.J., et al., *Attention-deficit/hyperactivity disorder, physical health, and lifestyle in older adults*. J Am Geriatr Soc, 2013. **61**(6): p. 882-7.
108. Ambrose, J.A. and R.S. Barua, *The pathophysiology of cigarette smoking and cardiovascular disease: an update*. Journal of the American college of cardiology, 2004. **43**(10): p. 1731-1737.
109. Malhotra, A. and J. Loscalzo, *Sleep and cardiovascular disease: an overview*. Progress in cardiovascular diseases, 2009. **51**(4): p. 279.
110. Powell-Wiley, T.M., et al., *Obesity and cardiovascular disease: a scientific statement from the American Heart Association*. Circulation, 2021. **143**(21): p. e984-e1010.
111. Einarson, T.R., et al., *Prevalence of cardiovascular disease in type 2 diabetes: a systematic literature review of scientific evidence from across the world in 2007–2017*. Cardiovascular Diabetology, 2018. **17**(1): p. 83.
112. Hedayatnia, M., et al., *Dyslipidemia and cardiovascular disease risk among the MASHAD study population*. Lipids in Health and Disease, 2020. **19**(1): p. 42.
113. Rasmussen, P. and C. Gillberg, *Natural outcome of ADHD with developmental coordination disorder at age 22 years: A controlled, longitudinal, community-based study*. Journal of the American Academy of Child and Adolescent Psychiatry, 2000. **39**(11): p. 1424-1431.
114. Adler, L.A., et al., *Functional outcomes in the treatment of adults with ADHD*. J Atten Disord, 2008. **11**(6): p. 720-7.
115. Safren, S.A., et al., *Life impairments in adults with medication-treated ADHD*. J Atten Disord, 2010. **13**(5): p. 524-31.
116. Gjervan, B., et al., *Functional impairment and occupational outcome in adults with ADHD*. J Atten Disord, 2012. **16**(7): p. 544-52.
117. Barkley, R.A. and K.R. Murphy, *Impairment in occupational functioning and adult ADHD: the predictive utility of executive function (EF) ratings versus EF tests*. Arch Clin Neuropsychol, 2010. **25**(3): p. 157-73.
118. Sobanski, E., et al., *Psychiatric comorbidity and functional impairment in a clinically referred sample of adults with attention-*

deficit/hyperactivity disorder (ADHD). Eur Arch Psychiatry Clin Neurosci, 2007. **257**(7): p. 371-7.
119. Makady, A., et al., *What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews*. Value in Health, 2017. **20**(7): p. 858-865.
120. Taipale, H., et al., *Representation and Outcomes of Individuals With Schizophrenia Seen in Everyday Practice Who Are Ineligible for Randomized Clinical Trials*. JAMA Psychiatry, 2022.
121. van Gelder, M.M.H.J., R.W. Bretveld, and N. Roeleveld, *Web-based Questionnaires: The Future in Epidemiology?* American Journal of Epidemiology, 2010. **172**(11): p. 1292-1298.
122. Grimberg, F., et al., *The Real-World Data Challenges Radar: A Review on the Challenges and Risks regarding the Use of Real-World Data*. Digital Biomarkers, 2021. **5**(2): p. 148-157.
123. Casey, J.A., et al., *Using Electronic Health Records for Population Health Research: A Review of Methods and Applications*. Annual review of public health, 2016. **37**: p. 61-81.
124. Hammerton, G. and M.R. Munafò, *Causal inference with observational data: the need for triangulation of evidence*. Psychological Medicine, 2021. **51**(4): p. 563-578.
125. Munafò, M.R. and G.D. Smith, *Robust research needs many lines of evidence*. 2018, Nature Publishing Group.
126. Hopper, J.L., D.T. Bishop, and D.F. Easton, *Population-based family studies in genetic epidemiology*. The Lancet, 2005. **366**(9494): p. 1397-1406.
127. Ludvigsson, J.F., et al., *The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research*. Eur J Epidemiol, 2009. **24**(11): p. 659-67.
128. Ludvigsson, J.F., et al., *Registers of the Swedish total population and their use in medical research*. Eur J Epidemiol, 2016. **31**(2): p. 125-36.
129. Ekbom, A., *The Swedish Multi-generation Register*. Methods Mol Biol, 2011. **675**: p. 215-20.
130. Ludvigsson, J.F., et al., *External review and validation of the Swedish national inpatient register*. BMC Public Health, 2011. **11**: p. 450.
131. Wettermark, B., et al., *The new Swedish Prescribed Drug Register—Opportunities for pharmacoepidemiological research and experience from the first six months*. Pharmacoepidemiology and Drug Safety, 2007. **16**(7): p. 726-735.
132. Ludvigsson, J.F., et al., *The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research*. Eur J Epidemiol, 2019. **34**(4): p. 423-437.

133. Frisell, T., P. Lichtenstein, and N. Långström, *Violent crime runs in families: a total population study of 12.5 million individuals*. Psychological Medicine, 2011. **41**(1): p. 97-105.
134. Brooke, H.L., et al., *The Swedish cause of death register*. European journal of epidemiology, 2017. **32**(9): p. 765-773.
135. Sun, S., et al., *Association of Psychiatric Comorbidity With the Risk of Premature Death Among Children and Adults With Attention-Deficit/Hyperactivity Disorder*. JAMA Psychiatry, 2019. **76**(11): p. 1141-1149.
136. Zhang, L., et al., *Prediction of treatment dosage and duration from free-text prescriptions: an application to ADHD medications in the Swedish prescribed drug register*. Evidence Based Mental Health, 2021: p. ebmental-2020-300231.
137. Weir, C.B. and A. Jan, *BMI Classification Percentile And Cut Off Points*. 2020: StatPearls Publishing, Treasure Island (FL).
138. Song, H., et al., *Stress related disorders and risk of cardiovascular disease: population based, sibling controlled cohort study*. BMJ, 2019. **365**: p. l1255.
139. Isomura, K., et al., *Risk of specific cardiovascular diseases in obsessive-compulsive disorder*. Journal of Psychiatric Research, 2021. **135**: p. 189-196.
140. Lundin, A., et al., *Unemployment and coronary heart disease among middle-aged men in Sweden: 39 243 men followed for 8 years*. Occupational and environmental medicine, 2014. **71**(3): p. 183-188.
141. Lundin, A., et al., *Unemployment and mortality—a longitudinal prospective study on selection and causation in 49321 Swedish middle-aged men*. Journal of Epidemiology & Community Health, 2010. **64**(01): p. 22-28.
142. Norrbäck, M., et al., *The association of mobility disability and obesity with risk of unemployment in two cohorts from Sweden*. BMC Public Health, 2019. **19**(1): p. 347.
143. Magnusson, P.K., et al., *The Swedish Twin Registry: establishment of a biobank and other recent developments*. Twin Res Hum Genet, 2013. **16**(1): p. 317-29.
144. Lichtenstein, P., et al., *The Swedish Twin Registry in the third millennium: an update*. Twin Res Hum Genet, 2006. **9**(6): p. 875-82.
145. Larsson, H., et al., *Genetic and environmental influences on adult attention deficit hyperactivity disorder symptoms: a large Swedish population-based study of twins*. Psychol Med, 2013. **43**(1): p. 197-207.

146. Del-Ponte, B., et al., *Dietary patterns and attention deficit/hyperactivity disorder (ADHD): A systematic review and meta-analysis.* J Affect Disord, 2019. **252**: p. 160-173.
147. Ahn, E. and H. Kang, *Introduction to systematic review and meta-analysis.* Korean journal of anesthesiology, 2018. **71**(2): p. 103-112.
148. Page, M.J., et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.* BMJ, 2021. **372**: p. n71.
149. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS medicine, 2009. **6**(7): p. e1000097.
150. Song, J.W. and K.C. Chung, *Observational studies: cohort and case-control studies.* Plastic and reconstructive surgery, 2010. **126**(6): p. 2234-2242.
151. Maret-Ouda, J., et al., *Nordic registry-based cohort studies: Possibilities and pitfalls when combining Nordic registry data.* Scandinavian journal of public health, 2017. **45**(17_suppl): p. 14-19.
152. Sahu, M. and J.G. Prasuna, *Twin Studies: A Unique Epidemiological Tool.* Indian journal of community medicine : official publication of Indian Association of Preventive & Social Medicine, 2016. **41**(3): p. 177-182.
153. Haidich, A.B., *Meta-analysis in medical research.* Hippokratia, 2010. **14**(Suppl 1): p. 29-37.
154. Riley, R.D., J.P.T. Higgins, and J.J. Deeks, *Interpretation of random effects meta-analyses.* BMJ, 2011. **342**: p. d549.
155. Hernán, M.A., *The Hazards of Hazard Ratios.* Epidemiology, 2010. **21**(1): p. 13-15.
156. Rijsdijk, F.V. and P.C. Sham, *Analytic approaches to twin data using structural equation models.* Briefings in bioinformatics, 2002. **3**(2): p. 119-133.
157. Shing, T.L., J.S. Preisser, and R.C. Zink, *GEECORR: A SAS macro for regression models of correlated binary responses and within-cluster correlation using generalized estimating equations.* Computer methods and programs in biomedicine, 2021. **208**: p. 106276.
158. Chen, Q., et al., *Attention-deficit/hyperactivity disorder and clinically diagnosed obesity in adolescence and young adulthood: a register-based study in Sweden.* Psychological medicine, 2019. **49**(11): p. 1841-1849.
159. Faraone, S.V. and H. Larsson, *Genetics of attention deficit hyperactivity disorder.* Molecular psychiatry, 2019. **24**(4): p. 562-575.

160. Granero, R., et al., *The Role of Iron and Zinc in the Treatment of ADHD among Children and Adolescents: A Systematic Review of Randomized Clinical Trials*. Nutrients, 2021. **13**(11): p. 4059.
161. Checa-Ros, A., et al., *Current Evidence on the Role of the Gut Microbiome in ADHD Pathophysiology and Therapeutic Implications*. Nutrients, 2021. **13**(1): p. 249.
162. Asherson, P. and H. Gurling, *Quantitative and molecular genetics of ADHD*. Curr Top Behav Neurosci, 2012. **9**: p. 239-72.
163. Larsson, H., et al., *Childhood attention-deficit hyperactivity disorder as an extreme of a continuous trait: a quantitative genetic study of 8,500 twin pairs*. J Child Psychol Psychiatry, 2012. **53**(1): p. 73-80.
164. Hasselbalch, A.L., et al., *Studies of twins indicate that genetics influence dietary intake*. J Nutr, 2008. **138**(12): p. 2406-12.
165. Meddens, S.F.W., et al., *Genomic analysis of diet composition finds novel loci and associations with health and lifestyle*. bioRxiv, 2018: p. 383406.
166. Meddens, S.F.W., et al., *Genomic analysis of diet composition finds novel loci and associations with health and lifestyle*. Molecular Psychiatry, 2021. **26**(6): p. 2056-2069.
167. Rios-Hernandez, A., et al., *The Mediterranean Diet and ADHD in Children and Adolescents*. Pediatrics, 2017. **139**(2).
168. Leppert, B., et al., *The Effect of Attention Deficit/Hyperactivity Disorder on Physical Health Outcomes: A 2-Sample Mendelian Randomization Study*. American Journal of Epidemiology, 2020. **190**(6): p. 1047-1055.
169. Fernández-Ruiz, I., *Immune system and cardiovascular disease*. Nature Reviews Cardiology, 2016. **13**(9): p. 503-503.
170. Hoekstra, P.J., *Attention-deficit/hyperactivity disorder: is there a connection with the immune system?* European Child & Adolescent Psychiatry, 2019. **28**(5): p. 601-602.
171. Ahmad Banday, A. and M.F. Lokhandwala, *Defective renal dopamine D1 receptor function contributes to hyperinsulinemia-mediated hypertension*. Clinical and experimental hypertension, 2006. **28**(8): p. 695-705.
172. Misener, V., et al., *Linkage of the dopamine receptor D1 gene to attention-deficit/hyperactivity disorder*. Molecular psychiatry, 2004. **9**(5): p. 500-509.
173. Jokinen, J. and P. Nordström, *HPA axis hyperactivity and cardiovascular mortality in mood disorder inpatients*. Journal of affective disorders, 2009. **116**(1-2): p. 88-92.
174. Corominas, M., et al., *Cortisol responses in children and adults with attention deficit hyperactivity disorder (ADHD): a possible*

marker of inhibition deficits. ADHD Attention Deficit and Hyperactivity Disorders, 2012. **4**(2): p. 63-75.
175. Hennekens, C.H., et al., *Schizophrenia and increased risks of cardiovascular disease*. The American heart journal, 2005. **150**(6): p. 1115-1121.
176. Hare, D.L., et al., *Depression and cardiovascular disease: a clinical review*. European Heart Journal, 2014. **35**(21): p. 1365-1372.
177. Swartz, H.A. and A. Fagiolini, *Cardiovascular Disease and Bipolar Disorder: Risk and Clinical Implications*. The journal of clinical psychiatry, 2012. **73**(12): p. 1563-1565.
178. Tully, P.J., et al., *Anxiety and Cardiovascular Disease Risk: a Review*. Current Cardiology Reports, 2016. **18**(12): p. 120.
179. Nielsen, R.E., J. Banner, and S.E. Jensen, *Cardiovascular disease in patients with severe mental illness*. Nature Reviews Cardiology, 2021. **18**(2): p. 136-145.
180. Halmøy, A., et al., *Occupational outcome in adult ADHD: impact of symptom profile, comorbid psychiatric problems, and treatment: a cross-sectional study of 414 clinically diagnosed adult ADHD patients*. Journal of attention disorders, 2009. **13**(2): p. 175-187.
181. Gjervan, B., et al., *Functional impairment and occupational outcome in adults with ADHD*. Journal of attention disorders, 2012. **16**(7): p. 544-552.
182. Adler, L.A., et al., *Functional outcomes in the treatment of adults with ADHD*. Journal of Attention Disorders, 2008. **11**(6): p. 720-727.
183. Shaw, M., et al., *A systematic review and analysis of long-term outcomes in attention deficit hyperactivity disorder: effects of treatment and non-treatment*. BMC Medicine, 2012. **10**(1): p. 99.
184. Chang, Z., et al., *Risks and benefits of attention-deficit/hyperactivity disorder medication on behavioral and neuropsychiatric outcomes: a qualitative review of pharmacoepidemiology studies using linked prescription databases*. Biological psychiatry, 2019. **86**(5): p. 335-343.
185. Boland, H., et al., *A literature review and meta-analysis on the effects of ADHD medications on functional outcomes*. Journal of Psychiatric Research, 2020. **123**: p. 21-30.
186. Armstrong, B.G., *Effect of measurement error on epidemiological studies of environmental and occupational exposures*. Occupational and environmental medicine, 1998. **55**(10): p. 651-656.
187. Buring, J.E., *Epidemiology in medicine*. Vol. 515. 1987: Lippincott Williams & Wilkins.

188. Merwood, A., et al., *Different heritabilities but shared etiological influences for parent, teacher and self-ratings of ADHD symptoms: an adolescent twin study*. Psychological medicine, 2013. **43**(9): p. 1973-1984.
189. Li, L., et al., *Maternal pre-pregnancy overweight/obesity and the risk of attention-deficit/hyperactivity disorder in offspring: a systematic review, meta-analysis and quasi-experimental family-based study*. International Journal of Epidemiology, 2020. **49**(3): p. 857-875.
190. Yazdanyar, A. and A.B. Newman, *The burden of cardiovascular disease in the elderly: morbidity, mortality, and costs*. Clinics in geriatric medicine, 2009. **25**(4): p. 563.
191. Jager, K.J., et al., *Confounding: What it is and how to deal with it*. Kidney International, 2008. **73**(3): p. 256-260.
192. Pourhoseingholi, M.A., A.R. Baghestani, and M. Vahedi, *How to control confounding effects by statistical analysis*. Gastroenterology and hepatology from bed to bench, 2012. **5**(2): p. 79-83.
193. Li, L., et al., *Associations of Prescribed ADHD Medication in Pregnancy with Pregnancy-Related and Offspring Outcomes: A Systematic Review*. CNS Drugs, 2020. **34**(7): p. 731-747.
194. Nazha, B., J.C.-H. Yang, and T.K. Owonikoko, *Benefits and limitations of real-world evidence: lessons from EGFR mutation-positive non-small-cell lung cancer*. Future Oncology, 2021. **17**(8): p. 965-977.
195. Rudrapatna, V.A. and A.J. Butte, *Opportunities and challenges in using real-world data for health care*. The Journal of Clinical Investigation, 2020. **130**(2): p. 565-574.
196. Cai, L. and Y. Zhu, *The challenges of data quality and data quality assessment in the big data era*. Data science journal, 2015. **14**.
197. Ross, M., W. Wei, and L. Ohno-Machado, *"Big data" and the electronic health record*. Yearbook of medical informatics, 2014. **23**(01): p. 97-104.
198. Ehrenstein, V., et al., *Clinical epidemiology in the era of big data: new opportunities, familiar challenges*. Clinical epidemiology, 2017. **9**: p. 245.
199. Simon, G.E., *Big data from health records in mental health care: hardly clairvoyant but already useful*. JAMA psychiatry, 2019. **76**(4): p. 349-350.