From the Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

HIGH PATERNAL AGE AND RISK OF PSYCHIATRIC DISORDERS IN OFFSPRING

Emma Frans

Karolinska Institutet

Stockholm 2013

Published by Karolinska Institutet. Printed by [Universitetsservice US-AB]

ISBN 978-91-7549-284-1

ABSTRACT

Parental ages at childbirth are increasing all over the world and later parenthood might have negative health outcomes for the offspring. During recent year, numerous studies report links between and high paternal age and psychiatric disorders such as schizophrenia and autism. However, the knowledge about this association is limited. The aim of this thesis was to gain valuable knowledge about the paternal age effect regarding its specificity, transmission and mechanism. The studies were conducted in an epidemiological setting by using multiple large population-based data sources that also enable controlling for a range of documented risk factors including parental, perinatal and socioeconomic variables.

In our first study we reported, for the first time, that high paternal age also is associated with bipolar disorder. The risk increased monotonically with age of the father and was, compared to younger men, highest in offspring of men in the oldest age category including men aged 55 years and older (odds ratio = 1.37) after adjustments. It was also evident that the paternal age effect was stronger when only analyzing individuals with an early disorder onset.

In study II, we confirmed an association between advancing paternal age and autism in the offspring. More importantly, we found an association between advancing paternal age and autism risk in the grandchild. Compared to younger fathers, men who fathered a daughter when they were 50 years or older were 1.79 times more likely to have a grandchild with autism, and men who father a son at 50 years of age or older were 1.67 times more likely to have an affected grandchild.

An increased burden of rare copy number variants (CNVs) has been found in individuals with schizophrenia and it has been suggested that CNVs can arise during replication. In study III, we used a sample consisting of individuals affected with schizophrenia and matched controls and examined paternal age in relation to rare CNVs. Although we found that rare CNVs were more common in individuals with schizophrenia and that their fathers were on average 0.75 years older than controls, we found no association between rare CNVs and paternal age.

In study IV, twin analyses showed that late fatherhood defined as becoming a father at age 40 years or above is under genetic influence (heritability = 0.33). However, a genetic liability for psychiatric disorders in men or their spouse was not associated with later fatherhood. Instead, a genetic liability for these disorders was generally associated with men having children at younger ages.

In conclusion, this thesis provides valuable knowledge about advanced paternal age as a risk factor for psychiatric disorders and might have important implications for clinicians, researchers, and those affected by the disorders.

LIST OF PUBLICATIONS

I. Frans EM, Sandin S, Reichenberg A, Långström N, Lichtenstein P, Hultman CM
Advancing Paternal Age and Bipolar Disorder
Arch Gen Psychiatry. 2008;65(9):1034-1040

II. Frans EM, Sandin S, Reichenberg A, Långström N, Lichtenstein P, McGrath JJ, Hultman CM
Autism Risk Across Generations – A Population-Based Study of Advancing Grandpaternal and Paternal Age
JAMA Psychiatry. 2003;70(5):516-521

III. Frans EM, Bergen SE, Grankvist A, Ruderfer DM, Purcell SM, Sklar P, Moran J, Scolnick E, Chambert K, Lichtenstein P, Hultman CM, Sullivan PF, Magnusson P
Rare copy number variation and the association between schizophrenia and advancing paternal age
Submitted

IV. Frans EM, Lichtenstein P, Hultman CM, Kuja-Halkola R
Paternal age: Heritability and associations to psychiatric disorders
Manuscript

CONTENTS

LIST OF ABBREVIATIONS

ASD	Autism spectrum disorder
BD / BPD	Bipolar disorder
CI	Confidence interval
CNV	Copy number variation
DNA	Deoxyribonucleic acid
DSM	Diagnostic and Statistical Manual of Mental Disorders
DZ	Dizygotic
FGFR	Fibroblast growth factor receptor
ICD	International Classification of Diseases
Kb	Kilobase
LISA	Longitudinal integration database for health insurance and labor market studies
MEN2	Multiple endocrine neoplasia type 2
Mb	Megabase
MZ	Monozygotic
OECD	Organisation for Economic Co-operation and Development
OR	Odds ratio
RAS	Reticular activating system
RET	Rearranged during transfection
RR	Relative risk
SES	Socio-economic status
SNP	Single nucleotide polymorfism

1 BACKGROUND

1.1 AGE OF THE PARENTS

Since the 1970s parental ages are increasing in Sweden. Today, the average man or woman in Sweden is approximately four years older when they have children as compared to mean parental ages 40 years ago (Figure 1). The high parental ages are mainly due to increased ages at the birth of the first child. A shift towards later parenthood is evident all over the world. In 2012, the Organisation for Economic Co-operation and Development reported on how the mean age of women at the birth of the first child changed from 1970 to 2009. This report showed that the mean age for mothers increased in all of the 22 countries included. The biggest increase was reported in Germany where mean maternal ages at first child birth increased with as much as six years.[1] There are several explanations to why both men and women are postponing parenthood: career and educational aspirations, increased life expectancy, the increased availability of contraception etc.[2] As a result, older parents are more likely to have a high education level and higher socio-economic status (SES) as compared to young parents and these factors are beneficial for the child.[3] From a public health perspective, however, the shift towards later parenthood might have negative consequences.

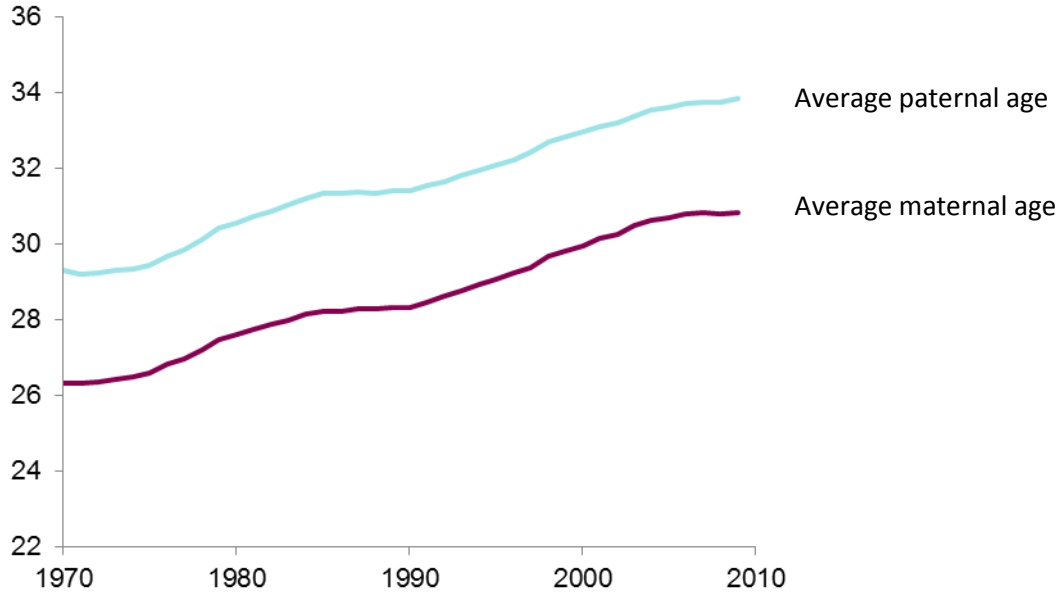

Figure 1: Average parental ages in Sweden since 1970

Regarding the negative effects of late parenthood on offspring health, the focus has traditionally been on maternal age. It is well established that the fecundity of women decline with age and the risk for adverse pregnancy and birth outcomes increase with maternal age.[4-6] In addition, offspring of older mothers have an increased risk of certain congenital anomalies, with Down syndrome (trisomy 21) as the most recognized example.[7] In recent years, the focus has shifted and the effects of late fatherhood are now the primary area of research. Late fatherhood has been associated with decreased

semen quality and decreased reproductive success.[8] Independent of maternal age, offspring of older fathers have an increased risk of a wide range of health conditions including adverse pregnancy outcomes, certain congenital disorders, childhood cancers, reduced neurocognitive ability and psychiatric disorders.[9] These findings have received much attention, especially since they challenge traditional perceptions on how male fertility is not affected by ageing. However, the links between advancing paternal age and negative outcomes have not always been replicable and the mechanism behind the paternal age effect is still unclear, at least regarding the more complex disorders linked to late fatherhood. In this thesis, we focus on the link between high paternal age and psychiatric disorders in the offspring.

1.2 PSYCHIATRIC DISORDERS OF INTEREST

1.2.1 Schizophrenia

Schizophrenia is a severe psychiatric disorder that is characterized by delusions, hallucinations, disorganized speech and behavior, as well as other symptoms that cause social or occupational dysfunction.[10] The disorder generally begins in early adulthood and, for both men and women, the lifetime risk is approximately 0.4 %.[11-13] It has long been recognized that schizophrenia coaggregates in families[12] and twin studies show that schizophrenia is strongly influenced by genes.[14] Based on Swedish register data, it has been estimated that the heritability of schizophrenia is 64 %.[15] Still, most schizophrenia cases are sporadic and the affected individual does not have relatives with schizophrenia.[12] Moreover, schizophrenia has in several studies been associated with reduced fertility in both affected individuals and their apparently healthy relatives.[16-18]

Researchers have struggled to identify common gene variants linked to schizophrenia. Only recently, genome-wide association studies on samples of substantial size have enabled identification of common genetic variation linked to schizophrenia.[19,20] In addition, new technologies have enabled the discovery of rare gene variants that are involved in the etiology of schizophrenia.

1.2.2 Autism

Autism spectrum disorders (ASDs) are a group of neural development disorders that are characterized by impaired communication and social interaction as well as restricted, repetitive or stereotyped behavior. These symptoms are evident in early childhood.[10] Males are more often affected by the disorder[21] and the population prevalence is approximately 0.6 %.[22] Autism is under a great genetic influence and twin studies on the more narrow definition of autism have estimated the heritability to be 80%.[23]Autism is also associated with a distinct reduction in fertility. A recent study using Swedish register data showed that men with ASD had a fertility ratio of 0.25 compared to healthy individuals.[17]

In approximately 10 % of ASD cases, the disorder is a part of a known genetic syndrome but most often, the cause is unknown. Moreover, specific rare genetic variants have been convincingly shown to cause autism.[24]

1.2.3 Bipolar disorder

Bipolar disorder (BD) is a mood disorder characterized by episodes of depression and mania/hypomania. Depending on severity, bipolar disorder is divided into type I and type II. Bipolar type I is the most severe type including severe depressions and episodes of mania that often requires hospitalizations.[25] Bipolar disorder type II has milder depressive periods and episodes of hypomania. In total, the disorder affects approximately 1 % of the population world-wide[26] with approximately equal distribution between genders.[25]

Overall, bipolar disorder is highly heterogeneous. There have therefore been efforts to dissect bipolar disorders into more well-defined subcategories. It has been suggested that early-onset BP should be considered a specific subtype[25,27-29] since early-onset BP often is characterized by worse prognosis, poor response to lithium, greater heritability and greater comorbidity problems.[29-32]

In 2009, Lichtenstein *et al.*[15] estimated, in a population-based study on Swedish register data, that the heritability for bipolar disorder is 59 % and that there is a large genetic overlap between bipolar disorder and schizophrenia. But except for a strong genetic influence, there are few well-established risk factors for bipolar disorder.

1.3 PSYCHIATRIC DISORDERS ASSOCIATED TO HIGH PATERNAL AGE

1.3.1 Schizophrenia

There are numerous studies showing a link between offspring schizophrenia and older fathers,[33-37] dating back to as early as 1958.[34] The interests in paternal age as a risk factor for schizophrenia become widespread following the publication by Malaspina *et al.* in 2001.[35] A recent meta-analysis[38] found that the relative risk (RR) for schizophrenia in offspring with the oldest fathers (≥50) was 1.66. It has also been estimated that 15 to 27 % of all schizophrenia cases are the result of advancing paternal age.[35,39] Potential confounding factors such as maternal age, parity of the mother, socioeconomic status, family history, and urbanicity have been examined and do not appear to account for the association. There have also been reports showing higher paternal ages in individuals with sporadic schizophrenia compared to individuals with familial schizophrenia, suggesting that the associations are causal.[40] However, a recent study on Danish population-based register data found that the association between advancing paternal age and schizophrenia in later born siblings disappeared when controlling for paternal age at the birth of the first child. The authors conclude that these results do not support the hypotheses of causal links, but instead suggests that factors related to postponing fatherhood are responsible for the association.[41]

1.3.2 Autism

In the 1970s, it was first suggested that there was a link between older paternal age and autism.[42,43] Since then, numerous studies have found associations between ASD and/or infantile autism and advancing paternal age. The results have been replicated in independent samples all over the world including Australia,[44] Denmark,[45] Israel,[46] USA[47] and Sweden.[48] The association between autism risk and paternal age is robust, although the magnitude of these associations has varied greatly. A meta-analysis on paternal age and autism risk showed that offspring of men aged 50 years or older were 2.2 times more likely to have autism than offspring of men younger than 30 years, after controlling for maternal age and documented risk factors for autism.[48] This report also included sibling-comparison analyses that showed that the risk of autism increased in later born children irrespective of the father's age when his first child was born, supporting the notion of a causal link.

1.4 OTHER HEALTH CONDITION LINKED TO HIGH PATERNAL AGE

1.4.1 Reduced fertility

In contrast to women, men can have children at very old ages and there have even been reports on men in their 90s fathering children.[49] Therefore, a common misperception is that male fertility is not affected by age. However, studies show that independent of maternal age, male fertility decreases with advancing paternal age.

It has been reported that semen quality declines with aging, including decreased semen volume and sperm motility, as well as altered sperm morphology.[8] It has also been reported that the time to pregnancy increases with higher paternal age, irrespective of maternal age.[50,51] Increased risks of miscarriage[52] and stillbirth[53] have also been linked to paternal age.

1.4.2 Congenital genetic disorders

In 1912, the German physician Wilhelm Weinberg first suggested that non-inherited cases of achondroplasia could be more common in last-born children than in older siblings. At this point, Weinberg made no distinction between paternal age, maternal age and birth order.[54] More than 40 years after Weinberg's observation, Penrose[55] showed that paternal age was the main, if not the only, cause of this association.

Since then, several genetic congenital disorders have been linked to high paternal age. The genetic conditions most strongly associated with advanced paternal age are autosomal dominant disorders caused by point mutations in the FGFR2, FGFR3, and RET genes, including Pfeiffer syndrome and Crouzon syndrome,[56,57] Apert syndrome,[58-61] achondroplasia and thanatophoric dysplasia,[62,63] as well as mutations in MEN2A and MEN2B.[54,64]

1.4.3 Other complex disorders

In addition to the links to schizophrenia and autism, advancing paternal age has been linked to other disorders or adverse behavioral trait with a complex etiology such as eating disorders,[65] childhood cancer,[66] congenital heart defects,[67] impairments in the general cognitive ability,[68] violent offending,[69] Alzheimers disease[70] and epilepsy.[71]

1.5 POSSIBLE EXPLANATIONS TO THE FINDINGS

The mechanism behind the paternal age effect and psychiatric disorder remains unclear. However, there is no shortage of hypotheses. Some of the hypotheses suggest that there is a causal link, while others claim that the associations can be explained by unmeasured confounding. The most noted hypotheses are discussed below as well as other hypotheses that form the basis of the studies in this thesis. It is also worth mentioning that the mechanism behind the paternal age effect might differ between the distinct disorders. Moreover, one possible explanation does not necessarily rule out another; it is also possible that the paternal age effect consist of a number of co-occurring events affecting the risk of disorders in the offspring.

1.5.1 *De novo* point mutations

The hypothesis that has received most attention suggests that the association between advancing paternal age and psychiatric disorders are, in similarity with its association to congenital genetic disorders, a result of the increased burden of *de novo* mutations in germ cells of older men.

Women are born with their full supply of germ cell and the oocytes only undergo a fixed number of cell divisions (23 chromosomal replications in total). By contrast, male germ cells are produced continuously through men's reproductive life.[72] More specific, spermatogonial cells replicate every 16th day, resulting in approximately 200 divisions by the age of 20 years and 660 divisions by the age of 40 years.[73] Each time the cell divides, the replication of the genome introduces the possibility of copy-error mutations. As a result of the large number of cell divisions during spermatogenesis, the mutation rate for base substitutions is much higher in men than in women, and increases with paternal age. These mutations may be inherited to the offspring and potentially have negative effects on their health. In humans, it has been confirmed that sperm from older men have significantly more mutations.[54,61,74] It has also been suggested that the gene replication process in older men is further compromised by how levels of DNA proof-reading and repair enzymes decline as a function of advancing paternal age[75,76] and that the DNA fragmentation increases.[77]

It is well established that the association between advancing paternal age and certain genetic disorders is caused by *de novo* mutations. However, it remains unclear whether new mutations are causing the links between advancing paternal age and psychiatric disorders or other complex traits. It is possible that mutations are of great importance for mental health since a large proportion of human genes are important for brain function. Human mental health depends on the functionality of a very large number of

genes and non-coding regulatory regions and thus, the mutational target size is large. Based on conservative estimates on human mutational load,[78-80] it has been estimated that the average human carries 500 mutations that have a negative effects on brain function, and the individual variability in mutation load is large.[81] Moreover, recent studies[82-84] using exome sequencing methods suggests that *de novo* point mutations have a major role in the etiology of schizophrenia. Similarly, several exome sequencing studies have reported an increased burden of *de novo* mutations in autism pedigrees and some of these studies suggest that most new mutations are paternal origin and occasionally even associated paternal age.[85-89] An Icelandic study,[90] published in August 2012, sequenced the whole genome of individuals with sporadic schizophrenia or autism and found that fathers, on average, pass on to their offspring 25 new point mutations at age 20, increasing to 65 mutations at age 40. They concluded that the rate of new mutations in relation to paternal age is 2 new mutations per year.

Although paternal age might increase the overall risk for new mutation and therefore specifically affect multingenous traits, it has also been suggested that spermatogonial cells carrying paternal age related mutations involved in the reticular activating system (RAS) pathway are positively selected and expand clonally in normal testes, leading to a relative enrichment of mutant sperm over time. Since the RAS pathway is important for brain function, this could explain the association between advanced paternal age and various neurodevelopmental disorders.[91]

1.5.2 Copy number variants

During the past decade, there has been a rapid development and expanded use of microarray technologies enabling whole-genome analysis with essentially unlimited resolution. These techniques resulted in the discovery of copy number variation (CNV) in the human genome.[92] It is well established that *de novo* copy number variation contributes to an increased risk for neurodevelopmental disorders such as schizophrenia,[93-95] bipolar disorder,[94] and autism.[96,97]

CNVs are structural alterations in the DNA resulting in an abnormal number of copies of DNA segments. If the alteration results in an increased number of copies these CNVs are referred to as duplications and if the alteration results in a loss of DNA the CNVs are instead called deletions. CNVs can be inherited or sporadic and both recombination- and replication-based mechanisms for CNV formation have been described. It has been suggested that CNVs that cause psychiatric disorders are more likely to be of large size and of *de novo* origin.[98]

As previously mentioned, germ cells of older men have undergone a larger number of replications compared to germ cells of younger men. As a result, mutations originating from replication are more likely to occur in germ cells of older men. It is therefore possible that the burden of CNVs in the offspring is correlated with the age of the father. Some studies show that structural variations such as altered copy number in repeat DNA and chromosome breakage occur in male germ cells and these variations have occasionally been linked to paternal age.[99-101] It has been reported that most new CNVs are paternal in origin,[102,103] and one of these studies also found an association

between advancing paternal age and *de novo* CNVs in individuals with intellectual disabilities.[102] By contrast, a recent study found no association between advancing and rare CNV load in 6,773 healthy individuals.[104] Moreover, genome-wide analyses on 834 Dutch schizophrenia patients found an increased frequency of CNV deletions in affected individuals but no association between CNV occurrence and paternal age.[105]

1.5.3 Epigenetic alterations

Alterations in gene expression that is not caused by DNA changes are referred to as epigenetic alterations. It has been suggested that epigenetic traits can by structural inheritance be transmitted to offspring.

Perrin *et al*.[106] and Sipos *et al*.[39] have suggested that epigenetic alterations that occur as paternal age advances, may be casually related to the susceptibility of schizophrenia. Interestingly, paternal exposure to toxicant and nutritional state as well as age, have been found to influence the development in offspring and sometimes even development of grand-offspring.[107]

1.5.4 Selection into late fatherhood

The most noted hypothesis in addition to the *de novo*-hypothesis suggests that the association between paternal age at birth and psychiatric disorder in offspring is confounded by psychiatric disorders or a genetic liability for psychiatric disorders in the father.[41,108] Individuals with genes predisposing for psychiatric illness are more likely to have children with similar disorders.[12] If a genetic liability for psychiatric disorders also is associated with a selection into late fatherhood this would result in a non-causal association between paternal age and psychiatric disorders in the child Moreover, it is possible that women with mental disorders (or a genetic liability for mental disorders) systematically have children with older men. Since these women are more likely to have children with similar disorders this mating pattern could result in an association between late fatherhood and mental disorders in the children. This hypothesis is supported by a study on a Finnish sample showing that advancing paternal age is associated with schizophrenia in the mother, but not in the father.[109]

These hypotheses are, however, opposed by the findings of higher paternal ages in individuals with sporadic schizophrenia.[40] In addition, numerous studies on advancing paternal ages and psychiatric disorders in the offspring show that the associations are robust even after adjusting for history of psychiatric disorders in the parents. Still, these adjustments might be influenced by missing data on parental psychiatric health history or subclinical traits of psychiatric disorder. Moreover, these adjustments do not include individuals that are unaffected but genetically related to individuals with mental disorders and thus, might transmit a genetic liability for psychiatric disorder to their offspring. In efforts to explore this hypothesis further, a study examining the effects of autism-related personality traits on paternal ages found no association between these personality traits and delayed fatherhood.

By using sibling-comparison analyses, there have been attempts to further explore this hypothesis. Since full siblings have the same parents, sibling-comparison studies automatically adjust for familial factors, for instance a genetic liability of psychiatric disorders in parents. As previously mentioned, the results of sibling analyses have been inconclusive.[41,48]

1.5.5 Characteristics of older fathers

It is also possible that some of the environmental characteristics of having an older father increase the risk of psychiatric disorder or alternatively; being diagnosed with a psychiatric disorders. The characteristics of men who became first-time fathers at an older age were recently described in a Norwegian population-based study. The study showed that older fathers had both higher and lower SES compared to younger fathers. These men were also more likely to engage in negative health behavior and be of poorer health.[110]

1.5.6 Maternal age

Paternal and maternal ages are highly correlated. It has therefore been difficult to distinguish their separate effects from each other. Especially for autism, there have been reports of a link between advancing maternal age after controlling for paternal age.[111] It is therefore considered that both maternal and paternal ages are independent risk factors for autism.

1.6 DEFINITION OF ADVANCING PATERNAL AGE

There is no universally accepted definition of advanced paternal age but within genetic counseling advancing paternal age is often referred to as fathers aged 40 years and older.[9] However, the genetic defects associated with advancing paternal age does not increase dramatically after this age. Instead, the risk for genetic changes increases linearly with paternal age and therefore, the cut-off at 40 years has no biological impact.

Similarly, studies on schizophrenia and autism shows no well-defined limit when high paternal age becomes a risk factor. A statistically significant increase is sometimes seen in offspring of men in their 30s,[112] although the effect sizes are rather small. The risk increase is not always linear but studies on paternal age in relation to psychiatric disorders show no evident threshold age at where the risk increases dramatically.

2 AIMS

The general aim of this thesis was to gain knowledge about the associations between high paternal age and psychiatric disorders. More specifically, each individual study aimed to answer the following research questions:

- Study I: Is advancing paternal age a specific risk factor for schizophrenia and autism or does it also affect the risk of bipolar disorder in the offspring?

- Study II: Is advancing paternal age only a risk factor for the offspring or does it also affect the mental health of grandchildren?

- Study III: Can the effect of advancing paternal age and schizophrenia be explained by the increased number of rare structural variations known as copy number variations found in individuals with schizophrenia?

- Study IV: Is advancing paternal age a heritable trait and can the influence of advancing paternal age on psychiatric disorders be explained by a link between individuals with a genetic liability for psychiatric disorders and late fatherhood?

3 MATERIALS AND METHODS

3.1 DATA SOURCES

All studies in this thesis include data from Swedish national registers. The primary key of the register linkage is the unique personal identification number (national registration number), a ten-digit code that is assigned to each Swedish citizen dating back to 1947.[113] For integrity reasons, these national registration numbers are replaced with unique sequence numbers when register data is used for research purposes.

Additionally, in study III, we used molecular genetic data in order to detect copy number variation (CNV) and linked this information to the register data. The genetic material was collected from a national sampling frame of individuals identified in the Swedish registers.

3.1.1 Register data

3.1.1.1 Patient Register

The Swedish Patient Register includes practically all psychiatric inpatient discharge diagnoses in Sweden since 1973 and includes discharge date, the main discharge diagnosis, and up to 8 secondary diagnoses assigned by the treating physician. The diagnoses are recorded according to the International Classification of Diseases (ICD) 8th, 9th, and 10th ICD editions.[114-116] Since 2001, the register also includes psychiatric outpatient care, although with incomplete coverage. Before outpatient care was included in this register, this register was known as the Hospital Discharge Register.

High validity of ICD diagnoses recorded in the Swedish Patient Register has been found by comparing diagnostic register code with medical records. The positive predictive value for most somatic and psychiatric inpatient diagnoses is approximately 85% to 95%.[117]

3.1.1.2 The Multi-Generation Register

The Swedish Multi-Generation Register contains information about biological parents of an index person and their birth dates, enabling assessment of parental ages. A prerequisite for being included in the register is that the index person was born after January 1, 1932, and ever registered as living in Sweden after 1960. For immigrants to Sweden, similar information exists for those who became citizens before age 18 years together with 1 or both parents. The biological father of the offspring is assumed to be the husband of the mother at the time of birth or identified “by acknowledgment” for unwed mothers.[118]

3.1.1.3 The Twin Register

The Swedish Twin Register is one of the world's largest twin databases. The register was established in the late 1950s[119] and enables identification of more than 194,000 twins, born since 1886. Currently, this register contains information on zygosity for 75,602 twin pairs. Zygosity is determined either by questions about intra-pair physical similarities in childhood, genotyping, or by being of opposite sex. [120]

3.1.1.4 The Education Register

The Swedish Register of Education includes highest obtained education level for an index person and is annually updated. The register was initiated in 1985 by Statistics Sweden but includes information about education level from decades prior obtained via census and other surveys. [121]

3.1.1.5 Longitudinal integration database for health insurance and labor market studies (LISA)

LISA integrates existing data from registers and is annually updated. The database includes information from the labor market, educational and social sectors. This database also includes information about highest education level obtained from the Education Register.[122]

3.1.1.6 Total Population Register

The total population register is the official register of the Swedish population administrated by the Swedish Tax Agency. The register contains for example place of birth, citizenship status, information about immigration and emigration, marital status and place of residence.[123]

3.1.1.7 Cause of Death Register

The Swedish cause of death register was established in 1961 and is annually updated. The register contains information on all deaths among Swedish residents including death dates and death causes according to ICD.[124]

3.1.2 Large-Scale Schizophrenia Association Study in Sweden (BROAD study)

The Broad sample includes DNA from schizophrenia cases and matched controls. Individuals affected with schizophrenia were identified by using the Patient Register. Cases were defined as adult individuals who during at least two separate occasions had been diagnosed with schizophrenia. For a subset of individuals, register diagnoses were confirmed by medical record reviews. The control subjects were randomly selected through population registers and frequency matched on age, gender and residential county. All included individuals were at least 18 years old with both parents born

within Scandinavia. Written informed consent was obtained from all study participants. In total, participation rates in the Broad sample are 53.3 % for cases and 58.3 % for controls.

3.2 MEASURES

3.2.1 Psychiatric disorders

Individuals with and without psychiatric disorders were identified by utilizing the Patient Register. Individuals with bipolar disorder were defined as having, on at least two separate occasions, an inpatient diagnosis of ICD-8 and ICD-9 codes 296 (except 296.2 and 296B for unipolar depression) as well as ICD-10 codes 30 and 31. In the second study, we defined autism cases as individuals that had received ICD-9 code 299.0 and ICD-10 code F84.0 during inpatient or outpatient care. In the final study we had a broader definition of psychiatric disorder requiring only one diagnosis of ICD-8 and ICD-9 codes 295-299 (except 296.2 and 296B) as well as ICD-10 codes F20-F29, F30-F31 and F84 given during inpatient or outpatient care.

3.2.2 Paternal ages

Paternal ages were defines as age of the biological father at the birth of the child. Similarly, grandpaternal ages were defined as age at the birth of the parent. Paternal ages were categorized into age group and compared to a reference group. Alternatively, mean paternal ages were calculated for all live children or at birth of the first child.

3.2.3 Copy number variations

A Genome-Wide Association study was conducted on DNA extracted from schizophrenia cases and matched controls (Broad Study). The samples were genotyped on Affymetrix Genome-Wide Human SNP 5.0 and 6.0 arrays at the Broad Institute of Harvard and MIT. After an initial SNP quality control, remaining samples were analyzed in regards to CNV by using the Birdsuite algorithm to identify rare CNVs via a hidden Markov model.[125] In order to be defined as rare, CNVs had to be rare among the HapMap CEU samples (< 3%) and be rare among our samples (< 1%). As a quality assessment, included CNVs had to span at least 100 kb and have a Birdsuite LOD score ≥ 10. Samples with CNVs spanning more than 10 Mb or that had more than 30 CNVs were also removed.

3.3 STATISTICAL ANALYSES

3.3.1 Linear regression

Linear regression is a commonly used method of analysis in epidemiological studies. For continuous outcome variables, for instance paternal age at childbirth, linear regression models the relationship between a dependent variable and a predictor

variable. The linear regression model assumes that the correlation is linear and that the data is normally distributed.[126]

3.3.2 Logistic regression

In epidemiological studies, logistic regression is a commonly used method of analysis. Logistic regression can describe the relationship between a categorical outcome (dependent variable) and a set of covariates (predictor variables). The categorical outcome may be binary (e.g., presence or absence of disease) or ordinal (e.g., normal, mild and severe). The predictor variables may be continuous or categorical. Logistic regression is often used to predict how binary outcomes are affected by a dependent variable, while controlling for potential confounding of other variables. By using logistic regression, we can estimate an odds ratio (OR) i.e. the odds of being a case if exposed compared to the odds of being a case if unexposed. Under certain conditions, for instance if a disorder is rare, OR can be interpreted as relative risk (RR). Logistic regression models assume that that the relationship between the outcome and the predictor variable is linear on the logit scale for continuous predictor variables and that the distribution of the residuals is normal.

For frequency matched case-controls studies, unconditional logistic regression can be used. By contrast, in case-control studies where controls are individually matched to the cases (stratified data), conditional logistic regression is the more suitable method of analysis. [126]

3.3.3 Poisson regression

Poisson regression is a regression model that is used to model count data. When analyzing CNV burden where an individual can be affected more than once, we utilize Poisson regression. Poisson regression assumes a Poisson distribution.[126]

3.3.4 Twin model

By using the twin model, we can estimate how much of the phenotypic variation of a trait that is due to genetic and environmental factors. This is done by comparing monozygotic (MZ) twins that are genetically identical with dizygotic (DZ) twins who share approximately half of their genes. The twin model assumes that the hereditability of a trait are due to three components: additive genetics (A), shared environmental factors (C) and non-shared environmental factors (E).

When similarities are greater between MZ than DZ twins this indicates a genetic effect. Similarities between twins that are not explained by genetic effects are assumed to be caused by shared environment and within-pair differences reflects the effect of non-shared environment.[127]

4 STUDY METHODS

4.1 STUDY I

By linking Swedish national registers we conducted a nested case-control study on advancing paternal age and bipolar disorder in the offspring. The cases were identified in the Swedish Patient Register and birth dates were obtained by the Multi-Generation Register. Inclusion criteria for cases were a discharge diagnosis of bipolar disorder on at least two separate occasions and the presence of both maternal and paternal ages. The patient register was followed until December 2001. In total, this resulted in a sample of 13 428 cases. For each case, we selected five healthy controls, matched on gender and birth year. Eligible controls were defined as individuals with parental age data who were alive and did not have an inpatient diagnosis of bipolar disorder in the Patient Register at the time of the first diagnosis of the corresponding case.

Paternal and maternal ages were calculated based on birth dates and categorized into 5-year intervals. Fathers and mothers aged 20 to 24 years at the offspring's birth constituted the reference category. Information about potential confounders was obtained from the Patient Register, the Education Register, and the Multi-Generation Register.

By using conditional logistic regression, we calculated ORs and 95 % confidence intervals (95 % CI). We used 4 different models: 1) unadjusted analyses on paternal and maternal ages (adjustment for the matching variables was included in the conditional regression model), 2) analyses on paternal and maternal ages adjusted for age of the other parent, 3) analyses on paternal and maternal ages adjusted for age of the other parent and for family history of psychotic disorders, and 4) analyses on paternal and maternal ages adjusted for age of the other parent, family history of psychotic disorders, SES, and parity. Highest education level within the family was used to reflect SES.

We performed separate analyses on early-onset cases defined as younger than 20 years at the time of their first diagnosis and their individually matched controls. In addition, we conducted specificity analyses when 1) excluding cases diagnosed as having schizophrenia after their last bipolar disorder diagnosis, and 2) defining cases as all individuals with a discharge diagnosis with bipolar disorder.

4.2 STUDY II

By linking Swedish national registers, we conducted a frequency matched case-control study on paternal ages and autism risk in both the child and the grandchild. We identified individuals diagnosed as having childhood autism in the Swedish Patient Register. We defined cases as all individuals with diagnoses given at discharge from inpatient care since 1987 when the specific diagnostic code for childhood autism was first introduced and diagnoses given during outpatient care since 2001. The Swedish Patient Register was followed up until December 31, 2009. Individuals who did not

meet our criteria for autism and were alive at the end of our observation time were eligible as controls. Five unaffected individuals for each affected were frequency matched for sex and year of birth. The matching was done at an early stage, enabling missing data comparison between cases and controls at an early state. After identifying cases and controls, we linked parents and grandparents by using the Swedish Multi-Generation Register. After birth date linkage, we calculated parental and grandparental ages. Ages of parents were defined as the parent's age at the time of the index person's birth. Ages of grandparents were defined as the grandparent's age at the time of the parent's birth. Information about potential confounders was retrieved from the Patient Register, the Longitudinal integration database for health insurance and labor market studies (LISA) and the Total Population Register.

By using logistic regression, we calculated ORs and 95 % CIs for parental ages in association to autism in the child or in the grandchild. Ages were categorized into 5-year intervals, with 20 to 24 years as the reference category. The analyses were performed in 4 steps controlling for different covariates including birth year, sex, age of spouse, family history, highest education level and residential county. The models were evaluated for goodness-of-fit by visual inspections of the model residuals. The final model was conducted on 5933 cases and 30904 controls.

We also performed the following sensitivity analyses: 1) only including inpatient diagnoses, 2) controlling grandparental ages for parental ages, and 3) analyzing parental ages for all individuals with present parental age data, i.e. before linking grandparental ages.

4.3 STUDY III

By linking the Broad sample and the Multi-Generation Register we conducted a study where we analyzed rare CNV occurrence in relation to paternal age in individuals affected with schizophrenia and healthy controls. Cases were defined as adult individuals with at least two diagnoses of schizophrenia given during inpatient care. Eligible controls were individuals without an inpatient diagnosis of schizophrenia or bipolar disorder. The DNA used in this study was extracted from venous blood collected from 1586 schizophrenia cases and 2092 controls. The samples were genotyped on Affymetrix arrays and after initial quality control, rare CNVs were identified using the Birdsuite algorithm. To increase the probability of including CNVs of *de novo* origin, CNVs had to be rare among the HapMap CEU samples (< 3 %) and be rare among our samples (< 1 %).

Information about parental identity was obtained by linking the genetic data to the Swedish Multi-Generation Register. As a result, we could calculate parental ages for the study participants. After this linkage, 1410 cases and 1850 controls remained and were included in the paternal age by CNV analyses. For analyses involving maternal age, 1393 cases and 1833 controls had information on ages of both parents and were included in these analyses.

We analyzed the association between parental age and schizophrenia with linear regression. The relationship between paternal age and genome-wide burden of rare CNVs was examined by using Poisson regression. We also conducted separate analyses on CNVs only observed once in our sample. Statistical testing of hypotheses was based on the two-sided 5 % level of significance.

4.4 STUDY IV

First, we calculated the heritability of late fatherhood by using the twin model. Twins were identified by the Swedish Twin Register and by linking these individuals to the Multi-Generation Register, we obtained information about the twins' children and could thereby link birth dates. In order to avoid censoring as well as reduce the chances of truncation of data we only included individuals born between 1920 and 1960 (14,679 male twins).

Advancing paternal age was defined as having a child at age 40 years or older. The analyses were conducted separately on individuals who had their first child at age 40 years or above. By using the twin model we estimated how much of the phenotypic variation of late fatherhood that was due to genetic and environmental factors. In order to estimate if the parameters describing the ACE-model changes over time we allowed the relative contributions of additive genetics (A), shared environment t (C) and non-shared environment (E) to vary with offspring birth year.[128]

The twin model estimating the impact of genes and environment on having a child at the ages 40 years or above was fitted in three different steps. The first, base model estimated the effects of genes and environment which were allowed to vary with time. In the second model, we assumed no effect of shared environment. In the final model we added the constraint that the ratio between the additive genetics and non-shared environmental factors had to be constant over time. We used likelihood ratio tests to examine fit of the different models. The heritability and environment parameters were estimated along with their 95 % CI. In the same manner, the analyses were separately conducted on individuals having their first child at age 40 years or above. The twin analyses were performed by R statistical software by employing OpenMx.[129]

Secondly, we conducted a cohort study where we examined paternal ages in families with or without genetic liability to psychiatric disorders. Affected individuals were defined as having at least one diagnosis in the Patient Register with schizophrenia, bipolar disorder or ASD. In order to collate a sample where we had adequate coverage of adult psychiatric and at the same time extensive coverage of their children, we selected a cohort of men born in 1955 through 1960. By linking the Swedish Multi-Generation Register, we identified children to these individuals and calculated mean age at fatherhood.

Mean paternal ages were examined in six groups: 1) healthy men, 2) men with psychiatric disorder, 3) healthy men with an affected sibling, 4) men with healthy spouses, 5) men with affected spouses, and 6) men with healthy spouses with an affected sibling.

We examined difference in mean ages for all children as well as conducting analyses only including ages at the birth of the first child. Moreover, we preformed logistic regression analyses after grouping paternal ages into 10-year categories and using men that had their children at the ages 20-29 years as the reference category.

5 RESULTS

5.1 STUDY I

In our main analyses, the unadjusted analyses showed an increased risk for bipolar disorder in offspring of both older men and older women. When controlling for the age of the other parent the association with maternal age largely disappeared but there were still a monotonically increasing risk with advancing paternal age. The analyses were controlled by potential confounding by family history of psychiatric disorders, education level and parity of the mother. These adjustments did not have a large impact on the OR, instead, the association to advancing paternal age remained. In the fully adjusted mode, the OR for bipolar disorder in offspring of men aged 55 years or older was 1.37.

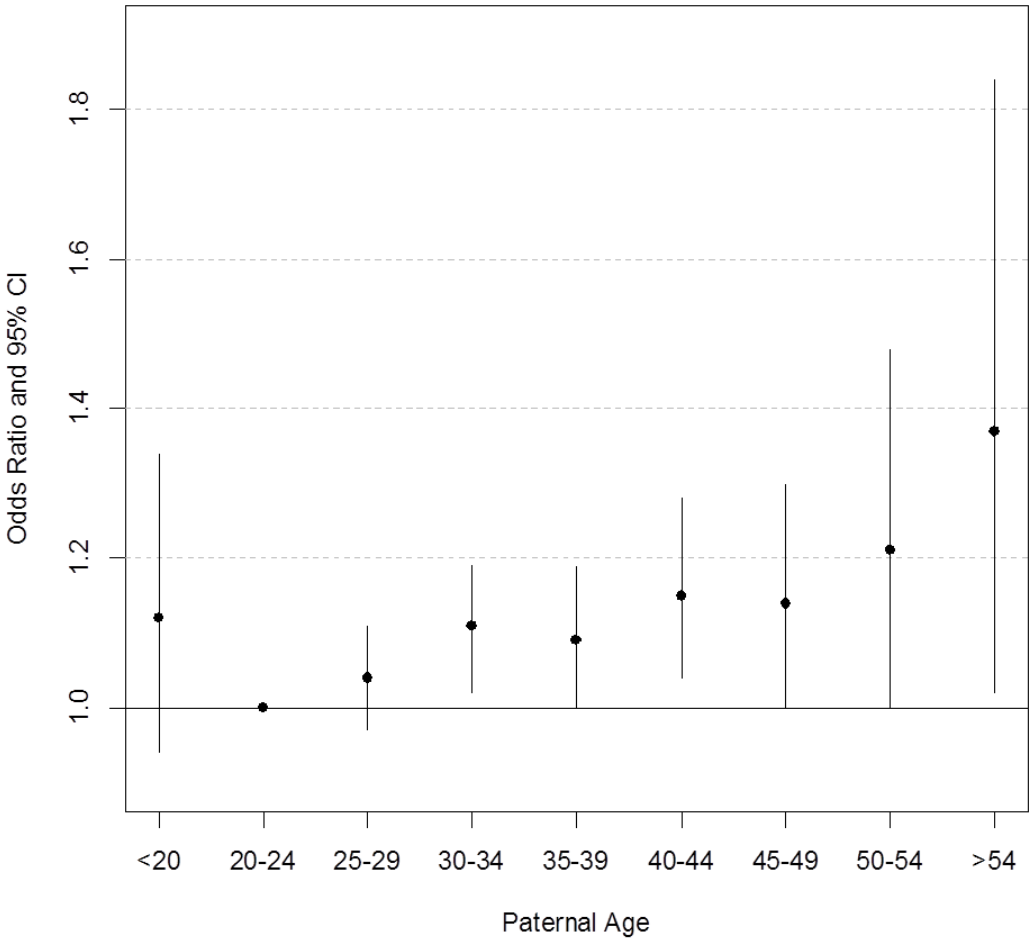

Figure 2: Results from conditional logistic regression analyses on paternal ages and bipolar disorder in the offspring – adjusted for maternal age, family history of psychiatric disorder, SES and parity

The subanalyses conducted on individuals with an early onset, defined as a discharge diagnosis of bipolar disorder before the age of 20 years, showed an even stronger association with advancing paternal age. Again, the risk increased with paternal age and

after adjusting for maternal age, the odds ratio reached 2.63 for offspring of men in the oldest age category. These analyses also showed an increased risk for bipolar disorder in offspring of teenage fathers resulting in a J-shaped association between paternal age and risk for bipolar disorder. In the early-onset analyses, we found no effect of maternal ages.

The results were consistent in the specificity analyses when we excluded cases diagnosed as having schizophrenia after their last bipolar disorder diagnosis as well as when we defined cases as all individuals with a discharge diagnosis with bipolar disorder.

5.2 STUDY II

After linking parental ages to our study subjects, 93 % of our cases and 90 % of the controls remained. Moreover, 60 % of cases and 63 % of controls had available grandparental data.

The result of the logistic regression analyses on advancing parental ages and risk for autism in the offspring confirmed previous studies showing a positive association. The risk increase was evident in all age categories over 30 years after adjusting for maternal age. Highest risk was found in the oldest paternal age category with an OR of 2.26 in the fully adjusted model. After controlling for paternal age, there was also an association between autism and having a mother in the oldest age category.

The logistic regression analyses showed that autism risk was positively correlated with age of both the maternal and the paternal grandfather, defined as the age when the grandfather became a father. Again, the highest risk was found in the oldest age category, defined as becoming a father at the age 50 years of above. In the highest age category the odds ratio of having a grandchild with autism was 1.79 for maternal grandfathers and 1.67 for paternal grandfathers in the fully adjusted model. By contrast, grandmaternal age was not associated to autism risk. The results of our sensitivity analyses were consistent with the main results.

Age	**Maternal grandfather**	**Paternal grandfather**
(years)	**OR (95 % CI)**	**OR (95 % CI)**
< 20	0.90 (0.73-1.11)	0.91 (0.73-1.12)
20-24	1.00	1.00
25-29	1.08 (0.99-1.18)	1.10 (1.00-1.20)
30-34	1.19 (1.07-1.32)	1.17(1.05-1.30)
35-39	1.31 (1.15-1.49)	1.15 (1.02-1.31)
40-44	1.32 (1.12-1.54)	1.23 (1.05-1.44)
45-49	1.34 (1.07-1.67)	1.60 (1.30-1.97)
≥ 50	1.79 (1.35-2.37)	1.67 (1.25-2.24)

Table 1: Results from logistic regression analyses on paternal ages and autism in the grandchild – adjusted for age of spouse, family history of psychiatric disorder, SES and residential county

5.3 STUDY III

We found that 65.1 % of cases and 60.9 % of controls had at least one rare CNV. The difference was statistically significant (p = 0.01). When comparing paternal ages of affected individuals and controls we saw that fathers of cases were, on average, 0.75 years older than fathers of the cases. Linear regression analyses showed that this age difference was statistically significant (p = 0.0024) However, when controlling for age of the mother, this age difference disappeared.

The result of the Poisson regression found no association between paternal age and rare CNV occurrence. The results were also negative when we analyzed deletions and duplication separately. Poisson regression analyses stratified on disease status showed consistency with previous analyses and the no association between paternal age and rare CNV burden in cases, controls, or when both groups were combined. The results were similar when we only included single-occurrence CNVs.

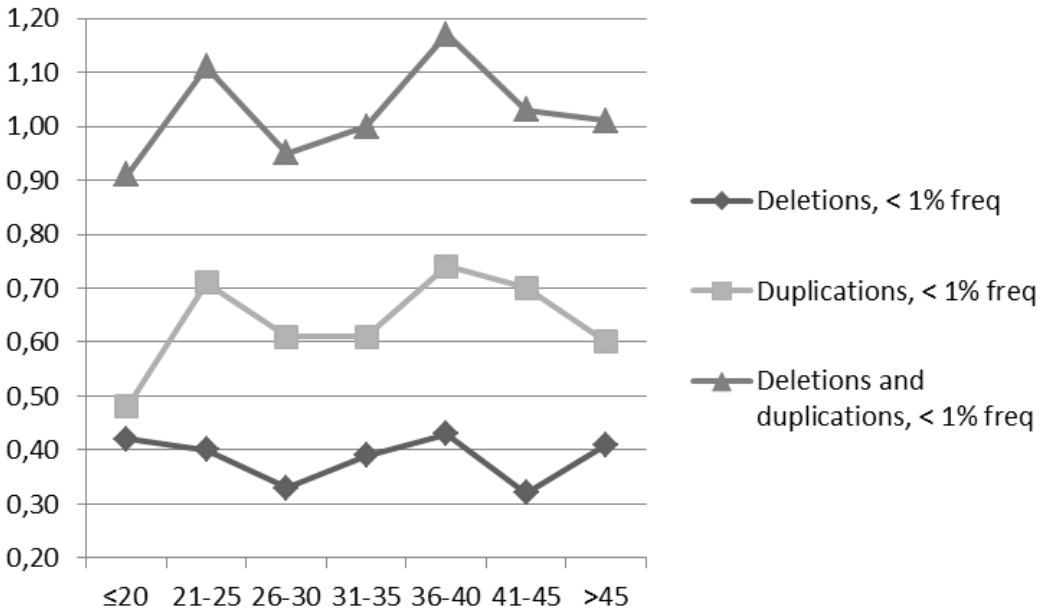

Figure 3. Mean number of rare CNVs per subject by paternal age group

5.4 STUDY IV

The twin model analyses showed that late parenthood, defined as having a child at age 40 years or above, is a heritable trait. More specifically, we estimated that the heritability of late fatherhood was 0.33 (95 % CI = 0.25-0.40) and that 0.67 (95 % CI = 0.60-0.75) of the phenotypic variance was explained by non-shared environment. We found no significant effect of shared environment and although the prevalence of late fatherhood changed over time, the relative impact of genes and environment remained constant.

When men who became fathers for the first time at age 40 years or older were analyzed separately the estimated heritability was 0.46 (95 % CI = 0.27-0.62) and the effects of

non-shared environment was 0.54 (95 % CI = 0.38-0.73). Again, we found no effect of shared environment and the impact of genes and environment was stable over time.

The analyses on the entire population born between 1955 and 1960 showed that men with psychiatric disorder had an average paternal age of 31.09 and healthy men from the same birth cohort were, on average, 31.62 years old when they had their children. Moreover, healthy men with an affected sibling had a mean paternal age of 31.10 years. Both men with psychiatric disorders and healthy brothers of affected individuals were statistically significantly younger than healthy individuals of the same birth cohort ($p<0.0001$) when they had their children.

Moreover, paternal ages in spouse of affected women were on average 31.34 years and the paternal ages in spouses of healthy women from the same birth cohort were on average 31.37 years. There was no significant age difference between these groups ($p = 0.57$). By contrast, spouses of women with an affected sibling had a statistically significantly lower mean paternal age (31.02 years) than the mean paternal age of the comparison group ($p < 0.0001$). When analyzing paternal ages at the birth of the first child, we found similar differences between the groups.

6 DISCUSSION

6.1 METHODOLOGICAL CONSIDERATIONS

The aim of an epidemiological study is often to assess how an exposure affects an outcome and more specifically to study risk factors for certain disorders. When studying the effects of an exposure on an outcome, the aim is often to assess causality. This can be done by conducting an experimental study where the exposure is randomized to the study participants. For both practical and ethical reasons, this study design is often impossible. Under these circumstances, we frequently rely on information from observational studies. In observational studies, such as the ones included in this thesis, we do not influence the population in any way or attempt to intervene in the study. Instead, we gather data and investigate correlations. The drawback of observational studies is, however, that they cannot prove causality and that unmeasured factors may affect the results.

It is obvious that paternal ages cannot be randomized to human study subjects. Therefore, the potential causal effects of advancing paternal age on psychiatric health in the child must be assessed by using other study designs. In this thesis, the aim of the first two studies was to explore correlations between paternal age (exposure) and psychiatric disorders in subsequent generations (outcome) by using large population-based data sources. The other two studies focus on the potential explanations behind the paternal age effect on psychiatric disorder in the offspring.

6.1.1 Study design

The study designs used in this thesis are cohort, case-control and nested case-control. In cohort studies, groups that differentiate regarding exposure are followed over time and compared on the basis of the outcome. Retrospective cohort studies are based on information that has already been collected, for instance information collected from national population and health registers. By using data that has previously been collected, we can avoid recall bias affecting the results. The advantage with cohort studies is that several different measures of outcome can be calculated. In some cases, for instance if the outcome is rare, cohort studies may have limited efficiency.

For rare outcomes, a case-control study is the more efficient alternative. In case-control studies two existing groups differing in outcome (case/control) are identified and compared on the basis of an exposure. If a case-control study includes all affected individuals in a population and randomly selected controls from the same population, the selection of controls will not bias the results. In nested cases-control studies, controls are ‘nested’ within the larger population cohort. The nested case-control design requires accurate registration of time-to-event for study subjects within a defined cohort. When an event occurs and a subject becomes a case, controls are randomly selected from the remaining subjects that are still at risk. As a result, nested case-control studies account for time at risk in a similar manner as cohort studies. In case-control studies we cannot estimate relative risks. Instead, odds ratios (ORs) are

commonly estimated. ORs are defined as the odds of being a case if exposed compared to the odds of being a case if unexposed. This measure might be difficult to interpret but if an outcome is rare, ORs can be interpreted as relative risks.[126]

The first study included in this thesis is a nested case-control study. Cases were defined as individuals with at least two separate hospitalizations with a diagnosis of bipolar disorder and information about ages of both parents, the exposure was paternal age at the birth of the proband. Controls were individually matched to a case from the population still at risk for developing bipolar disorder and with full information about parental age. Since bipolar disorder is a relatively rare disorder we interpreted the ORs from the analyses as relative risk estimates. This study could also have been designed as a cohort study, but by the nested case-control design, the results should be equivalent independent of study design.

In study II, we wanted at an early stage carefully examine how truncation of parental ages could potentially affect our results. We were particularly interested in how missing data on parental and grandparental age data differed between cases and controls. Therefore, we first identified individuals with autism and frequency-matched five times as many controls for each birth year and gender strata and subsequently linked parental and grandparental ages. By calculating the missing data frequency after each linkage, it was evident that the rates of missing data were similar for cases and controls. This study could also have been conducted by using a cohort design and later assessing the frequencies of missing data of parental ages. However, we chose this design to be able to validate these problems before analyzing the data.

The Broad sample that is used in study III is originally designed as a case-control study. However, our third study is not a case-control study *per se* and disease status was only used in order to stratify the study subjects in subgroups or considered the exposure in the linear regression analyses.

The cohorts in study IV was defined as male twins born in 1920 through 1960 and men born between January 1955 and December 1960 with at least one child. The study participants were followed until December 2009. The outcome measure was mean paternal age. In the second part of the study, the exposure was defined as having a diagnosis of psychiatric disorder, being a healthy sibling of an affected individual, or being a healthy individual. The narrow cohort limits the generalizability of the study. However, we consider the narrow cohort a prerequisite in order to avoid censoring problems affecting our results.

6.1.2 Random errors

The precision of a test is given by the difference of repeated measurements. Precision is defined as an absence of random errors. These errors are a result of chance and thus the likelihood of random errors affecting the results are reduced with increased sample sizes. In large epidemiological studies the risk of random errors is therefore small. Confidence intervals are estimated to indicate the amount of random error in the estimate. Commonly, the level of confidence is set at 95 %, meaning that if the analyses

are replicated over and over again, the confidence interval should include the correct value 95 % of the time. Confidence intervals are calculated by the same equations that are used to estimate the p-value. A significance level at $p< 0.05$ is equivalent to a 95 % CI not containing the null value. The studies included in this thesis are all of substantial size and the likelihood of random errors is therefore small.

6.1.3 Systematic errors

Systematic errors occur when the measurements have a tendency to deviate from the true value in a systematic way. A study with high internal validity is characterized by that we actually measure what we intend to measure. Full internal validity is defined as the absence of systematic errors. In contrast to random errors, systematic errors cannot be corrected by increasing the number of individuals included in the study and must therefore be considered in large epidemiological studies, such as the ones included in this thesis. The internal validity of a study is mainly affected by three sources of systematic error: selection bias, information bias and confounding.

6.1.3.1 Selection bias

Selection bias arises when factors influencing the selection or participation of study subjects cause distortions of the results. Although all studies in this thesis are based on national registers with full coverage of Swedish citizens, there are several potential sources of selection bias.

In the first study, cases are individuals that have been treated in inpatient care on at least two separate occasions. The intention of this narrow definition of the disorder was to avoid including misclassified schizophrenia patients. Therefore, cases are more likely to represent individuals with the more severe type of bipolar disorder: type I.

In the main analyses of study II, we only include individuals that had information about grandparental ages. This sample represents approximately 60 % of cases and control of the original sample. This restriction is a potential source of selection bias and generates a sample of patients that are more likely to have younger parents. These issues are further discussed in the section on *left truncation*.

Study III includes data from the Broad sample. In this sample, eligible study subjects were identified in the Patient Register or randomly selected from population-based register. In the description of the Broad sample, participation rates for cases and controls are given. Systematical differences between responders and non-responders could potentially affect our results. However, a particular strength of the Broad study is that it has been evaluated regarding potential non-response bias for both cases and controls. By using the Swedish population-based registers, demographic and medical information on all eligible subjects irrespective of participation status were assessed. The information indicated that cases and controls are well matched on sex and county but that there were a lower proportion of controls in younger ages. The participating cases was not significantly different compared to non-participating cases concerning

age, subdiagnosis or age at admission.

In the heritability analyses in study IV we only include twins with known zygosity and this restriction may lead to selection bias. These analyses were also restricted to only include male twin pairs and male twins with female co-twins were not included in the analyses. Moreover, the twins included in these analyses were required not only to have at least one child of their own but also to have a twin that had children. However, we consider it highly unlikely that male twins with sisters or with childless brothers are different from other male twin and this selection should therefore not affect the internal validity.

6.1.3.2 Information bias

Information bias arises when the study is subjected to measurement errors or misclassification. If the misclassification is the same across groups this is called non-differential misclassification. This type of misclassification will result in bias towards the null hypothesis. If the misclassification occurs more frequently in one of the study groups, then differential misclassification will occur, and the estimate of association can be overestimated or underestimated. In the present thesis the main sources of misclassification is the ascertainment of psychiatric disorders and discrepancy of paternal identity.

Misclassification of psychiatric disorders may lead to affected individuals being classified as unaffected (false negative) or non-affected individual being classified as affected (false positive). The diagnostic criteria for the disorders included in this thesis are based on the occurrence of certain clinical symptoms. In absence of biomarkers the accuracy of psychiatric diagnoses relies on the ascertainment of the treating physician. In psychiatric research the risk for misclassification is especially high and therefore the accuracy of psychiatric disorders has been examined in validation studies.

In our first study, it is possible that some individuals with an early-onset that does not require hospitalization are not included in our early-onset analyses. However, we consider the risk of false positives to be small and that the included individuals are truly patients with an early-onset disorder.

6.1.3.2.1 Validity of diagnoses

High validity of ICD diagnoses recorded in the Swedish Patient Register has been found by comparing diagnostic register code with medical records. The positive predictive value for most somatic and psychiatric inpatient diagnoses is approximately 85% to 95%.[117]

Symptomatically, bipolar disorder overlaps with both schizophrenia and depression. To avoid misclassification we only included individuals with two separate discharge diagnoses in our study on paternal age and bipolar disorder (study I). In 2011, Sellgren *et al*.[130] conducted a validation study on bipolar disorder in the Swedish Patient

Register. The study concluded that two inpatient episodes with a bipolar diagnosis were sufficiently sensitive and specific to reflect DSM-IV-TR definition of bipolar disorder.

To maximize diagnostic consistency across classification schemes and increase accuracy, our study on autism focuses on the narrow diagnosis of childhood autism (ICD-9 diagnostic codes 299A or ICD-10 diagnostic codes F84.0). We included diagnoses given at discharge from inpatient care since 1987 when the specific diagnostic code for childhood autism was first introduced and diagnoses given during outpatient care since 2001. We also conducted a separate analysis by only including individuals with diagnoses given during inpatient care and found that the results were consistent.

Ekholm *et al.*[131] conducted a validation study on patients with a discharge diagnosis of schizophrenia. The results of this study showed that the Swedish Patient Register had high positive power to predict a standard research DSM-IV diagnosis of schizophrenic psychosis obtained by using the Operational Criteria (OPRICT) algorithm.[132] Moreover, in study III we only included cases that had received schizophrenia diagnoses on two separate occasions. The sample used in study III has also been validated specifically by Dr. Hultman. This validation included a medical record review using a structured checklist, and 97.2 % (=106/109) met criteria for DSM-IV Schizophrenia.[133]

When we examined paternal ages when the proband, spouse or sibling had a diagnosis of schizophrenia, autism or bipolar disorder (study IV) the aim was not to look at the diagnoses specifically. Various studies show that these disorders are genetically related and we were interested in examining the overall paternal age pattern in individuals with a genetic liability for disorders previously linked to paternal age. Separating between these disorders was therefore not of interest, considering our study question. Therefore, we defined all individuals with a register diagnosis of any of these disorders as affected.

We carefully selected our sample with focus on the study objectives and took precautions to avoid misclassification. The high validity for psychiatric disorders in the patient register reduces the chances of false positive observations. However, it is possible that we include false negatives in our comparison groups. In the first and third study, diagnoses given during inpatient care were exclusion criteria for healthy control. It is therefore possible that we misclassify individuals that have milder variants of bipolar disorder or schizophrenia, not requiring hospitalization, as healthy. In study II and IV, we also include diagnoses given during both inpatient and outpatient care when defining cases and controls. However, the Patient Register only includes diagnoses in outpatient care since 2001 and does not have complete coverage of all individuals with psychiatric disorders treated in outpatient care in Sweden. The incomplete coverage of outpatient care is a potential source of information bias and this might lead primarily to false negatives. We assume that this misclassification is non-differential regarding paternal age and therefore, inclusion of false negatives in our analyses should only bias our results toward the null.

6.1.3.2.2 Accuracy of paternal age

In all studies, fathers were assumed to be the husband of the mother at the time of birth or by acknowledgment for unmarried mother. A study examining published data on levels of false paternity, the median discrepancy was estimated to be 3.7 % for studies based on populations chosen for reasons other than disputed paternity.[134] The accuracy of paternity in the Multi-Generation Register has not been investigated (to our knowledge), and it is possible that the discrepancy of paternal identity is higher among affected individuals. But assuming that this misclassification is non-differential regarding paternal ages, this type of misclassification should not affect the results to a large extent.

6.1.3.3 Confounding

A confounder is defined as a cause of confusion. In epidemiology a confounder can be thought as mixing of effects that may result in bias. Rothman[126] defines a confounder according following criteria:

- a confounder must be associated with the outcome
- a confounder must be associated with the exposure
- a confounder must not be an effect of the exposure

Statistical associations between two variables occur when one is the cause of the other, when they share a common cause, or both. Confounding is defined as a common cause to the exposure and the outcome. Factors associated with both the cause and the outcome that are effects of the exposure are either mediators (causing the outcome) or colliders (effect of the outcome). Controlling for mediators or collider may bias the results.

There are several ways to control for confounding. If a confounder is known, it is possible to restrict the study subjects to a specific category of this confounder. Confounding can also be accounted for by matching so that the potential confounder is equally distributed among exposed and non-exposed (cohort study) or cases and controls (case-control study). If study subjects are stratified into homogenous groups with regards to the potential confounder, the effects of this confounder can be evaluated by comparing differences between these groups. In multivariate analyses, it is also possible to adjust for potential confounders in the statistical models.[126]

By randomly assigning exposure to study subjects it is also possible to control for known confounders. This study design also enables control for unknown or unmeasured confounding. Randomization therefore enables identification of causal links. By contrast, associations in observational studies cannot be interpreted as causal, even after controlling for known or potential confounders. In these studies, it can never be fully established that the results are not due to unmeasured confounding.

In our first and second study we were interested in whether our results were affected by potential confounding. In the first study we adjusted for potential confounding by maternal age, family history of psychiatric disorders, parity, and socio-economic status.

We used highest education level to reflect SES. As compared to occupation, education level enables better comparability internationally and over time periods. The drawback might be that the general education level continuously increase, which could contribute to bias between birth strata. However, our matched case-control design reduces this bias since all cases are compared with controls born the same year. Parity was defined as the number of live born babies to the mother. We acknowledge the limitation that stillborn children do not receive a personal identification number and are therefore not included in the Multi-Generation Register. However, stillbirths in Sweden are rare (approximately 0.3 %)[135] and we therefore do not consider this a potential source of bias.

In the second study we also adjusted for maternal age, family history of psychiatric disorders, and SES (education level). We did not adjust for parity in this study since we do not consider it likely that parity is associated with age of grandparents at birth of the parent. We did, however, include adjustments of county of residence in theses analyses because it is possible that coverage of inpatient care varies across geographic areas.

By comparing models including different covariates we could examine the potential confounding effect of these variables. As previously mentioned we cannot rule out the chances of missing data of our included covariates affecting our results. However, the results of both these studies showed that maternal age/ age of spouse was the only covariate in our analyses that had a major effect on the association between paternal age and psychiatric disorder.

6.1.4 Left truncation

Left truncation refers to data that is retrospectively limited resulting in exclusion of study subjects.[136] The truncation is dependent on birth strata. Still, matching cases and controls on birth year does not automatically take care of problems due to left truncation. In these studies, the data might be truncated regarding parental ages.

Individuals born before the initiation of the Multi-Generation Register in 1932 are excluded due to left truncation of parental age data. However, the Multi-Generation Register includes parental information of individuals born since 1932 and ages of their parents are not extensively subjected to left truncation. We therefore do not consider left truncation to be a major issue in the studies where we only link parental ages to psychiatric disorders in the offspring (Study I and III).

In our second study, when we looked at paternal age over several generations, we were worried about potential problems with left truncation. We therefore selected a study design where we first identified all control and cases from our population before linking parental and grandparental ages. By doing this, we could detect if there were evident differences in missing data frequencies between affected and unaffected individuals. After the linkage, it was evident that the rates of missing data were similar among cases and controls. In order to collate a sample consisting of 3 generations, parents had to be born after 1932. As a result, older parents were more likely to be excluded than younger parents. To address this issue, we conducted sensitivity analyses

on a sample that not required grandparental age data linkage. The result of these analyses suggested a similar paternal age effect in comparison to the main analyses. We therefore concluded that there was no major truncation of parental ages in this study. Moreover, if we assume that parental ages and grandparental ages are correlated, selection bias excluding the older parents might lead to a selection of younger grandparents and might therefore bias our results. In subanalyses not included in the published material, we truncated parental ages even further, with regards to parental ages and found that the grandpaternal age effect was consistent.

In study IV, we select a rather late cohort regarding age at fatherhood and in this study we should have no problems with truncation of paternal ages.

6.1.5 Left censoring

When an event occurs before the individual comes under observation, this is called left censoring.[136] The studies on paternal age in association to psychiatric disorders might be subjected to left censoring regarding psychiatric diagnosis.

The Patient register does not cover inpatient diagnoses before 1973 or outpatient diagnoses before 2001. Therefore, individuals who only received diagnoses before these time points will be classified as non-affected. Left censoring of the patient register may therefore lead to misclassification and false negatives. The issues regarding misclassification were discussed in the previous section.

In study IV, we first selected a cohort were the oldest twins were born 1920 and thus, they were 12 years old when we could observe their potential children in the Multi-Generation Register. Therefore we consider it highly unlikely that censoring of earlier-born children would lead to a misclassification of paternal ages.

In the second part of study IV, we selected a narrow cohort in order to avoid censoring of psychiatric disorders and offspring information. By selecting individuals that were teenagers when the Patient register covers practically all psychiatric inpatient care we aimed to reduce the chances of misclassification.

In summary, the main focus in these studies was to minimize left censoring problem regarding our exposure and outcome or to assess their potential effects. However, it is possible that confounding variables such as family history of psychiatric disorder, educational level, and parity are subjected to left censoring.

6.1.6 Right censoring

When an event occurs after the observation time end-point, this is called right censoring.[136] In study I-III psychiatric diagnoses may be subjected to right censoring, meaning that a hospitalization might occur after the follow-up time of the registers.

In study IV we included a cohort of men born between January 1955 and December 1960. This selection aimed at identifying a sample where we had extensive coverage of

psychiatric diagnoses as well as extensive coverage over their children. For the youngest men in our study (born 1960), the follow-up time ends when these men are 49. It is possible that some men father children at 50 years or older. However, considering how rare fatherhood is at the age 50 years or above we do not think that it is likely that this censoring has a great impact on mean paternal ages.

6.1.7 Generalizability

External validity refers to the generalizability of a study and if the results are valid for populations beyond the subjects included in the study. A necessary prerequisite for external validity is internal validity. However, high internal validity does not constitute evidence for high external validity.

These studies are population-based and include data from national registers. Although the psychiatric diagnosis criteria are ascertained according to international standards, the results largely reflect the Swedish population. This population is an ethnically homogenous group and it is therefore unclear if the results are generalizable to other, non-Scandinavian populations.

When we estimated the heritability for late fatherhood we used a twin sample only including male twins with a twin brother. The analyses were also restricted to twin-pairs where both twins had at least on child. It is commonly questioned whether twins are representative members of the population.[137] If twins are different from the general population, the results of twin studies cannot automatically be generalized beyond the population in which they have been derived. However, we found no evident differences between twins and singletons regarding mean parental ages.

External validity also refers to if the results are valid for other birth cohorts. In study IV, we have a rather narrow cohort in regards to birth year, limiting the generalizability of this study. We know that parental ages have changed over time and that they are strongly correlated with socio-economic variables and the use of contraceptives. The offspring of individuals in this cohort are conceived in post-contraceptive times. We consider it likely that these results are not generalizable to pre-contraceptive times.

6.1.8 Ethical considerations

Medical research is often a balance act between gaining information that will increase the public health and protecting the welfare of the study participants. As previously mentioned, in observational studies we do not influence or intervene in any way when we gather data. From an ethical perspective, observational studies are therefore less problematic as compared to experimental studies. However, all studies in this thesis utilize sensitive personal information that may harm the personal integrity of the study participants.

According to Swedish law,[138] all studies on humans that include sensitive personal information must be approved by an ethical committee. In studies including biological material from a living individual, all study participants must give informed consent.

Research studies are only approved if they can be conducted with respect for human dignity. Human rights and fundamental liberties are considered during ethical vetting, and the welfare of people should always be given precedence over the needs of society and science. Moreover, ethical approval is only given if the potential risks the study participants are exposed to are counterbalanced by the scientific value of the research.

All studies in this thesis include individuals identified by national registers. To protect the integrity of the individuals, national registration numbers are replaced with a unique identification code ensuring that no specific individuals can be identified. In our third study, we also collected blood samples and this intervention might result in transient pain for the study participants. Studies including biological material from living individuals are potentially harmful to personal integrity. All studies included in this thesis were approved by regional ethic committees. Moreover, all participants in study III gave informed consent at inclusion.

The studies in this thesis all include individuals with psychiatric diagnoses, a particularly vulnerable patient group. At this point, our knowledge about the etiology of autism, schizophrenia and bipolar disorder is limited. The lack of knowledge about the mechanisms underlying these diseases hampers the development of effective treatments. These diseases are also associated with great suffering for those affected as well as their relatives. The studies included in this thesis can result in increased understanding within this field and thus reduce human suffering. We therefore consider that potential risks of these studies are small compared to their scientific value.

Regarding the effects of advancing paternal age, there is a major interest from the general public. Considering that the mechanism behind the paternal age effect is largely unknown and the rather moderate risk increase, we have stated in our publications that these findings should not discourage older men from fathering children but that our findings are interesting and may increase the understanding in the etiology of psychiatric disorders.

6.2 PERSPECTIVES ON FINDINGS AND IMPLICATIONS

6.2.1 The evolutionary paradox of common yet harmful psychiatric disorders

Considering how most psychiatric disorders have a negative impact on the affected individuals at a relatively young age, it is not surprising that most individuals with psychiatric disorders have a reduced fertility. In an evolutionary framework, traits that result in a reduced fertility are referred to as low fitness traits. A trait with low fitness is under a negative selection pressure and genes associated with these traits should be removed from the population. The rate of this selection is directly an effect on how harmful these genes are for an individual's fitness.[139] Considering that psychiatric disorders are largely influenced by genes, it is puzzling that psychiatric disorders are maintained in the general population. Psychiatric disorders are not only maintained in the population but are rather common, and it has even been suggested that these

disorders are increasing in prevalence. Evolutionary psychologists have struggled to explain this apparent paradox and several theories have been proposed including balancing selection, neutral evolution[140] and polygenic mutation-selection balance.[81]

Under the balancing selection hypothesis, genes associated with a disorder that have a negative impact on the fitness of an individual can be maintained in the population if these genes also are associated with a reproductive advantageous trait. In 2004, Burns suggested that schizophrenia is an unwanted byproduct of the evolution of the social human brain.[141] It has also been suggested that genes associated with psychiatric disorders also predisposes for creativity[142,143] and that the reduction in fertility for individuals with mental illness is counterbalanced by an increased fertility in their healthy relatives. However, several studies show that healthy individuals with affected relatives do not have more children compared to the general population. Instead, these individuals are also more likely to have a reduced fertility.[17,18]

In modern society, it is obvious that mental disorders have a negative impact on the fitness of the affected individuals. However, it has been suggested that this is merely a result of modern society and that mental disorders have historically been neutral or even advantageous for the fitness of an individual. Manias are often episodes of great productivity and it is possible that depressive episodes might have been an advantageous way to reduce energy expenditure and avoiding harmful events. It has been suggested that bipolar disorder may have been an adaptation to long, severe winters and short summers.[144] With autism and schizophrenia, it is not as obvious how these traits might have been neutral or advantageous in a historical setting. However, neutral evolution does not explain the persistence of psychiatric disorders in recent years.

Based on observations in a Swedish schizophrenia sample, Böök[145,146] proposed as early as 1953 that the incidence of schizophrenia was maintained at a state of equilibrium by new mutations. He suggested that these new mutations counteracted the losses due to natural selection. In 2009, Keller and Miller[81] conducted an extensive review and concluded that the most reasonable explanation for the paradox is that genetic variants increasing the risk for these disorders constantly arise as *de novo* mutations. As previously mentioned, harmful mutations are by natural selection removed from the population in a pace that is dependent on how negative the effects of these mutations are. However, if the mutation target size of the disorder is large, meaning that many genes are involved in mediating the disorder, new mutations predisposing for the disorder are introduced in a higher rate than the rate of the selection removing these genes.[147] In this case, a mutation-selection balance occurs where the natural selection removing harmful genes are counterbalanced by the rate of new harmful mutations that is introduced in the genome. A very large number of human genes are important for brain development and function and therefore, human mental health has a huge mutational target size.

The hypothesis of a mutation-selection equilibrium maintaining psychiatric disorders in the population is supported by recent genetic studies showing that psychiatric disorders are associated with new mutations - both CNVs and point mutations. This hypothesis is also supported by the association between paternal age and psychiatric disorders, since

we know that germ cells of older men have an increased number of *de novo* mutations. However, the mechanism behind the paternal age effect remains unclear and there are several other possible hypotheses suggesting that the effect is explained by epigenetic alterations, confounding by a genetic liability in the parents, environmental effects of having older parents etc.

6.2.2 Paternal age as a common risk factor for psychiatric disorders

To date, there are no biomarkers available for any of the psychiatric disorders. Instead, the diagnostic criteria for these disorders are based on the occurrence of certain clinical symptoms. Autism, schizophrenia and bipolar disorder are currently considered as distinctive psychiatric disorders. These disorders all have distinguishable, heterogeneous entities: for example, autism onsets in early childhood, schizophrenia is characterized by having both hallucinations and delusions, and bipolar disorder is composed by very opposite phases and symptoms (manic and depressive episodes).[10] However, these disorders also have common, homogenous clinical entities. Especially individuals with schizophrenia might initially be diagnosed with bipolar disorder and vice versa.[148] These disorder are, however, not only symptomatically homogenous, they also have a common etiology. Both genetic and family studies show genetic overlaps between schizophrenia, bipolar disorder and autism.

In 2009, Lichtenstein et al. published results from a large family study of schizophrenia and bipolar disorder based on Swedish national registers.[15] The results showed that schizophrenia and bipolar disorder partly share a common genetic cause. These findings challenge the current nosological dichotomy between schizophrenia and bipolar disorder, a dichotomous classification system that originates from more than 100 years ago when Emil Kraepelin split the non-organic (so-called functional) psychoses into two disorders; disorders that we nowadays refer to as schizophrenia and bipolar disorder. The evidence showing a common genetic cause for these disorders has resulted in debates about the scientific justifications for continued adherence to the Kraepelinian dichotomy.[149]

By contrast, ASD was historically regarded as childhood schizophrenia because of the similar features regarding impaired social interactions and bizarre behavior. In 1971, ASD was finally separated conceptually from schizophrenia based on symptomatic differences, with age of onset as the most distinguishable feature.[150] Moreover, the dichotomization was also motivated by differences in family history and differential treatment responses in individuals with suspected schizophrenia versus autism. However, this distinction may not be absolute, and there are also numerous studies showing that there are important genetic overlaps between ASD and schizophrenia. For instance, recent studies including populations from Denmark, Sweden and Israel show that a family history of schizophrenia is a risk factor for ASD,[151,152] indicating that these disorders share a common genetic factor. In addition, these studies show that bipolar disorder also coaggregates in families and thus, that they share a common genetic cause with ASD. These findings are confirmed by recent molecular genetic studies showing shared genetic variation across psychiatric disorders including schizophrenia, bipolar disorder and autism.[153]

In this thesis, we suggest that paternal age is a common risk factor for autism, schizophrenia and bipolar disorder. This notion is supported by the results from study I where we found that advancing paternal age was associated with bipolar disorder. This association remained after controlling for maternal age, family history of psychiatric disorder, education and parity and was even higher in subanalyses on patients with an early disorder onset. This study was the first to report an association between bipolar disorder and paternal age and since it was published there have been attempts to replicate our findings. The replication studies have, however, both confirmed[154] and refuted[155] our findings. A recent study by Grigoroiu-Serbanescu reported a paternal age effect for bipolar disorder type I that was exclusive for women and patients without a family history for psychiatric disorders.[156] As mentioned earlier, patients with bipolar disorder is a very heterogeneous group and this might explain the inconclusiveness of these studies. As our study suggests, the association between advancing paternal age and bipolar disorder in its broader definition is probably small to moderate. However, paternal age might be an important risk factor for more severe subtypes of bipolar disorder such as individuals with an early onset and episodes that require hospitalization.

Considering our findings, we provide additional evidence that these disorders have, in part, a common etiology. The homogenous characteristics of psychiatric disorders and their etiological similarities may have important implications for clinicians, researchers, and those affected by the disorders.

6.2.3 Transmission of the paternal age effect

We humans inherit, on average, half of our genes from our mothers and half of our genes from our fathers. By extension, this means that 25 % of our genes come from each one of our grandparents. Genes and mutations predisposing for a disorder can therefore be inherited to both children and grandchildren and the likelihood of transmission is reflected by genetic relatedness.

In 1997, population geneticist James F. Crow stated in the American journal *Proceedings of the National Academy of Science* that "the greatest mutational health hazard in the human population at present is fertile old males". He described in this review mutations that have a direct visible effect on the child's health and also mutations that can be latent or have minor visible effects on the child's health; many such mutations allow the child to reproduce, but cause more serious problems for grandchildren, greatgrandchildren and later generations.[157]

If a genetic disorder has high penetrance and results in a severe reduction in fertility, the possibility on transmission of this disorder to subsequent generations is very limited. For such disorders, we do not expect to see an effect of paternal age-related mutations in subsequent generations, such as grandchildren. However, older fathers also have an increased risk for having a child with new mutations on their X-chromosome. New mutations on the X-chromosome are usually not a cause for disorders in the children. Instead, these mutations are transmitted to daughters who are

at risk of having sons with X-linked diseases. This is an indirect paternal age effect; it is the effect of the age of the maternal grandfather. Example of X-linked disorders associated with high paternal age is Duchenne's muscular dystrophy[158] and Haemophilia A.[159]

Genetic disorders with incomplete penetrance or smaller fertility effects can be transmitted to further generations. As a result, it is feasible that some paternal age-related *de novo* mutations may not result in adverse health outcomes in the offspring, but still contribute the overall burden of mutations inherited by subsequent generations. If this hypothesis is valid, it would be predicted that both paternal and grandpaternal age could contribute to an increased risk of new mutations predisposing for psychiatric disorders in the offspring.

The second study showed that advancing paternal age did not only affect the autism risk in the child but that grandchildren of older men also had an increased risk for autism. These results were consistent even after controlling for potential confounders. Sensitivity analyses showed that the association with grandpaternal ages was independent of paternal age. It has previously been reported that age of the maternal grandfather is associated to risk of schizophrenia in the proband.[112] There are, however, no other studies showing a link between grandpaternal ages and autism. To our knowledge, this has only been examined in one previous study. The study showed no association between granpaternal age and autism risk. This study was, however, limited by the very small sample size (86 individuals).[160]

Considering the findings linking grandpaternal age and risk of autism, it is possible that a proportion of age-related *de novo* mutations are phenotypically silent in the offspring, but can still influence risk of autism in subsequent generations, perhaps via the interaction with other susceptibility factors. This indirect mechanism is consistent with the evidence that some mutations associated with neurodevelopmental disorders can occur in apparently healthy individuals.[161,162]

6.2.4 Maternal age

In the main analyses of the first study, there was a small association between bipolar disorder and some maternal age categories. When evaluating the paternal and maternal age effects by likelihood ratio tests, comparison of models showed statistical support for including paternal age to the model but not maternal age. Moreover, in study II it was evident that, independent of paternal age, women in the oldest age category were slightly more likely to have an affected child, supporting previous studies that show that autism is independently linked to both paternal and maternal age. By contrast, there was no effect of maternal age on autism risk in the grandchild, suggesting that the maternal age effect is not transmitted to subsequent generations.

6.2.5 The mechanism behind the paternal age effect

The first two studies in this thesis examine paternal age as a risk factor for psychiatric disorder in offspring and grandchildren. The last two studies are instead focused on gaining knowledge about the mechanism behind the paternal age effect.

The initial study showed an association between advancing paternal age and bipolar disorder, an association that was independent of maternal age, family history of psychiatric disorders, parity and SES. The association was even stronger when only including individuals with an early-onset disorder in the analyses; a suggested distinguishable subtype of bipolar disorder that is known to be under a greater genetic influence.

The second study confirmed earlier reports on a link between advancing paternal age and autism but more importantly, the study showed that grandchildren of older men also were at increased risk for developing autism, a finding that has not previously been reported. Again, these effects were not confounded by age of the spouse, family history of psychiatric disorder, SES or residential county.

Both the findings of a stronger link to early-onset bipolar disorder and the finding suggesting that the paternal age effect on autism risk can be transmitted to further generations suggest that the paternal age effect is genetically mediated. The analyses showed that this association was not confounded by a family history of psychiatric disorder. However, these findings might be affected by missing data.

In observational studies, such as the ones included in this thesis, it is never possible to exclude that the results are explained by unmeasured confounding. Since it is not possible to randomize paternal ages, other approaches are necessary in order to study the effects of paternal age. The aim of study III was to gain knowledge about the mechanism behind advancing paternal age by examining if there is a link between advancing paternal age and rare CNVs; structural variations in the genome that increase the risk of schizophrenia and autism. This study showed that, even though rare CNVs were more common in individuals with schizophrenia than in healthy controls and that individuals with schizophrenia had older fathers, there was no association between rare CNVs and paternal age. In conclusion, the study suggests that advancing paternal age and an increased burden of rare CNVs are independent risk factors for schizophrenia. It is, however, still possible that specific subtypes of CNVs are associated with paternal age, for instance *de novo* CNVs.

The final study showed that late fatherhood was indeed a heritable trait. However, paternal ages in individuals affected with ASD, schizophrenia and bipolar disorder or in individuals with affected family members were lower than paternal ages of healthy men. These findings contradict the commonly suggested hypothesis that the association between advancing paternal age and psychiatric disorders is explained by men with a genetic liability for psychiatric disorders systematically having children at older ages. Moreover, the associations between advancing paternal age and psychiatric disorders in not likely explained by women with at genetic liability for psychiatric disorder systematically having children with older men. Previous studies have found a reduced

fertility in individuals with psychiatric disorders. If the link between teenage fathers and psychiatric disorders was merely a result of reduced fertility after disorder onset it would not be evident that they have their first children earlier than the comparison group. Moreover, if the younger ages at fatherhood were explained by nothing more than the disorder onset, the same trend would not be evident in their healthy siblings.

The association between a genetic liability for psychiatric disorders and teenage fatherhood could possibly explain why many studies (including ours) report an association between psychiatric disorders and both older fathers and younger fathers. If genes predisposing for psychiatric disorders in healthy individuals also are associated with teenage fatherhood we would expect an association between teenage fatherhood and disorders in the child, even after controlling for psychiatric history in the parents.

6.2.6 Concluding remarks

In conclusion, the present studies show that paternal age at birth of the offspring is a not only a risk factor for schizophrenia and autism but also for bipolar disorder. These findings support previous reports of a partially common etiology for these psychiatric disorders. The studies also show that the paternal age effect might be particularly strong in certain subtypes of bipolar disorder such as disorders that has an early onset or requires hospitalization. In addition, this thesis show that high paternal age does not only increase the risk of autism in the child, but is also associated with an increased risk for autism in the grandchild, suggesting that the effect is genetically mediated. When studying the mechanism behind the paternal age effect, it was evident that although rare CNVs and high paternal age was linked to an increase risk of schizophrenia in our sample, they were not correlated to each other, suggesting that an increased burden of rare CNVs and advancing paternal age are independent risk factors for schizophrenia. Moreover, the final study showed that old age at fatherhood is indeed a heritable trait. However, individuals with a genetic liability for psychiatric disorders are not, in general, more likely to become fathers at older ages or to have children with an older man. These findings suggest that a psychiatric liability for psychiatric disorders in the parents is not likely the explanation behind the paternal age effect. The present studies indirectly support the hypothesis suggesting that *de novo* point mutations may be involved in mediating the paternal age effect.

Age at parenthood is increasing in many societies and this would predict an increased incidence of disorders that are causally related to advancing paternal age. This thesis provides new knowledge about the psychiatric disorders linked to advancing paternal age and how this effect can be transmitted to subsequent generations. This thesis also provides valuable clues about the mechanism behind the potential effect of late fatherhood. Although causality can only be confirmed by randomized studies, paternal ages are impossible to randomize to human subjects. However, by using different epidemiological study designs and genetic studies as well as potentially combining this knowledge with results from randomized animal studies, it is possible to gain knowledge about the potential mechanisms behind the paternal age effect.

More specifically, the rapid development in new techniques enabling whole genome sequencing of large samples will probably give more valuable clues, especially if analyzing study participants with psychiatric disorders and with genetic material in several generations. It is also possible that increased knowledge within the field of epigenetics will increase our understanding in the mechanism behind the paternal age effect. Increased knowledge within this field might have important implications for clinicians, researchers, and those affected by the disorders. It is also possible that this knowledge will be of importance from a public health perspective.

7 SVENSK SAMMANFATTNING

Över hela värden kan man se att hur föräldrars medelålder stiger. Traditionellt har fokus varit på de negativa hälsoeffekter som hör samman med att bli mamma vid en hög ålder. På senare år har dock en mängd studier visat kopplingar mellan pappans ålder och sjukdomar i avkomman, bland annat har individer med äldre pappor en högre risk att drabbas av autism eller schizofreni. Idag är kunskaperna om dessa associationer begränsade och syftet med denna avhandling var därför att öka dessa kunskaper. Studierna inkluderar data från populationsbaserade nationella register som möjliggör studier av väsentlig storlek och som även gör det möjligt att kontrollera för en mängd kända riskfaktorer bland annat parentala, perinatala och socioekonomiska variabler.

I vår första studie kunde vi, för första gången, visa att det finns ett samband mellan bipolär sjukdom och äldre pappor. Risken steg med pappans ålder och var högst i den äldsta ålderskategorin som inkluderade män som var 55 år och äldre. Jämfört med män som var 20-24 år hade barn till män i den äldsta ålderskategorin 1,37 gånger högre risk att drabbas av bipolär sjukdom. Vi kunde även se att sambandet var starkare när vi separat analyserade personer som insjuknat tidigt i livet.

I den andra studien kunde vi inte bara bekräfta tidigare fynd som visat på kopplingar mellan äldre pappor och autism hos barnet. Vi visade också att barnbarn till män som blev pappor vid en hög ålder har en ökad risk för autism. Återigen ökade risken med stigande ålder på pappan. I den äldsta ålderskategorin var risken för autism 1,79 om morfadern var äldre och 1,67 om farfadern var äldre, jämfört med referensgruppen.

Det har föreslagits att vissa strukturella variationer i genomet som ökar risken för schizofreni kan påverkas av pappans ålder. Vi studerade detta med hjälp av försöksdeltagare med eller utan schizofreni. Resultaten visade att trots att patienterna hade fler strukturella variationer och äldre pappor, fanns det ingen koppling mellan dessa två variabler.

Genom att analysera tvillingar kunde vi se att det fanns en genetisk komponent som påverkar huruvida män skaffar barn sent i livet. Vi kunde dock inte se att personer med en genetisk belastning för psykisk sjukdom oftare blev pappor senare i livet. Detsamma gällde för friska män vars partner hade en genetisk belastning för psykisk sjukdom. Den generella trenden var istället att dessa män blev pappor tidigare jämfört med andra män i samma födelsekohort.

Sammanfattningsvis ger studierna i denna avhandling nya värdefulla kunskaper om effekterna av att bli pappa vid en hög ålder samt etiologin bakom psykisk sjukdom och kan vara av betydelse för forskare, kliniker och de som är drabbade av dessa sjukdomar.

8 ACKNOWLEDGEMENTS

Christina Hultman, my main supervisor: For endless enthusiasm and support on both the professional and personal level. For giving me the opportunity to make this journey and for inspiring me to continue within the field of research.

Paul Lichtenstein, my co-supervisor: For letting me be a part of your excellent research group and for always making time for me in your busy schedule. For finding a perfect balance between keeping me grounded but also letting me use my own wings.

Niklas Långström, my co-supervisor: For being a great inspiration when it comes to communicating science beyond the academic community. For a clinical perspective and valuable input on my work.

Sven Sandin, co-author: For guiding me through the jungle of biostatistics, programming and data interpretation. For always taking time to give me you invaluable advice.

Patrik Magnusson, co-author: For having the perfect combination of scientific wisdom and wonderful sense of humor.

Ralf Kuja-Halkola, co-author: For being a great friend as well as a wonderful collaborator. For sharing my love for coffee and conversation.

My international co-authors **Avi Reichenberg** and **John McGrath**: For invaluable contribution to my studies.

Henrik Larsson: For great advice that goes beyond my work as a PhD student. For taking an interest in my future goals and aspirations.

Hasse Walum: For your friendship, support and inspiration. I am looking forward to making reality out of our high-flying plans.

Erik Pettersson: For providing me with graphs and snacks when I need them the most.

Anna Kähler: For keeping me up to date on what's happening in the world of molecular genetics and psychiatry and for being great company over lunch or coffee.

Therese Ljung: For connecting with me on many levels including work, pregnancies and motherhood.

Nancy Pedersen and **Henrik Grönberg**, past and present head of the department for Medical Epidemiology and Biostatistics (MEB): For creating such an inspiring and creative work environment at MEB.

My present and past colleagues at **MEB** for a friendly and creative work environment.

My best friends **Olivia**, **Susanna** and **Maria** for always listening to my ideas and giving me new perspectives in all aspects of life.

All other friends, too many to mention, that encourage me to pursue a career as a researcher. The last couple of months, I have been overwhelmed by your support.

Jocke, my brother: For being the best brother I could wish for and for bringing **Tilda**, **Maja** and **Greta** into my life.

Monica, my mother: For love and support on all levels. For your valuable proof-reading skills and input on this thesis.

Örjan, my father: For endless love and support. For encouraging me to peruse an academic career and for believing in my capability.

Pelle and **Beth**, my in-laws: For your support and for welcoming me into your family. Most of all, for being wonderful grandparents and for always making time for Klara and Valter.

My dearest children **Valter** and **Klara**, for keeping me grounded and giving me endless love and joy in life.

Hans, love of my life: For supporting my ambitions with never-ending enthusiasm. For your love and patience and for being the best husband and father in the world.

9 REFERENCES

1. Organisation for Economic Co-operation and Development (OECD). *OECD Family Database.* Paris2012.

2. Mills M, Rindfuss RR, McDonald P, te Velde E. Why do people postpone parenthood? Reasons and social policy incentives. *Hum Reprod Update.* 2011;17(6):848-860.

3. Hardy JB, Astone NM, Brooks-Gunn J, Shapiro S, Miller TL. Like mother, like child: intergenerational patterns of age at first birth and associations with childhood and adolescent characteristics and adult outcomes in the second generation. *Dev Psychol.* 1998;34(6):1220-1232.

4. Leridon H. Can assisted reproduction technology compensate for the natural decline in fertility with age? A model assessment. *Hum Reprod.* l 2004;19(7):1548-1553.

5. Leridon H. A new estimate of permanent sterility by age: sterility defined as the inability to conceive. *Popul Stud (Camb).* 2008;62(1):15-24.

6. Jacobsson B, Ladfors L, Milsom I. Advanced maternal age and adverse perinatal outcome. *Obstet Gynecol.* 2004;104(4):727-733.

7. Hassold T, Hunt P. Maternal age and chromosomally abnormal pregnancies: what we know and what we wish we knew. *Current opinion in pediatrics.* 2009;21(6):703-708.

8. Kidd SA, Eskenazi B, Wyrobek AJ. Effects of male age on semen quality and fertility: a review of the literature. *Fertil Steril.* 2001;75(2):237-248.

9. Toriello HV, Meck JM, Professional P, Guidelines C. Statement on guidance for genetic counseling in advanced paternal age. *Genetics in medicine : official journal of the American College of Medical Genetics.* 2008;10(6):457-460.

10. American Psychiatric Association., American Psychiatric Association. DSM-5 Task Force. *Diagnostic and statistical manual of mental disorders : DSM-5.* 5. ed. Arlington, Va.: American Psychiatric Association; 2013.

11. Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social psychiatry and psychiatric epidemiology.* 1998;33(12):587-595.

12. Lichtenstein P, Bjork C, Hultman CM, Scolnick E, Sklar P, Sullivan PF. Recurrence risks for schizophrenia in a Swedish national cohort. *Psychological medicine.* 2006;36(10):1417-1425.

13. Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med.* 2005;2(5):e141.

14. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry.* 2003;60(12):1187-1192.

15. Lichtenstein P, Yip BH, Bjork C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet.* 2009;373(9659):234-239.

16. Haukka J, Suvisaari J, Lonnqvist J. Fertility of patients with schizophrenia, their siblings, and the general population: a cohort study from 1950 to 1959 in Finland. *Am J Psychiatry.* 2003;160(3):460-463.

17. Power RA, Kyaga S, Uher R, et al. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. *JAMA Psychiatry.* 2013;70(1):22-30.

18. Svensson AC, Lichtenstein P, Sandin S, Hultman CM. Fertility of first-degree relatives of patients with schizophrenia: a three generation perspective. *Schizophr Res.* 2007;91(1-3):238-245.

19. Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet.* 2013.

20. Schizophrenia Psychiatric Genome-Wide Association Study C. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet.* 2011;43(10):969-976.
21. Chakrabarti S, Fombonne E. Pervasive developmental disorders in preschool children: confirmation of high prevalence. *Am J Psychiatry.* 2005;162(6):1133-1141.
22. Newschaffer CJ, Croen LA, Daniels J, et al. The epidemiology of autism spectrum disorders. *Annu Rev Public Health.* 2007;28:235-258.
23. Lichtenstein P, Carlstrom E, Rastam M, Gillberg C, Anckarsater H. The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am J Psychiatry.* 2010;167(11):1357-1363.
24. Persico AM, Napolioni V. Autism genetics. *Behavioural brain research.* 2013;251:95-112.
25. Goodwin FK, Jamison KR. *Manic-depressive illness : bipolar disorders and recurrent depression.* 2. ed. New York, N.Y.: Oxford University Press; 2007.
26. Merikangas KR, Jin R, He JP, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry.* 2011;68(3):241-251.
27. Bellivier F, Golmard JL, Henry C, Leboyer M, Schurhoff F. Admixture analysis of age at onset in bipolar I affective disorder. *Arch Gen Psychiatry.* 2001;58(5):510-512.
28. Bellivier F, Golmard JL, Rietschel M, et al. Age at onset in bipolar I affective disorder: further evidence for three subgroups. *Am J Psychiatry.* 2003;160(5):999-1001.
29. Schurhoff F, Bellivier F, Jouvent R, et al. Early and late onset bipolar disorders: two different forms of manic-depressive illness? *Journal of affective disorders.* 2000;58(3):215-221.
30. Carter TD, Mundo E, Parikh SV, Kennedy JL. Early age at onset as a risk factor for poor outcome of bipolar disorder. *Journal of psychiatric research.* 2003;37(4):297-303.
31. Grigoroiu-Serbanescu M, Martinez M, Nothen MM, et al. Different familial transmission patterns in bipolar I disorder with onset before and after age 25. *American journal of medical genetics.* 2001;105(8):765-773.
32. Somanath CP, Jain S, Reddy YC. A family study of early-onset bipolar I disorder. *Journal of affective disorders.* Jun 2002;70(1):91-94.
33. Byrne M, Agerbo E, Ewald H, Eaton WW, Mortensen PB. Parental age and risk of schizophrenia: a case-control study. *Arch Gen Psychiatry.* 2003;60(7):673-678.
34. Johanson E. A study of schizophrenia in the male: a psychiatric and social study based on 138 cases with follow up. *Acta psychiatrica et neurologica Scandinavica. Supplementum.* 1958;125:1-132.
35. Malaspina D, Harlap S, Fennig S, et al. Advancing paternal age and the risk of schizophrenia. *Arch Gen Psychiatry.* 2001;58(4):361-367.
36. Zammit S, Allebeck P, Dalman C, et al. Paternal age and risk for schizophrenia. *Br J Psychiatry.* 2003;183:405-408.
37. Bertranpetit J, Fananas L. Parental age in schizophrenia in a case-controlled study. *Br J Psychiatry.* 1993;162:574.
38. Miller B, Messias E, Miettunen J, et al. Meta-analysis of paternal age and schizophrenia risk in male versus female offspring. *Schizophr Bull.* 2011;37(5):1039-1047.
39. Sipos A, Rasmussen F, Harrison G, et al. Paternal age and schizophrenia: a population based cohort study. *BMJ.* 2004;329(7474):1070.
40. Malaspina D, Corcoran C, Fahim C, et al. Paternal age and sporadic schizophrenia: evidence for de novo mutations. *American journal of medical genetics.* 2002;114(3):299-303.
41. Petersen L, Mortensen PB, Pedersen CB. Paternal age at birth of first child and risk of schizophrenia. *Am J Psychiatry.* 2011;168(1):82-88.
42. Allen J, DeMeyer MK, Norton JA, Pontius W, Yang E. Intellectuality in parents of psychotic, subnormal, and normal children. *Journal of autism and childhood schizophrenia.* 1971;1(3):311-326.

43. Treffert DA. Epidemiology of infantile autism. *Arch Gen Psychiatry.* 1970;22(5):431-438.
44. Glasson EJ, Bower C, Petterson B, de Klerk N, Chaney G, Hallmayer JF. Perinatal factors and the development of autism: a population study. *Arch Gen Psychiatry.* 2004;61(6):618-627.
45. Lauritsen MB, Pedersen CB, Mortensen PB. Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study. *Journal of child psychology and psychiatry, and allied disciplines.* 2005;46(9):963-971.
46. Reichenberg A, Gross R, Weiser M, et al. Advancing paternal age and autism. *Arch Gen Psychiatry.* 2006;63(9):1026-1032.
47. Grether JK, Anderson MC, Croen LA, Smith D, Windham GC. Risk of autism and increasing maternal and paternal age in a large north American population. *Am J Epidemiol.* 1 2009;170(9):1118-1126.
48. Hultman CM, Sandin S, Levine SZ, Lichtenstein P, Reichenberg A. Advancing paternal age and risk of autism: new evidence from a population-based study and a meta-analysis of epidemiological studies. *Mol Psychiatry.* 2011;16(12):1203-1212.
49. Pasqualotto FF, Borges Junior E, Pasqualotto EB. The male biological clock is ticking: a review of the literature. *Sao Paulo medical journal = Revista paulista de medicina.* 2008;126(3):197-201.
50. Hassan MA, Killick SR. Effect of male age on fertility: evidence for the decline in male fertility with increasing age. *Fertil Steril.* 2003;79 Suppl 3:1520-1527.
51. Klonoff-Cohen HS, Natarajan L. The effect of advancing paternal age on pregnancy and live birth rates in couples undergoing in vitro fertilization or gamete intrafallopian transfer. *Am J Obstet Gynecol.* 2004;191(2):507-514.
52. Kleinhaus K, Perrin M, Friedlander Y, Paltiel O, Malaspina D, Harlap S. Paternal age and spontaneous abortion. *Obstet Gynecol.* 2006;108(2):369-377.
53. Andersen AMN, Hansen KD, Andersen PK, Davey Smith G. Advanced paternal age and risk of fetal death: A cohort study. *American Journal of Epidemiology.* 2004;160(12):1214-1222.
54. Crow JF. The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet.* 2000;1(1):40-47.
55. Penrose LS. Parental age and mutation. *Lancet.* 1955;269(6885):312-313.
56. Glaser RL, Jiang W, Boyadjiev SA, et al. Paternal origin of FGFR2 mutations in sporadic cases of Crouzon syndrome and Pfeiffer syndrome. *Am J Hum Genet.* 2000;66(3):768-777.
57. Passos-Bueno MR, Wilcox WR, Jabs EW, Sertie AL, Alonso LG, Kitoh H. Clinical spectrum of fibroblast growth factor receptor mutations. *Human mutation.* 1999;14(2):115-125.
58. Moloney DM, Slaney SF, Oldridge M, et al. Exclusive paternal origin of new mutations in Apert syndrome. *Nat Genet.* 1996;13(1):48-53.
59. Risch N, Reich EW, Wishnick MM, McCarthy JG. Spontaneous mutation and parental age in humans. *Am J Hum Genet.* 1987;41(2):218-248.
60. Tolarova MM, Harris JA, Ordway DE, Vargervik K. Birth prevalence, mutation rate, sex ratio, parents' age, and ethnicity in Apert syndrome. *American journal of medical genetics.* 1997;72(4):394-398.
61. Glaser RL, Broman KW, Schulman RL, Eskenazi B, Wyrobek AJ, Jabs EW. The paternal-age effect in Apert syndrome is due, in part, to the increased frequency of mutations in sperm. *Am J Hum Genet.* 2003;73(4):939-947.
62. Orioli IM, Castilla EE, Scarano G, Mastroiacovo P. Effect of paternal age in achondroplasia, thanatophoric dysplasia, and osteogenesis imperfecta. *American journal of medical genetics.* 1995;59(2):209-217.
63. Tavormina PL, Shiang R, Thompson LM, et al. Thanatophoric dysplasia (types I and II) caused by distinct mutations in fibroblast growth factor receptor 3. *Nat Genet.* 1995;9(3):321-328.
64. Jung A, Schuppe HC, Schill WB. Are children of older fathers at risk for genetic disorders? *Andrologia.* 2003;35(4):191-199.
65. Racine SE, Culbert KM, Burt SA, Klump KL. Advanced paternal age at birth: phenotypic and etiologic associations with eating pathology in offspring. *Psychological medicine.* 2013:1-13.

66. Yip BH, Pawitan Y, Czene K. Parental age and risk of childhood cancers: a population-based cohort study from Sweden. *International journal of epidemiology.* 2006;35(6):1495-1503.
67. Olshan AF, Schnitzer PG, Baird PA. Paternal age and the risk of congenital heart defects. *Teratology.* 1994;50(1):80-84.
68. Saha S, Barnett AG, Foldi C, et al. Advanced paternal age is associated with impaired neurocognitive outcomes during infancy and childhood. *PLoS Med.* 2009;6(3):e40.
69. Kuja-Halkola R, Pawitan Y, D'Onofrio BM, Langstrom N, Lichtenstein P. Advancing paternal age and offspring violent offending: a sibling-comparison study. *Development and psychopathology.* 2012;24(3):739-753.
70. Bertram L, Busch R, Spiegl M, Lautenschlager NT, Muller U, Kurz A. Paternal age is a risk factor for Alzheimer disease in the absence of a major gene. *Neurogenetics.* 1998;1(4):277-280.
71. Vestergaard M, Mork A, Madsen KM, Olsen J. Paternal age and epilepsy in the offspring. *European journal of epidemiology.* 2005;20(12):1003-1005.
72. Vogel F. A probable sex difference in some mutation rates. *Am J Hum Genet.* 1977;29(3):312-319.
73. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics.* 1998;148(4):1667-1686.
74. Bosch M, Rajmil O, Egozcue J, Templado C. Linear increase of structural and numerical chromosome 9 abnormalities in human sperm regarding age. *Eur J Hum Genet.* 2003;11(10):754-759.
75. Malaspina D. Paternal factors and schizophrenia risk: de novo mutations and imprinting. *Schizophr Bull.* 2001;27(3):379-393.
76. Tarin JJ, Brines J, Cano A. Long-term effects of delayed parenthood. *Hum Reprod.* 1998;13(9):2371-2376.
77. Wyrobek AJ, Eskenazi B, Young S, et al. Advancing age has differential effects on DNA damage, chromatin integrity, gene mutations, and aneuploidies in sperm. *Proc Natl Acad Sci U S A.* 2006;103(25):9601-9606.
78. Eyre-Walker A, Keightley PD. High genomic deleterious mutation rates in hominids. *Nature.* 1999;397(6717):344-347.
79. Fay JC, Wyckoff GJ, Wu CI. Positive and negative selection on the human genome. *Genetics.* 2001;158(3):1227-1234.
80. Sunyaev S, Ramensky V, Koch I, Lathe W, 3rd, Kondrashov AS, Bork P. Prediction of deleterious human alleles. *Hum Mol Genet.* 2001;10(6):591-597.
81. Keller MC, Miller G. Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behav Brain Sci.* 2006;29(4):385-404; discussion 405-352.
82. Girard SL, Gauthier J, Noreau A, et al. Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet.* Sep 2011;43(9):860-863.
83. Xu B, Ionita-Laza I, Roos JL, et al. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet.* 2012;44(12):1365-1369.
84. Xu B, Roos JL, Dexheimer P, et al. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet.* 2011;43(9):864-868.
85. Iossifov I, Ronemus M, Levy D, et al. De novo gene disruptions in children on the autistic spectrum. *Neuron.* 2012;74(2):285-299.
86. Michaelson JJ, Shi Y, Gujral M, et al. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell.* 2012;151(7):1431-1442.
87. Neale BM, Kou Y, Liu L, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature.* 10 2012;485(7397):242-245.
88. O'Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature.* 2012;485(7397):246-250.
89. Sanders SJ, Murtha MT, Gupta AR, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature.* 2012;485(7397):237-241.

90. Kong A, Frigge ML, Masson G, et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature.* 2012;488(7412):471-475.
91. Goriely A, Wilkie AO. Paternal age effect mutations and selfish spermatogonial selection: causes and consequences for human disease. *Am J Hum Genet.* 2012;90(2):175-200.
92. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annual review of medicine.* 2010;61:437-455.
93. Kirov G, Pocklington AJ, Holmans P, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry.* 2012;17(2):142-153.
94. Malhotra D, McCarthy S, Michaelson JJ, et al. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron.* 2011;72(6):951-963.
95. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet.* 2008;40(7):880-885.
96. Levy D, Ronemus M, Yamrom B, et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron.* 2011;70(5):886-897.
97. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science.* 2007;316(5823):445-449.
98. Rees E, Moskvina V, Owen MJ, O'Donovan MC, Kirov G. De Novo Rates and Selection of Schizophrenia-Associated Copy Number Variants. *Biol Psychiatry.* 2011.
99. Andreassen R, Lundsted J, Olaisen B. Mutation at minisatellite locus DYF155S1: allele length mutation rate is affected by age of progenitor. *Electrophoresis.* 2002;23(15):2377-2383.
100. Ellegren H. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.* 2000;16(12):551-558.
101. Singh NP, Muller CH, Berger RE. Effects of age on DNA double-strand breaks and apoptosis in human sperm. *Fertil Steril.* 2003;80(6):1420-1430.
102. Hehir-Kwa JY, Rodriguez-Santiago B, Vissers LE, et al. De novo copy number variants associated with intellectual disability have a paternal origin and age bias. *Journal of medical genetics.* 2011;48(11):776-778.
103. Sibbons C, Morris JK, Crolla JA, Jacobs PA, Thomas NS. De novo deletions and duplications detected by array CGH: a study of parental origin in relation to mechanisms of formation and size of imbalance. *Eur J Hum Genet.* 2012;20(2):155-160.
104. Buizer-Voskamp JE, Blauw HM, Boks MP, et al. Increased paternal age and the influence on burden of genomic copy number variation in the general population. *Human genetics.* 2013;132(4):443-450.
105. Buizer-Voskamp JE, Muntjewerff JW, Genetic R, et al. Genome-wide analysis shows increased frequency of copy number variation deletions in Dutch schizophrenia patients. *Biol Psychiatry.* 2011;70(7):655-662.
106. Perrin MC, Brown AS, Malaspina D. Aberrant epigenetic regulation could explain the relationship of paternal age to schizophrenia. *Schizophr Bull.* 2007;33(6):1270-1273.
107. Curley JP, Mashoodh R, Champagne FA. Epigenetics and the origins of paternal effects. *Hormones and behavior.* 2011;59(3):306-314.
108. Granville-Grossman KL. Parental age and schizophrenia. *Br J Psychiatry.* 1966;112(490):899-905.
109. Miller B, Suvisaari J, Miettunen J, et al. Advanced paternal age and parental history of schizophrenia. *Schizophr Res.* 2011;133(1-3):125-132.
110. Nilsen AB, Waldenstrom U, Rasmussen S, Hjelmstedt A, Schytt E. Characteristics of first-time fathers of advanced age: a Norwegian population-based study. *BMC pregnancy and childbirth.* 2013;13:29.
111. Sandin S, Hultman CM, Kolevzon A, Gross R, MacCabe JH, Reichenberg A. Advancing maternal age is associated with increasing risk for autism: a review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2012;51(5):477-486 e471.
112. Frans EM, McGrath JJ, Sandin S, et al. Advanced paternal and grandpaternal age and schizophrenia: A three-generation perspective. *Schizophr Res.* 2011.

113. Lunde AS, Lundeborg S, Lettenstrom GS, Thygesen L, Huebner J. The person-number systems of Sweden, Norway, Denmark, and Israel. *Vital and health statistics. Series 2, Data evaluation and methods research.* 1980(84):1-59.
114. World Health Organisation. *International Classification of Diseases (8th Revision).* Geneva: World Health Organisation;1967.
115. World Health Organisation. *International Classification of Diseases (9th Revision).* Geneva1977.
116. World Health Organisation. *International Classification of Diseases (10th Revision).* Geneva1992.
117. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC Public Health.* 2011;11:450.
118. Statistics Sweden. *Multi-Generation Register 2004 - A description of contents and quality.* 2005.
119. Lichtenstein P, Sullivan PF, Cnattingius S, et al. The Swedish Twin Registry in the third millennium: an update. *Twin Res Hum Genet.* 2006;9(6):875-882.
120. Magnusson PK, Almqvist C, Rahman I, et al. The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res Hum Genet.* 2013;16(1):317-329.
121. Statistics Sweden. The Swedish Register of Education. 2004.
122. Statistics Sweden. Longitudinal integration database for health insurance and labour market studies (LISA by Swedish acronym).
123. Swedish Tax Agency. Population registration in Sweden SKV 717B. Swedish Tax Agency, 4 ed. Stockholm 2007.
124. The National Board of Health and Welfare Causes of Death 2010. Stockholm2011.
125. Korn JM, Kuruvilla FG, McCarroll SA, et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet.* 2008;40(10):1253-1260.
126. Rothman KJ. *Epidemiology : an introduction.* Oxford: Oxford University Press; 2002.
127. Neale MC, Cardon LR, North Atlantic Treaty Organization. Scientific Affairs Division. *Methodology for genetic studies of twins and families.* Dordrecht: Kluwer Academic; 1992.
128. Purcell S. Variance components models for gene-environment interaction in twin analysis. *Twin Res.* 2002;5(6):554-571.
129. Boker S, Neale M, Maes H, et al. OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika.* 2011;76(2):306-317.
130. Sellgren C, Landen M, Lichtenstein P, Hultman CM, Langstrom N. Validity of bipolar disorder hospital discharge diagnoses: file review and multiple register linkage in Sweden. *Acta psychiatrica Scandinavica.* 2011;124(6):447-453.
131. Ekholm B, Ekholm A, Adolfsson R, et al. Evaluation of diagnostic procedures in Swedish patients with schizophrenia and related psychoses. *Nordic journal of psychiatry.* 2005;59(6):457-464.
132. McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry.* 1991;48(8):764-770.
133. International Schizophrenia C, Purcell SM, Wray NR, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature.* 2009;460(7256):748-752.
134. Bellis MA, Hughes K, Hughes S, Ashton JR. Measuring paternal discrepancy and its public health consequences. *Journal of epidemiology and community health.* 2005;59(9):749-754.
135. Fellman J, Eriksson AW. Stillbirth rates in singletons, twins and triplets in Sweden, 1869 to 2001. *Twin Res Hum Genet.* 2006;9(2):260-265.
136. Klein JP, Moeschberger ML. *Survival analysis : techniques for censored and truncated data.* New York: Springer; 1997.
137. Record RG, McKeown T, Edwards JH. An investigation of the difference in measured intelligence between twins and single births. *Annals of human genetics.* 1970;34(1):11-20.

138. Swedish Parliament. Lag (2003:460) om etikprövning av forskning som avser människor. 2003.
139. Fisher RA. *The genetical theory of natural selection.* Oxf.,1930.
140. Kimura M. *The neutral theory of molecular evolution.* Cambridge: Cambridge University Press; 1983.
141. Burns JK. An evolutionary theory of schizophrenia: cortical connectivity, metarepresentation, and the social brain. *Behav Brain Sci.* 2004;27(6):831-855; discussion 855-885.
142. Kyaga S, Landen M, Boman M, Hultman CM, Langstrom N, Lichtenstein P. Mental illness, suicide and creativity: 40-year prospective total population study. *Journal of psychiatric research.* 2013;47(1):83-90.
143. Kyaga S, Lichtenstein P, Boman M, Hultman C, Langstrom N, Landen M. Creativity and mental disorder: family study of 300,000 people with severe mental disorder. *Br J Psychiatry.* 2011;199(5):373-379.
144. Sherman JA. Evolutionary origin of bipolar disorder-revised: EOBD-R. *Medical hypotheses.* 2012;78(1):113-122.
145. Book JA. A genetic and neuropsychiatric investigation of a North-Swedish population with special regard to schizophrenia and mental deficiency. II. Mental deficiency and convulsive disorders. *Acta genetica et statistica medica.* 1953;4(4):345-414.
146. Book JA. Schizophrenia as a gene mutation. *Acta genetica et statistica medica.* 1953;4(2-3):133-139.
147. Haldane JBS, Jayakar SD. Equilibrium between Mutation and Selection in Bisexual Diploids. *J Genet.* 1972;61(1):1-13.
148. Pope HG, Jr, Lipinski JF, Jr. Diagnosis in schizophrenia and manic-depressive illness: A reassessment of the specificity of 'schizophrenic' symptoms in the light of current research. *Archives of General Psychiatry.* 1978;35(7):811-828.
149. Owen MJ, Craddock N. Diagnosis of functional psychoses: time to face the future. *Lancet.* 2009;373(9659):190-191.
150. Rutter M. Childhood schizophrenia reconsidered. *Journal of autism and childhood schizophrenia.* 1972;2(4):315-337.
151. Sullivan PF, Magnusson C, Reichenberg A, et al. Family history of schizophrenia and bipolar disorder as risk factors for autism. *Arch Gen Psychiatry.* 2012;69(11):1099-1103.
152. Larsson HJ, Eaton WW, Madsen KM, et al. Risk factors for autism: perinatal factors, parental psychiatric history, and socioeconomic status. *Am J Epidemiol.* 2005;161(10):916-925; discussion 926-918.
153. Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet.* 2013;45(9):984-994.
154. Menezes PR, Lewis G, Rasmussen F, et al. Paternal and maternal ages at conception and risk of bipolar affective disorder in their offspring. *Psychological medicine.* 2010;40(3):477-485.
155. Brown A, Bao Y, McKeague I, Shen L, Schaefer C. Parental age and risk of bipolar disorder in offspring. *Psychiatry research.* 2013;208(3):225-231.
156. Grigoroiu-Serbanescu M, Wickramaratne PJ, Mihailescu R, et al. Paternal age effect on age of onset in bipolar I disorder is mediated by sex and family history. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics.* 2012;159B(5):567-579.
157. Crow JF. The high spontaneous mutation rate: is it a health risk? *Proc Natl Acad Sci U S A.* 1997;94(16):8380-8386.
158. Yasuda N, Kondo K. The effect of parental age on rate of mutation for Duchenne Muscular dystrophy. *American journal of medical genetics.* 1982;13(1):91-99.
159. Herrmann J. [The effect of the age of the father at conception on the mutations to hemphilia A]. *Humangenetik.* 1966;3(1):1-16.
160. Golding J, Steer C, Pembrey M. Parental and grandparental ages in the autistic spectrum disorders: a birth cohort study. *PLoS One.* 2010;5(4):e9939.

161. Cooper GM, Coe BP, Girirajan S, et al. A copy number variation morbidity map of developmental delay. *Nat Genet.* 2011;43(9):838-846.
162. O'Donovan MC, Kirov G, Owen MJ. Phenotypic variations on the theme of CNVs. *Nat Genet.* 2008;40(12):1392-1393.