Psychotic disorders in the Lundby population 1947-1997: Incidence, life-time prevalence and predictors related to personality and behaviour

Bogren, Mats

2009

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
‘In three words I can sum up everything I’ve learned about life: it goes on.’

Robert Frost
5.4.5 Age-at-onset of nonorganic nonaffective acute remitting psychoses ............................................ 40
5.4.6 Age-at-onset of affective psychoses .......................................................... 40

5.5 Studies of the incidence of psychotic disorders ................. 41
5.5.1 Overview .................................................................................. 41
5.5.2 Incidence of psychotic disorder due to a general medical condition ........................................ 41
5.5.3 Incidence of substance-induced psychotic disorder .......... 41
5.5.4 Incidence of schizophrenia .......................................................... 42
5.5.5 Incidence of delusional disorder .................................................... 44
5.5.6 Incidence of nonorganic nonaffective acute remitting psychoses ........................................ 44
5.5.7 Incidence of affective psychoses ..................................................... 45

5.6 Studies of the life time prevalence of psychotic disorders ..... 45
5.6.1 Overview .................................................................................. 45
5.6.2 Lifetime prevalence of ‘psychotic disorder due to a general medical condition’ and ‘substance-induced psychotic disorder’ .......................................................... 46
5.6.3 Lifetime prevalence of nonaffective psychoses ........................................ 46
5.6.4 Lifetime prevalence of schizophrenia ............................................ 46
5.6.5 Lifetime prevalence of other nonaffective psychoses ............ 47
5.6.6 Lifetime prevalence of affective psychoses ..................................... 47

5.7 Studies of predictors of psychotic disorders ...................... 47
5.7.1 Overview .................................................................................. 47
5.7.2 Premorbid personality traits/behaviours as predictors of psychosis .................................................. 50

6. Aims of the thesis ............................................................................. 57

7. Material and methods ................................................................. 59
7.1 Overview .................................................................................. 59
7.2 The Lundby area ........................................................................ 60
7.3 The Lundby population ............................................................... 60
7.4 Case-finding method .................................................................... 65
7.5 Ascertainment of diagnosis ........................................................ 68
8. Epidemiological methods ................................................................. 73
  8.1 Paper I ...................................................................................... 73
  8.2 Paper II .................................................................................... 73
  8.3 Paper III ................................................................................... 73
  8.4 Paper IV ................................................................................... 74

9. Statistical methods ....................................................................... 75
  9.1 Paper I ...................................................................................... 75
  9.2 Paper II .................................................................................... 75
  9.3 Paper III ................................................................................... 76
  9.4 Paper IV ................................................................................... 77

10. Ethical approval ............................................................................. 79

11. Results .......................................................................................... 81
  11.1 Paper I .................................................................................... 81
  11.2 Paper II .................................................................................. 81
  11.3 Paper III ................................................................................. 83
  11.4 Paper IV .................................................................................. 85

12. General discussion ......................................................................... 89
  12.1 Overall incidence of psychotic disorders in males and females ................................................................. 89
  12.2 Age-at-onset and incidence by age of psychotic disorders ........................................................................ 90
  12.3 Period prevalence of psychotic disorders 1947-1997 .......... 91
  12.4 Lifetime prevalence of psychotic disorders 1997 .......... 91
  12.5 Predictors of psychosis related to personality and behaviour ....................................................................... 94
  12.6 Methodological issues ........................................................... 97
    12.6.1 Population representativity ............................................. 97
    12.6.2 Sampling ......................................................................... 98
12.6.3 Attrition ................................................................. 99
12.6.4 Case finding method and diagnostic ascertainment ..... 100

13. Conclusions................................................................. 103

14. Svensk populärvetenskaplig sammanfattning ............... 105

15. Acknowledgement....................................................... 109

16. References................................................................. 111
2. Abbreviations

**ABC Schizophrenia Study**: Age Beginning Course of Schizophrenia Study  
**ECA Study**: Epidemiologic Catchment Area Study  
**AESOP Study**: Aetiology and Ethnicity of Schizophrenia and other Psychoses Study  
**BPRS**: Brief Psychiatric Rating Scale  
**CIDI**: Composite International Diagnostic Interview  
**DIS**: Diagnostic Interview Schedule  
**DSM**: Diagnostic and Statistical Manual of Mental Disorders  
**GAF**: Global Assessment of Functioning  
**GHS**: German Health Interview and Examination Survey  
**ICD**: International Statistical Classification of Diseases and Related Health Problems  
**NARP**: Nonaffective Acute Remitting Psychoses  
**NCS**: National Comorbidity Survey  
**NEMESIS**: Netherlands Mental Health Survey and Incidence Study  
**NOS**: Not Otherwise Specified  
**PIF Study**: Psychosis in Finland Study  
**PSE**: Present State Examination  
**SCAN**: Schedules for Clinical Assessment in Neuropsychiatry  
**SCID-I**: Structured Clinical Interview for DSM Axis I Disorders  
**SPSS**: Statistical Package for the Social Sciences  
**WHO DOSMeD Study**: World Health Organization Determinants of Severe Mental Disorder Study (also referred to as the Ten Country Study)
3. Original papers

The thesis presents and discusses the following papers:

### 3.1 Paper I

### 3.2 Paper II

### 3.3 Paper III

### 3.4 Paper IV

The papers are reprinted with permission from the publishers.
4. Introduction

Psychosis refers to a complex group of experiences and behaviours such as: hallucinations, delusions, disorganized speech/thought and grossly disorganized or catatonic behaviour. Although the term – which is purely descriptive – is not equal to mental disorder or to some specific organic or psychological state, psychosis is nevertheless a salient feature of a group of severe mental disorders.

In the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) the major disorders that present with psychotic features are included under ‘schizophrenia and other psychotic disorders’ and ‘mood disorders’, respectively (American Psychiatric Association, 1994). The disorders are: ‘schizophrenia’, ‘schizophreniform disorder’, ‘schizoaffective disorder’, ‘delusional disorder’, ‘brief psychotic disorder’, ‘shared psychotic disorder’ ‘psychotic disorder due to a general medical condition’, ‘substance-induced psychotic disorder’ and ‘psychotic disorder not otherwise specified’, and ‘bipolar disorder with psychotic features’ and ‘major depressive disorder with psychotic features’, respectively. Thus, the group of mental disorders associated with psychosis is quite heterogeneous.

Although psychotic disorders have been known since ancient times (Evans et al, 2003; Angst and Marneros, 2001), their frequencies, distributions and determinants in the general population are still insufficiently described. Nevertheless, for schizophrenia there have been several studies analyzing its frequency in community populations and there is evidence for variability between: urban and rural areas, males and females, immigrants and native born people and across time, which is not explainable just by methodological differences between the studies (McGrath, 2006).

In recent reviews the median incidence of ‘schizophrenia’ in the general population has been estimated to be 15.2 per 100 000 person-years at risk (McGrath et al, 2004) and the median lifetime prevalence to be 0.4% (Saha et al 2005), although with quite big variation between different studies. For
the other psychotic disorders community studies are sparse. Nevertheless, these disorders are clearly also rare with probable rates around a few new cases per 100,000 person-years at risk (Kaplan and Sadock, 1994; Susser and Wanderling, 1994; Castagnini et al, 2008).

Notwithstanding the low incidences of the psychoses, in a lifetime perspective, psychotic disorders are more common than one might first guess. Community studies of lifetime prevalence suggest rates of any kind of psychotic disorder between 2.9% and 4.5% (Astrup, 1989; Perälä et al, 2007; van Os et al, 2001; Jacobi et al, 2004).

For the best researched psychotic disorder – ‘schizophrenia’ – the most typical age-at-onset is during the late teens and early 20s (Andreasen, 1999), but onset in older age also occurs (Castle and Murray, 1993; Henderson and Kay, 1997). The average onset of mania/bipolar I disorder is also generally regarded to occur during the early adult years (Lloyd and Jones, 2002), but there are some variation between studies indicating also later onsets. The age-at-onset for the other psychotic disorders has been insufficiently studied in community populations.

It has been shown that males, on average, develop ‘schizophrenia’ earlier in life than females (Hambrecht et al, 1992; Castle and Murray, 1993). Similarly, one community study suggested that the age-related risk for nonaffective acute remitting psychoses may also differ between the sexes in that the male risk was highest in the younger age groups whereas the female risk was highest in the older age groups (Castagnini et al, 2008). As opposed to studies of schizophrenia and other nonaffective psychoses most studies of affective psychoses indicate that the age-at-onset and incidence does not differ significantly between the sexes – although most studies have only investigated mania/bipolar disorder (Bland, 1977; Bland et al, 1988; Bebbington and Ramana, 1995; Hendrick et al, 2000; Baldwin et al, 2005; Kawa et al, 2005).

Studies from the last decades have indicated that males and females may be differently prone to develop nonaffective psychoses (Menezes et al, 2007). The overall risk to develop ‘schizophrenia’ may be greater in males than females (Aleman et al, 2003) – although previously it was generally held that ‘schizophrenia’ affects males and females equally (Bromet et al, 2002) – while the risk to develop nonaffective acute remitting psychoses (NARP)
– e.g. ‘schizotypic disorder’, ‘brief psychotic disorder’ and ‘psychotic disorder not otherwise specified’ – may be greater in females than in males (Susser and Wanderling, 1994).

Except for ‘psychotic disorder due to a general medical condition’ and ‘substance-induced psychotic disorder’ – which are defined by their causes – the causation of psychosis is unknown. Moreover, studies on the determinants of risk for nonorganic psychosis have mostly been related to ‘schizophrenia’ (Bromet et al, 2002). There are many factors known to be statistically associated with an increased risk to develop ‘schizophrenia’: a family history of psychosis, pregnancy and birth complications, various postnatal biologic and social exposures – e.g. childhood central nervous system infection, adolescent drug use, urban upbringing and migration – and certain biologic, neuropsychological and behavioural traits – e.g. enlargement of brain ventricles, delayed childhood development, impairments in attention and memory domains, social adjustment difficulties and schizotypal traits (Bromet et al, 2002). However, it is likely that some of these risk factors are actually part of the early manifestations of the illness or susceptibility markers rather than risk factors causally related to ‘schizophrenia’ (Compton, 2005). Furthermore, many of the identified risk factors are probably non-specific in the sense that they may increase the risk for mental disorder in general, not for psychosis – or specific types of psychosis – in particular (Weiser et al, 2005).

One model of the causation of psychosis – developed in the context of ‘schizophrenia’ research – is the neuro-developmental model. In its simple form it postulates that inherited genetic factors controlling brain development and/or environmental factors in early life – affecting genetic regulation and expression, or the brain directly – lead to deviant development of the brain, which in turn increases the vulnerability to psychosis (Marenco and Weinberger, 2000). The simple neuro-developmental model suggests that the vulnerability interacts with normal maturational aspects of the brain’s physiology during adolescence – including hormonal changes, neuronal proliferation and migration, synaptic pruning and myelination – and social stressors. However, the simple neuro-developmental model – although capable of explaining the biological, developmental, neuro-psychological and psycho-social abnormalities that are associated with an increased risk of ‘schizophrenia’ – fails to explain aspects of psychosis/’schizophrenia’ such as the different timing of the
onset of psychosis in different individuals and the biological process that is correlated to the onset of psychosis, the phenomenology of the psychotic experiences and fluctuations in symptoms over time (Broome et al, 2005). Therefore, recent formulations of the neuro-developmental model incorporate risk modulating factors such as: social risk factors, cognitive appraisal processes and dopamine dysregulation (Kapur, 2003).

The aetio-pathology of psychosis is obviously complicated and still not understood. To find the origins of the abnormalities that are associated with psychosis, research will be needed including: neuro-biology, psychology, sociology and classification. Epidemiological studies are needed to generate and test hypotheses. Moreover, since epidemiological studies can give information on the frequencies and distributions of psychotic disorders, they are also important for health planning.

This thesis analyzes the incidence, age-at-onset and lifetime prevalence of all DSM-IV psychotic disorders in a total community population followed for 50 years. Moreover, personality related predictors of a broad group of nonorganic (nonaffective and/or affective) psychoses and ‘schizophrenia’, respectively, will be analyzed.

The total population is the ‘Lundby population’, which has been investigated four times between 1947 and 1997. Altogether 3563 individuals have been followed on a personal level regarding the development of the mental health. The Lundby population originally lived in a defined area in the south of Sweden; but irrespective of whether the study subjects stayed there or moved they have been followed up. Since attrition has been low (1-6%) it offers the unique opportunity to analyze some epidemiological aspects of psychosis in a 50 year community perspective.
5. Background

5.1 Psychiatric epidemiology

Psychiatric epidemiology deals with the frequencies, distributions and determinants of mental disorders in specified samples of the population (Fleming and Hsieh, 2002). Below follows a presentation of some epidemiological concepts and terms relevant to the thesis.

Epidemiological studies can be experimental or observational. In observational studies a sample of the population is observed – without any other intervention than the study itself – in terms of occurrence of outcomes and exposures; and associations between exposures and outcomes. Observational studies try to describe the world as it is. Observational studies can be subdivided into descriptive and analytical.

In descriptive epidemiology frequencies and distributions of disorders – and exposures to putative risk factors for disorders – are described in relation to person, place and time. Descriptive epidemiology is about disorder and determinants of disorder related to the questions ‘who’, ‘where’ and ‘when’.

Descriptive epidemiology may generate hypotheses for aetiologic research. Analytical epidemiology is the branch that follows up on the hypotheses and searches for determinants (exposures) that influence disorders (outcomes). Analytical epidemiology compares groups of people that are exposed and unexposed, respectively, to certain factors, in order to analyze the associations between exposures and outcomes. Analytical epidemiology tries to answer the questions ‘why’ and ‘how’.

There are three basic designs of epidemiological observational studies of populations (Ejlertsson, 1984): cross-sectional, retrospective and prospective. In a prospective cohort study the study subjects are selected based on exposure (i.e. exposed versus unexposed) and then followed
forward with time to determine how many that subsequently develop the outcome under study in each group.

The basic epidemiological measures in observational studies are measures of frequency: incidence and prevalence. The incidence rate is the number of new cases in a population within a specified period of time divided by the total number of person observation years free of the disorder in question in the population during the period. It is to be noted that the denominator of the incidence rate is expressed as person-time at risk of the outcome event (not persons at risk) and consequently the incidence rate may theoretically vary between null and infinity and its unit is cases per person-time at risk. The person-time at risk is the pooled risk-time that all the study subjects at risk have been followed; a study subject stops being at risk when (s)he: dies, drops out of the study, gets the outcome (or some other outcome that prevents the outcome under study to happen) or when the study ends.

Prevalence, which is a measure of current (or previous) status rather than newly occurring outcome, is the proportion of a sample of the population that has (or has had) an outcome or some characteristic. The period prevalence refers to the proportion of the population which during a defined period possesses a condition. The lifetime prevalence is the proportion of subjects in a sample alive at a certain time point, which up to the time of the study - during their whole life – has had a condition.

Incidence is related to aetiology, whereas prevalence is related to incidence but also to the natural duration, curability, migration and mortality associated with the condition. Thus, prevalence may be harder to interpret than incidence. Also, the incidence rate has its drawbacks; e.g. the true incidence rate in a population may vary across time, which, however, the calculated incidence rate for the time period in question does not show since it is an average. Moreover, the incidence rate does not per se indicate if a large sample was followed for a short time or a small sample for a long time. Nevertheless, the incidence rate indicates at which speed the healthy part of the population becomes unhealthy (Ahlbom and Norell, 1987); hence indicating the risk for a randomly chosen individual of getting ill. The incidence rate measure is therefore used in studies of putative risk factors.
Since the incidence rate is associated with aetiological factors, it may be further explored to indicate the increased risk that is inherent in an exposure-outcome association. This may be achieved by comparing the incidence in a group of individuals that have been exposed to a factor to the corresponding incidence in a comparable group that has not been exposed to the factor. A relative comparative measure is the ratio of two incidences. In psychiatric analytical epidemiology such measures are used to find risk indicators of mental disorders (i.e. to answer the ‘why’, and ‘how’ questions). Preferably, comparison of incidence rates between exposed and unexposed subjects from prospective studies is employed, since the exposures (i.e. the hypothesized risk factors) in prospective studies are measured in subjects at risk before outcome has occurred, which increases the probability to correctly assess exposure (see 5.3.6 Bias and confounding). The ratio between the incidence rate of an outcome in a group that has been exposed to a certain factor and the corresponding incidence rate in a group that has not been exposed is called the relative risk (RR). Although the relative risk represents the strength (effect size) of an exposure-outcome association, it does not indicate whether this association is causal or not. It could also be coincidental or due to systematic error (bias) in the data. Therefore analytical epidemiology needs to rule out chance and systematic data error before an exposure-outcome association may be considered to be a probable causal association.

The confidence interval of the relative risk gives the probability that the association is due to chance and if the exposure-outcome association is statistically significant. But the possible presence of systematic error in the data must be interpreted in light of the study design (choice of study population, sampling method, case finding method and diagnostic ascertainment) and the size and type of the attrition (see 5.3 From population to result). If chance and bias can be ruled out as possible explanations of an exposure-outcome association, the assignment of causality is still in the end based on inference; underpinned by the data, previous research and models of causality.

If an exposure is causal it may be sufficient, necessary or contributing to the outcome. Most identified risk factors for psychotic disorders are probably contributing factors (van Os and Verdoux, 2003). A factor may be significantly associated with an outcome but not causally related to it. Such a factor may be a proxy that indicates the presence of one or more
causal risk factors or it may be a marker of susceptibility to the outcome. A marker of susceptibility with genetic underpinnings is called an endophenotype (Gottesman and Gould, 2003; Gottesman and Hanson, 2005; Compton, 2005; Weiser et al, 2005). A factor may also be an early manifestation of the outcome or a factor that modifies the expression of the outcome (i.e. a pathoplastic factor). Causality is complex and causal relationships are seldom on a one-to-one basis but more often indirect and multi-factorial. An exposure may be located up- or downstream a causal chain (Ahlbom and Norell, 1987; Fletcher et al, 1996). One exposure may modify the effect of another, so that the total effect of two exposures may be greater than the sum of the exposures’ individual effects. One and the same exposure may also be related to several outcomes and one outcome may be the result of different sets of exposures.

To be able to apply the frequency measures on data, psychiatric descriptive epidemiology needs valid concepts of psychiatric outcomes – i.e. diagnostic constructs (see 5.2 Diagnosis and classification) – and reliable methods for case finding and diagnostic ascertainment of the cases (see 5.3 From population to result).

5.2 Diagnosis and classification of psychotic disorders

5.2.1 Overview

Neither mental illness, nor psychosis are easily defined concepts. As pertains to mental illness, several definitions have been proposed: ‘statistical deviation from the average’, ‘a biologically disadvantageous deviation from the norm’, ‘distress, disability and/or impaired reality testing’ and ‘difference that arouse therapeutic concern’ (Farmer et al, 2002). Also for psychosis several definitions exist: ‘presence of certain symptoms’ (e.g. hallucinations), ‘significant loss of social/occupational function’ (e.g. few friends/unemployed), ‘loss of ego boundaries’ (e.g. distortion of the perspective of subjectivity) and ‘gross impairment in reality testing’ (e.g. delusional).

The medical model has greatly influenced the diagnosis and classification of mental disorders. Four versions of the medical model applied to mental disorder are: ‘the organic disease model’, ‘the altered function model’, ‘the
harmful dysfunction model’ and the ‘biopsychosocial model’ (Zachar and Kendler, 2007).

The ‘organic disease model’ states that mental disorders are the result of pathological processes in certain parts or systems of the brain. The processes are considered to result from specific aetiologies – external factors or internal dysfunctions – and represent the essence of the disorders.

The ‘altered function model’ states that a mental disorder is a condition of altered function that is a threat to health. The altered function may be on a normal physiological continuum or due to a pathological process.

The ‘harmful dysfunction model’ recognizes that a mental disorder has two components: a pathological process and harmfulness/maladaptiveness (not merely a threat to health).

Finally, the ‘biopsychosocial model’ (Engel, 1980) states that an integrated approach to human behaviour is necessary to adequately diagnose mental disorders. The biology, psychology and social environment all together influence the expression of mental disorders.

The definitions are based on different assumptions about what kind of criteria that should be used in building diagnostic constructs; assumptions which in turn are pinned on different ideas about the nature of causality in mental disorders. The different kinds of criteria that underly the different definitions of mental disorder have been outlined in six overlapping conceptual dimensions (Zachar and Kendler, 2007): ‘causalism-descriptivism’, ‘essentialism-nominalism’, ‘objectivism-evaluativism’, ‘internalism-externalism’, ‘entities-agents’ and ‘categories-continua’.

‘Causalism-descriptivism’ refers to whether a mental disorder should be categorized according to its causes or according to the clinical picture (because the causal relationships are so complex that it makes classification on the grounds of causes impossible).

‘Essentialism-nominalism’ refers to whether a mental disorder is an essential clearly delimited part of nature or an artificial construct.
‘Objectivism-evaluativism’ refers to whether a mental disorder should be looked upon as something that is objectively measurable or to be understood in relation to a person’s subjective and relative (value-laden) notions of health.

‘Internalism-externalism’ refers to the perspectives that range from inside the body/mind to the outside world, i.e. whether a mental disorder should be understood as the result of events taking place in the brains of people, their thoughts/emotions/self-constructs or in the outside environment.

‘Entities-agents’ refers to the conceptual dimension that ranges between regarding a mental disorder as a single unit which patients are struck by and a reaction which is subjectively unique and related to a person’s character.

‘Categories-continua’ refers to whether a mental disorder should be seen as a discrete qualitatively distinct category (separable from other disorders and from health; note: a discrete category may still be heterogeneous and broad) or as the extreme quantitative end on a continuous distribution of a normal trait/syndrome variation in the population.

5.2.2 DSM-IV

In the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) a mental disorder is conceptualized as a:

“clinically significant behavioural or psychological syndrome or pattern that occurs in an individual and that is associated with present distress (e.g., a painful symptom) or disability (i.e., impairment in one or more important areas of functioning) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom. In addition, this syndrome or pattern must not be merely an expectable and culturally sanctioned response to a particular event, for example, the death of a loved one. Whatever its original cause, it must currently be considered a manifestation of a behavioural, psychological, or biological dysfunction in the individual. Neither deviant behaviour (e.g., political, religious, or sexual) nor conflicts that are primarily between the individual and society are mental disorders unless the deviance or conflict is a symptom of a dysfunction in the individual, as described…” (American Psychiatric Association, 1994).
The definition, which is based on several concepts, does not specify a precise boundary for mental disorder but defines it on various levels of abstraction; e.g. by aetiology, structural pathology, norm deviance, symptom presentation, syndromal pattern, distress and disability. The DSM-IV employs a biopsychosocial model of mental disorders and holds an essentialistic, objectivistic and categorical position. However, as pertains to the dimensions of causalism-descriptivism, internalism-externalism and entities-agents the DSM-IV holds an intermediate position. The DSM-IV is concept driven and the different psychotic disorder categories are pinned on sets of criteria for inclusion and exclusion, which reflect a compromise between lumping and splitting, as the criteria on the one hand allow for some variation within the diagnostic constructs (i.e. lumping, allowing some heterogeneity and broadness within the categories) while on the other hand they establish boundaries between them (i.e. splitting, striving for homogeneity within the categories).

In ‘psychotic disorder due to a general medical condition’ and in ‘substance-induced psychotic disorder’, psychosis refers to the presence of delusions or hallucinations not accompanied by insight. In ‘schizophrenia’, ‘schizotypal disorder’, ‘schizoaffective disorder’ and ‘brief psychotic disorder’, psychosis refers to: the presence of delusions, prominent hallucinations – with or without insight – disorganized speech and/or disorganized or catatonic behaviour. In ‘delusional disorder’ and ‘shared psychotic disorder’, psychosis refers to the presence of delusions. In ‘bipolar disorder with psychotic features’ and in ‘major depressive disorder with psychotic features’, psychotic refers to the presence of delusions or hallucinations.

The DSM-IV criteria sets reflect on a pecking order introduced between the diagnoses, which is based on: whether there is evidence for the aetiology of the disorders, how pervasive the disorders are, the degree of functional impairment the disorders are associated with, whether some of the associated symptoms/signs of the disorders (e.g. as part of their course) are the defining features of other disorders, if other disorders were previously present and the duration of the disorders. Thus, there are disorders that present with psychosis which are due to: physical illness (‘psychotic disorder due to a general medical condition’), drugs of abuse, medicines, or toxins (‘substance-induced psychotic disorder’), and there are psychotic
disorders with unknown aetiology (from here on referred to as functional psychosis). The two first groups are diagnostically ranked higher than the third with reference to the known aetiologies.

Within the functional psychoses a distinction between affective and nonaffective psychoses is traditionally made based on whether psychosis has emerged within an affective symptom or syndrome pattern or not. The diagnostic relationship between affective and nonaffective psychotic disorders in the DSM-IV is mutually exclusive. The affective and nonaffective psychoses may each be further subdivided into mutually exclusive more specific diagnoses.

The DSM-IV recognizes that psychotic symptoms are diagnostically unspecific. Thus, psychosis may also be present as an accompanying feature (e.g. short-lived, intermittently or in an attenuated form) in other DSM-IV diagnoses beside those already mentioned. Interestingly enough, in general population surveys subclinical psychotic experiences have been reported by a small but not insignificant proportion of subjects, indicating that psychotic experiences are not exclusively associated with DSM-IV mental disorder. In fact, studies suggest that the community incidence of subclinical psychotic experiences ranges between 2 and 4.6 per 100 per year (Hansen et al, 2005; Wiles et al, 2006; Tien AY, 1991) which is about 100 times greater than the community incidence of schizophrenia (McGrath et al, 2004). Such data suggests that the ‘real’ community distribution of psychosis is not dichotomous but more probably continuous – a quantitative trait related to clinical caseness by degree of severity – or quasicontinuous – a quantitative trait related to clinical caseness by degree of severity but with sharply increasing risk over a threshold (van Os and Verdoux, 2003).

The uncertain construct validity of the DSM-IV psychotic disorder categories is a matter of great concern and there is a debate on whether psychotic disorders should be represented as discrete disease entities or as the extreme ends on continuously distributed dimensions or both. Up to five dimensions/syndromes simultaneously associated with psychosis have been identified: the depressive, manic, positive, negative and catatonic/disorganized dimensions (van Os and Verdoux, 2003). In the introduction to the DSM-IV you may read:
“It was suggested that the DSM-IV classification be organized following a dimensional model rather than the categorical...Although dimensional systems increase reliability and communicate more clinical information they also have serious limitations and thus far have been less useful than categorical systems in clinical practice and in stimulating research...Nonetheless, it is possible that the increasing research on, and familiarity with, dimensional systems may eventually result in their greater acceptance…” (American Psychiatric Association, 1994).

In the Lundby Study a simple conceptually based categorical diagnostic system of mental illness designed to fit an epidemiological field study was developed, which has been used throughout the study (Hagnell, 1966). In this ‘psychosis’ simply consisted of two diagnoses: ‘schizophrenia’ and ‘other psychoses’. However, in the present thesis the original Lundby psychoses have been re-diagnosed according to the DSM-IV (see 7.4 Ascertainment of diagnosis).

5.3 From population to result: methodological issues

5.3.1 Population

To be able to interpret and generalize prevalence and incidence estimates, as well as estimates of relative risk obtained from comparison, it must be made clear to which population the estimates apply and do not apply; i.e. the size and characteristics of the population must be described. The population must be adequately chosen to be relevant for the research question.

In epidemiological studies of the general population two principal methods exist: the census method and the generation (birth cohort) method (Essen-Möller, 1956). The census method involves all individuals, or a fraction thereof, living in a geographic area at a particular date. The generation method involves all newborn, or a fraction thereof, in an area during a defined period. The census method has a disadvantage not shared by the generation method in that elderly individuals in the population who contracted an outcome when they were young (i.e. before the inception of the study) cannot be studied prospectively for incidence and age-at-onset of that outcome; and those who died early cannot be studied for lifetime
prevalence of an outcome. Thus, there is a risk with the census method of underestimating the true frequencies (particularly in younger age groups) and overestimating the true age at onset in newly occurring cases of an outcome in the population under study. Another problem with the census method is that it misses the individuals who emigrated and might miss newly immigrated. However, an advantage of the census method is that it enables personal investigation of all the individuals in the population, including all ages. The generation method only provides information of the incidence and prevalence of an outcome up to the ages that the individuals in the birth cohort have reached.

5.3.2 Sampling

Since it is very resource demanding to study a complete population, usually a sample from the study population is drawn using some sampling method. The study population, from which the sample is drawn, may be called sample source (Fletcher et al, 1996; Eaton et al, 2007).

It is important that the sampling method results in a sample that is representative of the sample source so that the results of the study may be generalized. The extent to which this is the case is referred to as the external validity of a study. Representativeness may be accomplished using a complete population sample, random sampling or large samples. When non-random samples have been studied, such as convenience samples (e.g. the patients at an academic centre), results must be generalized with caution.

If the sample is to be followed prospectively for the onset of an outcome (i.e. a study of incidence) an ‘at risk’ cohort is defined. This cohort includes all individuals in the sample who have not yet manifested the outcome under study, and who may manifest it in the future.

5.3.3 Exposure and outcome

If comparative measures of an outcome involving individuals who are exposed and unexposed to some hypothesized risk factor(s), respectively, are to be assessed (i.e. a study of exposure-outcome associations) the individuals at risk in the cohort are, on entry in the cohort, classified as exposed versus unexposed. It is important that the exposures under study
are assessed accurately and reliably in all individuals of the study; and that
the cases manifesting the studied outcome are identified accurately and
reliably in all individuals over the complete study period. Should there be
under- or over-reporting of exposures and/or outcome the result will
naturally be misleading.

The identification of cases manifesting psychosis in the large psychiatric
epidemiologic surveys of the last decades (Robins and Regier, 1991;
Kessler et al, 1994; van Os et al, 2001) have been based on survey
instruments such as the DIS [Diagnostic Interview Schedule] (Robins et al,
1981) and the CIDI [Composite International Diagnostic Interview]
(Robins et al, 1988). These instruments are highly structured self-report
instruments, where lay interviewers encode the answers to structured
questions from respondents. However, validation-studies have shown the
instruments to be poor in identifying disorders such as schizophrenia and
bipolar disorder in the general population, as they rely on the judgement of
the respondent pertaining to the presence or absence of symptoms and
corresponding impairments (Eaton et al, 2000; Eaton et al, 2007).
Therefore, in some general population studies, a two stage procedure for
case identification has been applied: in the first stage a screening based on
e.g. CIDI – administered by lay interviewers – takes place, where after, in
the second stage, a semistructured interview of the screened individuals –
administered by health professionals – takes place. In the second stage
another armamentarium of instruments such as the SCID-I [Structured
Clinical Interview for DSM-IV (III-R) Axis I Disorders] (Spitzer et al,
1992), PSE [Present State Examination] (Wing et al, 1974) and SCAN
[Schedules for Clinical Assessment in Neuropsychiatry] (Wing et al, 1990)
has been used. Nevertheless, in one study, that specifically addressed
various methods of case finding of psychotic disorders in a general
population, it was shown that registers were the most important and reliable
source of information for identifying cases with psychosis, and that
screening based on multiple sources (i.e. combining treatment data and
interview data) was essential to achieve a high identification rate of
individuals with psychotic disorders (Perälä et al, 2007).

However, psychiatric case identification in population studies most often
has relied on interviews. The basic interview modality is pinned on the
notion that psychopathology may be assessed in an interview as symptoms
and signs indicative of mental disorder. Such an assessment is a
complicated procedure, which presupposes that the interviewer has experience of clinical psychopathology so that (s)he may recognize, elicit, and describe symptoms and signs in a respondent if present. Such experience comes from studying psychopathology, observing experienced clinicians and practicing one’s observational skills and cross-questioning technique. Thus, important integral parts of the traditional clinical interview are: clinical judgement, flexibility in interviewing and encoding responses.

In epidemiology there are three types of interviews: the unstructured (traditional clinical), the structured and the semi-structured (Brugha et al, 1999). In epidemiological research the traditional clinical approach has a problem; it is hard to standardize. Since standardization is necessary in epidemiological research, the two last mentioned types of interview have been developed. Thus, in fully structured interviews the standardization is total; questions are asked word by word in a fixed order and the answers from the respondents are encoded strictly according to the interview schedule. The encoded answers provide the basis of diagnosis (usually diagnosis is aided by a computer algorithm). Thus, in fully structured interviews the clinical judgement and flexibility – which lies at the heart of the traditional psychiatric interview – is omitted, and interviewers are typically lay men. However, in semi-structured interviews the component of clinical judgement, and flexible cross-questioning, is retained. Although questions are pre-worded also in the semi-structured variant, the interviewer – who is a health professional – may if (s)he feels that it is needed follow up with freely formed questions. Both in the fully structured interview and in the semistructured one the symptoms and signs are strictly defined and often the diagnostic mode is a computer algorithm. Examples of structured interview schedules for psychiatric epidemiology adapted for use by lay interviewers are the DIS and the CIDI. Examples of semi-structured schedules for use by clinicians are the PSE and the SCAN.

Additionally, to be able to identify cases the period of follow up must be long enough for the outcome under study to be expressed; some disorders (e.g. schizophrenia) may have long latency periods and disorders first assessed during follow-up may later turn out to be really a stage on the pathway to another diagnosis (e.g. initial depression that develops into dementia).
To separate the rare cases with current or previous psychosis from non-cases in the general population is a difficult task as individuals in the public have not been selected at all (subjectively or otherwise) as they have when seen in a practice or clinic. Moreover, thresholds for caseness are not easily defined, since a case defined by a threshold is a crude simplification of data. Nevertheless, such simplification may be justified out of practical reasons as all the details of data may not be necessary to guide the appropriate decision (e.g. as in the clinical context when to decide whether to treat or not to treat).

5.3.4 Diagnostic ascertainment

After a case has been identified as suffering currently, or previously from some psychotic disorder comes the ascertainment of diagnosis. Now the task is to assess, if possible, a specific diagnosis. This is presently in psychiatric epidemiology based on an international system of diagnosis, and classification such as the DSM-IV (American Psychiatric Association, 1994) or the ICD-10 (World Health Organization, 1992) that may be used in the general community as well as in the clinical setting. It is vital that diagnoses are accurate (valid) constructs that may be assessed with precision and reproducibility (i.e. reliably). Accurate pertains to the meaningfulness of the diagnosis; its content (i.e. that all necessary – but not the un-necessary – dimensions of the condition are included; in other words the criteria of the diagnosis should be necessary and sufficient) and its power to predict aspects of the condition such as: cause, associated symptom profile, course, outcome and response to treatment (van Praag, 1999). Moreover, it is also important that the diagnosis makes sense to the investigator using it (Kaplan and Sadock, 1994). Importantly, the validity of a diagnosis cannot be directly observed or measured; it must be inferred. Validity of a diagnosis is not simply present or absent; rather you may argue for or against it. The precision and reproducibility associated with a diagnostic construct pertains to whether or not the findings of the diagnostic procedure can be replicated by different assessors and at different times and places (Fletcher et al, 1996.).

5.3.5 Attrition

It is important that as many individuals as possible in the sample agrees to take part in the study. The sample is dependent on all its individuals to
retain its representativity of the source population. In prospective studies it is also important that participants are not lost to follow-up. If attrition is great a study may lose its representativity (see bias). Attrition may be due to refusal, migration and death.

5.3.6 Bias and confounding

Ideally, the individuals in a study sample should be very similar to each other (except for the exposures and outcomes studied) and to the source population. Bias is a process, related to the way information is collected and measured, that systematically affects the data. Thus, bias gives results that, more or less, deviate systematically from the true state of affairs. The most important types of biases are: selection bias and information bias. These biases may be seen as broad overlapping categories.

Selection bias may occur if groups are compared that systematically differ with regard to outcomes or determinants of outcomes other than the studied ones. Selection bias is a problem related to the choice of population, the sampling procedure and the attrition. Sampling bias – a form of selection bias – occurs when the sample systematically differs from the source population, e.g. due to incomplete or non-random sampling. Incomplete sampling, non-random sampling and selective attrition may seriously impair representativity.

Information bias refers to a systematic difference in how the information about exposures and outcomes was gathered in groups that are compared. One type of information bias is observation bias. This bias refers to a systematic difference between the studied groups in the number and types of data sources that are available (e.g. registers, case-files, key-informants and interviews). Observation bias is related to the case finding method.

Another type of information bias is recall bias. This refers to individuals giving distorted or completely wrong information due to erring memory. Recall bias may be related to a long time span between follow-ups.

A third type of information bias is measurement bias. Measurement bias may result when the groups that are compared differ in their response to measurement of an exposure/outcome. This may occur when the individuals in one group are easier to assess than the individuals in the
comparison group. Moreover, the frequency of an outcome may be systematically misrepresented due to an over- or underinclusive interview instrument. When interviewers assess respondents, the expectations and idiosyncracies of the former may also bias the assessments. Measurement bias is related to the case-finding method and the mode of diagnostic ascertainment. An example of measurement bias pertinent to psychosis epidemiology is that the CIDI has been shown to produce false positives when assessing bipolar disorder in the general population (Perälä et al, 2007).

Unknown factors may systematically be responsible for the association observed between the studied exposure and the outcome. This phenomenon – called confounding – represents one of the big problems in interpreting the results from epidemiological studies. Confounding is sometimes referred to as ‘confounding bias’ but confounding is not really a bias, although similarities between confounding and bias exist. Bias refers to error in the collection of information and measurement of a variable, whereas confounding refers to error in the interpretation of what may actually be an accurate piece of information/measurement, i.e. mistakenly applying the wrong model of causality to the observed data.

Biases cannot be eliminated from an epidemiological study, but the aim must be to keep biases at a minimum when designing studies. When the data has been collected biases cannot be corrected by statistical or other methods. At this point one should strive to identify possible biases, assess their potential impact and to take them into account when interpreting the results.

5.4 Studies of age-at-onset of psychotic disorders

5.4.1 Conceptual issues

The age-at-onset of a psychotic disorder may give important clues to the causes of the disorder on the individual level, provided that certain biological characteristics, processes or events can be shown to be associated with the typical age-at-onset of the disorder (DeLisi, 1992). Robust knowledge of typical age-at-onset would also aid in the process of constructing diagnostic criteria for psychotic disorders (i.e. age could be used as an inclusion and/or exclusion criterion). Thus, some important questions pertaining to the age-at-onset of psychotic disorders are:
i) Is the risk of onset of psychotic disorders related to age? Each age period may represent different components of the pathophysiology behind psychosis.

ii) Is the age-at-onset related to: heredity, sex, developmental disorder, neuropsychological deficits, premorbid personality, puberty, menopause and environmental factors (physical, biological and social)?

iii) Is the age-at-onset related to specific symptomatic pictures in the psychoses?

However, in research on the age-at-onset of psychotic disorders there is a fundamental problem regarding the conceptualization of age-at-onset itself. When does a psychotic disorder really begin? Ideally, one should distinguish the aetiology from the pathology. Aetiology and its corresponding characteristics, process(es) or event(s) include the time-period when the risk of developing pathology is increased, but aetiologies may be present before pathology emerges. Theoretically, the onset of a disorder would be when the process(es) or event(s) that are associated with the aetiology reaches the point of no return to develop the full criteria for the disorder (Eaton, 2002). As biological and neuropsychological markers with causal significance for psychosis onset are lacking (Häfner, 2003), psychotic disorders are still defined solely in clinical terms; i.e. based on the development over time of subjectively reported symptoms, observed signs/behaviours and their impact on individual function. However, the reported symptoms and observed behaviours cannot be ascribed to the disorder with perfect accuracy (Eaton, 2002). Consequently, there is not one generally accepted definition of age-at-onset. Many researchers and clinicians would date age-at-onset from the first appearance of psychotic symptoms (Clarke and O’Callaghan, 2003), but what is probably meant mostly with age-at-onset of a psychotic disorder is the age at which the full criteria of the disorder are met, and still this age may be considered to be the closest approximation to onset (Häfner, 2003). However, while such definitions may be pragmatic, they are not ‘true’ in the sense that the age of psychotic disorder onset also would indicate the age at which the disorders underpinning the psychotic experiences began, since these disorders probably started before the onset of the psychotic symptoms – moreover, not necessarily at the same time (DeLisi, 1992).
Research has shown that prodromal symptoms and early signs of schizophrenia – e.g. depression, anxiety, negative symptoms/signs, concentration difficulties and minor perceptual changes – often date back several years before the onset of frank psychosis (Häfner, 2003), indicating that the disorders behind schizophrenia may not be abrupt in onset and that they may appear years before psychosis sets in. In a retrospective study of the early phases of schizophrenia in a population based clinical first-illness sample (the ABC Schizophrenia Study) it was found that depressive symptoms had appeared 3-5 years before first admission due to schizophrenia, negative symptoms 2-4 years and positive symptoms about 1 year before first admission (Häfner, 2003). In the prospective Dunedin birth cohort study self-reported psychotic symptoms at age 11 years were associated with schizophreniform disorder at 26 years (Cannon et al, 2002) supporting the view that schizophrenia is a disorder that may have a long prodromal phase. It is, however, not yet possible to predict in the prodromal phase whether or not frank psychosis will develop (Häfner, 2003) as the early nonpsychotic symptoms and signs are unspecific in nature –although much research efforts are devoted to the development of instruments to measure ‘At Risk Mental States’ such as the Comprehensive Assessment of At Risk Mental States [CAARMS] (Yung et al, 1998).

In practice, still, prodromal phases of psychosis are always diagnosed retrospectively. Moreover, since it is still unclear if psychotic disorders are due to deviant neuro-development and/or some neuro-degenerative disease process, or both; and if psychosis develops in a sequence of phases, gradually or abruptly (Häfner, 2003) it is still not possible to define the true age-at-onset of psychotic disorders. Therefore, researchers have defined age-at-onset in different ways as: first-admission due to a psychotic disorder (Häfner, 2003), first treating contact due to a psychotic disorder, the reported first-onset of psychotic symptoms, the reported first-onset of prodromal symptoms and the first indication – e.g. as reported by relatives – of behavioural change eventually leading to psychosis. The three last types of definitions are all impaired by the often insidious onset of psychosis over many years. It may well be impossible to define one ‘gold standard’ of psychosis age-at-onset. The definition used should be clearly described to facilitate interpretation and comparison of results (Clarke and O’Callaghan, 2003).
Large community population surveys from recent decades regarding mental disorders – e.g. the ECA Study (Burke et al, 1990), the NCS (Kessler et al, 1994) and the NEMESIS (Bijl et al, 2002) – have been found to have questionable internal validity for estimations of incidence and age-at-onset of psychosis due to the reliance on structured interviews by lay interviewers (Kendler et al, 1996; Brugha et al, 1999; Eaton et al, 2000; Perälä et al, 2007). This being so, most current knowledge on age-at-onset of first-episode psychosis has been pinned on patient samples – mostly drawn from urban or mixed urban-rural mental health services – which largely have been restricted to schizophrenia spectrum psychoses and affective psychoses excluding older individuals – e.g. 55 or 65 years and older (Bromet and Fennig, 1999; Baldwin et al, 2005).

Regrettably, relevant epidemiological data about ‘psychotic disorder due to a general medical condition’ and ‘substance-induced psychotic disorder’ are lacking although one study of patients with alcohol dependence found a mean age-at-onset of alcohol psychosis of 47 years (Soyka, 2008). Consequently, studies on age-at-onset of psychosis do not represent the diversity of psychotic presentations in total community populations.

Furthermore, many of the studies that exist may have been subject to selection bias by socio-economic characteristics, ethnicity and substance-use in the patients studied. The external validity of many studies of age-at-onset of psychotic disorders may thus be questionable. To maximise the potential of first-episode studies of psychosis it has been suggested that epidemiologically complete and homogeneous populations, in which all first-episode psychosis cases are accrued over long periods, should be investigated (Baldwin et al, 2005; Scully et al, 2002).

5.4.2 Age-at-onset of nonaffective psychoses

Notwithstanding the limited scope of most epidemiological studies of the psychoses, in a recent review it was concluded that the median age-at-onset of nonaffective psychoses is in the range late teens through early twenties; although a limitation mentioned was that the estimates in the review were based on treated incidences (Kessler et al, 2007). However, these figures must be interpreted with some caution since many studies do not include all types of nonaffective psychoses and studies differ with regard to the age range included. Some studies have used the age range 15-45 or 15-54 years
which will naturally bias the results for age-at-onset downwards, whereas studies using a wider age range for the upper limit will find older median age-at-onset. For example, in a large study of first-contact psychosis for the age range 18-64 in a metropolitan area the median age-at-onset of nonaffective psychoses was 29.0 years (Menezes et al, 2007).

5.4.3 Age-at-onset of schizophrenia

The most typical age-at-onset of schizophrenia is generally considered to be the late teens and early 20s (Andreasen, 1999). However, epidemiological studies analyzing the age-at-onset of schizophrenia from different viewpoints have made some interesting findings. Thus, in most community based samples the mean age-at-onset of schizophrenia has been found to be on average 3-5 years earlier in males than females (Bland, 1977; Angermeyer and Kuhn, 1988; DeLisi, 1992; Jablensky et al, 1992; Hambrecht et al, 1992; Häfner et al, 1993; Castle and Murray, 1993; Brewin et al, 1997). This earlier onset of schizophrenia in males may have several explanations. Pregnancy and birth complications may be associated with an earlier age-at-onset of schizophrenia in both genders (Rosso et al, 2000; O’Callaghahan et al, 1992), but since such complications may be more common in males they could partly explain the earlier onset of schizophrenia in males than females (Clarke and O’Callaghahan, 2003).

The male proneness to get ill earlier than females may also be due to male-female biological dimorphism; i.e. the evolutionary based systematic sex difference in the brain size/form (Jablensky, 2000) and/or in the changes of the brain that are related to ageing – e.g. a slower rate of dopamine D2-receptor loss in females compared to males (Orr and Castle, 2003). A community study analyzing the incidence by age of hallucinations (irrespective of whether diagnostic criteria for some disorder were fulfilled) found some support for male-female biological dimorphism and/or differences in the aging brain (Tien, 1991) as the incidence of hallucinations by age distributions showed peaks at age 25 in males but before age 20 in females; and that females also had a second peak around age 40; and while the male rate of hallucinations trended down with increasing age the female increased after age 60.

Another hypothesis that could account for the later age-at-onset of schizophrenia in females and also a second incidence-peak around age 45,
which has been observed in females, is the oestrogen hypothesis; i.e. that oestrogen through its anti-dopaminergic properties may protect females from psychosis from menarche to menopause (Häfner et al, 1991; Häfner et al, 1993; Riecher-Rössler et al, 1997; Häfner et al, 1998). A protective effect of oestrogen may be supported by: delayed onset of schizophrenia in females compared to males, lower early incidence-peak in females compared to males, an association between earlier menarche and later onset of schizophrenia, frequent worsening of symptoms of schizophrenia in females during low oestrogen phases of the menstrual cycle, frequent symptom alleviation during pregnancy, frequent relapses postpartum, requirement of lower antipsychotic dosages in young females than following menopause and a second incidence-peak of female schizophrenia that has been shown in several studies (Grigoriadis and Seeman, 2002).

Although most studies support an earlier male than female average age-at-onset of schizophrenia it should nevertheless be emphasized that there are some findings of non-differing age-at-onset of schizophrenia in males and females (Murthy et al, 1998) as well as findings of earlier onset in females (Bland et al, 1988; Folinovovic and Folnegovic-Šmalc, 1994; Beiser et al, 1993). Actually, it has been suggested that the male-female age-at-onset difference for schizophrenia often found may be a confounded finding reflecting differences in e.g. marital status (onset-delaying effect), culture and premorbid personality traits (Jablenskys and Cole, 1997). Another possible confounding variable, that in part could contribute to the varying difference in age-at-onset, is heritability (Bromet and Fennig, 1999) as in familial schizophrenia the male-female age-at-onset does not seem to differ (DeLisi et al, 1987; DeLisi et al, 1992; DeLisis et al, 1994; Kendler and Walsh, 1995).

Nevertheless, the body of evidence supports that schizophrenia, on average, has an earlier onset in males than females; and in a large scale catchment area study that included older subjects, the mean age-at-onset for schizophrenia broadly defined was 31.2 years in males and 41.1 years in females, respectively (Castle and Murray, 1993). Moreover, in community schizophrenia the incidence by age pattern has been found to differ between the sexes. The ABC Schizophrenia Study showed that the male and female age distributions at the earliest sign of disorder differed in that males displayed an early peak at the age of 15-25, which was followed by a steady uninterrupted decline; whereas females displayed a later and smaller first
peak at age 20-29 followed by a subsequent decline, which was interrupted by a second smaller peak at age 45-49 which was not seen in males (Häfner et al, 1993). A recent study in a metropolitan area found a rather similar sex difference for first-contact due to any nonaffective psychosis in that the incidence by age between age 18-64 decreased sharply and consistently with increasing age in males from the clearly highest rate in the 18-24 interval, whereas the female rate was rather low and steady between 18-64 although it tended to peak slightly (but not nearly as much as in males) in the 25-29 year age interval (Menezes et al, 2007). Further, ‘late-onset schizophrenia’ (Harris and Jeste, 1988) and ‘late paraphrenia’ (Harris and Jeste, 1988; Henderson and Kay, 1997) – i.e. ‘schizophrenia’ and ‘schizophrenia-like’ conditions beginning after age 44 and 60, respectively – have been found to be more common in females than in males. However, the inclusion of late-onset cases in studies of the overall incidence and age-at-onset of schizophrenia presuppose that schizophrenia beginning in middle and old age belongs to the same disorder (or group of disorders) as schizophrenia beginning in young age.

Previously, non-organic and non-affective psychotic disorders with late onset were often referred to separate diagnostic categories such as: ‘paraphrenia’ (according to Kraepelin; with lesser disturbance of emotion and volition than dementia praecox), ‘late schizophrenia’ (according to Bleuler; clinically resembling schizophrenia, but onset after age 40 years), involutional psychotic reaction (according to the DSM-I; including depression in the involutional period), ‘late paraphrenia’ (according to Kay and Roth; onset after age 60 years and clinically encompassing a spectrum of paranoid-hallucinatory conditions including schizophrenia and delusional disorder), ‘involutional paraphrenia’ (according to the DSM-II; delusion formation in the involutional period in the absence of conspicuous thought disorder), ‘paranoid psychosis’ (according to the ICD-9; conspicuous hallucinations, but preserved personality), ‘paranoid disorders’ (according to the DSM-III; persistent persecutory delusions without prominent schizophrenia symptoms) and ‘late-onset schizophrenia’ (according to the DSM-III-R; onset after age 44 years and clinically resembling schizophrenia with early onset) (Harris and Jeste, 1988; Henderson and Kay DWK, 1997).
Consensus about terminology for schizophrenic symptoms appearing for the first time after age 44 or 60, and about the differentiation between e.g. ‘late-onset schizophrenia’ (>44), ‘late paraphrenia’ (>60) and ‘schizophrenia’ has been hard to reach. But, in the year of 2000 an international consensus document was produced (Howard, 2000a).

With the DSM-IV the separate subcategory of ‘late-onset schizophrenia’ in DSM-III-R was omitted since late and early-onset schizophrenia was judged not to differ substantially as regards signs and symptoms; and the varying presenting features of early and late-onset schizophrenia/late paraphrenia (see below) were interpreted as resulting from symptom shaping influences pertaining to age-related bio-psycho-social developmental stage rather than from differences in aetio-pathology (DeLisi, 1992; Riecher-Rössler et al, 1997; Jablensky, 2000).

Nevertheless, although evidence supports that schizophrenia-like psychoses may emerge at any age during the life course, there may – according to the consensus document – be three age-at-onset related patterns of schizophrenia, all of which fullfill DSM-IV criteria of schizophrenia: ‘early-onset (<40 years) schizophrenia’, ‘late-onset (40-60 years) schizophrenia’ and ‘very-late onset (>60 years) schizophrenia-like psychoses’ (Howard, 2000a,b). Importantly, the exact age cut-offs of these subgroups are – according to the consensus document – more or less arbitrary and should be seen primarily as a means to stimulate research on the fundamentally heterogeneous group of schizophrenias (Howard, 2000a).

In terms of symptoms, early-, late- and very-late onset schizophrenia are more similar than different. Moreover, brain imaging findings are also essentially similar regardless of age-at-onset (Howard, 2000a). Nevertheless, the varying presenting features and associated findings of early-, late- and very late-onset schizophrenia include that early- and late-onset cases have stronger heredity for schizophrenia than very-late onset cases, in which the lifetime morbid risk of schizophrenia in first-degree relatives actually is not increased compared to controls (Howard, 2000b). Furthermore, early-onset cases of schizophrenia have poorer premorbid adjustment and more premorbid abnormal personality traits than late- and very-late onset cases (although late- and very-late onsets are associated with premorbid personality traits of the paranoid and schizoid kind) (Howard, 2000a,b). Early onset schizophrenia is associated with lower
premorbid IQ than late- and very late onset schizophrenia (Castle and Murray, 1991; Castle and Murray, 1993). Early-onset schizophrenia cases have more formal thought disorder and affective blunting/inappropriate affect than late- and very-late onset schizophrenia (in very-late onset schizophrenia, formal thought disorder and negative symptoms are extremely rare) (Howard, 2000a,b). Moreover, early-onset cases are more likely to have a history of pre- or perinatal complications (Castle and Murray, 1991; Castle and Murray, 1993). Late- and very-late onset cases have more delusions — especially persecutory delusions — than early-onset cases and in very-late onset schizophrenia partition delusions are also very common (Howard, 2000a,b). In late- and very-late onset schizophrenia auditory hallucinations (accusatory, running commentary and third person hallucinations) are more common than in early-onset schizophrenia (Howard, 2000a,b). Late- and very-late onset cases are also more likely to have visual, olfactory, gustatory and tactile hallucinations than their early counterparts (Howard, 2000a). Affective features are more common in late- and very-late onset schizophrenia than in early-onset schizophrenia (Howard, 2000a,b; Castle and Murray, 1991; Castle and Murray, 1993). The late- and very-late onset cases may also have a stronger heredity of depression than early onset cases (Howard et al, 1997). Hearing- and visual impairments are more common in very-late than in early- and late-onset cases (Castle and Murray, 1991; Castle and Murray, 1993; Henderson and Kay, 1997; Howard, 2000a,b).

As males predominate among the early onset cases and females among the late and very late-onset cases the differences may be interpreted as males and females are differently prone to subtypes of schizophrenia with different aetiologies and that late and very late-onset schizophrenia may be a valid entity (or group of entities) (Harris and Jeste, 1988.). Males may more often have a more severe and early-onset form of disease due to pregnancy or birth complications and neurodevelopmental anomaly (i.e. the male fetus may be more prone to insult than the female), while females more often may have a later onset disorder related to affective psychosis (Hultman et al, 1999; Castle and Murray, 1991; Castle and Murray, 1993; Castle et al, 1995; Kirov et al, 1996). It has also been suggested that late-onset schizophrenia represents ‘true’ schizophrenia with a delayed onset, and that very-late onset schizophrenia represents a separate disorder (Howard, 2000b).
5.4.4 Age-at-onset of delusional disorder

Little is known about the incidence and age-at-onset of delusional disorder in the general population. Nevertheless, in a textbook the age-at-onset is estimated to be approximately 40 years, but with a wide range from 18 to the 90s (Kaplan and Sadock, 1994).

5.4.5 Age-at-onset of nonorganic nonaffective acute remitting psychoses

The previously quoted review on nonaffective psychoses based on treatment contacts concluded that the median age-at-onset of the whole nonaffective psychoses group (i.e. including schizophrenia, delusional disorder and other nonaffective psychoses) is in the range late teens through early twenties (Kessler et al, 2007).

However, general population studies on the age-at-onset of nonaffective psychotic disorders with an acute onset that remits promptly or within a few months (e.g. DSM-IV ‘brief psychotic disorder’, ‘schizophreniform disorder’ and ICD-10: ‘acute and transient psychotic disorder’) have been rare and robust knowledge of mean age-at-onset for these disorders are actually lacking. Hence, in a register study of ‘acute and transient psychotic disorder’ according to the ICD-10 the mean age at first-admission was significantly higher in females (46.2 years) than in males (37.8 years) and clearly outside the late teens and early twenties range in both sexes (Castagnini et al, 2008). Whereas, in a study on ‘nonaffective acute remitting psychosis’ (NARP) – approximately representing DSM-IV ‘schizophreniform disorder’, ‘brief psychotic disorder’ and ‘psychotic disorder NOS’ – that was part of the WHO cross-cultural Determinants of Severe Mental Disorder Study (DOSMED; also called the Ten Country Study) the mean age-at-onset did not differ significantly between males and females and the age-at-onset was in the early twenties (Susser and Wanderling, 1994).

5.4.6 Age-at-onset of affective psychoses

In contrast to schizophrenia for affective psychoses most studies have found non-differing age-at-onset in males and females (Bland, 1977; Bland et al, 1988; Baldwin et al, 2005; Hendrick et al, 2000; Kawa et al, 2005).
The reported mean age-at-onset of bipolar disorder in clinical studies has according to a review ranged from 20.7 to 33 years, and 21.5 to 31.6 years, in males and females, respectively (Bebbington and Ramana, 1995).

5.5 Studies of the incidence of psychotic disorders

5.5.1 Overview

There are a large number of community based incidence studies regarding first-episode ‘schizophrenia’ and ‘schizophrenia spectrum’, but for the other nonaffective psychoses (‘schizophreniform disorder’, ‘schizoaffective disorder’, ‘delusional disorder’, ‘brief psychotic disorder’ and ‘shared psychotic disorder’) studies are scarce. There are some community based first incidence studies on the general group of affective psychoses (Menezes et al, 2007), but most studies of affective psychoses only include ‘bipolar disorder’ or ‘mania’. Moreover, as most first-episode incidence studies of psychotic disorders have investigated treated samples and have restricted the age of the study subjects it is not possible to generalize the findings in the literature to the whole group of first-episode psychosis in the community; nor is it possible to utilize the findings to sufficiently understand the specifics of the different psychotic disorders that we currently diagnose (Baldwin et al, 2005).

5.5.2 Incidence of psychotic disorder due to a general medical condition

There have been no studies on the community incidence of psychotic disorder due to a general medical condition.

5.5.3 Incidence of substance-induced psychotic disorder

There have been very few studies analyzing the incidence of substance-induced psychotic disorder and the rate of psychosis among people with a substance-use disorder is not known (Caton et al, 2005). However, in Belarus the rate of psychosis due to alcohol (including delirium) was according to register data reported to have ranged from 6.8-23.7 per 100 000 per year from 1979 to 2005 (Razvodocsky, 2008). And in the recent UK AESOP (Aetiology and Ethnicity of Schizophrenia and Other
Psychoses) Study, which investigated the treated incidence in three mainly urban catchment areas (London, Bristol and Nottingham) in subjects aged 16-64, the overall incidence rate of substance-induced psychoses (excluding delirium) was 1.8 per 100 000 person-years (Kirkbride et al, 2006).

5.5.4 Incidence of schizophrenia

For schizophrenia there is a plethora of incidence studies and studies have analyzed the schizophrenia incidence from a number of perspectives; e.g. the incidence in males compared to females, in different age groups, in urban areas compared to rural and in immigrants compared to native born people.

5.5.4.1 Schizophrenia in males-females

A recent systematic review showed a median incidence of schizophrenia of 15.2 per 100 000 person-years at risk but with a median male/female incidence rate-ratio of 1.40 (McGrath et al, 2004). A meta-analysis also documented an apparent higher risk for males to develop schizophrenia, with a male/female incidence rate-ratio of 1.42 (Aleman et al, 2003). However, in the meta-analysis, the sex-difference was smaller in samples before 1980, which may be partly due to that the DSM-III (American Psychiatric Association, 1980) in 1980 introduced an age restriction on the schizophrenia diagnosis preventing it to be considered after age 44, although the age restriction was later removed by the DSM-III-R (American Psychiatric Association 1987) and the DSM-IV (American Psychiatric Association, 1994). There may actually have been a conceptual bias in some studies of schizophrenia rates before 1980 too, since the DSM-II (American Psychiatric Association, 1968), although not employing an age criterion for schizophrenia, had a separate category for psychotic disorders in the old age period that was unsharply delimited from schizophrenia – “involutional paraphrenia” – which may have prevented older subjects to be diagnosed with schizophrenia thereby biasing the sex ratio for schizophrenia. Furthermore, in the meta analysis mentioned (Aleman et al, 2003) there was no significant sex-difference of the schizophrenia incidence in studies from developing countries; and when the meta analysis was limited to studies with an age cut off not lower than 64 years the male/female rate-ratio decreased – although it was still 1.32. The authors of the previously mentioned review (McGrath et al, 2004) actually
pointed out that systematic exclusion of older age-groups in the reviewed studies behind the male/female effect could not be entirely ruled out.

5.5.4.2 Schizophrenia and schizophrenia like conditions with late onset

There are some incidence studies of late-onset psychosis, and some data based on registers have indicated a non negligible incidence-rate of 15-20 per 100 000 person-years for late paraphrenia (Kay, 1972; Holden, 1987; Castle and Murray, 1993; Howard et al, 1994; Henderson and Kay, 1997). Moreover, in a deprived inner city area sample the incidence of DSM-III-R schizophrenia was 12.6 per 100 000 person-years in the population aged 65 years and over (Castle and Murray, 1993). In a register study annual incidence rates of 12.6 per 100 000 population was reported for DSM-III-R schizophrenia with onset after age 44 years (ref…Copeland et al, 1998). Thus, if subjects with late-onset schizophrenia/late paraphrenia – which is more common in females than males (Henderson and Kay, 1997) – were not excluded in population based epidemiological studies the sex-ratio of the schizophrenia incidence would probably turn out differently (Jablensky, 2000).

5.5.4.3 Schizophrenia and urbanicity

Incidence studies of schizophrenia in urban areas have generated significantly higher rates in both males and females than studies in mixed urban-rural areas (very few studies of purely rural areas exist). In a review the median incidence of schizophrenia was 19 per 100 000 person-years in urban studies, whereas it was 13.3 in mixed urban-rural studies (McGrath et al, 2004). Thus, findings support that urban birth and/or urban upbringing/residence may be associated with an increased risk to develop schizophrenia. In a Swedish study the register based incidence of first-admission due to a nonaffective psychotic disorder increased by increasing population density in areas that were compared, with a 68-77% greater risk for those living in the most densely populated areas compared to those living in the least populated areas (Sundquist et al, 2004).

5.5.4.4. Schizophrenia in migrant groups

In a review comparing the incidence rates of schizophrenia in studies of migrant groups with rates in native-born groups, the median of the rates in
the migrant studies was 60.0 per 100 000 person-years, whereas the median in the native-born studies was 16.9 per 100 000 person-years (McGrath et al, 2004). The median of the migrant to native-born rate ratios was 4.6. A meta-analysis of studies of the risk of schizophrenia associated with migration showed a relative risk of 2.7 for first-generation migrants and 4.5 for second-generation migrants (Cantor-Graae and Selten, 2005). The magnitude of this difference is big and as it has not been explainable by: diagnostic bias, misclassification, a lower admission threshold in migrants, differential exposure to biological risk factors such as pregnancy and birth complications, influenza exposure, heredity, a higher incidence of schizophrenia in the countries from where the subjects emigrated, ethnicity (the migrant effect has been found in subjects of different ethnic origin) and drug abuse, other explanations such as stress, demoralisation and chronic outsider status due to discrimination or repeated disappointments in the new country have been considered to be associated with the increased schizophrenia incidence in migrants (Jablensky, 2000; Bromet et al, 2002; Boydell and Murray, 2003; Cantor-Graae and Selten, 2005).

5.5.5 Incidence of delusional disorder

Little is known about the incidence of delusional disorder in the community. In a textbook the annual incidence is estimated to range from 1-3 new cases per 100 000 population, with a slight preponderance of females (Kaplan and Sadock, 1994). However, in a random community sample of people listed at general practitioners the incidence of delusional disorder in persons aged 65 or more was 15.6 per 100 000 per year (Copeland et al, 1998).

5.5.6 Incidence of nonorganic nonaffective acute remitting psychoses

In the substudy of the WHO Determinants of Severe Mental Disorder (DOSMeD) Study on ‘nonaffective acute remitting psychosis’ (NARP) that was mentioned previously (Susser and Wanderling, 1994) the annual incidence in the developing country setting was about 10-fold that in the industrialized country setting for both males and females, respectively. Further, the incidence of ‘NARP’ was about twice as high in females as in males in both settings. In the developing country setting the male annual incidence was 4.86 per 100 000 people versus the female incidence 8.78 per
100,000. In the industrialized country setting the male annual incidence was 0.40 per 100,000 versus the female 1.04 per 100,000. In the previously quoted register study of ICD-10 ‘acute and transient psychotic disorder’ the incidence was 9.6 per 100,000 population (male rate 9.8, female rate 9.4). However, the incidences by age groups were higher in males before age 50, whereas they were higher in females after age 50 (Castagnini et al, 2008).

5.5.7 Incidence of affective psychoses

In a review the treated annual incidences of bipolar I disorder ranged from 2.6 to 20.8 per 100,000 population; male rates ranged from 3.0 to 15.2 per 100,000; and female rates from 2.0 to 32.5 per 100,000 (Bebbington and Ramana, 1995). The reported community annual incidence of bipolar disorder in the NEMESIS Study was 0.30 per 100,000 population (Bijl et al, 2002). Corresponding to the wider female incidence range seen in the review, some studies of affective psychoses have reported a greater incidence in females than in males (Bland, 1977; Brewin et al, 1997), but still most studies (Baldwin et al, 2005; Lloyd et al, 2005; Bijl et al, 2002), including one recent systematic review (Waraich et al, 2004), have reported no male/female difference in the general rates of bipolar disorder. However, a gender incidence difference has been noted in late-onset psychotic depression or mania - although less consistently than for late-onset schizophrenia/late paraphrenia - with a male/female rate ratio 1:1.1-1:1.5 (Henderson and Kay, 1997).

5.6 Studies of the life time prevalence of psychotic disorders

5.6.1 Overview

Most population based studies since around 1970 of the lifetime prevalence of psychotic disorder have used structured instruments for case-finding and diagnosis (e.g. DIS, CIDI). In most studies lay interviewers have done the field work and diagnoses have often been based on computer algorithms (e.g. generating DSM-III, DSM-III-R and DSM-IV diagnoses). Some studies have used two-stage procedures with clinicians following up the initial DIS/CIDI screen with PSE/SCID-I interviews (sometimes by telephone). A minority of the studies have supplemented interview data with data from other sources (e.g. registers, key-informants). Due to
differences between studies in study periods, populations (e.g. country/culture), sampling (national, geographical, urban, rural, socio-demographic structure, age-composition), case-finding methods, diagnostic ascertainment and attrition comparisons between studies are difficult. Nevertheless, the lifetime prevalence findings indicate that psychotic disorders are rather common in the lifetime perspective with estimates of any psychotic disorder ranging from 2.9% to 4.5% (Astrup, 1989; Jacobi et al, 2004; Perälä et al, 2007). However, most studies of lifetime prevalence of psychotic disorders have focused on schizophrenia and bipolar disorder/mania; thus, not giving estimates of all psychoses.

5.6.2 Lifetime prevalence of ‘psychotic disorder due to a general medical condition’ and ‘substance-induced psychotic disorder’

For the broad group organic-toxic psychoses estimates of the lifetime prevalence have varied from 0.5% to 0.8% (van Os et al, 2001; Astrup, 1989). The PIF Study that analyzed ‘psychotic disorder due to a general medical condition’ and ‘substance-induced psychoses’ separately found the lifetime prevalence of 0.21% and 0.42%, respectively (Perälä et al, 2007).

5.6.3 Lifetime prevalence of nonaffective psychoses

Some studies have analyzed the lifetime prevalence of the broad group nonaffective psychoses and found rates from 0.37% to 1.94% (Kessler et al, 1994; van Os et al, 2001; Kessler et al, 2005; Vicente et al, 2006; Cho et al, 2007; Perälä et al, 2007).

5.6.4 Lifetime prevalence of schizophrenia

For schizophrenia the lifetime prevalence in different studies have varied from 0.12% to 1.6% (Canino et al, 1987; Bland et al, 1988; Astrup, 1989; Hwu et al, 1989; Oakley-Browne et al, 1989; Lehtinen et al, 1990; Robins and Regier, 1991; Chen et al, 1993; Scully et al, 2004; Cho et al, 2007; Perälä et al, 2007). In one systematic review of the lifetime prevalence of schizophrenia the median lifetime prevalence was 0.4% (Saha et al, 2005). In another systematic review the pooled lifetime prevalence was 0.55% (Goldner et al, 2002).
5.6.5 Lifetime prevalence of other nonaffective psychoses

A few studies have estimated the lifetime prevalence of delusional disorder to be from 0.0 to 0.67% (Cho et al, 2007; Perälä et al, 2007; Hwu et al, 1989). Some studies have analyzed the lifetime prevalence of schizophreniform disorder to range from 0 to 0.2% (Canino et al, 1987; Bland et al, 1988; Hwu et al, 1989; Oakley-Browne et al, 1989; Robins and Regier, 1991; Chen et al, 1993; Cho et al, 2007; Perälä et al, 2007). The lifetime prevalence of brief psychotic disorder has been estimated to range from 0.05% to 0.9% (Cho et al, 2007; Perälä et al, 2007).

5.6.6 Lifetime prevalence of affective psychoses

For bipolar disorder/mania the lifetime prevalence in different studies have varied from 0.07% to 1.9% (Canino et al, 1987; Bland et al, 1988; Astrup, 1989; Oakley-Browne et al, 1989; Hwu et al, 1989; Robins and Regier, 1991; Chen et al, 1993; Kessler et al, 1994; van Os et al, 2001; ten Have et al, 2002; Hanssen et al, 2003; Scully et al, 2004; Vicente et al, 2006; Cho et al, 2007; Medina-Mora et al, 2007; Perälä et al, 2007). A systematic review of the lifetime prevalence of bipolar I disorder found rates from 0.15% to 1.8%, where the pooled rate was 0.8% (Waraich et al, 2004).

A few studies have analyzed the lifetime prevalence of major depressive disorder with psychotic features finding rates from 0.35% to 0.6% (Johnson et al, 1991; Perälä et al, 2007).

5.7 Studies of predictors of psychotic disorders

5.7.1 Overview

Research on risk factors for psychosis has focused primarily on schizophrenia (Bromet et al, 2002). Having a first-degree relative with schizophrenia or some other nonaffective psychotic disorder represents the strongest risk factor for schizophrenia known. Moreover, the risk for schizophrenia is also elevated in relatives of probands with affective psychotic illness, although not as much as for probands with nonaffective psychoses (Kendler et al, 1993; Bromet et al, 2002; Cardno and Murray, 2003).
As family, adoption and twin studies indicate that the familiarity of schizophrenia is due mainly to genetic effects, models to estimate the relative contribution of genetic and environmental effects have been constructed; and the heritability for the liability to schizophrenia has been estimated to be around 80% (Cardno and Murray, 2003; Sullivan, 2005), thus leaving about 20% for environmental or nongenetic effects. However, in spite of considerable research efforts and the identification of some ‘candidate genes’ – e.g. COMT (catechol-O-methyltransferase gene), DTNBP1 (Dystrobrevin Binding Protein 1), NRG1 (Neuregulin 1) and RGS4 (Regulator of G-protein Signalling 4), DISC1 (Disrupted In Schizophrenia 1) (Sullivan, 2005; Harrison and Weinberger, 2005) – no genes of major effect in affected families have yet been found, although replications for loci on several chromosomes that may be important in the schizophrenia aetiology have been reported – e.g. 22q11, 6p22, 8p12-21, 1q21-22, 1q42 (Bromet et al, 2002; Zammit et al, 2003). But no chromosomal loci finding has been replicated in all datasets (Zammit et al, 2003).

Environmental exposures related to pregnancy and birth complications that elevate the risk of psychosis include: high paternal age, prenatal famine, low maternal weight, prenatal infection, pre-eclampsia, maternal stress/depression during pregnancy (e.g. due to bereavement, unwantedness), birth in late winter/early spring, rhesus-incompatibility, hypoxia, perinatal brain damage, low birth weight, perinatal nutritional deprivation and urban birth (Cannon et al, 2003; Bromet et al, 2002).

Postnatal exposures that increase the risk for psychosis include: inadequate nutrition early in life, neonatal and early childhood central nervous system infection, frequent illnesses (e.g. respiratory infections), Cushing’s disease, steroid treatment, adolescent cannabis/amphetamine use, head injury, visual and hearing impairment, urban upbringing, migration and stress (Compton, 2005; Cannon et al, 2002a,b).

Studies of the risk for schizophrenia/schizophrenia spectrum disorders have indicated an overall odds ratio of approximately 2 for obstetric complications (all causes) (Cannon et al, 2003), 2-3 for moderate and heavy cannabis use, respectively (Smit et al, 2004; Zammit et al, 2002) and 2-4 for first and second generation of migrants, respectively (Cantor-Graae and
Selten, 2005). In Figure 1 the odds ratios of a selected set of risk factors for schizophrenia are summarized.

![Figure 1. Comparison of a selected set of relatively well-established risk factors for schizophrenia, focusing mainly on pre-and antenatal factors (abbreviations: CNS, central nervous system; depr, depression; Rh, Rhesus) [from Sullivan PF (2005) The genetics of schizophrenia PLoS Med 2(7):e212].](image)

Beside risk factors, which are thought to be possible antecedents with a causal role – or proxies to causal factors – in the development of psychosis, there are also endophenotypes (Gottesman and Gould, 2003; Gottesman and Hanson, 2005; Weiser et al, 2005). As previously mentioned, endophenotypes are biological or neuropsychological traits in the general population that indicate a genetic susceptibility toward developing psychosis. Alternative concepts with similar meaning, although not necessarily implying a genetic underpinning, are risk markers, biological markers, intermediate phenotypes, vulnerability markers and subclinical traits. An endophenotype cosegregates with psychosis in affected families; but it is independent of the psychotic state, and thus found also in nonpsychotic first degree relatives at a higher rate than in the general population. Endophenotypes lie closer to the causes of psychotic disorders than the psychotic symptoms of these disorders do; i.e. endophenotypes lie on the pathway between cause and effect. Endophenotypes may include...
features such as enlargement of brain ventricles, minor physical anomalies, dermatoglyphic abnormalities, delayed childhood development, smooth pursuit eye-tracking abnormalities, impaired pre pulse inhibition, olfactory identification deficits, deficits in motor skills, impairments in attention and memory domains and possibly also features of schizotypy – a behavioural endophenotype (Compton, 2005; Weiser et al, 2005). Some of the identified risk factors and suspected endophenotypes may however be early manifestations of the illness, which may be seen before psychosis develops.

Probable proxy factors include birth during winter or spring and urbanicity of place of birth or upbringing. The season of birth effect in schizophrenia may be related to something affecting fetal development, e.g. seasonal variation of maternal exposure to infectious agents, pregnancy or birth complications and nutritional deficiencies. Urban birth and upbringing could be proxies for physical and biological environment, lifestyle and social factors, but there is no conclusive evidence. The idea that causal agents are associated with urbanization has been called the ‘breeder hypothesis’ which contrasts to the ‘social drift hypothesis’ and the ‘social residue theory’, which have sought to explain the higher schizophrenia incidence associated with urbanization as resulting from people with latent or manifest schizophrenia moving into urban areas or that this group tend to remain in urban areas as a city grows while the mentally healthy tend to move (Boydell and Murray, 2003).

5.7.2 Premorbid personality traits/behaviours as predictors of psychosis

The neurodevelopmental model of psychosis involves biological vulnerability reflecting underlying brain pathology (Cornblatt et al, 2003). The resulting developmental disturbances, neuropsychological deficits and behavioural impairments are hypothesized to be contributing to development of psychosis; and to be the underlying causes of the disability often associated with psychosis. Furthermore, in cases in which psychotic symptoms do not emerge, the vulnerability is thought to contribute to the development of abnormal personalities such as: schizotypal, schizoid, avoidant personality disorder and other related disorders (Cornblatt et al, 2003). Thus, certain premorbid personality traits may betray an increased risk of psychosis. If so, the relationship between premorbid personality
traits and psychosis may be an expression of continuity between premorbid personality and psychosis – premorbid traits may be predictors of psychosis with causal significance (aetiological factors or endophenotypes), early attenuated illness manifestations or factors without causal significance that nevertheless shape the expressions of psychosis if psychotic illness starts. Finally, different premorbid personality traits may be related to different psychotic disorder diagnostic categories and/or to different symptom dimensions within disorders (Parnas et al, 1982; Parnas, 1999; Tsuang and Faraone, 1999; Tsuang et al, 2000; Johns et al, 2004; Weiser et al, 2005; Kendler, 2005).

The idea that premorbid personality may be a predictor of psychosis is an old one. Since the first description of schizophrenia the disorder was thought to be associated with premorbid anomalies of personality. Kraepelin and Bleuler identified peculiarities and eccentricities in relatives to dementia praecox/schizophrenia patients, which they believed were related to the full-blown disorder (Kraepelin, 1896; Bleuler, 1908). Early the concept of schizoidi was developed – a withdrawn, secluded, detached and irritable type of personality (Bleuler, 1924; Kretschmer, 1934; Essen-Möller, 1946). Bleuler and Kretschmer thought that accentuation of premorbid schizoidi produced schizophrenia.

A later contribution, in the same vein, came from Meehl who in 1962 hypothesised that a genetic neural deficit, which he called schizotaxia (a theoretical concept corresponding to a personality organization vulnerable to schizophrenia), usually would develop into schizotypy but occasionally, as a result of other unfavourable genetic influences, adverse life events and/or social learning, would develop into schizophrenia (Meehl, 1989). The traits associated with schizotaxia, according to Meehl, were three of the primary schizophrenia symptoms of Bleuler: anhedonia, cognitive slippage (associative dyscontrol) and ambivalence; but Meehl also included interpersonal aversiveness – i.e. social fear, expectation of rejection and conviction of unlovability – among the traits he believed to reflect liability to schizophrenia.

In the Roscommon Family Study schizophrenia was more common in relatives of probands with schizotypal personality disorder. Schizotypal, paranoid, schizoid and avoidant personality disorders were found to be significantly more common in relatives of probands with schizophrenia.
than in relatives of controls. Moreover, the occurrence of schizotypal personality disorder was also significantly elevated in relatives of probands with other non-affective psychoses, whereas not in relatives of probands with psychotic affective illness. Paranoid personality disorder was also significantly more common in relatives of probands with schizo-affective disorder (Kendler et al, 1993a; Kendler et al, 1993c).

Children, who later have developed schizophrenia, have often been observed to be anhedonic, socially isolated, mildly psychotic and antisocial; prospective and follow-back studies have shown them to be less responsive, give less eye contact and show less positive affect. Teenagers, who later have developed schizophrenia, have been shown to be more anxious, socially undesirable, withdrawn, suspicious and more often within the schizophrenia personality spectrum; self-reports of social anhedonia and magical ideation have been associated with increased risk for subsequent psychosis (Carter et al, 2002).

Thus, it seems that studies of premorbid personality may have a potential to give clues in the search of the aetiology of psychosis.

5.7.2.1 Prospective studies

There are not so many prospective studies addressing the issue of premorbid personality and psychosis, but there are some community studies on army conscripts, birth cohorts and general community samples; and some studies on genetic high-risk samples.

Prospective findings reflect a range of traits from all personality disorder clusters and multiple personality domains; e.g. social anxiety, socio-behavioural deficits, high neuroticism, low extraversion, emotional instability, aggression, schizotypal features and psychoticism; as well as signs of neuropsychological deviances – e.g. deficits in organizational abilities, intelligence, language and attention – and motor abnormalities.

In a study of Swedish male army conscripts during a 15-year follow-up variables reflecting interpersonal difficulties, such as feeling sensitive, were associated with schizophrenia and, although to a lesser extent, also with other psychoses (Malmberg et al, 1998). Premorbid neuroticism at the age of 16 years in the 1946 British national birth cohort was suggested to increase the risk of schizophrenia in a dose-response relationship (van Os
and Jones, 2001). In the NEMESIS Study of community residents in the Netherlands high neuroticism was a risk factor for BPRS rated psychosis (Krabbendam et al, 2002). In the New York High-Risk Project groups of offspring to parents with schizophrenia, affective disorder and no psychiatric disorder, respectively, were assessed for personality features and disorders at the mean age of 25 years. The subjects from both high-risk groups had higher rates of any DSM-III-R personality features and personality disorders from all clusters than the control group, but the high-risk groups did not differ between each other. Thus, the findings were consistent with the view that the major diagnostic categories may share aetiological components (Squires-Wheeler et al, 1993; Erlenmeyer-Kimling et al, 1995).

However, for the offspring of schizophrenic parents premorbid deficits in childhood concerning verbal memory, gross motor skills and attention identified 83%, 75% and 58%, respectively, of the subjects who went on to develop schizophrenia-related psychoses. Whereas, in contrast, for the offspring of affectively ill parents, such childhood deficits identified only 25%, 50% and 0%, respectively, of the subjects with later onset of schizophrenia-related psychoses. For the offspring of normal parents, deficits in the studied variables did not identify any of the subjects who developed schizophrenia-related psychoses. These findings were consistent with a neurodevelopmental model of schizophrenia and that childhood deficits in verbal memory, gross motor skills and attention may be relatively specific to schizophrenia risk, and that they may be indicators of the genetic liability to schizophrenia (Erlenmeyer-Kimling et al, 2000).

In a Danish birth cohort in which the children had a parent with schizophrenia, a nonpsychotic psychiatric disorder or no parent with psychiatric disorder, respectively, premorbid personality characteristics rated at average age 12 years and adult psychiatric outcomes – schizophrenia spectrum disorder, nonpsychotic psychiatric disorder and no psychiatric disorder – were studied. The groups who developed a schizophrenia spectrum disorder or a nonpsychotic disorder, respectively, had been rated lower than the healthy controls on concentration, extraversion, maturity, friendliness, cooperation and emotional stability. The group who developed a schizophrenia spectrum disorder had also been rated lower on intelligence and higher on aggression than the healthy controls, although they did not differ significantly from the group who
developed a nonpsychotic psychiatric disorder. These findings may thus also be suggestive of nonspecificity of premorbid personality deviations, at least for several traits, in groups who develop schizophrenia spectrum disorders and other psychiatric disorders, respectively, although intelligence and aggression may be more closely associated with risk to develop a schizophrenia spectrum disorder (Ekstrøm et al, 2006).

In the Copenhagen Schizophrenia High Risk Project in which adolescents (mean age 15) of schizophrenic mothers were followed for 25 years, it was found that the subjects who developed schizophrenia had scored higher on scales measuring unusual thoughts, experiences and psychoticism (Carter et al, 1999). In the Edinburgh High-Risk Study young subjects (age 16-24) with at least two relatives with schizophrenia and a control group with relatives with no history of psychotic illness were followed. Neuropsychological tests, neurodevelopmental variables, childhood behavioural traits and schizotypal features were assessed. At baseline, the high-risk sample differed from the control group on neuropsychological and neurodevelopmental variables but not on childhood behavioural traits and schizotypal features. After 10 years the high-risk sample was divided in three subgroups; subjects who had developed schizophrenia, partial psychotic symptoms and no psychotic symptoms, respectively. Premorbid schizotypal features provided separation between the high-risk subjects who had developed schizophrenia from both the other high-risk groups and from the normal-risk control group. Furthermore, the high-risk group who had developed schizophrenia differed from the high-risk group who had not developed any psychotic symptoms and the normal-risk control group, respectively, on childhood anxiety-depression and attention problems. The high-risk subjects who had developed schizophrenia also differed from the controls (but not from the other high-risk groups) on neuropsychological tests, ocular hypertelorism, dermatoglyphics and left thalamic nucleus volume.

It was concluded that, among subjects at increased genetic risk of schizophrenia, partial and transient symptoms reflecting a state of vulnerability can be found in more subjects than will develop frank schizophrenia. Furthermore, neuropsychological tests and neurodevelopmental measures were better at separating subjects at enhanced genetic risk for schizophrenia from healthy controls at normal risk, than to separate, among high-risk subjects those who will develop
The central finding was that measures of schizotypal features and social anxiety were the best measures to distinguish the high-risk subjects who will develop schizophrenia from those who will not (Johnstone et al, 2005).

To sum up, prospective studies of army conscripts, national birth cohorts, general population samples and high risk cohorts support that it may be feasible to screen subjects who are vulnerable to psychosis. Impairments of neuropsychological functions, such as memory and attention, and developmental variables, such as motor skills and morphometric variables, may be relatively specific indicators of the genetic susceptibility to schizophrenia-related disorders. Whereas the findings of premorbid personality related traits and behaviours from most studies may rather suggest nonspecificity of premorbid personality deviations – e.g. social anxiety, high neuroticism, personality disorder and emotional instability – in groups who develop schizophrenia-spectrum and other psychiatric disorders, respectively. However, schizotypal features may have some specificity for the liability to schizophrenia or schizophrenia-spectrum disorders.

5.7.2.2 Retrospective studies

Some studies that have tried to elucidate the associations between premorbid personality and psychosis have employed retrospective designs, e.g. interviewing relatives to assess premorbid personality in probands after the onset of psychotic illness. Such studies may suffer from difficulties of differing premorbid personality traits from prodromal signs of psychosis due to recall bias. Moreover, the samples, which consist of patients, may not be representative due to selection. Nevertheless interesting findings have emerged. In a retrospective study of adult premorbid personality, premorbid schizoidi tended to be common in both schizophrenia (56%) and other non-organic psychoses (35%). However, premorbid explosive and paranoid traits were significantly more common in subjects with schizophrenia than in subjects with other non-organic psychoses, and histrionic traits were significantly more common in the other psychoses than in schizophrenia (Dalkin et al, 1994). Another study (Rodríguez Solano and González De Chávez, 2000) found that 85% of schizophrenic patients fulfilled personality disorder criteria before the onset of psychosis (avoidant- 32.5%, schizoid- 27.5%, paranoid- 20%, dependent- 20% and schizotypal personality disorder 12.5%). In studies on samples with recent-
and first-onset of broadly diagnosed functional psychosis, the premorbid schizoid dimension was found to be correlated to the negative schizophrenia dimension, the premorbid schizotypal dimension to the positive schizophrenia dimension and the premorbid sociopathic dimension to the disorganization schizophrenia dimension (Cuesta et al, 1999; Cuesta et al, 2002).
6. Aims of the thesis

i. To analyze and compare male and female overall first incidence of psychotic disorders in the Lundby population 1947-1997 (paper I).

ii. To analyze and compare male and female mean and median age-at-onset and incidence by age of psychotic disorders in the Lundby population 1947-1997 (paper I).

iii. To analyze the period prevalences of psychotic disorders in the Lundby population 1947-1997 (paper II)?

iv. To analyze the lifetime prevalences of psychotic disorders in the Lundby population 1997 (paper II).

v. To analyze whether certain constructed clusters of premorbid behavioural and personality-related signs and symptoms were predictors of functional (nonaffective and/or affective) psychosis and schizophrenia, respectively, in the Lundby population 1947-1997 (paper III).

vi. To describe general methodological difficulties in the papers (paper I, II, III, IV).
7. Material and methods

7.1 Overview

The Lundby Study was initiated in 1947 by Erik Essen-Möller (1901-1992). It was the theory by Henrik Sjöbring (1879-1956) on ‘normal and lesional personality traits’ (Sjöbring, 1904; Sjöbring, 1958) – which had been largely based on observations of mentally disordered patients – that inspired Essen-Möller, who wanted to investigate the distribution of the personality variants according to Sjöbring in an ordinary, unselected population (thus, including mostly healthy subjects). It was this research question that from the beginning oriented the Lundby Study globally towards studying the whole spectrum of mental states, personality traits, mental disorders and their possible forerunners (including substance abuse, physical illness and socio-demographic factors) that may be found in a general population. Originally the Lundby Study was meant to be a cross-sectional study (Essen-Möller et al, 1956). However, it developed into a longitudinal investigation; and so far there have been three follow-ups – 1957 (Hagnell, 1966), 1972 (Hagnell et al, 1990) and 1997 (Nettelbladt et al, 2005) – with a research methodology including individual interviews, data from registers, case-files and interviews of close relatives and other informants such as nursing staff.

At the latest follow-up the Lundby Study covered a period of 50 years. Such a long follow-up of a general population is unprecedented in psychiatric epidemiology, and it offers an opportunity to study several research questions. However, the long time that has passed also means that many things have changed in society and in the population that has been followed. Moreover, psychiatry has also changed. Thus, in analyzing the Lundby data attention has to be paid to several methodological issues.
7.2 The Lundby area

Originally, the Lundby area consisted of two neighbouring parishes – Dalby and Bonderup – in the south part of Sweden. The area, which is situated about 15 km from the medium-sized university town Lund and about 25 km from the larger city Malmö, is situated around the village Dalby. On July 1st 1947, when the Lundby area was delineated, altogether 2550 subjects were on the population registers of the two parishes, and at that time the central village – which since long had been the traditional meeting place of the surrounding population – contained a court, a medieval church and about 1000 inhabitants. In 1947, the Lundby area was largely rural with farming being the dominant trade. However, there were also self-employed artisans, one large industry and some smaller ones altogether employing a couple of hundred workers. The Lundby area was a thriving place. Importantly, the nearby cities – to which the communications were good – had large and specialized hospitals.

In 1947, the study area was judged to be representative of a rural area in the south of Sweden. Moreover, the Lundby area has ever since 1947 continued to follow the pattern of other developing areas in the western world, and may thus represent many of the changes that have taken place in rural western districts after the 2nd world war. Thus, gradually from around the late 50s the area acquired a more suburban character and the two parishes became part of the municipality of Lund in 1974. Although some parts of the Lundby area were still in 1997 best described as rural or semi-rural, most of the area has since 1972 been suburban with most people of working age commuting to other places in the neighbouring city areas; a development that had become even more pronounced in 1997. Also, from 1947 to 1997 the population living in the study area more than doubled, including a considerable expansion of the village Dalby. Moreover, an aggregation into hamlets and villages, in the previously thinly populated areas surrounding Dalby, took place (Hagnell et al, 1990; Hagnell et al, 1994; Nettelbladt et al, 2005).

7.3 The Lundby population

The original study population consisted of all the subjects who were on the parish registers of Dalby and Bonderup on the cut-off date 1st July 1947; and this population may be referred to as the 1947 cohort (n=2550; males, 1312; females, 1238). Moreover, on the 1st July 1957 another cohort (the
‘geographical’ 1957 cohort) was formed by the individuals who were by then registered in the area; i.e. were still alive and had not moved or had been added to the area since 1947 (n=2612). However, if also those subjects from the 1947 cohort who by 1957 had moved out of the study area and were still alive in 1957 (n=698) were included an ‘extended’ 1957 cohort contained 3310 subjects (males, 1696; females, 1614). Between 1947 and 1957 253 subjects had died and 1013 had been born (n=228) or migrated (n=785) into the Lundby area. The two differently defined 1957 cohorts are both overlapping with the 1947 cohort; with 1599 subjects being in both the 1947 (n=2550) and the ‘geographical’ 1957 (n=2612; males, 1335; females, 1277) cohorts, and 2297 subjects being in both the 1947 (n=2550) and the ‘extended’ 1957 (n=3310) cohorts, respectively.

At inception the subjects in the 1947 cohort were 0-92 years old (median age 34 years), whereas, the subjects who were incepted in the ‘geographical’ 1957 cohort were 0-96 years old (median age 36 years); the 1013 newcomers in 1957 were 0-95 years old (median age 22 years). The total study population counts 3563 subjects (males, 1823; females, 1740) as no new subjects have been incepted in it since 1957. From the perspective of the total population the subjects were 0-95 years old (median age 31 years) when they were incepted for the first time in 1947 or 1957. Between 1957 and 1972 another 483 subjects had died, thus reducing the living study population in 1972 to 2827 subjects aged 15-97 years (males, 1425; females, 1402). Between 1972 and 1997 1030 more subjects died, further reducing the population in 1997 to 1797 subjects aged 40-96 years (males, 851; females, 946). In Figure 2 a flow chart of the total Lundby population between 1947 and 1997 may be seen.
Figure 2. The Lundby population 1947-1997
From the start migration of the study subjects from the study area to other places have been marked. Thus, in 1957 about 30% (698/2297) of the 1947 cohort still alive had moved out; by 1972 almost 50% (1403/2827) of the total population still alive lived outside the Lundby area; and by 1997 the corresponding figure had risen to 66% (1195/1797). Taken together, about 50% of the males (904/1823), and 52% of the females (913/17450) of the total Lundby population migrated from the Lundby area from 1947 to 1997. Some subjects moved a long way – e.g. to Stockholm or even abroad – but most of the migrants stayed in the south of Sweden; mostly in areas neighbouring the Lundby area.

The Lundby population has been exposed to a great change of societal structure leading from rurality to what may be called partial ‘urbanization’ – including that the Lundby area was transformed into a suburban area and that many study subjects moved to suburbs or cities – but still, the majority of the population did not live in large cities but rather in semirural areas, suburban areas or small towns (of the subjects alive in 1997 about 70% lived in communities with less than 10 000 inhabitants); and certainly the great majority of the subjects were rurally born and raised. Thus, the Lundby population – although exposed to partial ‘urbanization’ – has certainly not been exposed to the kind of urbanicity that is associated with metropolitan life. In all longitudinal studies attrition may seriously impair the validity of the result due to selection bias. However, there are many aspects of attrition: some subjects may actively drop-out due to lack of motivation to continue in the study, others may passively fall off because they could not be tracked (e.g. due to migration) and others still may not be possible to follow up as they died between two waves of investigation. Furthermore, subjects who were nonresponders in one wave – because they were untraceable – may be responders in the next, if found. The attrition in the Lundby Study of subjects who could not be tracked at follow-up or had died between follow-ups may, at least in part, have been balanced by the use of other data sources (which may also have partly balanced out recall bias at the personal interviews). Thus, depending on whether attrition is defined as lack of a follow-up interview or lack of any data at all the fall-off rate may vary. For interviews at follow-up of living subjects the attrition was 1.2% (30/2550), 1.5% (50/3310), 1.8% (50/2827) and 13% (238/1797) in the follow-ups in 1947, 1957, 1972 and 1997, respectively. However, if all sources of
information were used (interviews, registers, case-files and informants) the number of subjects rose for whom information was sufficient to allow a psychiatric epidemiologic evaluation at follow-up. If sufficient information in subjects, who were alive at follow-up, was defined as obtained from a personal interview (supplemented or not by other data sources) or – in the absence of an interview – from at least two other data sources the attrition was lowered to 0.7% (17/2550), 0.6% (19/3310), 0.1% (4/2827) and 8.7% (156/1797) in 1947, 1957, 1972 and 1997, respectively.

If all sources of information are used, and attrition is analyzed including both the subjects who were alive at the follow-ups and the subjects who died between the follow-ups, the attrition was 0.5% (17/2550), 1.1% (39/3563), 0.2% (6/3310) and 5.9% (168/2827) at the follow-ups in 1947, 1957, 1972 and 1997, respectively. Table 1 details the data sources and the attrition.

<table>
<thead>
<tr>
<th>Year of follow-up</th>
<th>All subjects</th>
<th>Alive</th>
<th>Dead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Data obtained from</td>
<td>Data obtained from</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interview</td>
<td>Other sources only</td>
</tr>
<tr>
<td>1947</td>
<td>2550</td>
<td>2520</td>
<td>13</td>
</tr>
<tr>
<td>1957</td>
<td>3563*</td>
<td>3260</td>
<td>31</td>
</tr>
<tr>
<td>1972</td>
<td>3310</td>
<td>2777</td>
<td>46</td>
</tr>
<tr>
<td>1997</td>
<td>2827</td>
<td>1559</td>
<td>82</td>
</tr>
</tbody>
</table>

*Original cohort (2550) + newcomers (1013).

Overall, including all subjects and utilizing all sources of information, the longitudinal attrition between 1947 and 1972 was about 1% (26/3563); and the corresponding attrition between 1972 and 1997 was about 6% (168/2827). Importantly, the attrition by age 1972-1997 was not equally distributed as it was higher in the subjects aged under 50 years in 1972 –
7% in males and 10% in females – and lower in the subjects aged 50 or more in 1972 – 2% in males and 3% in females.

To sum up, the Lundby population is a complete population that consists of an ethnically homogeneous group of Nordic people (Hagnell et al, 1990). All age-groups and all subjects – including those institutionalized – have been studied. At the beginning (1947 and 1957) the population was rural but by 1972 and 1997 the major part of the population may be described as suburban. Moreover, during the study period the population has been exposed to societal changes associated with ‘urbanization’ – including increased welfare, development of public health care and education, entrance of women into the labour market, introduction of birth control, changes in family structure, changing roles of males and females, changing roles of social classes and lessening of the cohesive power of family, church and community. The general societal changes that the Lundby population has been exposed to have also been described as increasingly manifesting themselves in a process of individualization (Beck, 1994; Lindbladh and Lyttkens, 2002; Bogren et al, 2007). Furthermore, several changes of biological and physical factors may be subsumed in the socioenvironmental changes (such as smoking habits, diet, medication, environmental toxins and pollution).

The Lundby population represents people born between 1854 and 1957 who have been exposed to the changes in society that followed after the 2nd World War. The study population is not representative for generations born after 1957, people born, raised and living in large cities and it does not contain any migrants from outside the Nordic countries. Moreover, it is not representative regarding effects on a population of widespread use of illegal drugs such as cannabis, amphetamine, cocaine and heroin, as there have been very few users of drugs other than alcohol in the Lundby population.

### 7.4 Case-finding method

Psychiatrists trained and working at the psychiatric clinic in Lund have carried out all field-investigations in the Lundby Study. In 1947 four clinicians (Essen-Möller, Larsson, Uddenberg, White) led by Essen-Möller conducted the study. In 1957 Hagnell alone (supervised by Essen-Möller) and in 1972 Hagnell together with Öjesjö did the follow-ups. In 1997 a team (Nettelblad, Mattisson, Bogren, Hofvendahl, Toråker, Öjesjö,
Hagnell) lead by Nettelbladt conducted the field work (Nettelbladt et al, 2005). The same case finding method – built on the principle to collect as much relevant data as possible from as many sources as possible – has been used throughout the Lundby Study. This principle draws on the assumption that bits of information from various sources are complementary and increases reliability and validity. The field work included face to face interviews with the study subjects of varying duration (from a few minutes to several hours), the majority of which took place in the homes of the study subjects or at their work places. The interviews, which were semi-structured and combined with an observation of the subject and his/her environment, basically kept the same form from 1947 to 1997. The interviews covered about 150 items including questions about: the mental and physical health of the subjects since the previous investigation, contacts with mental and other health services and reported subjective complaints (e.g. feelings of tension, worry and tiredness). Moreover, suicidal ideation and suicide attempts were probed for, as well as for psychotic ideas. An assessment of alcohol and drug problems was included. Data about family and socio-demographic characteristics were also recorded (Hagnell et al, 1990; Nettelbladt et al, 2005). The investigation also included an assessment according to a check list of whether certain behaviours – e.g. lowered mood, restlessness, sensitivity and eccentric behaviour – could be observed or not in the study subject during the interview. At the follow-up in 1997 the Mini Mental Test was administered if there was a clinical suspicion of cognitive impairment (Folstein et al, 1975).

Beside the semi-structured format the interview also included a free conversation. It was actually rather common that additional important interview information came under the free conversation – either from the study subject him/herself or from the spouse (as it was not unusual that married couples were interviewed together). Subjects with different educational background and social circumstances tended to use different concepts when they tried to explain their experiences and it was important for the interviewers to be flexibly attentive as not to miss important information or hints. The free interview partly filled this purpose. Generally, it was also during the free part that the field-workers got the best opportunity to evaluate general behaviour and non-verbal communication in the interpersonal contact. Apart from answering and grading the responses from the study subjects on the interview schedule, the interviewing psychiatrist also, since 1957, formulated a description of
his/her impressions from the interview, including impressions of the environment.

Importantly, the case finding method was a multi-source method as it also included extensive collection of data from outside sources; mostly from various local and national health registers, archives of nearby hospitals and out-patient clinics (including psychiatric clinics, other clinics, private practitioners and primary care units). Via the registers and archives case-files were procured. Another important source of information – particularly at the investigations in 1947 and 1957 – was reports from informants (e.g. relatives, nursing staff and local authorities like clerics and teachers). At the 1972 and 1997 investigations key-informant information was more difficult to obtain, due to the dispersal of the study subjects, but nevertheless such information was sometimes obtained during the interviews when parents told us about their children and children about their parents or other relatives, since the study subjects were sometimes related to each other. The field work was in some respects a detective’s job, as one data source sometimes gave clues leading to other sources where important information was eventually discovered. All relevant ‘hints’, including register-hits, were followed up with requests for case-files and the subsequent extracting of relevant clinical data. In the case of subjects who had died between two study waves the outside sources were the only way to achieve a follow-up, but also in cases where interviews were performed the outside sources often proved very valuable in balancing biased self-reports (e.g. in giving information on hospitalizations which the subject had forgotten about, in modifying the symptomatic picture described by the subject of some previous disorder and in improving the assessment of the time of onset of a disorder).

There were, taken together, more sources of outside information available between 1947-1972 than between 1972-1997 (e.g. a register over social insurance office data, more reports from informants and information from county temperance boards). The clinical data between 1947-1972 were based on access to the archives of the nearby hospitals and extensive contacts with local outpatient clinics and private practitioners. The clinical data 1972-1997 were based on the national patient register covering all hospitalizations (all causes) in Sweden 1972-1997, a local out-patient register covering the Lundby Study area and contacts with relevant clinics when indicated by information in the registers, interviews or other sources.
7.5 Ascertainment of diagnosis

A psychiatric categorical diagnosis which refers to some pattern of symptoms and signs may be assessed by degrees of severity. The related concept of caseness refers to the dichotomy of being assessed as ill or not, i.e. to the threshold for clinical diagnosis. Thus, criteria for a diagnosis may be fulfilled without criteria for caseness being met, in case of which the diagnosis is judged to be subclinical (e.g. not associated with significant functional impairment, suffering or need for intervention). In epidemiology, various definitions of caseness may be used depending on the purpose of the study, the amount and quality of data that are to be assessed and the study sample.

In the Lundby Study one of the aims was to study the broad categories of mental disorder that are frequent in a normal population using data with varying quality from several sources that had been collected during a long follow-up period. The definition for caseness that was judged to best suit these aims and circumstances was the diagnosis of mental illness made by the evaluators, based on information from all kinds of available sources (Hagnell et al, 1990). To aid the evaluators in their assessments of diagnosis/caseness, ‘mental illness’ was defined as:

‘a deviation from the person’s usual way of functioning so obvious that it is easy to recognize. It should show itself either as a real suffering for the individual, a disease, or as an interference with his work capacity, or a combination of these’ (Hagnell et al, 1990).

Moreover, in the Lundby Study each episode of a diagnosed disorder was rated along with the degree of dysfunction that was judged to be caused by it (minimal, mild, medium, severe or very severe) according to the criteria defined by Leighton et al (Leighton et al, 1963). This enabled the researchers to define caseness flexibly (e.g. requiring different degrees of impairment for a diagnosis to qualify for caseness). For psychotic disorders, due to the inherent severity of these disorders, the impairment level will tend to be at least medium, which roughly corresponds to the GAF score.
60-1 (although this will of course also depend on how the criteria for the psychotic disorder are defined).

Thus, in practice in the Lundby Study the ascertainment of diagnosis was built on all data that had been procured. Furthermore, data from interviews, key-informants, registers and case-files was, in principle, equally important. Each individual assessment was based on an attempt to integrate all that was known about a study subject. The diagnoses were based on best estimates. Moreover, since 1957 the diagnoses were based on consensus between the researchers – in 1957 between Hagnell and Essen-Möller, in 1972 between Hagnell and Öjesjö and in 1997 between Nettelbladt, Mattisson, and Bogren. During the 1997 follow-up the field workers had regular meetings with the previous field workers Hagnell and Öjesjö for supervision and support to improve diagnostic reliability. During the Lundby Study the diagnostic tradition, laid down by Essen-Möller and Hagnell, has thus been carried on.

In the first wave of the Lundby Study diagnoses were assessed according to an original system created by Essen-Möller (Essen-Möller et al, 1956). However, at the second study wave, the original classification was further developed by Essen-Möller and Hagnell (Hagnell et al, 1990), and the cases of mental illness recorded in 1947 were then re-diagnosed according to the revised system (by Hagnell and Essen-Möller). The Lundby 1957 classification is practical and adapted to the field-conditions of the Lundby Study; but it is also a reflection of the predominant way of thinking about ‘mental illnesses’ in the year of 1957.

The classification is conceptually based and it divides mental disorders into three main classes: ‘neuroses’, ‘psychoses’ and ‘organic brain disorders’. The ‘neuroses’ and ‘psychoses’, as opposed to the ‘organic brain disorders’, refer to disorders without obvious underlying gross organic aetiology. The ‘neuroses’, as opposed to the ‘psychoses’, refer to disorders with preserved insight. However, the demarcation lines between the classes are unclear. The main classes may be subdivided into main categories (essentially named according to their dominating symptom): ‘depression’, ‘anxiety’, ‘tiredness’, ‘mixed neurosis’, ‘schizophrenia’, ‘other psychoses’, ‘organic syndrome’ and ‘dementia’. The main categories may be further subdivided into less broad categories; e.g. ‘depression’ may be split into ‘depression proper’ (reflecting a well demarcated depressive syndrome without other
prominent psychiatric symptoms) and ‘depression plus other psychiatric symptoms’ (reflecting an episode with marked depressive symptoms combined with a significant – although not dominant – portion of non-depressive symptoms – e.g. anxiety or delusions). The grouping of categories was based on the principle to lump disorders according to likeness of psychopathological symptoms and signs. The construction of the nomenclature is hierarchical, thus allowing only one diagnosis per episode of mental illness; and in the hierarchy the organic brain disorders rule out the psychoses, which in turn rule out the neuroses.

In this schema disorders with psychotic symptoms as main features belong to the class ‘psychosis’ (categories: ‘schizophrenia’ [diagnosed according to Bleuler] or ‘other psychoses’). Psychotic states associated with a primary depression belong to the class ‘neurosis’ [the psychotic symptoms associated with depression are not viewed as main pathological features but as secondary symptoms] (category: ‘depression plus other psychiatric symptoms’). Moreover, psychotic states associated with dementia belong to the class ‘organic brain disorders’ (category: ‘dementia’). Manic states – which are viewed as psychotic per se – organic and toxic psychoses (separable from dementia) are included in the heterogeneous category ‘other psychoses’.

At the 1997 study wave the Lundby system for diagnosis was kept to make comparisons over time feasible. However, diagnoses according to the DSM-IV were also assessed simultaneously for the period 1972-1997 (best-estimate consensus grouping). Furthermore, to be able to compare DSM-IV diagnoses of psychotic disorders over the whole study period 1947-1997 the episodes of psychotic Lundby mental disorders recorded between 1947-1972 were re-diagnosed according to the DSM-IV. This included episodes with Lundby ‘schizophrenia’, ‘other psychoses’ and ‘depression plus other psychiatric symptoms’. For the papers in this thesis psychotic disorders diagnosed according to the DSM-IV have been used. Obviously this is not ideal as the DSM-IV disorders between 1947 and 1972 were diagnosed retrospectively using DSM-IV rules to prior data. Nevertheless, we regarded the risk for misclassification as sufficiently small to justify the use of a contemporary diagnostic system, which is considered to be important, to be able to communicate data.
The DSM-IV diagnoses were allocated to five groups [a, b, c, d, e]: a) ‘psychotic disorder - including delirium - due to a general medical condition’, b) ‘substance-induced psychotic disorder – including delirium and intoxication with psychotic features’, c) ‘schizophrenia – including schizoaffective disorder’, d) ‘other non-affective psychotic disorders’, and e) ‘affective psychotic disorders – including major depressive disorder with psychotic features, and bipolar disorder’. A hierarchy was used to settle questions related to boundary problems. Thus, group a) and b) are ranked in the top of the diagnostic hierarchy. Moreover, between the groups c), d), and e) the assessment of c) ‘schizophrenia/schizoaffective disorder’ meant that d) ‘other non-affective psychoses’, and e) ‘affective psychotic disorder’ could no longer be diagnosed. The definition used for age-at-onset of a disorder was the age of the subject when the diagnostic criteria for the disorder were met for the first time.
8. Epidemiological methods

8.1 Paper I
The first paper – ‘Incidence of psychotic disorders in the 50-year follow-up of the Lundby population’ – is an observational prospective study describing the first incidence rate of psychotic and bipolar disorders diagnosed according to the DSM-IV in the total Lundby population under risk for psychosis at the inception in the study. Comparisons between males and females for overall and age-specific incidence rates are made, as well as for age-at-onset of psychotic disorders. The differences between the sexes are discussed in view of previous studies and hypotheses that may be helpful in explaining the effects.

8.2 Paper II
The second paper – ‘How common are psychotic and bipolar disorders? A 50-year follow-up of the Lundby population’ – also draws on the observational prospective method. However, this paper describes the 50-year prevalences (1947-1997) of the DSM-IV psychotic and bipolar disorders in the Lundby population and the corresponding 1997 lifetime prevalences. The findings are compared to findings from other community studies; differences and similarities are discussed.

8.3 Paper III
The third paper – ‘Predictors of psychosis: a 50-year follow-up of the Lundby population’ – is an analytical (aetiological) study proceeding from the prospective background factors that were collected in the 1947 and ‘extended’ 1957 Lundby cohorts, respectively. The background factors, which describe clusters of premorbid personality-related behaviours and self-reported traits, are analyzed with regard to their relationship to the subsequent development of psychotic disorders. In the study the background factors were aggregated in dichotomous predictor variables.
(e.g. ‘nervous-tense’, ‘blunt-deteriorated’, and ‘paranoid-schizotypal’) and time to onset of nonaffective and/or affective psychosis and schizophrenia, respectively, was assessed and compared in subjects exposed and unexposed, respectively, to the predictors.

8.4 Paper IV

The fourth paper – ‘Does it make sense to do repeated surveys? – the Lundby Study 1947-1997’ – describes the Lundby Study and the difficulties in achieving a reliable and valid 50 year follow-up. The paper describes the study population, the collection of data, and diagnostic evaluations. It discusses various problems inherent in longitudinal research – e.g. related to changing sample representativeness, changing data sources, increasing attrition, adding of biases, difficulties maintaining diagnostic inter-rater reliability over time, and the impact of the definition of caseness for reliability and validity of the results. Methodological considerations and problems are exemplified by a presentation of first incidence rate data for neurosis of varying degrees of severity for the period 1947-1997.
9. Statistical methods

9.1 Paper I

Calculations of first incidence rates were based upon subjects who at the start of the follow-up (either in 1947 or in 1957) had been free from the outcome under study and outcomes being ranked higher in the diagnostic hierarchy. Male-female differences of mean age-at-onset of disorders were tested by constructing 95% confidence intervals (CI) for the differences. A difference not having a 95% CI including 0 was considered to be statistically significant. First incidence rates were calculated as the number of first onsets of a disorder in study subjects aged 15 years or more divided by the accumulated number of person-years at risk for that disorder. Male/female differences of IR were tested by constructing 95% CI for the male/female IR ratios. A difference not having a 95% CI including 1 was considered to be statistically significant (Rothman and Greenland, 1998).

SPSS for windows, version 13.0 was used for the calculations.

9.2 Paper II

For the period 1947-1997 prevalences of DSM-IV psychotic disorder categories were calculated. In 1997, for all subjects alive, lifetime prevalences of the DSM-IV categories were calculated along with their 95% confidence intervals. The statistical significance of a difference between the prevalences in males and females for the major diagnostic groups was tested. A difference not having a 95% CI including 0 was considered to be statistically significant (Altman, 1991).

SPSS for windows, version 13.0 was used for the calculations.
9.3 Paper III

To reduce the number of the personality-related behaviours and self-reported traits that had originally been assessed in subjects at risk for psychosis, factor analysis, along with consensus decisions in the research team, were used. However, the factor analysis was only guiding and it was in the consensus group that the final decisions were taken on how to group the various original items into predictor variables. Thus, the original items ‘paranoid’, ‘schizoid’ and ‘bizarre’ were grouped together to constitute the predictor ‘paranoid-schizotypal’, and the original items ‘sensitive’, brittle’ and ‘frail’ were grouped to constitute the predictor ‘sensitive-frail’ and so on. In all 11 predictors were constructed, six of which were based on the original items that were assessed in the follow-ups of both the 1947 and the ‘extended’ 1957 cohort, but five of which were based on items assessed only in the ‘extended’ 1957 cohort as these items were added to the semi-structured interview in 1957. Actually, it was this incongruence pertaining to assessment of personality-related items that made it necessary to analyze the 1947 and the ‘extended’ 1957 cohorts separately, as not to lose information. A table of all constructed predictor variables in both cohorts may be seen in paper IV (Table 3).

The predictors were dichotomous and assessed positively if one or more of the original items in the group were rated as medium or high on its scale, whereas a predictor was assessed negatively if all of the original items belonging to it were rated with low severity or if no original item was present. As the original items, on which the predictors were based, were assessed at the face-to-face interviews in 1947, 1957 and 1972 the predictors were time-dependent in the analysis, i.e. they could change their values at follow-ups. However, between follow-ups the predictors had to be regarded as invariable.

The relationship between the constructed predictor variables and time to first onset of nonaffective and/or affective psychosis and schizophrenia, respectively, was analyzed through Cox regression with time-dependent risk factors in both cohorts. The Cox regression procedure yielded hazard ratios, confidence intervals and p-values in simple and multivariate models for exposures to the predictor variables and time to the psychotic outcomes. To adjust for sex and age these variables were included in the simple analyses and were kept in the final multivariate models. P=0.05 was set as threshold for statistical significance in the regression procedures. Study
subjects stopped providing person-years at risk for the outcomes in the models (i.e. were censored) if they died, contracted dementia, dropped off the study or as the study period ended.

SPSS for windows, version 13.0 was used for the calculations.

9.4 Paper IV

Calculations of first incidence rates for neurosis were based upon subjects who at the start of the follow-up (either in 1947 or in 1957) had been free from neurosis and outcomes being ranked higher in the diagnostic hierarchy (e.g. schizophrenia and dementia). For the period 1947-1997 age-standardised incidence rates for ten five-year periods (1947-1952, 1952-1957, 1957-1962 etc) were calculated for neurosis with medium and severe impairment as threshold for caseness, respectively, in males and females, respectively.

The inter-rater reliability for the main Lundby classes of mental illness with impairment level medium for caseness (neurosis, psychosis, organic brain syndrome and no diagnosis) between the research teams in 1957-1972 (Hagnell, Öjesjö) and 1997 (Mattisson, Bogren), were calculated as Cohen’s kappa (Landis and Koch, 1977).

SPSS for windows, version 11.1.1 was used for the calculations.
10. Ethical approval

The ethics committee of the medical faculty at Lund university hospital approved the 1997 follow-up of the Lundby Study and the participants provided written consent.
11. Results

11.1 Paper I

In paper I – ‘Incidence of psychotic disorders in the 50-year follow-up of the Lundby population’ – the overall 50-year incidence rate in males was higher than the female for ‘substance-induced psychotic disorder’, but for the other psychotic disorders the overall rates did not differ significantly between the sexes. The male mean age-at-onset was lower than the female for ‘any psychotic or bipolar disorder’, ‘psychotic disorder due to a general medical condition’, ‘nonaffective psychotic disorder’, and ‘schizophrenia’, respectively. The mean age-at-onset also tended to be lower in males for ‘other nonaffective psychoses’ and ‘affective psychoses’, whereas the mean age-at-onset of ‘substance-induced psychoses’ in males and females were more or less equal. Males and females had different incidence by age patterns for ‘any psychotic or bipolar disorder’, ‘nonaffective psychotic disorder’, ‘schizophrenia’ and ‘other nonaffective psychotic disorder’, respectively; with a male preponderance among early onset cases and a female preponderance among late onset cases. It was concluded that the differences in age-at-onset and incidence by age between the sexes may indicate psychotic disorder delaying mechanisms in females and/or different aetiologies of psychosis in males and females.

11.2 Paper II

In paper II – ‘How common are psychotic and bipolar disorders? A 50-year follow-up of the Lundby population’ – the period prevalence 1947-1997 and the lifetime prevalence for the surviving study subjects in 1997 were analyzed.

The 50-year period prevalence was: 4.24% for ‘any psychotic or bipolar disorder’, 2.25% for ‘nonaffective psychotic disorder’, 1.43% for ‘schizophrenia’ (including ‘schizoaffective disorder’), 0.76% for ‘psychotic
disorder due to a general medical condition’, 0.62% for ‘affective psychotic disorder’, and 0.59% for ‘substance-induced psychotic disorder’.

The lifetime prevalence in 1997 was: 2.82% for ‘any psychotic or bipolar disorder’, 1.38% for ‘nonaffective psychotic disorder’, 0.84% for ‘schizophrenia’ (there were no cases of lifetime ‘schizoaffective disorder’ in 1997), 0.54% for ‘psychotic disorder due to a general medical condition’, 0.48% for ‘substance-induced psychotic disorder’, and 0.42% for ‘affective psychotic disorder’.

For the 50-year period prevalences the sexes differed significantly only for ‘substance-induced psychotic disorder’ with 0.99% in males, and 0.17% in females, but the period prevalences for the other psychotic disorders did not differ significantly between the sexes.

For the lifetime prevalences in 1997 the sexes differed significantly for the general group ‘any psychotic or bipolar disorder’ with 3.75% in males, and 1.96% in females, but for the separate psychotic disorder groups the lifetime prevalences in 1997 did not differ between the sexes.

There were no other studies found to compare the 50-year period prevalence estimates in the Lundby Study with, but for the lifetime prevalences there were several studies with which comparisons were feasible (see table 4 in paper III). The lifetime prevalences obtained in the Lundby Study in 1997 were higher than in most recent studies for ‘psychotic disorder due to a general medical condition’, ‘substance-induced psychotic disorder’, ‘schizophrenia’, and ‘delusional disorder’, whereas they were lower for ‘brief psychotic disorder’, ‘schizophreniform disorder’, and ‘affective psychotic disorder’.

It was concluded, however, that the lifetime prevalences obtained in the Lundby Study cannot be generalized to other populations, and comparisons with other studies are problematic as other studies have often assessed the lifetime prevalence in samples with different age structures (e.g. 18-65 years) than the Lundby population in 1997, and also the case-finding methods and the diagnostic ascertainment have been different. Moreover, there may in the Lundby data in 1997 be a bias problem based on a relatively higher selection of non-psychotic than psychotic survivors, as
subjects who have contracted a psychotic disorder at a young age may die earlier than subjects who have been free from psychosis.

Nevertheless, in spite of the methodological issues it was concluded that the prevalence findings in the Lundby Study suggests that psychotic disorders are common in the general community.

11.3 Paper III

In paper III – ‘Predictors of psychosis: a 50-year follow-up of the Lundby population’ – the relationships between dichotomized predictors and the subsequent development of first incident ‘nonaffective and/or affective psychosis’ and ‘schizophrenia’, respectively, were analyzed. The predictors were related to clusters of premorbid personality traits, enduring subjective subclinical symptoms and behaviours. The exposure-outcome associations were analysed by simple and multivariate Cox regression models including both males and females.

It was found that the predictors ‘nervous-tense’ and ‘blunt-deteriorated’ were significantly related to an increased risk to develop ‘nonaffective and/or affective psychosis’ in simple and multivariate models of the 1947 cohort, while the predictors ‘nervous-tense’, ‘paranoid-schizotypal’ and ‘tired-distracted’ were significantly related to psychosis in the ‘extended’ 1957 cohort (note: the predictor ‘tired-distracted’ could not be tested in the 1947 cohort).

Moreover, in simple models for ‘nonaffective and/or affective psychosis’ in the ‘extended’ 1957 cohort the predictors ‘down-semidepressed’, ‘sensitive-frail’ and ‘easily hurt’ were also significantly related to an increased risk to develop ‘nonaffective and/or affective psychosis’ (note: these predictors could not be tested in the 1947-cohort).

When ‘schizophrenia’ was analyzed separately the predictor ‘nervous-tense’ remained significant in the simple and multivariate models for the ‘extended’ 1957 cohort, although not in the 1947 cohort, but ‘blunt-deteriorated’, ‘paranoid-schizotypal’ and ‘tired-distracted’ did not reach significance in any of the models of the cohorts.
However, for ‘schizophrenia’ the predictor ‘abnormal-antisocial’ reached significance in the simple and multivariate models of the 1947 cohort. Also, in a simple model for ‘schizophrenia in the 1947 cohort the predictor ‘blunt-deteriorated’ bordered to significance.

In an analysis of sex x predictor interactions one predictor (but only in the 1947 cohort) – ‘nervous-tense’ – was differently associated with risk to develop ‘nonaffective and/or affective psychosis’ in males and females, respectively. In the male simple and multivariate Cox regression models ‘nervous-tense’ was significantly associated with ‘nonaffective and/or affective psychosis’ risk, whereas it was not significantly associated with psychosis in any of the female models.

Since all of the other predictor-outcome associations did not differ significantly between the sexes, and because the sample was small, separate analyses for the sexes were not undertaken.

For the predictors that could be studied in both cohorts only one predictor – ‘nervous-tense’– turned out to be significantly associated with an increased risk to develop ‘nonaffective and/or affective psychosis’ in both cohorts (simple and multivariate models), whereas ‘blunt-deteriorated’ was significant only in the 1947-cohort; and ‘down-semidepressed’ and ‘paranoid-schizotypal’ were significant only in the ‘extended’ 1957 cohort (simple model and simple and multivariate models, respectively).

It was hypothesized that the findings of different predictors associated with psychosis in the 1947- and 1957 cohorts, respectively, could be the result of differences in the cohorts pertaining to the types of psychosis outcomes during the follow-ups; e.g. different psychosis types related to age-at-onset and severity/pervasiveness of symptom patterns. Support for this hypothesis was found as there was a somewhat higher proportion in the 1947 cohort compared to the 1957 cohort with first onset cases of psychotic disorders that are usually severe, i.e. ‘schizophrenia’ (29% versus 24%), and ‘affective psychosis’ (33% versus 26%), and a somewhat lower proportion of psychotic disorders that are usually less severe, i.e. ‘delusional disorder’, ‘brief psychotic disorder’ and’ psychotic disorder NOS’ (16% versus 26%). Additional support for the hypothesis was found as the subjects in the 1957 ‘extended’ cohort who represented the overlapping members from the 1947 cohort (n=2220) – and actually constituted the majority of the ‘extended’
1957 cohort – had developed ‘nonaffective and/or affective psychosis’ at the median age 53.5 years, whereas the members of the 1947 cohort had developed psychosis at the somewhat earlier median age 48.0 years. Taken together these findings suggested that the types of psychotic disorders that had emerged in the 1947 cohort as compared to the ‘extended’ 1957 cohort may represent a group of cases with a higher proportion of subjects more seriously and pervasively affected. A reasonable explanation for this would be a selection effect resulting from the inclusion of the still psychosis free survivors of the 1947 cohort into the 1957 cohort (among whom some subsequently developed ‘milder’ psychotic disorders, which thus tended to have a later onset). The different predictor profiles in the 1947 compared to the 1957 cohort may thus be the biased result of selecting a somewhat different set of psychotic outcomes in the latter cohort. This hypothesis is compatible with the more serious and pervasive predictor ‘blunt-deteriorated’ being significant in the 1947 cohort and the somewhat less serious and pervasive predictors ‘down-semidepressed’ and ‘paranoid-schizotypal’ being significant in the 1957 cohort.

As for the difference between the sexes regarding the association of the predictor ‘nervous-tense’ and ‘nonaffective and/or affective psychosis’, it was suggested that measurement bias is capable of explaining the effect, as male respondents who at face-to-face interviews reported and/or displayed signs such as worry, nervous tension and restlessness may well represent a subset of males more seriously assailed than the female respondents who reported/displayed such phenomena, due to a culturally sanctioned lower threshold in females than in males to reveal such feelings in words and/or behaviour.

11.4 Paper IV

Paper IV – ‘Does it make sense to do repeated surveys? – the Lundby Study 1947-1997’ – describes the materials and methods of the Lundby Study focusing on the difficulties to achieve a reliable and valid follow-up after 50 years.

It is concluded that to make sense repeated surveys must deal with several problems that are inherent in longitudinal studies. Thus, with time a population sample will grow older and more and more subjects will die. A
sample will also be reduced due to migration and to refusal in continuing of participation in further follow-ups.

By the last follow-up in 1997, the Lundby population had dwindled to 1797 living study subjects aged 40-96 years, compared to the 2550 subjects aged 0-92 years initially recruited in 1947 and the 1013 subjects aged 0-95 incepted in 1957. Moreover, about half of the study population had migrated from the Lundby area during the follow-up.

In the paper the Lundby case-finding method is described and it is concluded that the multi-source method has been crucial to achieve a follow up of the Lundby population. For instance, at the follow-up in 1997 1030 of the study subjects that were to be followed-up from 1972 had died, and the very long follow-up period between 1972 and 1997 made it difficult for the subjects who were interviewed in 1997 to remember episodes of mental disorder (especially of milder degree), clearly necessitating the use of outside sources to achieve a follow-up and to counteract attrition and recall bias. When all sources were used for all subjects, including those who died between follow-ups, the overall attrition was about 1% (26/3563) between 1947 and 1972 and about 6% (168/2827) between 1972 and 1997. Whereas the attrition specific for the interviews of the living subjects at the 1997 follow-up was considerably higher, at 13% (238/1797), and the attrition based on the total cross-sectional information (all sources) for the subjects alive on July 1st 1997 was also rather high at 9% (156/1797).

A problem associated with the multi-source method in a study of long duration is that data sources change over time, which may reduce diagnostic consistency during a long follow-up. Nevertheless, if several data sources are available the consistency problem may be partly counterbalanced. Unfortunately however, in the Lundby Study there were more data sources between 1947 and 1972 than between 1972 and 1997, which is a potential source of bias that must be taken into account. One more factor that may influence data consistency is that with the passing of time vernacular use may change and moreover the way people conceptualize, experience and express symptoms may also change (Jorm et al, 2006).

Furthermore, in the paper the diagnostic evaluation according to the Lundby model is described and discussed, including caseness, symptom patterns and impairment ratings. The problems of attrition, bias and inter-
rater reliability over time are analyzed and discussed. Again the use of a multi-source method is concluded to be a strength in the Lundby Study as it may balance attrition, bias and poor inter-rater agreement.

In the paper incidence rate data for neurosis by five year periods from 1947 to 1997 is presented to illustrate that a raised caseness level may increase diagnostic consistency over time. As for representativeness it is concluded that after 50 years of follow-up, the Lundby population has naturally lost its original representativeness. By 1997 it represented an initially rural population that had been exposed to the changes in society between 1947 and 1997 while ageing.

The strengths of the Lundby Study include that the population at the start represented a homogeneous complete rural population including all age-groups from a defined area. Furthermore, the population has been followed for a long time; meaning that many subjects by the last follow-up had passed through their ages at risk for several disorders. As diagnostic and contextual information collected form several sources have underpinned consensus-based diagnoses the validity and reliability of the diagnoses are essentially satisfactory, although re-evaluations were deemed necessary to increase diagnostic consistency in some cases.

The weaknesses in the Lundby Study include that the population followed is rather small, especially for analyzing rare disorders like psychotic disorders. Furthermore, as no new subjects have been incepted after 1957 the study population has been more and more reduced during its follow-up. The reduction of the population and the increasing dispersal of the subjects over the study period have with time made the population less representative for the contemporary Lundby area and the society as a whole. For instance, there were no subjects under the age of 20 left after 1977 and there are no immigrants from outside the Nordic countries in the study population. One more problem is the increasing attrition over time, especially after 1972. Further, the attrition between 1972 and 1997 was larger in subjects aged under 50 in 1972 than in those aged 50 or more, which may have introduced selection bias. In males and females aged under 50 in 1972 the subsequent attrition was 6-8% and 7-13%, respectively.
12. General discussion

12.1 Overall incidence of psychotic disorders in males and females

The first aim of the thesis was to compare the overall male and female first-incidence of psychotic disorders in the Lundby population (paper I).

As opposed to several recent studies (Aleman et al, 2003; McGrath et al, 2004) a non-differing overall male–female incidence of nonaffective psychoses – including schizophrenia – was found. Actually, in the Lundby population the male and female overall incidence rates were not significantly separated in any of the psychotic disorder groups, except for substance-induced psychoses. The reason for the non-differing overall male and female rates of nonaffective psychoses and schizophrenia in the Lundby population was due to the inclusion of old age groups – the higher nonaffective psychoses/schizophrenia rate in males compared to females in the 15-29 year age band was balanced by a more steady rate in females across the life course plus an increasing rate of nonaffective psychoses in females after 65, whereas the rates in males decreased consistently with increasing age. Most studies that have found an increased incidence of schizophrenia in males compared to females have excluded older subjects.

The findings in the Lundby population suggest that the overall risk for functional (nonaffective and/or affective) psychotic disorders in males and females may be rather equal, although males seem to consume their risk earlier in life than females.
12.2 Age-at-onset and incidence by age of psychotic disorders

The second aim of the thesis was to analyze and compare male and female mean and median age-at-onset and incidence by age of psychotic disorders in the Lundby population (paper I).

It was found that the male mean age-at-onset was lower than that for females for ‘any psychotic disorder’, ‘psychotic disorder due to a general medical condition’, ‘nonaffective psychotic disorder’ and ‘schizophrenia’. Moreover, males and females had different incidences by age patterns for ‘any psychotic disorder’, ‘nonaffective psychotic disorder’, ‘schizophrenia’ and ‘other nonaffective psychotic disorder’, with a male dominance among early onset cases and a female dominance among late onset cases.

The male–female differences of age-at-onset and incidence by age – which has been observed in several studies of schizophrenia – have been suggested to be due to: male–female biological dimorphism (Jablensky, 2000), different ageing processes in males and females (Tien, 1991; Orr and Castle, 2003) and a protective effect of oestrogen in females (Häfner et al, 1993; Häfner et al, 1998; Grigoriadis and Seeman, 2002).

It has also been suggested that males and females are differently prone to develop subtypes of schizophrenia, with males more often contracting a neurodevelopmental type of disorder with early onset and females more often a type of disorder related to affective psychosis with late onset (Castle and Murray, 1993; Hultman et al, 1999; Castle and Murray, 1991; Kirov, et al 1996; Howard et al, 1997).

The findings in the Lundby Study adds to the evidence that there is a link between the life course and the tendency to contract psychosis with differences between males and females in their propensities to contract psychotic disorders – at least for nonaffective psychoses – which may indicate underlying biological sex differences such as psychotic disorder-delaying mechanisms in females or different aetiologies of psychosis in males and females. Interestingly enough, the male average age-at-onset of ‘psychotic disorder due to a general medical condition’ was also significantly earlier than in females and the male age-at-onset of ‘affective
psychotic disorder’ tended to be earlier in males too, suggesting a general difference between the sexes.

12.3 Period prevalence of psychotic disorders 1947-1997

The third aim of the thesis was to analyze the proportion of the Lundby population that had a psychotic or bipolar disorder 1947-1997 (paper II).

It was found that the period prevalence 1947-1997 in all subjects was 4.24 % for any psychotic or bipolar disorder, 0.76 % for psychotic disorder due to a general medical condition, 0.59 % for substance-induced psychoses, 2.25 % for nonaffective psychoses and 0.62 % for affective psychoses.

There was a statistically significant difference between males and females for the period prevalence of ‘substance-induced psychosis’ with a higher prevalence in males. For the other disorders the period prevalences did not differ significantly between the sexes. There was no study found in the literature to compare with.

As the Lundby population is a normal unselected homogeneous community population that has been largely unexposed to urbanicity, migration, drug use and big socioeconomic differences, it suggests that psychotic disorder rates may be rather high also in the absence of these factors. Moreover, the 50 year follow-up suggests that the male and female probabilities to develop psychotic disorders at some point during their life courses – in the absence of the mentioned factors – may essentially be equal.

12.4 Lifetime prevalence of psychotic disorders 1997

The fourth aim of the thesis was to analyze the proportion of the Lundby population that was alive in 1997 and had experienced a psychotic or bipolar disorder (paper II).

The lifetime prevalence in 1997 was 2.82% for ‘any psychotic or bipolar disorder’, 1.38% for ‘nonaffective psychotic disorder’, 0.84% for ‘schizophrenia’, 0.54% for ‘psychotic disorder due to a general medical condition’, 0.48% for ‘substance-induced psychotic disorder’ and 0.42% for
‘affective psychotic disorder’. Males and females differed significantly for ‘any psychotic or bipolar disorder’ with a lifetime prevalence of 3.75% and 1.96% in males and females, respectively. For the specific psychotic disorder groups the rates did not differ significantly between the sexes.

Compared with other studies the lifetime prevalences were greater in the Lundby Study for ‘psychotic disorder due to a general medical condition’, ‘substance-induced psychotic disorder’, ‘schizophrenia’ and ‘delusional disorder’, but they were smaller for ‘brief psychotic disorder’, ‘schizophrheniform disorder’ and ‘affective psychotic disorder’. The higher lifetime prevalences of ‘schizophrenia’ and ‘delusional disorder’ found in the Lundby population than in other studies may be due to that several sources of information were included in the Lundby Study, whereas other studies have mostly relied on DIS or CIDI/SCID-I without the supplementary information from other sources.

A possible reason for the the lower rates in the Lundby Study of ‘brief psychotic disorder’ and ‘schizophreniform disorder’ may be the long follow-up period, as this may have allowed diagnostic transitions from remitting psychotic disorders of short duration to chronic psychotic disorders. This would also be consistent with the high lifetime prevalence of ‘schizophrenia’ and ‘delusional disorder’ that was found in the Lundby population. However, comparisons with other studies are difficult to make due to methodological differences between studies, such as different age structures in the samples studied and different methods for case-finding and diagnosis. Moreover, the Lundby lifetime prevalence data in 1997 is biased due to selective attrition.

In addition, the lifetime prevalences of the surviving study population in 1997, which were lower than the corresponding period prevalences 1947-1997, must also be seen in the light of different age structures in the samples on which the analyses of lifetime prevalences 1997 and period prevalences 1947-1997 were based, respectively, besides the biased attrition of cases with psychotic disorders during the follow-up. Thus, one may speculate on that the differing lifetime prevalences between males and females in 1997 for ‘any psychotic disorder’ (which was not seen for the corresponding period prevalences of ‘any psychotic disorder’ 1947-1997) may be due to differential selective attrition in the sexes during the follow-up (e.g. a higher survival rate of psychotic males than females and/or fewer
psychotic males than females being lost to follow-up in 1997). However, it may also reflect a period effect with increasing incidence rates of psychosis in males and/or with decreasing rates in females during the latter part of the study. Actually, in an analysis of the period specific incidences in 1947-1972 and 1972-1997, respectively, it was found that the age adjusted incidence of ‘any psychotic disorder’ had increased significantly in males (from 8 to 16 cases per 10 000 person years at risk), whereas it had decreased minimally in females (from 8 to 7 cases per 10 000 person years at risk) (Bogren et al, 2007).

Interestingly enough, in a further subanalysis (not published data) it was found that the male increase of the general psychosis incidence from 1947-1972 to 1972-1997 mirrored an increase in all psychotic disorder subcategories – psychotic disorder due to a general medical condition (from 2.2 to 3.8 cases per 10 000 person years at risk), substance-induced psychotic disorder (from 1.9 to 4.3 cases per 10 000 years at risk) and functional (nonaffective and/or affective) psychotic disorder ( from 4.4 to 7.5 cases per 10 000 years at risk) – whereas in females behind the slight decrease of any psychotic disorder from 1947-1972 to 1972-1997 there was a clear trend downwards for functional psychotic disorder (from 7.5 to 3.8 per 10 000 years at risk) although this had been obscured by increasing rates of psychotic disorder due to a general medical condition (from 0 to 3.0 cases per 10 000 years at risk) and substance-induced psychotic disorder (from 0.2 to 0.7 cases per 10 000 years at risk). Moreover, the female decrease of the incidence of functional psychotic disorder from 1947-1972 to 1972-1997 was almost reaching statistical significance – the 95% confidence interval of the ratio of the female functional psychotic disorder rates in 1947-1972 and 1972-1997, respectively, was almost separated from zero (0.96-4.04).

It is tempting to interpret these findings as suggestive of a period effect reflecting increasing and decreasing incidence of functional psychotic disorder in males and females, respectively. However, we must be tempered by the small size of the sample which may implicate that the findings are due to chance. Moreover, there is the question of the external validity of the findings – if a ‘period effect’ was actually present in the Lundby population we could not with certainty generalize it to the community at large; just to the pre-war generations in the rural part of southern Sweden.
12.5 Predictors of psychosis related to personality and behaviour

The fifth aim of the thesis was to analyze whether premorbid traits related to personality and behaviour were associated with an increased risk to develop psychosis in the Lundby population (paper III).

At the interviews in 1947, 1957 and 1972 psychiatrists had rated several observable behaviours such as tension, gloominess and sensitivity in the study subjects. Furthermore, the study subjects had been asked to rate themselves on several subjective dispositions such as if they were habitually nervously disposed, sensitive or tired. Following the interviews, the psychiatrists had integrated the personality related information to assess schizoidi and abnormal personality according to Bleuler and Schneider, respectively, in due cases. Based on the rated behaviours, self-evaluated traits and the assessments of schizoidi and abnormal personality, dichotomous predictors meant to be related to premorbid personality were constructed (see paper III, Table 3). In multivariate and/or univariate Cox regression models it was found that some of the constructed predictors were significantly associated with a subsequent first onset of functional (nonaffective and/or affective) psychosis and/or ‘schizophrenia’.

The constructed predictors associated with later functional psychosis were: nervous-tense, down-semidepressed, blunt-deteriorated, paranoid-schizotypal, sensitive-frail, easily hurt and tired-distracted. When schizophrenia was analyzed separately the predictors associated with a later onset of disorder were: nervous-tense and abnormal-antisocial. A person rated as nervous-tense may be conceptualized as a person with DSM-IV cluster C traits and a high degree of neuroticism. A person rated as down-semidepressed and/or blunt-deteriorated have more or less of the traits belonging to the negative symptom dimension often seen in schizophrenia. A person rated as paranoid-schizotypal displays DSM-IV cluster-A traits. A person rated as sensitive-frail and easily hurt has displayed interpersonal sensitivity/frailness in the interview situation and rated him/herself as an easily hurt person, respectively. A person rated as abnormal-antisocial has a severe personality disorder with conspicuous traits such as: suspiciousness, fanaticism, indolence, emotional lability, aggressiveness and explosiveness.
The traits, associated with an increased risk of psychosis in the Lundby population, accorded with findings from several earlier studies which have used different methodologies to study premorbid personality and psychosis – most often however in the form of schizophrenia (see paper III, Table 1). Thus, some prospective community studies have shown that impaired emotional reactivity, impaired attention and anxiety and other traits of neuroticism are associated with psychosis vulnerability (Bearden et al, 2000; Krabbendam et al, 2002). Moreover, high risk schizophrenia studies have shown that anxiety, tension, social sensitivity, suspiciousness, schizotypal traits and concentration deficits may be associated with psychosis vulnerability (Amminger et al, 1999; Carter et al, 1999; Ekstrøm et al, 2006; Johnstone et al, 2005; Niemi et al, 2005; Parnas et al, 1982; Schiffman et al, 2004). Retrospective studies have also shown that neurotic, schizoid, paranoid and schizotypal traits are associated with psychosis vulnerability (Dalkin et al, 1994; Rodríguez Solano and González De Chávez, 2000).

The predictors blunt-deteriorated, tired-distracted, down-semidepressed, sensitive-frail and easily hurt – although they were associated with the broader functional psychosis outcome in the Lundby population – seem to echo some aspects of schizotaxia, which – according to Meehl – is associated with cognitive slippage, ambivalence, hypohedonia and interpersonal aversiveness (Meehl, 1989). Additionally, nervous-tense resembles Meehl’s concept of anxiety readiness. Moreover, the findings in the Lundby population lend support to the idea that there may be premorbid personality traits expressed by many individuals in the general population – the distributions of the constructed predictors in the Lundby cohorts ranged from 3.1% to 54.2%, see paper III, Table 4 – that are associated with an increased (but far from definite) risk to develop functional psychosis.

The main finding in paper III was related to the study of the broad psychosis group – functional psychoses. But a separate analysis was also carried out for schizophrenia. However, as the study sample may have been to small to study a rare outcome like schizophrenia, statistical power may have been insufficient to study schizophrenia separately (there were only 14 and 10 cases of schizophrenia in the 1947 and 1957 cohorts, respectively). In analyzing the broad group of functional psychoses we may simply have achieved a larger sample size at the expense of specificity of interpretation. But on the other hand, as psychotic symptoms may be part of several DSM-
IV diagnoses and as the traditional diagnostic categories of the psychoses have been found to have limited value in predicting social and clinical outcomes (van Os et al, 1999), and as many of the identified risk factors for psychosis may not be related to specific diagnostic categories but rather to be shared between disorders with psychotic symptoms (Cannon et al, 1997; van Os et al, 1998; Kelly and Murray, 2000; Cannon et al, 2003; Broome et al, 2005; Weiser et al, 2005; Lichtenstein et al, 2009; Owen and Cradock, 2009; The International Schizophrenia Consortium, 2009), it may be questioned whether the traditional diagnostic constructs (like schizophrenia) really delineate disease entities with clear boundaries. And from such a view point it may actually be accurate to study a broad group of psychoses such as the functional psychoses in paper III in relation to presumed risk factors.

Thus, the clusters of traits found in paper III to be associated with an increased risk of functional (nonaffective and/or affective) psychosis may then represent vulnerability markers or behavioural endophenotypes contributing to an increased risk to develop psychosis. However, it must be emphasized that it is an important limitation that the predictors were constructed concepts based upon a heterogeneous collection of clinically and self rated data which may question their accuracy and reliability. Moreover, the predictors – which were meant to reflect personality traits – in effect probably reflected a blend of personality traits in the proper sense and varying affective symptoms which may be unrelated to personality traits per se.

There is also the question whether the identified predictors may be specifically related to psychosis or rather to mental disorder in general. In a separate Cox regression analysis of the 1947 cohort (not published data) it was found that paranoid-schizotypal was significantly associated with neurosis in males, whereas abnormal-antisocial reached significance in females. In an analysis of the 1957 cohort (not published data) nervous-tense was associated with neurosis in males but no other predictors reached significance for neurosis. Moreover, previously in the Lundby Study it has been shown that the predictors nervous-tense, abnormal-antisocial, tired-distracted and easily hurt are associated with an increased risk to develop depression in males or females (Mattisson et al, 2009).
Thus, the predictors found to be associated with neurosis and/or depression in males or females are: nervous-tense (associated with neurosis in males and depression in both sexes), paranoid-schizotypal (associated with neurosis in males), abnormal-antisocial (associated with neurosis and depression in females), easily hurt (associated with depression in females) and tired-distracted (associated with depression in females). The predictors that were only associated with psychosis were: down-semidepressed, blunt-deteriorated and sensitive-frail. Again the core traits of schizotaxia according to Meehl seem to be echoed – anhedonia, cognitive slippage, ambivalence and interpersonal aversiveness.

Meehl coined the term schizotaxia for the hypothesis that the genetic vulnerability to schizophrenia manifests as a neural integrative defect (Meehl, 1962). Later research on first-degree relatives of schizophrenic patients have suggested a reformulation of Meehl’s schizotaxia in that it may not always, or mostly, develop into either schizophrenia or schizotypy, as Meehl initially thought, but often into a related spectrum state, i.e. a non-psychotic, non-schizotypal and non-prodromal syndrome with negative symptoms, neuropsychological deficits and psychosocial dysfunction. In this reformulated version, schizotaxia, alternatively viewed as schizotypal personality disorder minus the positive symptoms (i.e. negative schizotypy), has been incorporated into a multifactorial polygenic and environmental vulnerability-stress hypothesis of schizophrenia aetiology (Faraone et al, 2001; Tsuang et al, 2002).

12.6 Methodological issues

The sixth aim of the thesis – to describe general methodological difficulties in the Lundby Study – resulted in the recognition of several issues concerning population representativity, sampling, attrition, case finding method and diagnostic ascertainment.

12.6.1 Population representativity

At the beginning of the Lundby Study the Lundby population was a rural population which was considered to be representative of other such populations. Naturally, in 1997 the survivors in the study were no longer representative of the current demographic structure of the Lundby area as new generations had been added to it. Additionally, in 1997 the study
population was aged, largely dispersed outside the Lundby area and also reduced due to attrition. As for the generalizability of the findings from the 50-year follow-up it is important to keep in mind that the results are primarily valid for those generations that lived in the rural part of the south of Sweden in the 40’s and 50’s. However, the findings may also be generalized more broadly as the change in society that the study population was subject to during the 50-year follow-up in much followed the same pattern as in many other European countries after the Second World War.

Nevertheless, the psychotic disorder rates in the study population do not necessarily apply to the people living in the Lundby area today or to the current general Swedish population (although the incidence rate of nonaffective psychoses in the Lundby population 1947-1997 was consistent with the rate in a register study of the total Swedish population 1997-1999; see paper I). Rather the rates of psychotic disorders in the study population are more likely to represent what may be seen as ‘base rates’, since the study population has been largely unaffected by several factors that are known to increase the risk for nonaffective psychoses in contemporary communities – urbanicity, migration and drug use. The Lundby rates may thus be suggestive of what you may expect in a ‘pristine’ community population – although rates in such populations may of course vary due to other factors than urbanicity, migration and drug use. Moreover, the incidence rates of psychotic disorders may of course have varied in the Lundby population during the follow-up, and the incidence rates for the 50 year period presented in the thesis are thus averages of possible fluctuations.

The lesson learned is that the representativity of a sample in a longitudinal study changes with time.

12.6.2 Sampling

A strength in the Lundby Study is that it is based on the investigation of a complete population sample according to the census method, which implies a low degree of sampling bias; although this method may systematically underrate incidence and overrate age-at-onset of disorders. Nevertheless, a problem regarding the study of the rates of the psychotic disorders was poor precision of the estimates due to the small population sample. Actually, when the study sample was stratified by sex and age the number of cases in
the strata became so small that it was not possible to analyze time trends of the incidence of psychotic disorders in the population – e.g. to compare the incidence 1947-1972 and 1972-1997. However, the small study population does not invalidate the findings of psychotic disorder rates 1947-1997, although it urges us to interpret the data with caution.

The Lundby Study underscores the importance of investigating sufficiently large samples to be able to estimate psychotic disorder rates with some precision.

12.6.3 Attrition

Due to the cumulative attrition associated with the follow-ups, selection bias was probably introduced with time. Attrition due to refusal of subjects to continue in the study and complete loss to follow-up of subjects, was small but it increased between 1972 and 1997; especially in subjects who were young or middle aged in 1972 (see paper IV, Table 1). Attrition due to deaths was large; especially between 1972 and 1997 (by 1997 approximately 50% of the study subjects had died, see Table 1). Fortunately – thanks to the case finding method which included several data sources – it was feasible to obtain a follow-up of the subjects who had died between study waves, almost eliminating mortality associated attrition (out of the 1766 subjects who died between follow-ups 1732 could be assessed through outside sources, see Table 1). Moreover, the supplementary data sources also reduced the attrition of the living subjects that were lost to follow-up by interview (see Table 1). Of course, data from outside sources such as registers, case files and key informants do not give the same kind of information as data obtained from face to face interviews. But for the identification of psychotic disorder cases they may be adequate.

The lesson learned is that in epidemiological longitudinal community studies of psychotic disorders, to counteract attrition, it is important to use multiple data sources including both interviews and treatment data. Attrition, due to migration and death, may also be further reduced if the time interval between follow-ups is not too long.
12.6.4 Case finding method and diagnostic ascertainment

Of great importance in psychiatric epidemiology is that the case finding method and the diagnostic ascertainment allow interpretation and comparison of studies. Moreover, in longitudinal studies, data from different follow-ups must be comparable. Thus, it is vital that the methods used have accuracy and precision. However, despite the development of structured and semistructured interview schedules (DIS, CIDI, PSE, SCAN) and the development of systems for diagnosis and classification that has increased reliability (DSM-IV, ICD-10), a perfect definition of a psychiatric case and a gold standard for diagnosis do not exist.

As opposed to the large psychiatric field surveys from the 80’s and onward – which have relied on lay interviewers administrating structured schedules underpinning diagnostic algorithms – the Lundby Study has been pinned on a case finding approach more similar to the traditional clinical method with psychiatrists collecting and combining semistructured and free data from interviews and data from multiple outside sources to reach consensus decisions on psychiatric diagnoses. Admittedly the ‘clinical’ Lundby approach has its drawbacks as it is not standardized, but also the ‘structured’ methods have their limitations when applied to the study of psychotic disorders in community populations as they have been shown to under- and overestimate psychotic and bipolar disorders, respectively, which have raised questions about their validity. With data from several sources to synthesize a best estimate consensus approach was deemed to be the most suitable method for diagnostic ascertainment in the Lundby Study.

A factor in the Lundby Study, which may have affected the case finding and diagnostic ascertainment, was the very long time period between the follow-ups in 1972 and 1997, which probably introduced a considerable amount of recall bias in the interviews in 1997. The sharp decline in the first incidence rate of neurosis around the years 1972-1977 (see paper I, Figure 3) may – at least in part – be explained by recall bias. Although other factors capable of explaining the fall of the neurosis incidence after 1972 may also exist: selection bias (for instance due to that the subjects who dropped out in 1997 represented the first incident cases of neurosis after 1972), observation bias (due to the availability of fewer outside data sources between 1972-1997 than 1947-1972, and a relatively high proportion of deaths between 1972 and 1997 making follow-up in a high proportion of subjects in 1997 totally dependent on outside data).
Nevertheless, for the identification and diagnosis of psychotic disorders in the Lundby Study, recall and observation biases were probably – for most cases – counterbalanced by the multiple sources of treatment data available, as the majority of the cases who experienced a psychotic disorder probably came to medical attention. Treatment data were available throughout the Lundby study; and found in 103 of the 108 psychosis cases that were identified.

In a longitudinal study with several follow-ups during 50 years including a relay of field workers inter-rater reliability will be a problem. To assess the diagnostic agreement over time the two main field-workers from the 1997 follow-up (Cecilia Mattisson and Mats Bogren) blindly diagnosed 200 randomly selected subjects according to the Lundby classification of mental illness using the data from 1947-1972. Thereafter comparisons were made with the diagnoses that had been assessed by the previous field team in 1957 and in 1972 (Olle Hagnell and Leif Öjesjö). Calculations of kappa values indicated substantial to moderate agreement for the general neurosis and organic brain syndrome categories, but the result was inconclusive for psychosis due to few cases (see paper I, Table 2). Nevertheless, the kappa calculations indicated a general problem with inter-rater agreement. To solve this for the study of psychotic disorders all cases with a suspected psychotic disorder 1947-1972 according to the Lundby classification (‘schizophrenia’, ‘other psychoses’ and ‘depression plus other psychiatric symptoms’) were rediagnosed according to the DSM-IV. This may of course have introduced some misclassification since the early data naturally had not been collected with the structure of the DSM-IV in mind. Nevertheless, a reevaluation was feasible thanks to the availability of the original data.

The lesson learned is that multiple data sources and short intervals between follow-ups may improve the case finding and diagnostic ascertainment in longitudinal community studies.
13. Conclusions

In the 50 year follow-up of the Lundby population the sexes did not differ significantly in their overall risks to contract psychotic disorders, except for substance-induce psychoses. Similar overall incidences were found, as well as period prevalences of psychotic disorder due to a general medical condition, nonaffective psychoses and affective psychoses. However, the sexes did differ in their average age-at-onsets for psychotic disorder due to a general medical condition and nonaffective psychoses (and tended to differ for affective psychoses), with earlier onsets in males than females. They also differed in their incidences by age patterns for nonaffective psychoses, with an early and sharp peak in males followed by a consistent decline with increasing age versus a later and lower peak in females followed by a slight decline and a later peak in old age. Thus, the results indicate probable differences between the sexes in their vulnerabilities to contract psychosis across the life course. Such differences may be due to disorder delaying mechanisms in females or different aetiologies affecting the sexes in different life phases.

In an analysis of putative associations between premorbid behaviours and traits aggregated in dichotomous variables and functional (nonaffective and/or affective) psychosis it was found that anxiety proneness, affective/cognitive blunting, poor concentration, personality cluster-A like traits and interpersonal sensitivity all were associated with an increased risk to develop psychosis. The findings support the hypothesis that certain premorbid trait characteristics may be associated with psychosis vulnerability. There were some similarities between the findings and the concept of schizotaxia.

There were many methodological difficulties associated with the 50 year follow-up of the Lundby population and the analyses of psychotic disorders. Thus, the results are primarily representative for the rural pre-war generations in the south of Sweden rather than the current general population – although the estimated rates may be suggestive of rates
‘inherent’ in a normal population as the Lundby population has been minimally affected by urbanicity, migration and drugs. Nevertheless, rates ought to be seen as minimum estimates as the census method may underestimate rates to a certain degree and there may also have been some attrition – although this was probably largely balanced out by the use of many data sources. Moreover, rates must also be judged in view of the small sample; thus, primarily allowing interpretation of the broad diagnostic categories (e.g. nonaffective psychoses). For the analyses of the predictors it was a limitation that the predictors were concepts constructed within the Lundby Study and not based on ‘validated’ scales.

manodepressiv sjukdom). Man delar ofta in de funktionella psykoserna i tillstånd som har ett klart samband med påtagliga variationer i sinnesstämningen (t.ex. depression och mani) och de som uppträder utan någon tydlig påverkan på sinnesstämningen. Den första gruppen kallas för affektiva psykoser och den andra för icke-affektiva psykoser.

Förekomsten av de olika psykostillstånden är bristfälligt studerat i den allmänna befolkningen. Det finns visserligen många studier som undersökt hur vanligt schizofreni är och ganska många som undersökt förkomsten av bipolärt syndrom, men oftast har studierna byggt på sjukvårdsnäck och bara undersökt vissa åldersgrupper, t.ex. 18-64 år. Icke desto mindre har man gjort intressanta fynd; t.ex. att schizofreni brukar drabba män tidigare i livet än kvinnor (även om risken att insjukna i schizofreni för båda könen är högst i ungdomsåren) medan åldern vid insjuknande i bipolärt syndrom verkar vara lika hos könen. Studier på senare år har talat för att män har en högre risk än kvinnor att drabbar av schizofreni under livet, medan risken för bipolärt syndrom verkar vara lika hos män och kvinnor.


Studier av kopplingarna mellan personlighetsdrag, som fanns innan en psykossjukdom debuterade, och själva sjukdomen är svåra att genomföra.
Personligheten hos dem som blivit sjuka har oftast inte studerats före insjuknandet. Men i Lundbystudien har man, allt sedan den startade, gjort bedömningar av personlighetsdrag och noterat förekomsten av t.ex. nervositet, misstänksamhet, blödighet, känslomässig avtrubbning och irritabilitet. Man har också under uppföljningen kunnat se vem som drabbats av psykostillstånd.

I den aktuella avhandlingen redovisas studier av frekvensen av olika psykosstörningar i den studerade befolkningen samt en studie av sambandet mellan personlighetsdrag och senare utveckling av psykos.

Vad gäller frekvenserna av psykosinsjuknanden i befolkningen visade det sig att män och kvinnor insjuknat ungefär lika ofta i de olika psykoskategorier som analyserades med undantag för substansrelaterade tillstånd som var vanligare hos männen. Under hela perioden diagnosticerades någon sorts psykos hos 4.24% av studiepersonerna. Den enskilt vanligaste diagnosen var schizofreni, vilket diagnosticerades hos 1.43%. Till skillnad från många studier på senare år skilde sig inte männen och kvinnornas insjuknande i schizofreni nämnvärt åt.

Vad gäller åldern vid sjukdomsdebuten insjuknade männen i genomsnitt klart tidigare än kvinnorna i de flesta olika psykoskategorier (dock inte substansrelaterade psykoser). Mönstret för insjuknande över livsloppet skilde sig mellan könen genom att männen tenderade att insjukna i icke-affektiva psykoser i unga år, varefter risken avtog med ökad ålder, medan kvinnorna började insjukna i icke-affektiva psykoser något senare än männen. Å andra sidan behöll kvinnorna en högre insjuknandefrekvens med ökad ålder och fick dessutom en ökad risk efter 65 års ålder. Skillnaderna mellan män och kvinnor skulle kunna avspeglas att det finns faktorer som skyddar mot psykos hos kvinnor som sedan faller bort, t.ex. östrogen. Det skulle också kunna bero på att åldrandet hos män och kvinnor inte följer samma mönster eller att män och kvinnor utsetts för olika riskfaktorer under livsloppet.

Vad gäller studiet av personlighet och psykos studerades inte män och kvinnor separat, men för den samlade gruppen var vissa konstellationer av drag associerade med en riskökning att drabbas av funktionell (icke-affektiv och/eller affektiv) psykos. Nervositet-spänningsbenägenhet, lustlöshet-känslomässig avtrubbning, schizotypala drag, mellanmänsklig känslighet.
och en tendens att vara uttröttbar och spänningsbenägen föll ut som riskfaktorer. Fynden stöder hypotesen att det finns personlighetsdrag som är kopplade till en ökad psykosrisk. Det var också intressant att de personlighetsdrag som ökade risken för psykos liknar de drag som andra forskare föreslagit skulle kunna vara kopplade till den genetiska risken för psykos.

Fynden är intressanta men skall tolkas med försiktighet. Metodproblemen handlar om representativitet, bortfall, diagnostisk säkerhet och slumpmässig variation. En stor del av avhandligen diskuterar därför hur resultaten skall tolkas i ljuset av de metodproblem som finns. Resultaten är inte generaliserbara till befolkningen i allmänhet men är ändå intressanta eftersom Lundbybefolkningen är så grundligt undersökt och följd så länge; och att den faktiskt exponerats för de samhällsförändringar som typiskt förknippas med efterkrigstiden. Dessutom kan fynden av frekvenser för psykos ge en uppfattning om hur vanliga sådana tillstånd kan vara i en befolkning som inte i nämnvärda utsträckning utsatts för ett antal av de riskfaktorer som anses öka risken för psykos i befolkningen; nämligen storstadsliv, migration och drogmissbruk. Vad gäller fynden av personlighetsdrag som var förknippade med en ökad risk för psykos i Lundbybefolkningen bevisar fynden inget, men de ger stöd för hypoteser som redan finns och skulle därmed kunna stimulera till ny forskning inom området.
15. Acknowledgement

I would like to express my gratitude to everybody who has contributed to the thesis. Thank You:

All people and families in the Lundby population for your openness and friendliness; we are obliged to you for taking part in the Lundby Study.

Associate professor Per Nettelbladt – my supervisor – for scientific guidance, enthusiasm and unwavering support.

Associate professor Leif Öjesjö – my co-supervisor – for practical advice and cheerful encouragement.

Professor emeritus Olle Hagnell for sharing your unique experience of human nature.

Dr Cecilia Mattisson for friendship and mind nurturing discussions.

Research assistant Lena Otterbeck for your kind introduction into the Lundby world.

Research assistant Anders Odensten for tireless keeping of the database and provision of data.

Dr Erik Hofvendahl and Dr Per Toråker for excellent field work.

Research assistant Eva Spännare for orderly sorting of the old file cabinets.

Professor Lil Träskman-Bendz for encouragement and support of the Lundby Study.

Professor Povl Munk-Jørgensen for inspiration and believing in the Lundby Study.
Professor Kristian Tambs for great help in matters of scientific methodology.

Professor Dinesh Bhugra for support of the Lundby Study.

Statisticians Vibeke Horstmann, Per-Erik Isberg and Agneta Sternerup for invaluable statistical assistance.

Hanna-Britt Franzén and Ulla Persson for kind help in practical matters.

Gunnel Fogelström for excellent help with English grammar, idioms and spelling.

Dr Lars Nordström – head of the Habilitation-rehabilitation unit at the psychiatric clinic Lund – for giving me the time to complete the thesis.

Dr Björn Nylander for sharing outstanding knowledge and clinical wisdom.

My family for everything.

To put things in perspective it ought also to be mentioned that there had never been a 50 year follow-up of the Lundby population had not the late Professor Erik Essen-Möller – a true pioneer and a visionary in psychiatric epidemiology – launched the Lundby Study in 1947. A special tribute is also due to Professor emeritus Olle Hagnell, who carried Essen-Möller’s vision forward.

The thesis was supported by grants from:

the Swedish Council for Planning and Coordination of Research, the Swedish Council for Social Research, the Swedish Research Council, the Swedish Medical Research Council, the Medical Faculty of Lund University, the Provincial Government of Skåne, the Gohrton Foundation, the OM Persson Foundation, the Pfannenstill Foundation, the Söderström-Königska Foundation, the Professor Bror Gadelius Foundation, the Ellen and Henrik Sjöbring Foundation and the Division of Psychiatry, Lund University.
16. References


Burke KC, Burke JD, Regier DA, Rae DS (1990) Age at onset of selected mental disorders in five community populations. Arch Gen Psychiatry 47: 511-8.


Robins LN, Regier DA (1991) Psychiatric disorders in America, the epidemiological catchment area study. New York, the Free Press.


