Chapter

Contribution of genetic epidemiology to our understanding of psychiatric disorders

Kathleen Ries Merikangas and Anibal Cravchik

Introduction

This chapter will: (1) provide a summary of the background disciplines and approaches to understanding the role of genetic factors in mental disorders; (2) review the current knowledge in the genetic epidemiology of mental disorders; and (3) summarize the role of epidemiology in the current generation of genome-wide association studies of mental disorders.

Genetic epidemiology

The pioneering work of Böök, Sjögren, Angst, Perris, and others in Europe and Kallman, Heston, Rosenthal, Wender, and Kety in the United States firmly established the important role of genetic susceptibility factors in psychiatric disorders. Heston's [1] original finding that adult offspring of hospitalized schizophrenic mothers had significantly higher rates of schizophrenia than offspring of parents with no mental illness was confirmed and extended by Kety, Rosenthal, and Wender's [2, 3] studies in a much larger sample of adopted away offspring of schizophrenics in Denmark. These studies demonstrated clearly that the presence of schizophrenia in birth parents, independent of the rearing environment, significantly increases offspring's risk for the development of the disease [4, 5]. During the last decades of the twentieth century, the study of genes in psychiatric disorders expanded beyond hospital settings to outpatient treatment settings, particularly in the United States [6]. With the introduction of epidemiology to the study of psychiatry, systematic control groups were included in family studies and methods for incorporating population base rates and risk assessment were developed.

The field of genetic epidemiology is defined as the study of the distribution of, and risk factors for

diseases and genetic and environmental causes of familial resemblance. Genetic epidemiology focuses on how genetic factors and their interactions with other risk factors increase vulnerability to, or protection against, disease [7]. Genetic epidemiology employs traditional epidemiological study designs to explain aggregation in groups as closely related as twins or as loosely related as migrant cohorts. Epidemiology has developed sophisticated designs and analytic methods for identifying disease risk factors. With increasing progress in gene identification, these methods have been extended to include both genetic and environmental factors [8, 9]. In general, study designs in genetic epidemiology either control for genetic background while letting the environment vary (e.g. migrant studies, half siblings, separated twins) or control for the environment while allowing variance in the genetic background (e.g. siblings, twins, adoptees nonbiological siblings). Investigations in genetic epidemiology are typically based on a combination of study designs including family, twin, and adoption studies.

Family studies

Familial aggregation is generally the first source of evidence that genetic factors may play a role in a disorder. The most common indicator of familial aggregation is the relative risk ratio, computed as the rate of a disorder in families of affected persons divided by the corresponding rate in families of controls. The patterns of genetic factors underlying a disorder can be inferred from the extent to which patterns of familial resemblance adhere to the expectations of Mendelian laws of inheritance. The degree of genetic relatedness among relatives is based on the proportion of shared genes between a particular

1

Principles of Psychiatric Genetics, eds John I. Nurnberger, Jr. and Wade H. Berrettini. Published by Cambridge University Press. © Cambridge University Press 2012.

Chapter 1: Contribution of genetic epidemiology

relative and an index family member or proband. First-degree relatives share 50% of their genes in common; second-degree relatives share 25% of their genes in common, and third-degree relatives share 12.5% of their genes in common. If familial resemblance is wholly attributable to genes, there should be a 50% decrease in disease risk with each successive increase in degree of relatedness, from first to second to third, and so forth. This information can be used to derive estimates of familial recurrence risk within and across generations as a function of population prevalence (λ) [10]. Whereas λ tends to exceed 20 for most autosomal dominant diseases, values of λ derived from family studies of many complex disorders tend to range from 2 to 5. Diseases with strong genetic contributions tend to be characterized by 50% decrease in risk across successive generations. Decrease in risk according to the degree of genetic relatedness can also be examined to detect interactions between several loci. If the risk to secondand third-degree relatives decreases by more than 50% this implies that more than a single locus must contribute to disease risk and that no single locus can largely predominate.

The major advantage of studying diseases within families is that disease manifestations are more likely to result within families than they are across families from the same underlying etiological factors. Family studies are therefore more effective than between family designs in examining the validity of diagnostic categories because they more accurately assess the specificity of transmission of symptom patterns and disorders. Data from family studies can also provide evidence regarding etiological or phenotypic heterogeneity. Phenotypic heterogeneity is suggested by variable expressivity of symptoms of the same underlying risk factors, whereas etiological heterogeneity is demonstrated by common manifestations of expression of different etiological factors between families. Moreover, the family study method permits assessment of associations between disorders by evaluating specific patterns of co-segregation of two or more disorders within families [11].

Twin studies

Twin studies that compare concordance rates for monozygotic twins (who share the same genotype) with those of dizygotic twins (who share an average of 50% of their genes) provide estimates of the degree to which genetic factors contribute to the etiology of a disease phenotype. A crude estimate of the genetic contribution to risk for a disorder is calculated by doubling the difference between the concordance rates for monozygous and dizygous twin pairs. Modern genetic studies employ path analytic models to estimate the proportion of variance attributable to additive genes, common environment, and unique environment. There are several other applications of the twin study design that may inform our understanding of the roles of genetic and environmental risk factors for disease. First, twin studies provide information on the genetic and environmental sources of sex differences in a disease. Second, environmental exposures may be identified through comparison of discordant monozygotic twins. Third, twin studies can be used to identify the genetic mode of transmission of a disease by inspection of the degree of adherence of the difference in risk between monozygotic and dizygotic twins to the Mendelian ratio of 50%. Fourth, twin studies may contribute to enhancing the validity of a disease through inspection of the components of the phenotypes that are most heritable. The twin family design is one of the most powerful study designs in genetic epidemiology because it yields estimates of heritability but also permits evaluation of multigenerational patterns of expression of genetic and environmental risk factors. Several recent updates of findings of twin studies of psychiatric disorders are available [12, 13].

Adoption studies

Adoption studies have been the major source of evidence regarding the joint contribution of genetic and environmental factors to disease etiology. Adoption studies either compare the similarity between an adoptee and his or her biological versus adoptive relatives, or the similarity between biological relatives of affected adoptees with those of unaffected, or control adoptees. The latter approach is more powerful because it eliminates the potentially confounding effect of environmental factors. Similar to the familial recurrence risk, the genetic contribution in adoption studies is estimated by comparing the risk of disease to biological versus adoptive relatives, or the risk of disease in biological relatives of affected versus control adoptees. These estimates of risk are often adjusted for sex, age, ethnicity, and other factors that

may confound the links between adoption status and an index disease.

With the recent trends towards selective adoption and the diminishing frequency of adoptions in the United States, adoption studies are becoming less feasible methods for identifying genetic and environmental sources of disease etiology [14]. However, the increased rate of reconstituted families (families comprised of both siblings and half siblings) may offer a new way to evaluate the role of genetic factors in the transmission of complex disorders. Genetic models predict that half siblings should have a 50% reduction in disease risk compared to that of full siblings. Deviations from this risk provide evidence for either polygenic transmission, gene–environment interaction, or other complex modes of transmission.

Migration studies

Migrant studies are perhaps the most powerful study design to identify environmental and cultural risk factors. When used to study Asian immigrants to the United States, this study design demonstrated the significant contribution of the environment to the development of many forms of cancer and heart disease [15]. One of the earliest controlled migrant studies evaluated rates of psychosis among Norwegian immigrants to Minnesota compared to native Minnesotans and native Norwegians [16]. A higher rate of psychosis was found among the immigrants compared to both the native Minnesotans and Norwegians and was attributed to increased susceptibility to psychosis among the migrants who left Norway. It was found that migration selection bias was the major explanatory factor, rather than environmental exposure in the new culture. The application of migration studies to the identification of environmental factors is only valid if potential bias attributed to selection is considered. Selection bias has been tested through comparisons of factors that may influence a particular disease of interest in a migrant sample and a similar sample that did not migrate.

Genetic epidemiology of psychiatric disorders

The wealth of data from family, twin, and adoption studies of the major psychiatric disorders exceeds that of all other chronic human diseases. The increased

Chapter 1: Contribution of genetic epidemiology

 Table 1.1 Risk ratios and heritability estimates for major mental disorders.

Disorder	Risk ratios	Heritability estimates
<i>Mood disorders</i> Bipolar disorder Major depression	7–10 2–3	60–70 28–40
<i>Anxiety</i> All Panic disorder	4–6 3–8	30–40 50–60
Schizophrenia	8–10	80-84
Substance dependence	4–8	30–50

recognition of the role of biological and genetic vulnerability factors for psychiatric disorders has led to research with increasing methodological sophistication over the course of the second half of the twentieth century. There are numerous comprehensive reviews of genetic research on specific disorders of interest as well as on psychiatric genetics in general [12, 17–20].

Table 1.1 presents a summary of the relative risks (i.e. proportion affected among first-degree relatives of affected probands versus those of relatives of controls) derived from controlled family studies of selected psychiatric disorders. The risk ratios comparing the proportion of affected relatives of cases versus controls are greatest for autism, bipolar disorder, and schizophrenia; intermediate for substance dependence and subtypes of anxiety, particularly panic; and lowest for major depression. The estimates of heritability (i.e. the proportion of variance attributable to genetic factors) are derived from twin studies, which compare rates of disorders in monozygotic and dizygotic twins. These findings reinforce the notion that genes play a major role in the extent to which mental disorders run in families. The heritability estimates for specific disorders shown in Table 1.1 are parallel to the risk ratios derived from family studies. Furthermore, adoption and half-sibling studies also support a genetic basis for the observed familial aggregation.

Schizophrenia

More is known about the genetic basis of schizophrenia than perhaps any other psychiatric disorder, with genetically informative studies stemming from

Chapter 1: Contribution of genetic epidemiology

early in the last century. There are numerous reviews of this extensive body of research [21-24]. Despite wide differences in methods, samples, and geographic locations, controlled family studies yield a remarkably similar average relative risk of 8.9 to first-degree relatives. The four-fold greater proband-wise concordance rate of schizophrenia in monozygotic compared to dizygotic twins, found in 12 studies to date, demonstrates the role of genetic factors in the familial aggregation of schizophrenia. The average heritability in liability to schizophrenia across 12 studies is 0.81 [25]. Similarly, adoption studies using traditional paradigms and modern diagnostic criteria (if available) demonstrates that the average risk to first-degree relatives was 15.5% compared to 3.6% for controls, yielding a relative risk of 4.3.

Despite evidence regarding the importance of genetic risk factors for schizophrenia, the lack of expected Mendelian risk ratios in the difference in risk of schizophrenia as a function of genetic similarity suggests that schizophrenia is a genetically complex disorder [10]. Recent reviews of the genetic epidemiology of schizophrenia also converge in demonstrating the multifactorial etiology of this condition [25–29]. The largest and most recent cross-fostering study of schizophrenia showed that adoptive family environment was associated with schizophrenia spectrum disorders among genetically vulnerable individuals [30], implying the contributions of nonspecific environmental factors (i.e. multiple factors that may affect brain development) to schizophrenia's etiology.

Another important clue about potential environmental risk factors is the increased risk for the development of schizophrenia among immigrants in several different countries including East African immigrants to Sweden [31], Surinamese immigrants to the Netherlands [32], Afro-Caribbean immigrants to the UK [33], Finnish immigrants to Sweden [34], and European immigrants to Canada [35]. Although selective migration may be one explanation, there is converging evidence that socially disrupted environments may trigger the onset of schizophrenia in susceptible individuals.

Children at high risk for schizophrenia (children with an affected parent) show an increased incidence of numerous neurodevelopmental abnormalities as compared to offspring of parents without schizophrenia [36, 37]. This discrepancy has led to a focus on early developmental factors in the etiology of schizophrenia. Several recent studies have focused on

genomic copy number variants (CNVs) potentially affecting the expression or function of genes that are relevant to brain development [38]. Of particular interest is the velo-cardio-facial syndrome caused by a deletion CNV in chromosome 22q, which confers a 25% risk for schizophrenia [39]. Some of the specific environmental risk factors currently under investigation include obstetric complications [40], childhood trauma [41], prenatal factors such as nutritional deficiencies [42], increased paternal age [43], family interactions [28], maternal infections [44], maternal cytokines [45], gluten sensitivity [46], and cannabis use [47, 48]. In summary, schizophrenia is now widely viewed as a neurodevelopmental disorder comprised of a confluence of vulnerability genes and environmental exposures [49].

Mood disorders

A heterogeneous group of syndromes, of which major depression and bipolar disorder (manic depression) are major subtypes, comprise mood disorders. Bipolar disorder is one of the psychiatric disorders most widely studied from a genetic perspective [50, 51]. Both major depression and bipolar disorder have important genetic components. Controlled family studies show a five-fold risk to relatives of major depression, and greater than a ten-fold risk to firstdegree relatives of bipolar patients for developing these disorders. The concordance rate for bipolar monozygotic twins is over five times that of dizygotic twins, and twin concordance for depression shows less dramatic but still notable differences. A summary of five methodologically comparable twin studies of major depression yielded an average estimate of the heritability of major depression of 0.37, with the remainder (0.63) nearly totally attributable to environmental factors unique to the individual [50]. The relative risks based on the few existing adoption studies also confirm that the familial recurrence cannot be attributed solely to environmental factors [51].

The aggregate adoption study data on mood disorders reveal a moderate increase in rates of mood disorders among the biological compared to adoptive relatives of adoptees with mood disorders [52]. With respect to bipolar disorder, there is little evidence for differential risk among biological compared to adoptive relatives of adoptees with bipolar disorder. However, the small numbers of bipolar adoptees who have been studied (i.e. less than 50) do not provide an

adequate test of genetic and environmental influences. The most compelling finding from adoption studies, however, is the dramatic increase in completed suicide among biologic, as opposed to adoptive, relatives of mood disorder probands [2, 53].

Anxiety disorders

At present, relatively few studies have examined anxiety disorders from the perspective of genetic epidemiology, and there is virtually no data from certain paradigms, such as adoption studies [54, 55]. However, the existing research indicates that most anxiety disorders aggregate in families and several investigations have offered specific support for genetic etiology.

Panic disorder

Panic disorder has the strongest degree of familial aggregation of any of the anxiety disorder subtypes. A review of 13 family studies of panic disorder by Gorwood [56] shows a seven-fold relative risk of panic among relatives of panic probands compared to controls. In addition, early-onset panic, panic associated with childhood separation anxiety, and panic associated with respiratory symptoms have each been shown to have a higher familial loading than other varieties of panic disorder [57]. Although there has been some inconsistency reported among twin studies of panic disorder, recent studies using contemporary diagnostic criteria show that panic disorder has the highest heritability of all anxiety disorders (44%) [58].

Phobic states

Although there are far fewer controlled family and twin studies of the anxiety subtypes other than panic disorder, all of the phobic states (i.e. specific phobia, agoraphobia) have also been shown to be familial [59, 60]. The average relative risk of phobic disorders in the relatives of phobics is 3.1, with greater familial aggregation for the generalized subtype of social phobia. The heritability of phobias according to twin studies is about 0.35% [61].

Generalized anxiety disorder

A limited number of studies also provide evidence for both the familial aggregation and heritability of generalized anxiety disorder [62]. The average familial odds ratio for the disorder is approximately 5 [60, 63] and the heritability is 0.32 among female twin pairs [64]. Chapter 1: Contribution of genetic epidemiology

Obsessive-compulsive disorder

Controlled family studies of obsessive-compulsive disorder reveal an elevated familial risk in probands with obsessive-compulsive disorder compared to controls, with greater familial aggregation associated with early age of onset and obsessions rather than compulsions [65–67]. Twin studies of obsessive-compulsive disorder, however, have yielded only weak evidence for heritability [54, 68, 69].

Substance use disorders

A positive family history of a substance use disorder is a consistent and robust risk factor for substance use in first-degree relatives (for comprehensive reviews of alcoholism see [70–73]). Controlled family studies of alcohol use disorder reveal a three-fold increased risk of alcoholism and two-fold increased risk of drug abuse among the relatives of probands with alcoholism as compared to those of controls [74]. Both alcohol abuse and dependence appear to be familial among females, whereas only dependence aggregates among the relatives of males with alcohol dependence [75].

Twin studies have demonstrated the contribution of both genetic and environmental risk factors to both alcoholism and drug abuse [76]. Heritability of alcoholism (narrowly defined) has been estimated at 59% by some researchers [77], while the heritability of problem drinking (using broad definitions) has been estimated at 8-44% in females and 10-50% in males [70]. Several adoption studies conducted in Scandinavia demonstrated the importance of genetic factors underlying alcoholism [78-80]. Adoption study paradigms have shown not only that a disturbed adoptive family environment interacts with a genetic predisposition for alcoholism to increase the risk for the disorder [81], but also that the adoptive family environment can predict alcohol abuse or dependence independent of genetic vulnerability [82]. A recent "quasi-adoption" study that investigated the association between the biological family background (genetic factors), and a history of exposure to alcoholism during childhood (environmental factors) revealed greater effect of genetic risk factors among men than among women. The study also showed common genetic and environmental risk factors contributing to alcohol dependence in both men and women [83]. The importance of the environment as a mediating factor in the transmission of substance use disorders

Chapter 1: Contribution of genetic epidemiology

was demonstrated in a recent study of adoptive and step families [84].

Although there has been less systematic research on the familial aggregation of drug use disorders, numerous family history studies and uncontrolled and controlled family studies have demonstrated that rates of substance use disorders are elevated among relatives of drug abusers compared to those of controls and compared to population expectations [85, 86]. One controlled family study of drug use disorders using contemporary family study data [75] showed an eight-fold increased risk of substance use disorders (opioids, cocaine, cannabis, and alcohol) among relatives of probands with drug disorders compared with relatives of people with psychiatric disorders and normal controls. Family, twin, and adoption studies have also demonstrated common genetic and environmental factors that contribute to cannabis use disorders and other drug use disorders [87]. The results of numerous twin studies of substance use disorders in general as well as those of specific drugs have shown that there are genetic factors underlying drug abuse in general [88], as well as unique genetic factors associated with specific drugs of abuse including nicotine and cannabis [72, 77, 89, 90].

Sources of complexity in mental disorders

Two factors which contribute to the complexity of the patterns of inheritance of psychiatric disorders are the *lack of validity of the classification* of psychiatric disorders (e.g. phenotypes, or observable aspects of diseases) and the *complexity of the pathways from geno-types to psychiatric phenotypes* (i.e. heterogeneity).

Lack of validity of the classification system

The development of structured interviews has enhanced comparability of diagnostic methods within the United States and worldwide. There is now an exciting venture designed to collect information on the prevalence of mental disorders using comparable diagnostic tools in more than 30 countries under the auspices of the World Health Organization and Harvard University [91]. The lack of conclusive evidence for the validity of classifications of psychiatric disorder phenotypes, because they are based solely on clinical manifestations without pathognomonic markers, continues to impede advances in psychiatry [92, 93]. Growing research on the dimensional classification of disorders further demonstrates the difficulties in creating a valid classification system for psychiatric disorders because of the prevalence of subthreshold diagnostic categories and diagnostic spectra, and the pervasive comorbidity between purportedly distinct diagnostic entities; there is widespread agreement that the categorical classification system in psychiatry lack validity [27, 94, 95].

The greater complexity of psychiatric disease, as compared to other types of disease explains the continued reliance on the descriptive approach as the sole basis for diagnosis in psychiatry. The difficulty in classifying human cognition, behavior, and emotion is understandable in light of the complex psychological and physiological states underlying mental function, which is the culmination of human adaptation to the environment up to the current point in time. Progress in neuroscience that reveals information about the biological pathways underlying psychiatric disorders should also advance our understanding of the classification of psychiatric phenotypes.

Complex patterns of transmission

The application of advances in genomics to mental disorders is still limited by the complexity of the process through which genes influence the development and progression of mental disorders. There is substantial evidence that a lack of one-to-one correspondence between the genotype and phenotype exists for most of the major psychiatric disorders. Psychiatric disorders, like numerous other complex disorders for which susceptibility alleles have been identified, are characterized by phenomena such as incomplete penetrance (i.e. probability of phenotypic expression among individuals with susceptibility gene), variable expressivity (i.e. variation in clinical expression associated with a particular gene), geneenvironment interaction (i.e. expression of genotype only in the presence of particular environmental exposures), pleiotropy (i.e. capacity of genes to manifest several different phenotypes simultaneously), genetic heterogeneity (i.e. different genes leading to indistinguishable phenotypes), gene-environment correlation [96] and polygenic and oligogenic modes of inheritance (i.e. simultaneous contributions of multiple genes rather than Mendelian single gene models) [10, 97]. Other proposed mechanisms of transmission include mitochondrial inheritance, imprinting, and epigenetic phenomena [98].

Chapter 1: Contribution of genetic epidemiology

Comorbidity

The high magnitude of *comorbidity* and co-aggregation of index disorders with other major psychiatric disorders (i.e. bipolar disorder and alcoholism, major depression and anxiety disorders, schizophrenia and drug dependence), in part induced by the classification system, has been demonstrated in both clinical and community studies [11, 86, 99, 100]. For example, alcoholism, a well-established complication of bipolar illness, may mask the underlying features of bipolarity, leading to phenotypic misclassification in genetic studies [101]. Nonrandom mating is also a common phenomenon in mental disorders that impedes evaluation of patterns of familial transmission [102]. Assortative mating is particularly pronounced for substance use disorders for which substance dependence among spouses of substance dependent probands may be as high as 90% [103].

These phenomena serve to decrease the signal to noise ratio in defining psychiatric disorders for genetic studies. Studies that attempt to identify the impact of these phenomena on phenotypic expression in individuals and families will bring us closer to understanding the role of the underlying genes on the components of psychiatric disorders.

Applications of genetic epidemiology to gene identification

There is a widespread consensus among geneticists and epidemiologists on the importance of epidemiology to the future of genetics and on the conclusion that the best strategy for susceptibility risk factor identification for common and complex disorders will ultimately involve large epidemiological studies from diverse populations [73, 104-108]. It is likely that population-based association studies will assume increasing importance in translating the products of genomics to public health. There are several reasons that population-based studies are critical to current studies seeking to identify genes underlying psychiatric disorders. First, the frequency of newly identified polymorphisms, whether single nucleotide polymorphisms (SNPs) or other variants such as copy number variation (CNVs), especially in particular population subgroups, is not known. Second, current knowledge of genes as risk factors is based nearly exclusively on clinical and nonsystematic samples. Hence, the significance of the susceptibility alleles that have been identified for cancer, heart disease, diabetes, and other common disorders is unknown in the population at large. In order to provide accurate risk estimates, the next stage of research needs to move beyond samples identified through affected individuals to the population as a whole. Third, identification of risk profiles will require large samples to assess the significance of vulnerability genes with relatively low expected population frequencies. Fourth, similar to the role of epidemiology in quantifying risk associated with traditional disease risk factors, applications of human genome epidemiology can provide information on the specificity, sensitivity, and impact of genetic tests to inform science and individuals [105]. Below we review the role of the tools of epidemiology in ongoing and future studies designed to identify genes underlying mental disorders.

Samples

The shift from systematic large-scale family studies to linkage studies in psychiatry has led to the collection of families according to very specific sampling strategies (e.g. many affected relatives, affected sibling pairs, affected relatives on one side of the family only, availability of parents for study, etc.) in order to maximize the power of detecting genes according to the assumed model of familial transmission. Despite the increase in power for detecting genes, these sampling approaches have diminished the generalizability of the study findings, and contribute little else to the knowledge base if genes are not discovered. Recent genome-wide association studies of psychiatric disorders have included probands from families previously collected for linkage studies and single cases collected more recently from hospital admissions, almost all of self-reported European descent. Future studies will attempt to collect both families and controls from representative samples of the population so that results can be used to estimate population risk parameters, to examine the specificity of endophenotypic transmission and so results can be generalized to whole populations.

Selection of controls

The most serious problem in the design of association studies is the difficulty of selecting controls that are comparable to the cases on all factors except the disease of interest [109, 110]. Controls should be

Chapter 1: Contribution of genetic epidemiology

drawn from the same population as cases, and must have the same probability of exposure (i.e. genes) as cases. Controls should be selected to ensure the validity rather than the representativeness of a study. Failure to equate cases and controls may lead to confounding (i.e. a spurious association due to an unmeasured factor that is associated with both the candidate gene and the disease). In genetic casecontrol studies, the most likely source of confounding is ethnicity because of differential gene and disease frequencies in different ethnic subgroups. Recent genome-wide association studies of psychiatric disorders have included control samples recruited from the general population using self-administered psychiatric screens and from blood bank samples that exclude donors reporting major psychiatric diagnoses or taking psychiatric medications. The matching of controls to cases on ethnic background is largely based on self-report; several methods are used to screen for and exclude subjects with substantial differences in ancestry.

Risk estimation

Because genetic polymorphisms involved in complex diseases are likely to be nondeterministic (i.e. the marker neither predicts disease nor nondisease with certainty), traditional epidemiological risk factor designs can be used to estimate the impact of these genetic polymorphisms. Increased attention to alleles as a part of risk equations in epidemiology will likely resolve the contradictory findings from studies that have generally employed solely environmental risk factors, such as diet, smoking, alcohol use, etc. Likewise, the studies that seek solely to identify small risk alleles will continue to be inconsistent because they do not consider the effects of nongenetic biological parameters or environmental factors that contribute to the diseases of interest.

There are several types of risk estimates that are used in public health. The most common is *relative risk*, defined as the magnitude of the association between an exposure and disease. It is independent of the prevalence of the exposure. The *absolute risk* is the overall probability of developing a disease in an individual or in a particular population [111]. The *attributable risk* is the difference in the risk of the disease in those exposed to a particular risk factor compared to the background risk of a disease in a population (i.e. in the unexposed). The population attributable risk relates to the risk of a disease in a total population (exposed and unexposed) and indicates the amount the disease can be reduced in a population if an exposure is eliminated. The population attributable risk depends on the prevalence of the exposure, or in the case of risk alleles, the allele frequency. Genetic attributable risk would indicate the proportion of a particular disease that would be eliminated if a particular gene or genes were not involved in the disease. For example, the two vulnerability alleles for Alzheimer's disease include the very rare, but deterministic alleles in the β -amyloid precursor, presentiin-1, and -2 genes, which signal a very high probability of the development of Alzheimer's disease, particularly at a young age, and the susceptibility allele $\varepsilon 4$ in the apolipoprotein-E gene (APOE ɛ4) [112]. The apolipoprotein-E £4 (APOE £4) allele has been shown to increase the risk of Alzheimer's disease in a dosedependent fashion. Using data from a large multiethnic sample collected by more than 40 research teams, Farrer [113] reported a 2.6-3.2 greater odds of Alzheimer's disease among those with one copy, and 14.9 odds of Alzheimer's disease among those with two copies of the APOE ɛ4 allele. Moreover, there was a significant protective effect among those with the $\varepsilon 2/\varepsilon 3$ genotype. As opposed to the deterministic mutations, the APOE £4 allele has a very high population attributable risk because of its high frequency in the population. The APOE £4 allele is likely to interact with environmental risk and protective factors [114, 115]. The population risk attributable to these mutations is quite low because of the very low population prevalence of disease associated with these alleles. This model of combination of several rare deterministic alleles in a small subset of families and common alleles with lower relative risk to individuals but high population attributable risk is likely to apply to many of the psychiatric disorders as well, and may in part explain some of the discrepancies in findings across studies to date.

Recent genome-wide association studies have uncovered risk alleles associated with coronary artery disease, Crohn's disease, rheumatoid arthritis, type 1 and type 2 diabetes [116], and schizophrenia. Those genetic variants appear to confer only modest increases in disease risk (odd ratios [ORs] between 1.2 and 1.5) compared with other established risk factors for common chronic diseases.

Use of endophenotypes for classification

Numerous studies have begun to deconstruct psychiatric phenotypes by their component features or subtypes including bipolar disorder [117, 118], general anxiety disorder [119], obsessive-compulsive disorder [120], schizophrenia [121], and panic disorder [122]. Identification of phenotypic traits or markers, which are themselves heritable, and which may represent intermediate forms of expression between the output of underlying genes and the broader disease phenotype, have been termed "endophenotypes" [123]. Studies of the role of genetic factors involved in these systems may be more informative than studies of the aggregate psychiatric phenotypes since they may more closely represent the expression of underlying biological systems. To the extent that particular endophenotypes more clearly represent expression of genotypes, they may help to unravel the complexity of transmission of the mental disorders. For example, some of the endophenotypes that may underlie mood disorders include circadian rhythm, stress reactivity, and mood, sleep and appetite regulation [95]. Cognitive, neurophysiologic, and structural measures continue to be tested as potential schizophrenia endophenotypes [124, 125]. However, before applying endophenotypes in gene identification studies, there should be evidence that the endophenotype has a stronger genetic signal than the broader phenotype. A recent meta-analysis of psychiatric endophenotypes [126] and a review of the genetic architecture of traits in model organisms do not provide evidence that endophenotypes are superior to current phenotypic disease definitions [127].

Identification of environmental factors

In parallel with the identification of susceptibility alleles, it is important to identify environmental factors that operate either specifically or nonspecifically on those with susceptibility to psychiatric disorders in order to develop effective prevention and intervention efforts. Langholz et al. [128] describes some of the world's prospective cohort studies that may serve as a basis for studies of gene-disease associations or gene-environment interactions. There is increasing evidence that gene-environment interaction will underlie many of the complex human diseases. Some examples include inborn errors of metabolism, individual variation in response to drugs

Chapter 1: Contribution of genetic epidemiology

[129], substance use disorders [71, 130] and the protective influence of a deletion in the CCR5 gene on exposure to the human immunodeficiency virus (HIV) [131].

In prospective studies, however, few environmental exposures have been shown to have an etiological role in psychiatric disorders [132]. Over the next decades, it will be important to identify and evaluate the effects of specific environmental factors on disease outcomes and to refine measurement of environmental exposures to evaluate specificity of effects. Study designs and statistical methods should focus increasingly on gene-environment interaction [106, 133, 134]. Although numerous recent studies have reported gene-environment interaction between several genes that interact with nonspecific environmental exposures such as life stress or childhood adversity and a range of outcomes including depression, cannabis dependence, and conduct disorder [135], replication of these findings has not been consistent [136]. Increased knowledge of the developmental pathways of emotion, cognition, and behavior will expand our ability to identify specific environmental factors such as infection, poor diet, prenatal environment, and early life experiences that interact with the genetic architecture of mood regulation and cognition [137].

Impact of genomics on psychiatric science and practice

Progress in genomics has far outstripped advances in our understanding of psychiatric disorders and their etiologies. Technical advances and availability of rapidly expanding genetic databases provide extraordinary opportunities for understanding disease pathogenesis. However, the application of psychiatric genetic research to study diagnostic heterogeneity, course and/or treatment outcome is still limited due to the lack of consistent genetic findings to date. Over the next decade increasing understanding of the complex mechanisms through which genetic risk factors influence disease should enhance the clinical utility of psychiatric genetics.

The goal of genomics research is ultimately prevention, the cornerstone of public health. An understanding of the significance of genetic risk factors and proper interpretations of their meaning for patients and their families will ultimately become part of clinical practice. Clinicians will become

Chapter 1: Contribution of genetic epidemiology

increasingly involved in helping patients to comprehend the meaning and potential impact of genetic risk for both psychiatric and nonpsychiatric disorders. As our knowledge of the role of genetic risk factors in psychiatric disorders advances, it will be incumbent

References

- Heston LL. Br J Psychiatry 1966; 112(489):819–825.
- 2. Wender PH, et al. *Arch Gen Psychiatry* 1986;43(10):923–929.
- 3. Wender PH. *Annu Rev Med* 1972;23:355–371.
- 4. Kety SS, et al. *Schizophr Bull* 1976;**2**(3):413–428.
- Kety SS, et al. Genetic Relationships within the Schizophrenia Spectrum: Evidence from Adoption Studies. New York: Raven Press; 1978.
- Gershon ES, et al. (eds.). Genetic Approaches to Mental Disorders. New York: American Psychopathological Association Press; 1994.
- Beaty TH, et al. *Epidemiol Rev* 2000;**22**(1):120–125.
- MacMahon B, et al. Epidemiology: Principles and Methods. Boston: Little Brown and Company; 1996.
- 9. Kuller LH. *Prog Clin Biol Res* 1979;**32**:489–495.
- Risch N. Am J Hum Genet 1990; 46(2):222–228.
- Merikangas KR. Comorbity for Anxiety and Depression: Review of Family and Genetic Studies. Washington, DC: American Psychiatric Press; 1990.
- 12. Shih RA, et al. *Int Rev Psychiatry* 2004;**16**(4):260–283.
- 13. Kendler KS, et al. *Psychol Med* 2007;**37**(5):615–626.
- National Adoption Information Clearinghouse. *Families 2007 Edition*. Washington, DC: US Department of Health and Social Services; 2007.
- 15. Kolonel LN, et al. *Nat Rev Cancer* 2004;4(7):519–527.

- Ödegaard Ö. Emigration and Insanity: A Study of Mental Disorders among the Norwegian Born Population of Minnesota. Copenhagen: Levin and Munksgaards; 1932.
- 17. Merikangas K, et al. *Epidemiol Rev* 1996;19:1–12.
- 18. Kendler KS. *Arch Gen Psychiatry* 1995;**52**(11):895–899.
- 19. Tandon K, et al. *Eur J Neurosci* 2002;**16**(3):403–407.
- 20. Tsuang MT, et al. *Methods Mol Med* 2003;77:251–265.
- 21. Zerbin-Rudin E. *Nervenarzt* 1980;51(7):379–391.
- 22. Gottesman, II. Acta Psychiatr Scand Suppl 1994;**384**:26–33.
- 23. Kendler KS, et al. *Schizophr Bull* 1993;**19**(2):261–285.
- 24. Sullivan PF. *PLoS Med* 2005; 2(7):e212.
- Sullivan PF, et al. Arch Gen Psychiatry 2003;60(12): 1187–1192.
- Hallmayer J. Aust N Z J Psychiatry 2000;34 Suppl:S47–55; discussion S56–47.
- 27. Tsuang MT. Am J Med Genet 2001;**105**(1):8–10.
- 28. McGuffin P. *Psychiatr Prax* 2004;**31** Suppl 2: S189–193.
- 29. Portin P, et al. *Acta Psychiatr Scand* 1997;**95**(1):1–5.
- 30. Tienari P, et al. *Br J Psychiatry* 2004;**184**:216–222.
- 31. Selten JP, et al. *Am J Psychiatry* 2002;**159**(4):669–671.
- 32. Hanoeman M, et al. *Schizophr Res* 2002;**54**(3):219–221.
- 33. Cooper B. *Epidemiol Psichiatr Soc* 2005;14(3):137–144.

upon clinicians to become familiar with knowledge gleaned from genetic epidemiological and genomics research. In the meanwhile, use of recurrence risk estimates from family studies best predicts the risk of the development of mental disorders.

- 34. Leao TS, et al. J Nerv Ment Dis 2006;**194**(1):27–33.
- 35. Smith GN, et al. Schizophr Res 2006;87(1-3):205-211.
- 36. Owens DG, et al. *Psychol Med* 2006;**36**(11):1501–1514.
- Brewer WJ, et al. Schizophr Bull 2006;32(3):538–555.
- 38. Cook EH, Jr., et al. *Nature* 2008;455(7215):919–923.
- 39. Bassett AS, et al. *Hum Mol Genet* 2008;17(24):4045–4053.
- 40. Clarke MC, et al. *Schizophr Bull* 2006;**32**(1):3–8.
- 41. Morgan C, et al. *Schizophr Bull* 2007;**33**(1):3–10.
- 42. Ludvigsson JF, et al. Scand J Gastroenterol 2007;49:179.
- 43. Malaspina D, et al. CNS Spectr 2002;7(1):26–29.
- 44. Buka SL, et al. *Arch Gen Psychiatry* 2001;**58**(11):1032–1037.
- 45. Buka SL, et al. *Brain Behav Immun* 2001;15(4):411-420.
- 46. Kalaydjian AE, et al. *Acta Psychiatr Scand* 2006;113(2):82–90.
- 47. Arseneault L, et al. *BMJ* 2002;**325** (7374):1212–1213.
- 48. Dean K, et al. *Dialogues Clin Neurosci* 2005;7(1):69–80.
- 49. Dealberto MJ. *Med Hypotheses* 2007;**68**(2):259–267.
- 50. Sullivan PF, et al. *Am J Psychiatry* 2000;157(10):1552–1562.
- 51. Merikangas KR, et al. *Clin Neurosci Res* 2002;**2**:127.
- 52. Faraone SV, et al. *Psychol Bull* 1990;**108**(1):109–127.
- 53. Mendlewicz J, et al. *Nature* 1977;**268**(5618):327–329.
- 54. Carey G, et al. In Rabkin J (ed.). *Anxiety: New Research and*