Published in final edited form as:
Int Rev Psychiatry. 2007 October ; 19(5): 509–521.

The challenges of using epidemiology to inform clinical practice

Ronald C. Kessler, Ph.D.
The Department of Health Care Policy, Harvard Medical School, Boston, USA.

Abstract

This paper discusses challenges and prospects for increasing the clinical relevance of psychiatric epidemiological research. The discussion begins with a review of the structural determinants of the fact that current psychiatric epidemiological research has less clinical relevance that epidemiological research in other areas of medicine. The discussion then turns to ways in which the focus of psychiatric epidemiological research might be changed to increase its clinical relevance. A review is then presented of recent innovations in community psychiatric epidemiological research that were designed to increase clinical relevance. An argument is then made that the full clinical value of psychiatric epidemiology will only be realized when community epidemiology becomes better integrated with clinical epidemiology and the latter takes on a more prominent role than it currently has in psychiatric research. Existing initiatives to realize an integration of community psychiatric epidemiology with clinical epidemiology are then reviewed. Finally, an agenda is proposed for an expansion of clinical psychiatric epidemiology to include a focus on both naturalistic and quasi-experimental studies of illness course and treatment response in diverse clinical samples.

Keywords

Mental illness; Psychiatric epidemiology; epidemiology; clinical epidemiology; World Mental Health (WMH) Survey Initiative

I was asked to prepare a paper on the challenges of using epidemiology to inform clinical practice in psychiatry because of my involvement in recently published exchanges about the clinical value (or lack thereof) of the WHO World Mental Health (WMH) Survey Initiative (Jorm, 2006; Kessler, 2006; Kessler et al., in press-b; Weich and Araya, 2004). The WMH Survey Initiative (henceforth WMH) is a project of the Assessment, Classification, and Evaluation (ACE) Group at the WHO that involves the implementation, coordination, and centralized analysis of data from a series of psychiatric epidemiological surveys in more than two dozen countries around the world. (www.hcp.med.harvard.edu/wmh.) I am one of the WMH co-directors.

Psychiatric epidemiologists have done quite a bit over the past two decades to increase the clinical relevance of their research. The vast majority of this work has been carried out in general population samples. The potential for clinical relevance is much greater, though, for epidemiological research carried out in clinical samples. In this paper I review recent developments in community psychiatric epidemiology related to clinical relevance. I also sketch out a research agenda for clinical psychiatric epidemiology aimed at increasing clinical relevance. Before turning to these topics, though, it is important to define some basic terms and to discuss the overarching practical challenges faced by psychiatric epidemiologists as they attempt to increase the clinical relevance of their research.

Contact information: RC Kessler Department of Health Care Policy Harvard Medical School 180 Longwood Avenue Boston, MA, USA 02115 617-432-3587 (Voice) 617-432-3588 (Fax) kessler@hcp.med.harvard.edu (email)..

I. Defining terms

Epidemiology is "the study of the distribution and determinants of health-related states or events in specified populations and the application of this study to control of health problems" (Last et al., 1995). Although mental health professionals might think of epidemiology as being synonymous with *descriptive community* epidemiology, it is important to realize that epidemiological research is often concerned with causal inferences rather than mere description and that epidemiological studies in other branches of medicine are often carried out in clinical samples rather than in community samples (Rothman and Greenland, 1998).

Risk factor assessments in what are generally referred to as *analytical* epidemiological studies (i.e., epidemiological studies designed to discover causes) often lead to *experimental* epidemiological studies. The latter typically investigate the effects of presumed causal risk factors on the onset and course of illness (Fletcher and Fletcher, 2005). These experiments are considered *epidemiological* rather than *clinical* because they are generally carried out in community samples with the goal of determining the generalizability of effects. Epidemiology is additionally used in clinical settings to evaluate the validity of diagnostic tests and to study predictors of treatment response that might be targeted in subsequent interventions (Greenberg et al., 2000). This sort of *clinical epidemiology* is under-developed in psychiatry.

II. Overarching Practical challenges

Based on the definitions provided above, it is obvious that epidemiology can have enormous value for clinicians. But does psychiatric epidemiological research, as currently practiced, has all the clinical value it could have? The answer is clearly no. A number of reasons can be cited for this. One is that there are more disagreements about fundamental aspects of diagnoses in psychiatry than in other branches of medicine. Some of these disagreements concern the very existence of particular diagnoses (e.g., Faraone et al., 2006; Gadow and Weiss, 2001). Other disagreements concern the correct operational criteria for a particular disorder (e.g., Ruscio et al., 2005). These disagreements have led psychiatric epidemiologists to focus much of their attention on basic descriptive investigations to the neglect of analytical and clinical studies.

Another important reason for the limited clinical value of psychiatric epidemiology is that the processes of care used to treat mental disorders are much more diverse than in other branches of medicine. This makes it more difficult than otherwise to carry out clinical epidemiological studies of naturalistic variation in treatment response. As a result of this problem, we know much less than we should know about the number of people in the population who would profit from treatment or about the actual effectiveness of these treatments as they are delivered in clinical practice.

These gaps in our knowledge have created widespread lack of recognition about the enormous societal costs of mental disorders. This lack of recognition, in turn, has created perhaps the most important problem for the advancement of epidemiological research to inform clinical practice in psychiatry: the absence of funds for such research. This problem is much more severe than in areas of chronic disease epidemiology concerned with physical disorders that cause much less overall societal impairment than mental disorders. Numerous large-scale prospective epidemiological studies of risk factors for cancer, diabetes, heart disease, and other chronic physical disorders, for example, exist that dwarf even the largest psychiatric epidemiological study.

III. Community psychiatric epidemiology

A. Descriptive community psychiatric epidemiology

As noted above, a much larger proportion of epidemiological research in psychiatry than other areas of medicine is descriptive rather than analytical or experimental. I believe that this is due to the more widespread disagreements in psychiatry than in other branches of medicine about basic descriptive issues. Most descriptive psychiatric epidemiological research is based on community samples because we know that many people with mental disorders do not receive treatment, making it important to obtain broadly representative data that can be used to estimate the prevalence and correlates of unmet need for treatment.

Descriptive community psychiatric epidemiology has gone through an unprecedented period of growth over the past twenty years. Starting with the Epidemiologic Catchment Area (ECA) study in the USA (Robins and Regier, 1991), large community surveys of mental disorders have been carried out in a number of countries throughout the world. An important innovation of the ECA was the development of a fully structured research diagnostic interview known as the Diagnostic Interview Schedule (DIS) (Robins et al., 1981) that could be used by trained lay interviewers to generate diagnoses approximating those produced by clinicians. Methodological studies demonstrated that the DIS yielded reasonably reliable and valid diagnoses (Helzer et al., 1985), a result that was very important in promoting the ECA–DIS methodology.

The first expansion of the ECA–DIS methodology was carried out by the World Health Organization (WHO) in collaboration with the US Alcohol, Drug, and Mental Health Administration to add questions that would operationalize International Statistical Classification of Diseases (ICD) criteria and to produce versions of the interview in many different languages. The resulting instrument, the WHO Composite International Diagnostic Interview (CIDI) (World Health Organization, 1990), first became available in 1990. WHO technical support led to an unprecedented number of major epidemiological surveys using the CIDI over the next decade in countries as diverse as Brazil (Andrade, 1996), Canada (Offord et al., 1994), Germany (Wittchen et al., 1992), Mexico (Caraveo et al., 1998), the Netherlands (Bijl et al., 1998), and Turkey (Kylyc, 1998). In 1997, WHO created the International Consortium in Psychiatric Epidemiology (ICPE) to coordinate the comparative analysis of these data (Kessler, 1999). The ICPE also provides technical assistance to researchers planning new CIDI surveys. The WMH initiative is an outgrowth of these technical assistance activities.

Several important results have consistently emerged from the DIS and CIDI surveys. One is that mental disorders are among the most prevalent chronic diseases in the general population, with lifetime prevalence close to 50% of the population in some countries and 12-month prevalence often in the 15–25% range (Robins and Regier, 1991). Another important result is that mental disorders typically have much earlier ages of onset than other chronic diseases. For example, anxiety disorders and impulse-control disorders have median ages of onset in the early to late teens in most surveys, while mood and substance use disorders have median ages of onset in the early to mid twenties (Kessler et al., 2005c; WHO International Consortium in Psychiatric Epidemiology, 2000).

DIS and CIDI surveys have also shown that mental disorders are among the most impairing of all chronic diseases (Kessler et al., 2001b). Furthermore, only a minority of the respondents with a mental disorder in most DIS-CIDI surveys report receiving treatment (Alegria et al., 2000). In developed countries, the measures of disorder severity included in the surveys are consistently associated both with high probability of service use and with high likelihood that service use occurred in the specialty sector. Severity is also generally associated with high intensity of treatment, documenting that there is some rationality both in help-seeking and in

the allocation of treatment resources. However, the surveys also show that only a minority of patients even in the richest countries describe a course of treatment that meets minimal criteria for adequacy in terms of currently available treatment guidelines (Katz et al., 1998).

B. Responding to the challenges: Validity of diagnoses

The high rates of disorder found in the DIS and CIDI surveys have led some commentators to raise questions about the plausibility of the prevalence estimates generated by these surveys (Jorm, 2006; Regier et al., 1998; Weich and Araya, 2004). At least two separate challenges exist here. One is to determine whether the fully-structured diagnostic interviews used in these surveys yield valid assessments of mental disorders among respondents who *are willing* to disclose information about their emotional functioning to interviewers. Willingness is an important issue to consider because a major difference between community epidemiology and clinical epidemiology is that respondents in clinical studies can generally be relied on much more than respondents in community surveys to provide honest and thoughtful responses to interviewers. This is not to say that respondents in clinical interviews never dissemble. They do. However, respondents in community epidemiological surveys can be expected to be less engaged in the process and more likely to provide inaccurate responses because they often think of their participation in the survey as a lark rather than something serious, while respondents in clinical studies are more serious due to the fact that they are seeking professional help. The second challenge in developing fully-structured assessments of mental disorders, then, is to create a protocol for survey administration that encourages respondents in community epidemiological surveys to provide thoughtful and honest responses to fully structured questions after information is available that these questions yield clinically relevant information when they are answered seriously and honestly.

A good deal of work has been carried out by psychiatric epidemiologists to address these challenges. The DIS was initially developed in such a way as to address the first of the two challenges enumerated in the last paragraph by mimicing DSM diagnostic criteria as closely as possible. Clinical reappraisal studies showed, though, that the concordance of the DIS and the initial version of the CIDI with independent clinical diagnoses was often only modest (Wittchen, 1994). Based on these results, refinements were made in the CIDI to improve validity in assessing DSM-III-R disorders (Kessler et al., 1998). Generally good concordance with blinded clinical diagnoses was found in this revised version of CIDI in the general population of the US (Kessler et al., 1998; Wittchen, 1994; Wittchen et al., 1995; Wittchen et al., 1996). However, the results of CIDI clinical reappraisal studies varied across other settings. Some studies showed the CIDI diagnoses to have poor agreement with diagnoses based on the SCAN clinical interview (Wing et al., 1990) in a community sample (Brugha et al., 2001) and others showing agreement to be either good (Andrews et al., 1995) in a patient sample or excellent (Jordanova et al., 2004) in a primary care provider sample. This variation in results raises, with much higher concordance in patient samples than community samples, raises the possibility that respondent motivation is more of an issue than problems with question wording.

Based on these results, the CIDI was revised to generate diagnoses using DSM-IV criteria (Kessler and Ustun, 2004). In doing this, an attempt was made to address the motivation problem. This was done by using principles taken from the methodological literature on cognitive survey research to increase respondent understanding, motivation, and ability to provide accurate survey responses. Although these methods are too complex to review in the current report, they are discussed in detail elsewhere (Kessler et al., 1999; Kessler and Ustun, 2004; Kessler et al., 2000; Kessler et al., 1998).

The WMH surveys were the first ones to administer the revised version of CIDI for DSM-IV (CIDI Version 3.0) in large community samples. CIDI clinical reappraisal studies to validate the diagnoses generated by this instrument against blinded clinical reappraisal interviews were

carried out in probability sub-samples of the WMH samples in four countries: France, Italy, Spain, and the US (Haro et al., in press). Good individual-level concordance of diagnoses based on the CIDI with diagnoses based on clinical interviews was found for lifetime prevalence estimates of most disorders, with an area under the ROC curve (AUC, a measure of classification accuracy that is not influenced by disorder prevalence (Hanley and McNeil, 1982); of.76 for the dichotomous classification of having any of the lifetime DSM-IV disorders assessed in the surveys and in the range .62-.93 for individual disorders, with an inter-quartile range (IQR) of .71-.86. Concordance increased when CIDI symptom-level data were added to predict clinical diagnoses in logistic regression equations. AUC for individual disorders in these equations was in the range .74-.99, with an IQR of .87-.96.

CIDI lifetime prevalence estimates were generally conservative relative to clinical estimates in the WMH clinical reappraisal studies. Concordance of CIDI diagnoses and clinical diagnoses for 12-month prevalence could be studied powerfully only for two disorder classes, any anxiety disorder (AUC = .88) and any mood disorder (AUC = .83). As with lifetime prevalence, 12-month concordance improved when CIDI symptom-level data were added to predict clinical diagnoses. CIDI 12-month prevalence estimates were unbiased relative to clinical estimates. It is important to note that the validity of the CIDI was likely to be under-estimated in these comparisons due to the fact that the reliability of the clinical diagnoses, which was presumably less than perfect, set a ceiling on maximum concordance.

In interpreting these results it is important to recognize that all the countries in which the WMH clinical reappraisal studies were completed are developed countries in the West. There is no guarantee that the same good validity of the CIDI will be found in other parts of the world. Indeed, there is good reason to believe that CIDI validity will be considerably worse in less developed countries based on the implausibly low prevalence estimates that have been found in CIDI surveys that were carried out in some such countries (Gureje et al., 2006; Shen et al., 2006). Based on these results, the WMH investigators in the countries with implausibly low prevalence estimates are collaborating in a methodological initiative that is using their samples as natural laboratories to carry out in-depth clinical reappraisal interviews. These interviews are designed, in part, to determine whether the CIDI symptom-level information can be recalibrated to generate diagnoses that are more consistent with clinical judgments than the current CIDI diagnostic algorithms. The larger goal of these methodological studies, though, is to discover ways in which the CIDI questions and procedures might be modified for future research so as to improve concordance with clinical judgments in these countries.

It is important to note that the CIDI diagnostic data described up to now are dichotomous yes-no classifications of whether or not respondents meet criteria for a particular DSM or ICD disorder. Simple dichotomous assessments, while useful, are much less relevant to clinical practice than dimensional measures that provide information about clinical severity. This is especially so in light of the fact that the prevalence of mental disorders in the general population is now known to be quite high. With lifetime prevalence as high as 50% in some countries, health policy planners find themselves increasingly interested in disorder severity rather than in the mere presence of a mental disorder.

In order to advance our understanding of disorder severity, it is important to include assessments of clinical severity in community surveys. The WMH surveys did this by embedding fully structured versions of standard clinical severity measures into the assessments of specific disorders. For example, the Quick Inventory of Depressive Symptoms Self-Report (QIDS-SR) (Rush et al., 2003) was used to assess the severity of 12-month major depressive episodes, a fully-structured version of the Young Mania Rating scale was used to assess the severity of 12-month manic episodes (Young et al., 1978), and the Panic Disorder Severity Scale was used to assess the severity of 12-month panic disorder (Shear et al., 2001). Expansion

of this sort of dimensional evaluation in conjunction with categorical assessment is important to increasing the clinical relevance of community epidemiological studies.

More recent epidemiological work on the validity of psychiatric diagnostic assessments has focused on the development of short global screening scales that can be used to monitor trends in the overall prevalence of mental disorders in government tracking surveys (Kessler et al., 2003a) and disorder-specific screening scales that can be used to screen for specific disorders in primary care samples, workplace samples, and as part of general population mental health education campaigns (e.g., Kessler et al., 2005b; Kessler et al., in press-a). There is a good deal of work going on, additionally, to develop fairly detailed disorder-specific computer-assisted self-administered diagnostic assessments based on the CIDI for use in primary care settings and as first-phase batteries in specialty mental health settings and in large-scale evaluation setting (e.g. military induction). There is also a good deal of interest in developing comparable tools for assessing children and adolescents (Achenbach et al., 1987; Angold and Costello, 2000; Goodman et al., 2000; Shaffer et al., 2000).

C. Responding to the challenges: Documenting the burden of illness

A good argument could be made that the most valuable short-term direction for descriptive community psychiatric epidemiology would be to maintain the focus on societal burden of mental disorders in order to expand the political mandate for research and treatment of mental disorders. As noted in the introduction, the WMH Survey Initiative is designed to do exactly that. Even though a number of WMH surveys are still underway and results will not be available for several more years, preliminary results already have definitively confirmed several facts that were reviewed above in the section on descriptive epidemiology: that mental disorders are highly prevalent in all regions of the world; that people with mental and physical disorders in the general population consistently attribute higher disability to their mental disorders than to their physical disorders in all regions of the world; that mental disorders are associated with a substantial proportion of all disorder-specific disability in the general population in all regions of the world; and that the resources allocated by society to treat mental disorders are much lower than those allocated to treat chronic physical disorders in relation to the societal burdens of these different classes of disorder in all regions of the world.

In order to make sure these results are heard by policymakers, the results need to be repeated time and again in addressing diverse audiences with a number of policy perspectives. Perhaps the most useful perspective in this regard is the perspective of the employer, as employers have an enormous influence on political processes in almost all countries and have a special interest in human capital investments (Kessler and Stang, 2006). Considerable research from the WMH surveys has shown that mental disorders have enormous workplace costs in terms of sickness absence days, low productivity while on the job, disability, and workplace accidents-injuries (Kessler et al., 2005a; Kessler et al., 2006a). Indeed, comparative studies have shown that because of their high prevalence and serious impact on role performance, mental disorders are associated with more work impairment than almost any other class of health problems (Kessler et al., 2001a; Stewart et al., 2003). For example, untreated ADHD is estimated to be associated with nearly $20 billion in lost productivity each year in the US labor force (Kessler et al., 2005a).

Statistics such as these can capture the attention of employers and lead to concerns being raised by employers about the cost-effectiveness of treatments for mental disorders that, in turn, lead to implementation of experimental epidemiological interventions designed to evaluate the return-on-investment (ROI) from the employer perspective of innovative workplace screening, outreach and treatment initiatives. Once initiatives of this sort become the subject of serious discussion, a number of more technical epidemiological questions will take on new importance.

For example, we cannot carry out screening programs to detect mental disorders among workers unless we have valid screening scales.

This means that expansion of research on cost-effective treatment could well lead to a new need for improved short screening scales that can be administered in workplace health risk appraisal surveys (Kessler et al., 2005b; Kessler et al., in press-a). New descriptive epidemiological studies will also be needed to evaluate the implications of modifying diagnostic boundaries to determine whether cases with sub-threshold symptoms experience clinically significant impairments in role functioning that can be modified with treatment (Judd and Akiskal, 2003). Interest in preventive interventions could also lead to a more active program of research than currently exists on the modifiable epidemiological predictors of onset and course of illness.

D. Analytical community psychiatric epidemiological research

In addition to having political value in documenting the enormity of the societal burden of mental illness, community epidemiological research can have valuable clinical implications in several other ways. Perhaps the most obvious of these is that community surveys can be used to generate comparative information on treatment adequacy, modifiable barriers to receiving treatment, and predictors of treatment dropout (Wang et al., 2005).

Community epidemiological research can also be used to search for risk factors that predict onset and course of illness. Epidemiological studies of smoking and cancer are, of course, the classic example of successful epidemiological risk factor research (Peto et al., 2000). Studies of prenatal exposure to famine and offspring schizophrenia are a recent example of important risk factor research from psychiatric epidemiology (Hulshoff Pol et al., 2000; St. Clair et al., 2005). Although such naturalistic risk factor studies are limited in their ability to document causes of complex diseases (Buchanan et al., 2006), they can be very useful in disconfirming otherwise plausible hypotheses and in narrowing the range of testable possibilities to the point where critical tests can be carried out with experimental manipulation. Genetically informative research designs greatly expand the value of epidemiological studies in this regard (Kendler and Prescott, 2006), as do designs that allow researchers to examine broad contextual influences (Susser et al., 2006).

E. Experimental community psychiatric epidemiological research

Although most community epidemiological research on risk factors is non-experimental, opportunities occasionally present themselves to investigate natural experiments or quasi-experiments. Psychiatric epidemiologists need to be more entrepreneurial than they have up to now in exploiting these opportunities. Social policy interventions are the most obvious cases where such opportunities present themselves. An example of considerable current interest is the “black box” warning imposed by the FDA in the fall of 2004 on manufacturers of antidepressants to notify potential users of an increased risk of suicidality associated with the use of antidepressant medications among depressed children and adolescents. (www.fda.gov/bbs/topics/news/2004NEW01124.) This warning was based on post-marketing surveillance data and epidemiological research that documented evidence consistent with the possibility that antidepressant therapy might cause suicidality among some children and adolescents (Jick et al., 2004; Olfson et al., 2006). Other epidemiological research, though, concluded that no such adverse effect of antidepressant medications existed (Gibbons et al., 2005; Valuck et al., 2004).

The black box warning was a potent intervention in that it led to as much as a one-third drop in the number of US youth who are treated with antidepressant medications within months of the warning going into effect. A question that presents itself based on this fact is whether this

intervention had the intended effect of decreasing the number of child and adolescent suicides in the population. Controversy exists on this question, with some mental health professionals fearing that the black box warning led to an increase rather than to a decrease in suicide because so many youth who would otherwise have been in treatment failed to receive treatment. Epidemiological research is currently underway to examine this question by using interrupted time series analysis (McDowall et al., 1980) to study trends in child and adolescent suicide trends before and after the introduction in the black box warning as mediated by disaggregated changes in sales levels of antidepressants for youth.

Policy interventions not originally targeted at mental health outcomes can also sometimes be used to increase out understanding of the structural determinants of mental disorders. Kling and his associates (in press), for example, are evaluating the effects of neighborhood disorganization on the mental health of neighborhood residents by carrying out an epidemiological survey of low-income single mothers who applied for HUD housing vouchers that were allocated on the basis of a lottery. A random half of the applicants were awarded the vouchers, creating a unique opportunity to study neighborhood effects on the mental health of high-risk women and their children. A follow-up survey of the participants in the experiment is currently being launched to evaluate the mental health of intervention and control subjects to address this question.

Experiments of nature can also be used in a similar way. For example, Costello and her associates (Costello et al., 2003) studies the effects of the opening of a casino on an American Indian reservation midway between the two waves of an epidemiological survey to evaluate the impact of increased parental income on child mental health. A great many experiments of nature exist that could be used in a similar way to expand our understanding of the environmental determinants of mental disorders.

F. The costs and benefits of large-scale community epidemiological research

Before leaving the discussion of community psychiatric epidemiological research, it is important to address the concern of critics that psychiatric epidemiological surveys are very expensive and that this high expense may not be justified by the limited importance of the results generated by the surveys. A similar concern can be raised about any sort of large-scale health survey. Yet governments throughout the world carry out such surveys on an ongoing basis in order to obtain information needed to make healthcare policy planning decisions. It is an ongoing challenge to develop more efficient ways to implement these surveys at a reduced cost. There can be no doubt, though, that the costs of healthcare are so high throughout the world that the planning information obtained in health tracking surveys is worth the effort even if this information is able to improve efficiency of healthcare spending by only a small fraction of a percent.

The WMH Survey Initiative has been the subject of related concerns about cost-effectiveness (Weich and Araya, 2004) and about the possibility that the inundation of the literature with WMH reports will lead to a reification of the idiosyncratic decisions made in that study regarding conceptualization and measurement (Jorm, 2006). I agree with these concerns, but not with the implication that large-scale community psychiatric epidemiological research is not cost-effective because of them. The problem with the latter conclusion as it applies to the WMH Initiative is that it is based on an inaccurate assessment of counter-balancing advantages. The two main advantages seen by the critics were (i) that the WMH surveys have been able to generate prevalence estimates of mental disorders that can be used by policy planners and (ii) that the WMH surveys generate estimates of the societal burden of mental disorders. The first of these two presumed advantages has been criticized, though, on the grounds that categorical models of mental disorder lack validity and the second presumed advantage has been criticized

on the grounds that the methods used to estimate the global burden of disease in previous WHO studies are imprecise (Musgrove, 2003).

The criticism of strict adherence to invalid categorical systems is misplaced because the WMH assessment was carried out in such a way as to assess sub-threshold cases in an explicit effort to explore the validity of the diagnostic boundaries currently specified in the ICD and DSM systems. The criticism of using controversial methods to estimate disease burden is misplaced because the main criticism of these methods has been that they rely on imputation rather than empirical analysis. The WMH surveys are carrying out precisely the kind of empirical analysis called for by the critics of previous disease burden estimates.

The critics of the WMH Initiative also overlook a number of other advantages from a cost-effectiveness perspective. (i) The infrastructure costs of the WMH Initiative were provided to participating countries without cost due to generous grants from several foundations that built the infrastructure and shared this resource through dissemination activities. This allowed participating countries to carry out high quality large-scale mental health epidemiological needs assessment surveys at a much lower cost than if they had attempted to launch such surveys on their own. In most cases, local investigators would have been unable to replicate this infrastructure regardless of cost, which means that WMH made it possible to begin a tradition of community mental health needs assessment in these countries. (ii) The cross-sectional WMH surveys are serving as baselines for a number of prospective studies of exactly the sort called for by the critics. (iii) The experiences gained in WMH are helping to train a new generation of psychiatric epidemiologists in countries that lack a strong epidemiological infrastructure. This cadre of trained researchers will be of great value to healthcare planners as evidence-based methods are introduced into health policy planning in the coming years. Based on these considerations, I believe that the WMH Initiative is, on balance, quite a positive development for the field, albeit one that has to be seen as only a foundational development that focuses on description. As noted above, the field needs to go beyond mere description to analytical and experimental epidemiological studies to realize its potential.

IV. Clinical psychiatric epidemiology

A. An agenda for clinical psychiatric epidemiology

As noted earlier, the recent growth in descriptive community psychiatric epidemiological research aimed at studying societal costs will likely lead to a call for interventions and an evaluation of the cost-effectiveness of these interventions from a societal perspective. The latter, in turn, would inevitably lead to a more general growth of interest in clinical epidemiology. One can already see signs of this increased interest in research on the workplace costs of mental disorders and the potential return-on-investment (ROI) from the employer perspective of innovative workplace screening, outreach and treatment initiatives aimed at addressing the problem of mental disorders in the workplace. Several interventions of this sort are currently underway in Australia, Canada, and the US to evaluate the ROI of interventions to treat workers with depression (Wang et al., 2003). Based on positive preliminary results from the depression interventions, related workplace treatment initiatives are currently being developed to treat adult ADHD and bipolar spectrum disorder.

Although there are occasions when experimental methods can be applied to evaluate novel interventions targeted by earlier descriptive and analytical epidemiological studies, the opportunities for true experimentation are limited. However, other important clinical epidemiological studies are also needed to gather and analyze descriptive information that estimates the proportion of patients who recover, the proportion who improve but do not recover, and the proportion who fail to improve in response to existing or new treatments. As

a high proportion of patients drop out of treatment (Edlund et al., 2002), such studies could also valuably document prevalence, timing, and correlates of treatment dropout.

The reason few opportunities exist for genuine experimental evaluation of the effects of routine treatment is that substantial constraints exist in clinical settings on the creation of control groups. Naturalistic studies of clinical cases can nonetheless be very valuable in providing information about the success of real-world treatments, especially when the comparison groups are made up of patients who receive different types of treatments. Merely knowing the percent of patients who experience complete remission, partial remission, and no symptom relief in response to a particular intervention is in itself of considerable value. This value is increased considerably when information is available on the determinants of selection into a particular type of treatment in situations where alternative treatment options exist. In situations of this sort, quasi-experimental inferences can sometimes be drawn. As with natural experiments in community settings, opportunities to make such comparisons need to be sought out actively.

A good example of the latter opportunity can be found in the work of McClellan, McNeil and Newhouse (McClellan et al., 1994), who were interested in evaluating the effects of increased treatment intensity on long-term survival of elderly patients with an acute myocardial infarction (AMI). These researchers discovered that intensity of treatment of AMI was related to severity of illness, making it impossible to examine the effects of treatment intensity in a simple bivariate analysis. However, the researchers also discovered that distance of patient residence from the nearest hospital that treated a high volume of AMI patients was unrelated to severity of illness but strongly related to intensity of AMI treatment, making it possible to use information about distance from residence as an instrumental variable (Hernan and Robins, 2006) to carry out a quasi-experimental analysis of the effects of treatment intensity on survival of AMI patients. The analysis using this approach produced convincing evidence that while intensive acute treatment procedures in the first 24 hours of admission substantial increased survival of elderly AMI patients, the benefits of expensive catheterization and revascularization procedures were minimal.

Another interesting and more generalizable example can be found in recent pharmacoepidemiological studies in the US of the effects of high medication co-payments on treatment intensity (Gibson et al., 2006). These studies show that high co-payments are associated with low treatment adherence and poor clinical outcomes. In cases where before-after changes in treatment associated with co-payment increases can plausibly be assumed to be random with respect to baseline individual differences, as they typically can be other than in marginal cases of adverse selection, causal inferences can legitimately be made from such data (Hernan and Robins, 2006). Research using this logic has documented a number of powerful treatment effects, such as a two-fold decrease in asthma-related hospitalizations associated with a 25% increase in the proportion of time patients with asthma use inhaled corticosteroids (Williams et al., 2004). Although this same research paradigm could be used to carry out clinical epidemiological studies on the effectiveness of commonly used treatments for mental disorders, we are unaware of any published clinical epidemiological research of this sort that has focused on mental disorders.

In addition to studying the aggregate magnitude of treatment effects, clinical epidemiological studies are needed to study the predictors of individual differences in treatment response. This type of work would ideally involve investigating baseline (i.e., as of the onset of treatment) predictors of course of illness in broadly representative clinical samples. Assuming variation in treatment, it would be a simple step from the investigation of predictors to the examination of treatment response in a comparative perspective, with the aim of determining whether the preferred type of treatment varies for patients who differ in baseline characteristics. Research

of this sort could be especially valuable in psychiatry in light of the wide range of treatment alternatives available.

Valuable research of this sort has already been carried out using an experimental paradigm to examine the differential effects of alternative therapies for the treatment of alcoholism based on patient characteristics (e.g., Cutler and Fishbain, 2005). Non-experimental clinical epidemiological research of a similar sort could be useful in providing suggestive information about patient-treatment matching for a wide range of mental disorders where alternative treatment possibilities exist. Recent research on gender differences in response to different psychological and pharmacological treatments for depression provides an excellent illustration of the potential value of such research (Gorman, 2006; Kornstein et al., 2002; Thase et al., 2005). Opportunities exist to dovetail such studies with emerging studies of targeted pharmacological therapies keyed to the results of genetic tests (Malhotra et al., 2004). Promising developments in such studies would presumably lead to broader interest in expanded clinical epidemiological research on comorbidity and disorder sub-typing.

B. Clinical epidemiological studies of long-term course and treatment

All of the proposed investigations discussed in the last section implicitly focus on episode resolution. Clinical epidemiological research could also be very useful in studying the long-term course of mental illness. Little systematic information exists about long-term course of mental illness other than among patients with chronic psychosis, as so much treatment is short-term and based on acute episodic symptom relief. Most of our systematic knowledge about course of more common mental disorders comes from a small number of landmark long-term prospective clinical epidemiological studies, most notably the NIMH Collaborative Depression Study (Coryell et al., 2003; Mueller et al., 1999) and the Harvard/Brown Anxiety Disorders Research Program (Bruce et al., 2005; Rodriguez et al., 2005). Serious questions can be raised, though, about the extent to which information about illness course from these studies can be generalized due the fact that the patients in these studies were recruited from tertiary care facilities. The latter means that the study members over-represent refractive cases that typically had a higher quality of care than the majority of patients in the population. There could be great value in expanding the number and breadth of long-term clinical epidemiological studies to include broader and more representative samples of patients who were exposed to a more heterogeneous set of treatments.

One question of enormous importance to the study of long-term illness course is whether expansion of the use of maintenance medication or long-term psychological treatment would have a significant impact on the course of mental illness. Maintenance medication is the norm for some disorders, such as non-affective psychosis and bipolar disorder (Chou and Fazzio, 2006), and has been suggested as the best approach for other disorders, such as recurrent major depression (Lenze et al., 2002). Yet considerable uncertainty continues to exist about the effects of maintenance treatments on recurrence risk. Clinical epidemiological research in long-term prospective samples could make valuable contributions to reducing this uncertainty. The most obvious value of clinical epidemiological studies in this regard would be to provide descriptive information about actual patterns of illness course among patients with versus without maintenance therapy. Such studies might be even more valuable, though, in providing information about potentially modifiable environmental triggers of episode recurrence (Malkoff-Schwartz et al., 2000; Monroe et al., 2006).

V. Overview

I have tried to show in this review that psychiatric epidemiology involves far more than descriptive investigation of prevalence estimates in community surveys. Descriptive community epidemiology is, of course, important and is in many ways the foundation on which

more complex analytical and experimental investigations are build, but epidemiology goes well beyond simple description. Psychiatric epidemiology has focused on description in recent years because of the continuing debates that exist in the mental health field on what constitutes a “case” (Kendell and Jablensky, 2003; Kessler et al., 2003b; Regier et al., 1998; Wakefield and Spitzer, 2002). Descriptive community epidemiological studies can be useful in producing information about internal consistency, taxonicity, family aggregation, and other things that can be used to help make clinical decisions about the definition of a case (e.g., Blanchard et al., 2005; Vasey et al., 2005). Descriptive community epidemiological studies also can provide useful information about the implications of alternative decisions on how to define a case with regard to changes in prevalence and correlates (e.g., Hudson et al., 2006; Kessler et al., 2006b). Descriptive epidemiological data are, in fact, being heavily used in these ways to help support revisions of the DSM and ICD diagnostic systems. These are only first steps, though, in realizing the full potential of clinical value of psychiatric epidemiology.

The full clinical value of psychiatric epidemiology can only be realized when community epidemiology becomes integrated with clinical epidemiology and the latter takes on a more prominent role than it has up to now. The process of integration was started in the WMH surveys by building standard clinical measures of disorder severity into community assessments with the goal of creating a cross-walk with the results of clinical studies. Further integration is occurring in the various community experimental interventions that are beginning to emerge in response to epidemiological information about the enormous societal burden of mental illness. Although, as noted earlier in the paper, interventions focused on the workplace are currently the most common of these experiments, additional important interventions exist in other populations, such as school children (e.g., Furr-Holden et al., 2004) and residents of nursing homes (e.g., Chao et al., 2006).

Perhaps the most important areas of integration needed for future development of psychiatric epidemiology are naturalistic and quasi-experimental epidemiological studies of illness course and treatment response in clinical samples. The emerging recognition that mental disorders are often chronic-recurrent and require lifelong treatment rather than only acute care will inevitably increase clinical interest in research on course of illness. Advances in the development and use of electronic medical records will presumably make it easier to carry out this type of research than it has been in the past. To the extent that research on course of illness can be implemented in a way that cuts across particular clinical settings, opportunities will emerge for quasi-experimental evaluation of treatment effects using innovative paradigms similar to those described earlier in the paper based on research in other branches of medicine. When this occurs, the lessons of epidemiology will become much more clear and immediate than currently, resulting in a substantial increase in the clinical relevance of epidemiological research.

Acknowledgements

This paper was prepared in conjunction with the author's involvement as co-director of the WHO World Mental Health Survey Initiative, which is supported by the Alden Trust, AstraZeneca, Eli Lilly and Company, GlaxoSmithKline, the Robert Wood Johnson Foundation, the John D. and Catherine T. MacArthur Foundation, OrthoMcNeil, the Pan American Health Organization, the Pfizer Foundation, Pfizer, Inc., the US Public Health Service (R13MH066849, R01MH069864, R01DA016558), the National Institute of Mental Health (R01 MH070884), and Wyeth-Ayerst. More information about the WMH surveys can be obtained at www.hcp.med.harvard.edu/wmh. The author thanks his WMH colleagues Sergio Aguilar-Gaxiola, Jordi Alonso, Sing Lee, and Bedirhan Ustun for helpful comments on previous drafts of this paper.

REFERENCES

Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychological Bulletin 1987;101:213–232. [PubMed: 3562706]

Alegria, M.; Kessler, RC.; Bijl, R.; Lin, E.; Heeringa, S.; Takeuchi, DT.; Kolody, B. Comparing mental health service use data across countries. In: Andrews, G., editor. Unmet Need in Mental Health Service Delivery. Cambridge University Press; Cambridge: 2000. p. 97-118.

Andrade, L. Lifetime prevalence of mental disorders in a catchment area in São Paulo, Brazil; Paper presented at: VII Congress of the International Federation of Psychiatric Epidemiology; Santiago, Chile. August 1996; 1996.

Andrews G, Peters L, Guzman AM, Bird K. A comparison of two structured diagnostic interviews: CIDI and SCAN. Australian and New Zealand Journal of Psychiatry 1995;29:124–132. [PubMed: 7625961]

Angold A, Costello EJ. The Child and Adolescent Psychiatric Assessment (CAPA). Journal of the American Academy of Child and Adolescent Psychiatry 2000;39:39–48. [PubMed: 10638066]

Bijl RV, van Zessen G, Ravelli A, de Rijk C, Langendoen Y. The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. Social Psychiatry and Psychiatric Epidemiology 1998;33:581–586. [PubMed: 9857790]

Blanchard JJ, Horan WP, Collins LM. Examining the latent structure of negative symptoms: Is there a distinct subtype of negative symptom schizophrenia? Schizophrenia Research 2005;77:151–165. [PubMed: 15916881]

Bruce SE, Yonkers KA, Otto MW, Eisen JL, Weisberg RB, Pagano M, Shea MT, Keller MB. Influence of psychiatric comorbidity on recovery and recurrence in generalized anxiety disorder, social phobia, and panic disorder: A 12-year prospective study. American Journal of Psychiatry 2005;162:1179–1187. [PubMed: 15930067]

Brugha TS, Jenkins R, Taub N, Meltzer H, Bebbington PE. A general population comparison of the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). Psychological Medicine 2001;31:1001–1013. [PubMed: 11513368]

Buchanan AV, Weiss KM, Fullerton SM. Dissecting complex disease: The quest for the Philosopher's Stone? International Journal of Epidemiology 2006;35:562–571. [PubMed: 16540539]

Caraveo J, Martinez J, Rivera B. A model for epidemiological studies on mental health and psychiatric morbidity. Salud Mental 1998;21:48–57.

Chao SY, Liu HY, Wu CY, Jin SF, Chu TL, Huang TS, Clark MJ. The effects of group reminiscence therapy on depression, self esteem, and life satisfaction of elderly nursing home residents. Journal of Nursing Research 2006;14:36–45. [PubMed: 16547904]

Chou JC, Fazzio L. Maintenance treatment of bipolar disorder: Applying research to clinical practice. Journal of Psychiatric Practice 2006;12:283–299. [PubMed: 16998416]

Coryell W, Solomon D, Turvey C, Keller M, Leon AC, Endicott J, Schettler P, Judd L, Mueller T. The long-term course of rapid-cycling bipolar disorder. Archives of General Psychiatry 2003;60:914–920. [PubMed: 12963673]

Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Archives of General Psychiatry 2003;60:837–844. [PubMed: 12912767]

Cutler RB, Fishbain DA. Are alcoholism treatments effective? The Project MATCH data. BMC Public Health 2005;5:75. [PubMed: 16018798]

Edlund MJ, Wang PS, Berglund PA, Katz SJ, Lin E, Kessler RC. Dropping out of mental health treatment: patterns and predictors among epidemiological survey respondents in the United States and Ontario. American Journal of Psychiatry 2002;159:845–851. [PubMed: 11986140]

Faraone SV, Biederman J, Spencer T, Mick E, Murray K, Petty C, Adamson JJ, Monuteaux MC. Diagnosing adult attention deficit hyperactivity disorder: Are late onset and subthreshold diagnoses valid? American Journal of Psychiatry 2006;163:1720–1729. [PubMed: 17012682]

Fletcher, RH.; Fletcher, SW. Clinical Epidemiology: The Essentials. Lippincott Williams & Wilkins; Philadelphia: 2005.

Furr-Holden CD, Ialongo NS, Anthony JC, Petras H, Kellam SG. Developmentally inspired drug prevention: Middle school outcomes in a school-based randomized prevention trial. Drug and Alcohol Dependence 2004;73:149–158. [PubMed: 14725954]

Gadow KD, Weiss M. Attention-deficit/hyperactivity disorder in adults: Beyond controversy. Archives of General Psychiatry 2001;58:784–785. [PubMed: 11483145]

Gibbons RD, Hur K, Bhaumik DK, Mann JJ. The relationship between antidepressant medication use and rate of suicide. Archives of General Psychiatry 2005;62:165–172. [PubMed: 15699293]

Gibson TB, Mark TL, McGuigan KA, Axelsen K, Wang S. The effects of prescription drug copayments on statin adherence. American Journal of Managed Care 2006;12:509–517. [PubMed: 16961440]

Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. Journal of Child Psychology and Psychiatry 2000;41:645–655. [PubMed: 10946756]

Gorman JM. Gender differences in depression and response to psychotropic medication. Gender Medicine 2006;3:93–109. [PubMed: 16860269]

Greenberg, RS.; Daniels, SR.; Flanders, WD.; Eley, JW.; Boring, JRNM-H. Medical Epidemiology. Lange; Norwalk: 2000.

Gureje O, Lasebikan VO, Kola L, Makanjuola VA. Lifetime and 12-month prevalence of mental disorders in the Nigerian Survey of Mental Health and Well-Being. British Journal of Psychiatry 2006;188:465–471. [PubMed: 16648534]

Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36. [PubMed: 7063747]

Haro JM, Arbabzadeh-Bouchez S, Brugha TS, de Girolamo G, Guyer ME, Jin R, Lepine JP, Mazzi F, Reneses B, Vilagut Saiz G, Sampson NS, Kessler RC. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. International Journal of Methods in Psychiatric Research. in press

Helzer JE, Robins LN, McEvoy LT, Spitznagel EL, Stoltzman RK, Farmer A, Brockington IF. A comparison of clinical and diagnostic interview schedule diagnoses. Physician reexamination of lay-interviewed cases in the general population. Archives of General Psychiatry 1985;42:657–666. [PubMed: 4015307]

Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology 2006;17:360–372. [PubMed: 16755261]

Hudson JI, Hiripi E, Pope HG Jr. Kessler RC. The Prevalence and Correlates of Eating Disorders in the National Comorbidity Survey Replication. Biological Psychiatry. 2006

Hulshoff Pol HE, Hoek HW, Susser E, Brown AS, Dingemans A, Schnack HG, van Haren NEM, Ramos LMP, Gispen-de Wied CC, Kahn RS. Prenatal exposure to famine and brain morphology in schizophrenia. American Journal of Psychiatry 2000;157:1170–1172. [PubMed: 10873931]

Jick H, Kaye JA, Jick SS. Antidepressants and the risk of suicidal behaviors. JAMA 2004;292:338–343. [PubMed: 15265848]

Jordanova V, Wickramesinghe C, Gerada C, Prince M. Validation of two survey diagnostic interviews among primary care attendees: a comparison of CISR and CIDI with SCAN ICD-10 diagnostic categories. Psychological Medicine 2004;34:1013–1024. [PubMed: 15554572]

Jorm AF. National surveys of mental disorders: Are they researching scientific facts or constructing useful myths? Australian and New Zealand Journal of Psychiatry 2006;40:830–834. [PubMed: 16959008]

Judd LL, Akiskal HS. The prevalence and disability of bipolar spectrum disorders in the US population: re-analysis of the ECA database taking into account subthreshold cases. Journal of Affective Disorders 2003;73:123–131. [PubMed: 12507745]

Katz SJ, Kessler RC, Lin E, Wells KB. Medication management of depression in the United States and Ontario. Journal of General Internal Medicine 1998;13:77–85. [PubMed: 9502366]

Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. American Journal of Psychiatry 2003;160:4–12. [PubMed: 12505793]

Kendler, KS.; Prescott, CO. Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders. Guilford Press; New York: 2006.

Kessler RC. The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): Initial work and future directions -- the NAPE Lecture 1998. Nordic Association for Psychiatric Epidemiology. Acta Psychiatrica Scandinavica 1999;99:2–9. [PubMed: 10066001]

Kessler RC. Comments on the initial papers from Te Rau Hinengaro: The New Zealand Mental Health Survey. Australian and New Zealand Journal of Psychiatry 2006;40:823–826. [PubMed: 16959006]

Kessler RC, Adler L, Ames M, Barkley RA, Birnbaum H, Greenberg P, Johnston JA, Spencer T, Ustun TB. The prevalence and effects of adult attention deficit/hyperactivity disorder on work performance in a nationally representative sample of workers. Journal of Occupational and Environmental Medicine 2005a;47:565–572. [PubMed: 15951716]

Kessler RC, Adler L, Ames M, Demler O, Faraone S, Hiripi E, Howes MJ, Jin R, Secnik K, Spencer T, Ustun TB, Walters EE. The World Health Organization Adult ADHD Self-Report Scale (ASRS): A short screening scale for use in the general population. Psychological Medicine 2005b;35:245–256. [PubMed: 15841682]

Kessler RC, Akiskal HS, Ames M, Birnbaum H, Greenberg P, A RM, Jin R, Merikangas KR, Simon GE, Wang PS. Prevalence and effects of mood disorders on work performance in a nationally representative sample of U.S. workers. American Journal of Psychiatry 2006a;163:1561–1568. [PubMed: 16946181]

Kessler RC, Akiskal HS, Angst J, Deason-Guyer M, Hirschfeld RMA, Merikangas KR, Stang PE. Validity of the Assessment of Bipolar Spectrum Disorder in the WHO CIDI 3.0. Journal of Affective Disorders. in press-a

Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand SL, Manderscheid RW, Walters EE, Zaslavsky AM. Screening for serious mental illness in the general population. Archives of General Psychiatry 2003a;60:184–189. [PubMed: 12578436]

Kessler RC, Berglund P, Demler O, Jin R, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry 2005c;62:593–602. [PubMed: 15939837]

Kessler RC, Coccaro EF, Fava M, Jaeger S, Jin R, Walters E. The prevalence and correlates of DSM-IV intermittent explosive disorder in the National Comorbidity Survey Replication. Archives of General Psychiatry 2006b;63:669–678. [PubMed: 16754840]

Kessler RC, Greenberg PE, Mickelson KD, Meneades LM, Wang PS. The effects of chronic medical conditions on work loss and work cutback. Journal of Occupational and Environmental Medicine 2001a;43:218–225. [PubMed: 11285869]

Kessler RC, Haro JM, Heeringa SG, Pennell BE, Ustun TB. The World Health Organization World Mental Health Survey Initiative. Epidemiologia e Psichiatria Sociale. in press-b

Kessler RC, Merikangas KR, Berglund P, Eaton WW, Koretz DS, Walters EE. Mild disorders should not be eliminated from the DSM-V. Archives of General Psychiatry 2003b;60:1117–1122. [PubMed: 14609887]

Kessler, RC.; Mickelson, KD.; Barber, CB.; Wang, PS. The association between chronic medical conditions and work impairment. In: Rossi, AS., editor. Caring and Doing for Others: Social Responsibility in the Domains of Family, Work, and Community. University of Chicago Press; Chicago: 2001b. p. 403-426.

Kessler, RC.; Mroczek, DK.; Belli, RF. Retrospective adult assessment of childhood psychopathology. In: Shaffer, D.; Lucas, CP.; Richters, JE., editors. Diagnostic Assessment in Child and Adolescent Psychopathology. Guilford Press; New York: 1999. p. 256-284.

Kessler, RC.; Stang, PE. Health and Work Productivity: Making the Business Case for Quality Health Care. University of Chicago Press; Chicago: 2006.

Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). International Journal of Methods in Psychiatric Research 2004;13:93–121. [PubMed: 15297906]

Kessler, RC.; Wittchen, H-U.; Abelson, J.; Zhao, S. Methodological issues in assessing psychiatric disorder with self-reports. In: Stone, AA.; Turrkan, JS.; Bachrach, CA., et al., editors. The Science of Self-Report: Implications for Research and Practice. Lawrence Erlbaum Associates; Mahwah: 2000. p. 229-255.

Kessler RC, Wittchen H-U, Abelson JM, McGonagle KA, Schwarz N, Kendler KS, Knäuper B, Zhao S. Methodological studies of the Composite International Diagnostic Interview (CIDI) in the US National Comorbidity Survey. International Journal of Methods in Psychiatric Research 1998;7:33–55.

Kling JR, Liebman JB, Katz LF. Experimental analysis of neighborhood effects. Econometrica. in press

Kornstein SG, Sloan DM, Thase ME. Gender-specific differences in depression and treatment response. Psychopharmacology Bulletin 2002;36:99–112.

Kylyc, C. Mental Health Profile of Turkey: Main report. Ministry of Health Publications; Ankara: 1998.

Last, JM.; Abramson, JH.; Friedman, GD.; Porta, M.; Spasoff, RA.; Thuriaux, M. A Dictionary of Epidemiology. Oxford University Press; New York: 1995. 1995

Lenze EJ, Dew MA, Mazumdar S, Begley AE, Cornes C, Miller MD, Imber SD, Frank E, Kupfer DJ, Reynolds CF 3rd. Combined pharmacotherapy and psychotherapy as maintenance treatment for late-life depression: effects on social adjustment. American Journal of Psychiatry 2002;159:466–468. [PubMed: 11870013]

Malhotra AK, Murphy GM Jr. Kennedy JL. Pharmacogenetics of psychotropic drug response. American Journal of Psychiatry 2004;161:780–796. [PubMed: 15121641]

Malkoff-Schwartz S, Frank E, Anderson BP, Hlastala SA, Luther JF, Sherrill JT, Houck PR, Kupfer DJ. Social rhythm disruption and stressful life events in the onset of bipolar and unipolar episodes. Psychological Medicine 2000;30:1005–1016. [PubMed: 12027038]

McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA 1994;272:859–866. [PubMed: 8078163]

McDowall, D.; McCleary, R.; Meidinger, EE.; Hay, RA, Jr.. Interrupted Time Series Analysis. Sage; Beverly Hills: 1980.

Monroe SM, Torres LD, Guillaumot J, Harkness KL, Roberts JE, Frank E, Kupfer D. Life stress and the long-term treatment course of recurrent depression: III. Nonsevere life events predict recurrence for medicated patients over 3 years. Journal of Consulting and Clinical Psychology 2006;74:112–120. [PubMed: 16551148]

Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, Warshaw M, Maser JD. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. American Journal of Psychiatry 1999;156:1000–1006. [PubMed: 10401442]

Musgrove P. Judging health systems: reflections on WHO's methods. Lancet 2003;361:1817–1820. [PubMed: 12781555]

Offord, DR.; Boyle, M.; Campbell, D.; Cochrane, J.; Goering, PN.; Lin, E.; Rhodes, A.; Wong, M. Mental Health in Ontario: Selected Findings from the Mental Health Supplement to the Ontario Health Survey. Queen's Printer for Ontario; Toronto: 1994. 1994

Olfson M, Marcus SC, Shaffer D. Antidepressant drug therapy and suicide in severely depressed children and adults: A case-control study. Archives of General Psychiatry 2006;63:865–872. [PubMed: 16894062]

Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. British Medical Journal 2000;321:323–329. [PubMed: 10926586]

Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC, Norquist GS. Limitations of diagnostic criteria and assessment instruments for mental disorders. Implications for research and policy. Archives of General Psychiatry 1998;55:109–115. [PubMed: 9477922]

Robins LN, Helzer JE, Croughan JL, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics and validity. Archives of General Psychiatry 1981;38:381–389. [PubMed: 6260053]

Robins, LN.; Regier, DA. Psychiatric Disorders in America: the Epidemiologic Catchment Area Study. The Free Press; New York: 1991.

Rodriguez BF, Bruce SE, Pagano ME, Keller MB. Relationships among psychosocial functioning, diagnostic comorbidity, and the recurrence of generalized anxiety disorder, panic disorder, and major depression. Journal of Anxiety Disorders 2005;19:752–766. [PubMed: 16076422]

Rothman, KL.; Greenland, S. Modern Epidemiology. Lippincott Williams & Wilkins; Philadelphia: 1998.

Ruscio AM, Lane M, Roy-Byrne P, Stang PE, Stein DJ, Wittchen HU, Kessler RC. Should excessive worry be required for a diagnosis of generalized anxiety disorder? Results from the US National Comorbidity Survey Replication. Psychological Medicine 2005;35:1761–1772. [PubMed: 16300690]

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, Keller MB. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. Biological Psychiatry 2003;54:573–583. [PubMed: 12946886]

Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. Journal of the American Academy of Child and Adolescent Psychiatry 2000;39:28–38. [PubMed: 10638065]

Shear MK, Rucci P, Williams J, Frank E, Grochocinski V, Vander Bilt J, Houck P, Wang T. Reliability and validity of the Panic Disorder Severity Scale: Replication and extension. Journal of Psychiatric Research 2001;35:293–296. [PubMed: 11591432]

Shen YC, Zhang MY, Huang YQ, He YL, Liu ZR, Cheng H, Tsang A, Lee S, Kessler RC. Twelve-month prevalence, severity, and unmet need for treatment of mental disorders in metropolitan China. Psychological Medicine 2006;36:257–267. [PubMed: 16332281]

Clair D, Xu M, Wang P, Yu Y, Fang Y, Zhang F, Zheng X, Gu N, Feng G, Sham P, He L. Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959-1961. JAMA 2005;294:557–562. [PubMed: 16077049]

Stewart WF, Ricci JA, Chee E, Morganstein D. Lost productive work time costs from health conditions in the United States: Results from the American Productivity Audit. Journal of Occupational and Environmental Medicine 2003;45:1234–1246. [PubMed: 14665809]

Susser, E.; Schwartz, S.; Morabia, A.; Bromet, E. Psychiatric Epidemiology: Searching for the Causes of Mental Disorders. Oxford University Press; New York: 2006.

Thase ME, Entsuah R, Cantillon M, Kornstein SG. Relative antidepressant efficacy of venlafaxine and SSRIs: Sex-age interactions. Journal of Women's Health 2005;14:609–616.

Valuck RJ, Libby AM, Sills MR, Giese AA, Allen RR. Antidepressant treatment and risk of suicide attempt by adolescents with major depressive disorder: A propensity-adjusted retrospective cohort study. CNS Drugs 2004;18:1119–1132. [PubMed: 15581382]

Vasey MW, Kotov R, Frick PJ, Loney BR. The latent structure of psychopathy in youth: A taxometric investigation. Journal of Abnormal Child Psychology 2005;33:411–429. [PubMed: 16118989]

Wakefield JC, Spitzer RL. Lowered estimates--but of what? Archives of General Psychiatry 2002;59:129–130. [PubMed: 11825132]

Wang PS, Lane M, Kessler RC, Olfson M, Pincus HA, Wells KB. Twelve-month use of mental health services in the U.S.: Results from the National Comorbidity Survey Replication (NCS-R). Archives of General Psychiatry 2005;62:629–640. [PubMed: 15939840]

Wang PS, Simon G, Kessler RC. The economic burden of depression and the cost-effectiveness of treatment. International Journal of Methods in Psychiatric Research 2003;12:22–33. [PubMed: 12830307]

Weich S, Araya R. International and regional variation in the prevalence of common mental disorders: Do we need more surveys? British Journal of Psychiatry 2004;184:289–290. [PubMed: 15056570]

WHO International Consortium in Psychiatric Epidemiology. Cross-national comparisons of the prevalences and correlates of mental disorders. WHO International Consortium in Psychiatric Epidemiology. Bulletin of the World Health Organization 2000;78:413–426. [PubMed: 10885160]

Williams LK, Pladevall M, Xi H, Peterson EL, Joseph C, Lafata JE, Ownby DR, Johnson CC. Relationship between adherence to inhaled corticosteroids and poor outcomes among adults with asthma. Journal of Allergy and Clinical Immunology 2004;114:1288–1293. [PubMed: 15577825]

Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Regier D, Sartorius N. SCAN. Schedules for Clinical Assessment in Neuropsychiatry. Archives of General Psychiatry 1990;47:589–593. [PubMed: 2190539]

Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): A critical review. Journal of Psychiatric Research 1994;28:57–84. [PubMed: 8064641]

Wittchen H-U, Essau CA, von Zerssen D, Krieg CJ, Zaudig M. Lifetime and six-month prevalence of mental disorders in the Munich Follow-up Study. European Archives of Psychiatry and Clinical Neuroscience 1992;241:247–258. [PubMed: 1576182]

Wittchen HU, Kessler RC, Zhao S, Abelson JM. Reliability and clinical validity of UM-CIDI DSM-III-R generalized anxiety disorder. Journal of Psychiatric Research 1995;29:95–110. [PubMed: 7666382]

Wittchen HU, Zhao S, Abelson JM, Abelson JL, Kessler RC. Reliability and procedural validity of UM-CIDI DSM-III-R phobic disorders. Psychological Medicine 1996;26:1169–1177. [PubMed: 8931163]

World Health Organization. Composite International Diagnostic Interview (CIDI, Version 1.0). World Health Organization; Geneva: 1990.

Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. British Journal of Psychiatry 1978;133:429–435. [PubMed: 728692]