Diagnostic and genetic overlap of three common mental disorders in structured interviews and health registries

Fartein Ask Torvik (1,2), Eivind Ystrom (1,2,3), Kristin Gustavson (1,2), Tom Henrik Rosenström (1), Jørgen G. Bramness (4), Nathan Gillespie (5), Steven H. Aggen (5), Kenneth S. Kendler (5,6), & Ted Reichborn-Kjennerud (1,7)

Running title: Overlap between interviews and health registries

Affiliations:

(1) Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway; (2) Department of Psychology, University of Oslo, Norway; (3) PharmacoEpidemiology and Drug Safety Research Group, School of Pharmacy, University of Oslo, Norway; (4) Norwegian National Advisory Unit on Concurrent Substance Abuse and Mental Health Disorders, Brumunddal, Norway; (5) Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA; (6) Department of Human and Molecular Genetics and Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA; (7) Institute of Clinical Medicine, University of Oslo, Norway

Corresponding author: Fartein Ask Torvik, Norwegian Institute of Public Health

E-mail: fato@fhi.no, telephone (+47) 21078307

Abstract

Objective – To investigate whether diagnostic data from structured interviews, primary care and specialist care registries on major depressive disorder (MDD), anxiety disorders (AD), and alcohol use disorder (AUD) identify the same individuals, yield comparable comorbidity estimates, and reflect the same genetic influences.

Methods – Registry data from primary and specialist care were available for 11,727 twins and diagnostic interview data for 2,271 of these. We used logistic regression analyses and biometric modelling to investigate the overlap between the data sources.

Results – Most individuals meeting diagnostic criteria at interview were not registered with a corresponding diagnosis. The rates of registration were higher for MDD (36% in primary care and 15% in specialist care) and AD (21% and 18%) than for AUD (3% and 7%). Comorbidity estimated as odds ratios, but not polychoric correlations, were higher in the registries than in the interviews. Genetic influences on the disorders were highly correlated across data sources (median r = 0.81), bordering unity for MDD and AD.

Conclusion – Prevalence and comorbidity estimates differ between registries and population-based assessment. Nevertheless, diagnoses from health registries reflect the same genetic influences as common mental disorders assessed in the general population, indicating generalizability of etiological factors across data sources.

Keywords (MeSH) – Alcoholism; Anxiety disorders; Depressive Disorder; Genetics, Behavioral; Registries

Significant outcomes

- Genetic findings for major depressive disorder and anxiety disorders are likely to be generalizable between general population samples and data from health care registries.
- Genetic findings from primary and specialist care registries can be used interchangeably for major depressive disorder, anxiety disorders, and alcohol use disorder, but primary care registries usually include more cases and yield higher power.
- Prevalence and comorbidity estimates cannot always be generalized between the different types of samples, but comorbidity estimated as polychoric correlation was more generalizable than odds-ratio estimates.

Limitations

- The interviews are influenced by measurement error and concerns about proper cut-offs for defining caseness.
- The generalizability of the results may be limited to young adults and populations with culture and health care system as in Norway.
- We did not study the proportion of individuals that has ever received treatment, or that will eventually show up in the registries, but rather generalizability between different sources of information on three common mental disorders.

Introduction

Major depressive disorder (MDD), anxiety disorders (AD), and alcohol use disorder (AUD) are the most common mental disorders and inflict a large burden of disease (1). Use of data from health registries to study these disorders is becoming increasingly popular due to their availability, coverage, and large sample sizes. Registry data sets have in recent years provided novel insights into mental disorders, especially when they are augmented with kinship data (2-8). However, health registry studies are limited to individuals who have been in contact with health services, whereas many individuals with MDD (9-17), AD (12, 17-20), and AUD (17, 21) go untreated, according to self-reported data. This can limit the generalizability of findings from health registry studies. Population-based studies can avoid this limitation by contacting and interviewing individuals regardless of service use. However, such studies are costly and prone to selective and increasing non-response (22). Future research is therefore likely to become more dependent on the use of registry data. Some studies have evaluated the reliability of various diagnoses in health care registries versus clinical reassessment (23-25). Large-scale epidemiological studies on MDD, AD, and AUD typically rely on interviews or self-report, rather than clinical assessment. It is therefore important to clarify the degree to which one can generalize findings between registry data and population based samples.

Although common mental disorders are often untreated (9-21), the rates of treatment among individuals identified in diagnostic interviews are associated with uncertainty because the previous studies have relied on self-reported health service use. Recall bias is prevalent (27), and registry data can be used to obtain more precise estimates. In addition, symptom severity and comorbidity are the most important predictors of health service use (9, 11, 19, 28-32), and individuals with long-lasting disorders are more likely to be sampled (33). Elevated comorbidity resulting from the higher probability of being in treatment among individuals with two compared to one disorder is known as Berkson's bias (34, 35). It is therefore likely that health registries have higher rates of comorbidity than individuals in the general population. We are not aware of any studies that have estimated the magnitude of such possible biases in health registry data.

The most important question is, however, whether the same etiological factors are identified in different kinds of samples, that is, whether there are qualitative or only quantitative differences in causal factors. If separate heritable factors influence health service use, the correlation between genetic risk factors for a disorder assessed in the general population and in treatment will be lower than unity. In that case, results from molecular genetic studies relying on clinically recruited samples (36) may not generalize to the population at large. In addition, quantitative genetic studies based purely on registry data could be biased, because treatment seeking among depressed individuals is associated with having depressed relatives (37, 38). Alternatively, a high genetic correlation would indicate that the same etiological factors are identified by both the registry and population based samples, and that one can generalize findings from general population studies to patient populations and vice versa. This is supported by one study finding that co-twins had comparable risk of affective illness whether they were assessed via hospitalization or questionnaire (39). Otherwise, it is unknown whether etiological findings from one sample type can generalize to another.

To describe the generalizability between registry and population based samples, we have linked data from structured diagnostic interviews in a population-based twin sample with diagnoses given by the treating physicians from Norwegian primary and specialist care registries. This is the first study to use objective indicators of health service use and the first study to investigate whether the same etiological factors are identified for three common mental disorders in population-based samples, primary care, and specialist care.

Aims of the study

The aims were to investigate i) the level of primary and specialist health service utilization among individuals satisfying diagnostic criteria for major depressive disorder, anxiety disorders, and alcohol use disorder at interview, ii) the level of comorbidity in data from interviews, primary and specialist care registries, and iii) the degree to which the same genetic and environmental factors are identified in a) diagnostic interviews and b) primary and c) secondary health care registries for these disorders.

Materials and methods

Sample

The data for this study consist of registry data on 11,727 twins and diagnostic interview data on 2,271 of these. Twins were identified through the mandatory Norwegian Medical Birth Registry (MBR), established in 1967. Twins born in Norway between 1967 and 1991 were in 2013-2014 invited to be permanently registered in the Norwegian Twin Registry (NTR). In total 21,517 twins were invited. Among these, 433 had unknown address and 11,608 (53.9%) consented to registration and linkage to health registries. In addition, three twins born abroad self-recruited, and 116 twins participated in an interview study, but did not respond to the invitation to permanent registration. The full sample thus consists of 11,727 individuals (6,985 women, 59.6%) with an average age of 29.2 years (SD=7.6; range 16-40) in January 2008. The NTR is extensively described elsewhere (40).

The diagnostic interviews were carried out in a subsample born between 1967 and 1979. Between 1999 and 2004, psychiatric disorders were assessed at interview of 2,801 twins (44.4% response rate). We do not use these data here because the health registries were not person-identifiable at the time. Between October 2010 and November 2011, 2,758 of the responders were invited to a new wave of diagnostic interviews. After two written reminders and a final telephone contact to non-responders, 2,284 twins were interviewed (82.8% of the eligible). Of these, 2,272 (1,474 women) could be linked to registry data. The mean age of the respondents at the time of interview was 38.4 years (SD = 3.8, range 31 – 44).

Zygosity

Zygosity was determined by questionnaire items and genotyping. In the full sample, there were 1889 complete monozygotic (MZ) pairs, 2256 dizygotic (DZ) pairs, and 3,437 individuals in incomplete pairs. Among the interviewed twins, there were 510 complete MZ pairs, 469 DZ pairs, and 313 individuals in incomplete pairs.

Ethics

The study was approved by the Regional Ethical Committee for Medical and Health Research Ethics (case 2014/1527), and written informed consent was obtained from all participants.

Measures

Diagnostic interview. To assess ICD-10 diagnoses of MDD, AD, and AUD, we used a computerized Norwegian version of the Composite International Diagnostic Interview (CIDI) (41). This is a structured diagnostic interview developed by the World Health Organization (WHO) for the assessment of the DSM-IV and ICD-10 diagnoses. Interviewers were mostly senior clinical psychology graduate students, experienced psychiatric nurses, and clinical psychologists. All interviewers received a standardized training program administered by teachers certified by the WHO, and were closely supervised during the data collection process. All interviews were conducted by telephone, and different interviewers assessed each twin in a pair. The recency of symptoms was also reported. The CIDI assigns subthreshold diagnoses in cases where all but one of the criteria of the full disorder are met. In the present study, we used diagnoses of MDD, AD, and AUD that were currently present or with reported recency in or after 2008, which is the time from which we have complete registry data. AD included generalized anxiety disorder, social anxiety disorder, panic disorder, and agoraphobia. Specific phobia was not included because a majority (60.9%) of individuals with anxiety disorders had specific phobia only. Preliminary analyses suggested that specific phobia was weakly related to service use, and none of the participants were registered in specialist care with only specific phobia. Skipping patterns prevent rescoring of the interviews based on audiotaped records. Previous studies have shown that CIDI has good interrater (42) and test–retest reliability, with Kappa values of 0.68 for MDD, 0.81 for AD, and 0.78 for AUD (43).

Primary care. All persons who legally reside in Norway are members of the National Insurance Scheme (NIS) and assigned a general practitioner. General practitioners and other health service providers, such as emergency rooms, send billing information to NIS along with a diagnosis or reason for the visit in order to receive reimbursements. Due to economic incentives, it is unlikely that

health visits go unreported. Diagnostic information is coded according to the International Classification of Primary Care (ICPC-2) (44) and registered in the database Control and Payment of Health Reimbursements (CPHR). The ICPC-2 contains both diagnoses and complaints. In the present study, we calculated the number of visits that were registered with i) a code in the chapter P "Psychological", ii) "Depressive disorder" (P76), iii) "Anxiety disorder" or "Phobic/obsessive-compulsive disorder" (P74 and P79), and iv) "Chronic alcohol abuse" or "Acute alcohol abuse" (P15 and P16).

Specialist care. The Norwegian Patient Registry (NPR) includes information on all publicly funded in- and out-patient treatment in specialist care. Since January 1, 2008, information from this registry, including ICD-10 diagnoses (45), has become person-identifiable, and can therefore be used for research. Linkage between data sources has been made possible via the unique national identity number. In the present study, we calculated the number of entries that were registered with diagnoses in "ICD-10 Chapter V: Mental and behavioural disorders". For depression, we calculated the number of entries that were due to "Depressive episode" (F32) and "Recurrent depressive disorder" (F33); for anxiety entries due to F40 "Phobic anxiety disorders" or F41 "Other anxiety disorders"; and for alcohol use disorder entries due to F10 "Mental and behavioural disorders due to use of alcohol". Several diagnoses can be registered for each visit. On average, each visit had 1.3 diagnoses (SD=0.7), range 1-21.

Control variables. Information on sex and year and month of birth were available from MBR, whereas educational attainment was provided by Statistics Norway. Educational attainment was organized into four categories: Primary education only (10.1%), completed high school (40.1%), lower level higher education (36.7%), and master's degree or equivalent (13.1%).

Statistical analyses

In order to utilize the longest overlapping timeframe for the interview and registry data, we used diagnoses from the interviews with recency after 2008 and entries in the registries due to the various diagnoses from January 1, 2008 until the day of interview in 2010 or 2011. This way, we

ensure that interview and registry data cover the same episodes. The average follow-up time was 3.16 years (SD = 0.22 years; range 2.76 – 3.87 years). For individuals with only registry data, we used registry data covering 3.16 years, starting January 1, 2008.

For MDD, AD, and AUD assessed at interview we estimated the association with four outcomes: 1) The corresponding disorder (MDD, AD, or AUD) registered in primary care, 2) the corresponding disorder registered in specialist care, 3) all types of mental disorders and complaints registered in primary care, and 4) all types of mental disorders registered in specialist care. We estimated i) the proportion of individuals who were registered with each outcome among individuals who received a diagnosis at interview and ii) among individuals who did not receive an interview diagnosis. We also calculated iii) the proportion of individuals who got a diagnosis in the interview among those with a diagnosis in the registries and iv) among those with no diagnosis in the registries. These numbers can be used to characterise the fidelity with which one can translate from one type of assessment to the other. If one wants to translate from registries to interviews, the proportion registered among individuals with a disorder according to the interview is the sensitivity of the registries, whereas the proportion registered among those without the disorder according to the interview is the false discovery rate of the registries, or 1 – specificity. If one wants to translate from interviews to registries, interview diagnoses among registered and unregistered would correspond to the sensitivity and false positive rates of the interviews, respectively.

We further present the observed phenotypic polychoric correlations and odds ratios for being registered if one was diagnosed in the interview. We investigated comorbidity in the interview data, primary care, and specialist care by estimating polychoric correlations and odds ratios between MDD, AD, and AUD in the three data sources. We obtained odds ratios by fitting logistic regression analyses and adjusted for statistical dependence between twins using generalized estimating equations.

In the description of overlap, all variables are coded as dichotomous, i.e. a person is either registered or not registered with a disorder. When estimating polychoric correlations in the twin

analyses, we used the counts of visits as ordinal variables with five categories in order to increase statistical power. The first category indicated zero diagnostic entries, and the cut-offs between the remaining categories were set so that each included approximately the same number of individuals. We also included subthreshold diagnoses as intermediate categories on the interview based AD and AUD variables.

We investigated whether the disorders as assessed by interviews and recorded in the registries reflected the same genetic and environmental risk factors by using standard twin methods (46) for ordinal data (47). Variance in traits are assumed to arise from three latent sources: additive genetic factors (A), which MZ twins share 100% and DZ twins 50%; common or shared environmental factors (C), which contribute equally to twin similarity among MZs and DZs; and individual-specific environmental factors (E), which contribute to differences between twins and include measurement error.

We ran three sets of twin analyses, one for MDD assessed by interview, primary care and specialist care, one model for ADs assessed by these same three methods, and one for AUD again assessed by these three methods. We used the Cholesky decomposition to obtain estimates of the correlations between genetic influences on the three measures of each disorder, and similarly for the environmental influences. We tested the presence of qualitative and quantitative sex differences (48), sex differences in thresholds, and the significance of A and C factors. Finally, we tested whether the correlations between the genetic influences on a disorder across the three data sources could be set to unity, and whether the disorders were equally heritable in data from the three sources. We determined goodness of fit using likelihood ratio chi-square tests and by comparing the Akaike information criterion (AIC). By the principle of parsimony, models with the lowest AIC were preferred (49).

Results

According to the diagnostic interview data, the lifetime prevalence of MDD was 17.8%, of AD 9.6%, and of AUD 6.6%. Between 2008 and the time of interview, 11.0% was diagnosed with MDD,

7.0% with AD, and 4.0% with AUD, according to interview data. In addition, 13.1% had subthreshold AD and 7.4% had subthreshold AUD. The health registries had excellent coverage – 99.1% of individuals had at least one registry entry since the registries were started, 43.8% with at least one psychological entry. During the overlapping timeframe of the interviews and registries, 93.3% of the sample had one or more registry entries in primary care for any complaint or disorder, and 66.4% had entries in specialist care for any disorder, whereas 23.0% were registered in primary care with a mental disorder or complaint and 8.0% in specialist care with a mental disorder. Table 1 shows the distribution of interview and registry variables and the frequency of different combinations of each diagnosis at interview and in registries.

--- Insert Table 1 approximately here ---

Correspondence between interview and health registries

Table 2 shows the correspondence between interview and health registry data for MDD, ADs, and AUD assessed during the overlapping timeframe. A majority of individuals with MDD or ADs according to the interview data had consulted primary care for a mental disorder or complaint, but most of these did not receive the same diagnosis as in the interview. Most were not in contact with specialist care. The discrepancy between primary and specialist care was lower for AD than for MDD. Interview-based AUD was associated with being registered with mental disorders but not as strongly as MDD or AD. The vast majority of individuals with AUD according to interview data did not show up in the health registries with any alcohol related diagnoses. Whereas neither data source can be considered an objective assessment, these numbers indicate that the registries have moderate sensitivity as indicator of interview-based diagnoses, but excellent specificity because of the few false positives. On the other hand, the interviews only identified roughly half of individuals diagnosed with a particular disorder in specialist care, which would indicate moderate sensitivity of the interview. Moderate proportions of unregistered individuals were diagnosed in the interview.

--- Insert Table 2 approximately here ---

The correlations and ORs indicate that there is high correspondence between diagnoses assessed at interview and in the two registries, although many individuals did not show up in the registries. MDD and AD were also good indicators of the tendency to get registered due to any mental disorder. Although individuals with MDD or AD were more often in contact with primary care than with specialist care, diagnoses from the two lines of treatment were equally strongly related to interview-based diagnoses.

Comorbidity

Table 3 shows the associations between the three disorders the three data sources. P-values for the differences in correlations, crude, and adjusted log odds are available in the supplement and Table S1. The correlation between MDD and AD was not higher in either of the health registries compared to the population sample, although it was higher in specialist than in primary care (p = 0.002). The correlations between MDD and AUD could be set to equal in across data sources (p = 0.135), but not for AD and AUD (p = 0.013). Polychoric correlations are, unlike odds ratios, theoretically independent of specific thresholds of severity for setting a diagnosis on a disorder continuum. In primary care, we found higher ORs between MDD (p = 0.031) or AD (p = 0.018) on one hand and AUD on the other, compared to the interviews. All ORs were significantly elevated in specialist care compared to the interviews ($p_{MDD\text{-}AD}$ = 0.010; $p_{MDD\text{-}AUD}$ = 0.001; $p_{AD\text{-}AUD}$ < 0.001).

--- Insert Table 3 approximately here ---

Genetic risk factors

Model fit indices for the biometric models are presented in Table S2-S4. We found no evidence for any qualitative or quantitative sex differences, or for effects of shared environmental factors. Table 4 summarizes the genetic and environmental correlations between diagnoses based on the three sources of diagnoses and heritability estimates. Figure S1 shows the models underlying these numbers. The genetic correlations ranged from 0.30 to 1.00, with a median of 0.81.

For MDD, genetic correlations between interview, primary and specialist care were all high (0.80 or above). The genetic correlation between diagnoses at interview and in specialist care, which

both use the ICD system, was 1.00. Setting all three genetic correlations to unity or constraining the heritabilities to be equal in the three data sources slightly worsened the model fit.

--- Insert Table 4 approximately here ---

For AD, genetic correlations ranged from 0.81 to 0.93. Setting all three genetic correlations to unity and the heritabilities to be equal (h^2 = 0.44) improved the model fit.

For AUD, genetic correlations between diagnoses based on interview data and registry data were estimated at 0.30. Since a low proportion of individuals with AUD were registered, the confidence intervals were large and consistent with a range of point values. The genetic correlation between diagnoses based on the two registries was estimated at 1.00, with high accuracy. The model fit improved when we set all genetic correlations to unity and the heritability of AUD to be the same in all three data sources (h^2 = 0.45).

For all three disorders, the non-shared environmental correlations were relatively high, ranging from 0.36 to 0.90, with a median of 0.55. One must consider these correlations high, because this source of phenotypic variance includes measurement error.

Discussion

The present study demonstrates that whether MDD, AD, or AUD are measured in interviews, primary care registries, or in specialist health care registries, the genetic risk factors identified are largely overlapping across data source. Hence, etiological factors are likely to be largely generalizable across mode of assessment. This is also the study first to demonstrate that a substantial proportion of individuals with one or more of three common mental disorders according to interview go untreated according to registries, and that the rates of comorbidity between these disorders are higher in health registries than in a population-based sample.

Diagnoses based on interview, primary care, and specialist care data primarily reflected the same genetic factors. For MDD, all three types of information showed strong genetic correlations (0.80 or above). The genetic correlation between primary and specialist care could not be set to unity, possibly due to differing diagnostic manuals and practices. For AD, the genetic factors of all

three measures correlated at unity in the best fitting model, and even the lower bounds of the confidence intervals indicated high genetic correlations. The low proportion of registrations among individuals with AUD resulted in low statistical power. The genetic correlation between interview and registry data could be set to unity, but the data were consistent with a range of values. Statistical power was higher for the genetic correlation between AUD primary and specialist care, which was estimated at unity. If separate genetic factors influence service use, besides those that influence interview assessed disorders, the genetic correlation between interview and registries will be below unity. Our finding suggest that there are small or no effects of such genetic factors for MDD and AD, and possibly also for AUD. Our findings imply that genetic studies on these disorders are likely to identify the same genetic risk factors regardless of sample type. Thus, registry studies of genetic risk factors and molecular genetic findings from clinical samples are likely to be generalizable to the general population that includes less severe and less comorbid cases. Likewise, studies of genetic risk factors based on non-clinical samples, such as twin samples, are likely to be generalizable to clinical settings.

Our study is the first to examine the genetic and environmental overlap between data obtained via personal interviews and national health registries, and should be viewed as a starting point for further investigations. Primary and specialist care data had very high or unity genetic correlations for all three disorders. Genetic findings from primary and specialist care registries can therefore be used interchangeably, but primary care registries are likely to have higher power due to the higher number of cases. As most previous health registry studies have used specialist care data, future studies could benefit by utilizing primary care data.

AD and AUD had equal heritability across all three data sources, whereas MDD showed a lower heritability in interview than in registry data. This should be interpreted with caution, because the fit of the models with equal and varying heritability were close. All the environmental correlations were high, although repeated measures would be required to separate the effects of environmental factors influencing service use from random measurement error.

Previous interview studies have found low rates of treatment using self-report data (9-21, 32, 51). We confirmed this using registry data. A majority of individuals with MDD, AD, or AUD, as identified in diagnostic interviews, were not registered with the corresponding diagnosis in primary care and only a small fraction were registered with the corresponding diagnosis in specialist care. This was true for both MDD and AD, and to a greater extent for AUD. These low rates of formal diagnosis occurred despite very high rates of overall contact with the health services, and despite that a majority of individuals identified as having any of these mental disorders reported psychological symptoms to primary care providers.

Neither the interview nor the registries can be considered a "gold standard", but the different assessment methods have different strengths. Many individuals with common mental disorders never receive treatment, for a variety of reasons including attitudinal barriers such as a perceived lack of efficacy (52). These individuals can only be identified in population-based samples. However, the diagnostic criteria could include individuals with low levels of clinical impairment (53) and could be considered too low for identifying individuals in need of treatment (54). False positives in the interview would lead to an overestimation of unmet needs, but the 'true' overlap between etiological factors for interview and registries would be higher than our estimates. The interviews are influenced by measurement error (58), but were conducted using standardized methods. The clinicians, on the other hand, can have varying practices for diagnosing and referring patients. In addition, pragmatic issues and misreporting of symptoms could influence diagnostic practices, particularly in primary care, and lead to discrepancy between interview and registry data. For example, a patient may 'require' a diagnosis to be granted a sick leave. The expert assessment in specialist care presumably have considerably fewer false positives, but rather many false negatives, because only a small proportion of the sample were ever evaluated in specialist care. The interview's ability to identify the same diagnosis as specialist care in roughly half of the cases may indicate lack of precision in the interview. The large potential for false negatives in the registries imply that they can not be used alone to estimate prevalence for these disorders, but they may prove useful as

weighting variables in future prevalence studies. Yet despite the differences in ascertainment, the polychoric correlations, which are independent of threshold or prevalence, were very high between all three data sources, indicating that the interview and health registries index the same phenomena including the genetic and environmental risk factors, and that general practitioners and specialists are largely diagnosing the same conditions.

The comorbidity between MDD or AD with AUD, expressed as ORs, increased for each level of care, that is, from interviews to primary care and further to specialist care. The OR between MDD and AD was higher only in specialist care. The comorbidity estimates may be artificially inflated in the registries, even when the health personnel diagnose their patients correctly. Comorbidity is an important predictor of entering the formal treatment system (9, 11, 19, 28-31). If other personal characteristics predict service use independent of disorder status (55), that would further enhance estimated comorbidity. In addition, the duration of a disorder is related to severity, and thereby comorbidity, and proportional to the probability that it will appear in a clinical setting (33). Berkson (34) described how sampling into clinical settings can increase rates of comorbidity (35). This effect was less pronounced for comorbidity expressed as polychoric correlations, probably because polychoric correlations are theoretically unaffected by thresholds, or severity. The higher rates of comorbidity in health registry data imply that comorbidity estimates obtained from registry data cannot be generalized to the population when they are expressed as ORs, but that polychoric correlations are more generalizable. Whereas this finding is limited to the three included disorders, similar processes can influence all comorbidity rates derived from registry data.

Limitations

Although we have used health registry data with excellent coverage and diagnostic interview data in a population based twin-sample, some limitations are noteworthy. First, the sample was based on informed consent, and thus subject to non-response and possibly associated biases. Mental disorders are typically associated with lower response rates in health studies (22). Previous analyses on Norwegian twin data have shown that participation was predicted by female sex, monozygosity

and higher educational status, but not statistically significantly by symptoms of psychiatric disorders or substance abuse (56). Non-response can reduce statistical power and bias prevalence estimates. However, estimates of associations between variables are more robust (57). The findings concerning diagnostic and genotypic associations between different data sources are therefore less likely to be biased. Second, ICPC-2 used in primary care relies on broader diagnostic categories than the ICD-10 used in the interviews and in specialist care. Thus, obsessive-compulsive disorder and specific phobia were included only in the primary care AD variable. Third, neither of our data sources provide an optimal assessment of mental disorders. The interviews are influenced by measurement error, problems with accurate recall (58), and concerns about proper cut-offs for defining caseness (54), whereas the health registries are likely to miss many true cases that have not been in treatment. The high genetic correlations between the different assessments indicate high generalizability between genetic studies using the three different sources of information, despite these differences in assessment. Fourth, the generalizability of the results may be limited to populations with similar age, culture, and health care systems as in Norway. Nevertheless, the lifetime prevalences in our interviews were similar to those observed in National Comorbidity Study Replication in USA (9). Fifth, we did not study the proportion of individuals that has ever received treatment, or that eventually will show up in the registries, but rather generalizability between different methods when the disorders are measured over a fixed time frame. The results may be different for other mental disorders and timeframes than those examined here.

In conclusion, only a minority of individuals who met criteria for MDD, AD, and AUD according to interview data were registered with a corresponding diagnoses in health registries. Rates of comorbidity between these three disorders can be elevated in health registries. Despite these differences, the genetic risk factors for MDD, AD, and AUD seems to be generalizable between population-based samples and health registries.

Acknowledgements

Kristian Tambs, deceased 18 June 2017, contributed with critical and constructive comments to an early version of this manuscript. This project was supported by Research Council of Norway (RCN) (grant 240061). Previous collections and analyses of twin data from this project were in part supported by the RCN (196148) and the National Institutes of Health (grant MH-068643), and the National Institute on Drug Abuse (1R01DA037558-01A1).

References

1. GBD 2015 DALYS AND HALE COLLABORATORS. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;**388**:1603-58.

2. CHANG Z, LARSSON H, LICHTENSTEIN P, FAZEL S. Psychiatric disorders and violent reoffending: a national cohort study of convicted prisoners in Sweden. Lancet Psychiat 2015;**2**:891-900.

3. CHANG Z, LICHTENSTEIN P, LARSSON H, FAZEL S. Substance use disorders, psychiatric disorders, and mortality after release from prison: a nationwide longitudinal cohort study. Lancet Psychiat 2015;**2**:422-30.

4. FAZEL S, WOLF A, CHANG Z, LARSSON H, GOODWIN GM, LICHTENSTEIN P. Depression and violence: a Swedish population study. Lancet Psychiat 2015;**2**:224-32.

5. NORDSLETTEN AE, LARSSON H, CROWLEY JJ, ALMQVIST C, LICHTENSTEIN P, MATAIX-COLS D. Patterns of nonrandom mating within and ccross 11 major psychiatric disorders. JAMA Psychiatry 2016;**73**:354-61.

6. PETTERSSON E, LARSSON H, LICHTENSTEIN P. Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population. Mol Psychiatr 2016;**21**:717-21.

7. POWER RA, KYAGA S, UHER R, et al. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. JAMA Psychiatry 2013;**70**:22-30.

8. SONG J, BERGEN SE, KUJA-HALKOLA R, LARSSON H, LANDÉN M, LICHTENSTEIN P. Bipolar disorder and its relation to major psychiatric disorders: a family-based study in the Swedish population. Bipolar Disord 2015;**17**:184-93.

9. KESSLER RC, BERGLUND P, DEMLER O, et al. The epidemiology of major depressive disorder - Results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003;**289**:3095-105.

10. KESSLER RC, MERIKANGAS KR, WANG PS. Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the twenty-first century. Ann Rev Clin Psychol 2007;**3**:137-58.

11. HAMALAINEN J, ISOMETSA E, LAUKKALA T, et al. Use of health services for major depressive episode in Finland. J Affect Disorders 2004;**79**:105-12.

12. RONESS A, MYKLETUN A, DAHL AA. Help-seeking behaviour in patients with anxiety disorder and depression. Acta Psychiat Scand 2005;**111**:51-8.

13. WANG PS, LANE M, OLFSON M, PINCUS HA, WELLS KB, KESSLER RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiat 2005;**62**:629-40.

14. BYERS AL, AREAN PA, YAFFE K. Low use of mental health services among older Americans with mood and anxiety disorders. Psychiat Serv 2012;**63**:66-72.

15. OLFSON M, BLANCO C, MARCUS SC. Treatment of adult depression in the United States. JAMA Intern Med 2016;**176**:1482-91.

16. THORNICROFT G, CHATTERJI S, EVANS-LACKO S, et al. Undertreatment of people with major depressive disorder in 21 countries. Brit J Psychiat 2016.

17. ALONSO J, ANGERMEYER MC, BERNERT S, et al. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiat Scand 2004;**109**:47-54.

18. KESSLER RC, KELLER MB, WITTCHEN H-U. The epidemiology of generalized anxiety disorder. Psychiatr Clin North Am 2001;**24**:19-39.

19. RUSCIO AM, BROWN TA, CHIU WT, SAREEN J, STEIN MB, KESSLER RC. Social fears and social phobia in the USA: results from the National Comorbidity Survey Replication. Psychol Med 2008;**38**:15-28.

20. JOHNSON EM, COLES ME. Failure and delay in treatment-seeking across anxiety disorders. Community Ment Hlt J 2013;**49**:668-74.

21. HASIN DS, STINSON FS, OGBURN E, GRANT BF. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States - Results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiat 2007;**64**:830-42.

22. TORVIK FA, ROGNMO K, TAMBS K. Alcohol use and mental distress as predictors of non-response in a general population health survey: the HUNT study. Soc Psych Psych Epid 2012;**47**:805-16.

23. BYRNE N, REGAN C, HOWARD L. Administrative registers in psychiatric research: a systematic review of validity studies. Acta Psychiat Scand 2005;**112**:409-14.

24. DAVIS KAS, SUDLOW CLM, HOTOPF M. Can mental health diagnoses in administrative data be used for research? A systematic review of the accuracy of routinely collected diagnoses. BMC Psychiatry 2016;**16**.

25. OIESVOLD T, NIVISON M, HANSEN V, SKRE I, OSTENSEN L, SORGAARD KW. Diagnosing comorbidity in psychiatric hospital: challenging the validity of administrative registers. BMC Psychiatry 2013;**13**.

26. SMOLLER JW. The use of electronic health records for psychiatric phenotyping and genomics. Am J Med Genet B Neuropsychiatr Genet 2017.

27. MOFFITT TE, CASPI A, TAYLOR A, et al. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. Psychol Med 2010;**40**:899-909.

28. DEW MA, DUNN LO, BROMET EJ, SCHULBERG HC. Factors affecting help-seeking during depression in a community sample. J Affect Disorders 1988;**14**:223-34.

29. OLFSON M, KLERMAN GL. Depressive symptoms and mental health service utilization in a community sample. Soc Psych Psych Epid 1992;**27**:161-7.

30. TEN HAVE M, DE GRAAF R, VOLLEBERGH W, BEEKMAN A. What depressive symptoms are associated with the use of care services? Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). J Affect Disorders 2004;**80**:239-48.

31. WANG PS, AGUILAR-GAXIOLA S, ALONSO J, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. The Lancet 2007;**370**:841-50.

32. WANG PS, BERGLUND P, OLFSON M, PINCUS HA, WELLS KB, KESSLER RC. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. Arch Gen Psychiat 2005;**62**:603-13.

33. COHEN P, COHEN J. The clinician's illusion. Arch Gen Psychiat 1984;**41**:1178-82.

34. BERKSON J. Limitations of the application of fourfold table analysis to hospital data. Biometrics 1946;**2**:47-53.

35. ROBERTS RS, SPITZER WO, DELMORE T, SACKETT DL. An empirical demonstration of Berkson's bias. J Chronic Dis 1978;**31**:119-28.

36. SULLIVAN PF, DALY MJ, RIPKE S, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatr 2013;**18**:497-511.

37. KENDLER KS. Is seeking treatment for depression predicted by a history of depression in relatives? Implications for family studies of affective disorder. Psychol Med 1995;**25**:807-14.

38. PROKOFYEVA E, MARTINS SS, YOUNES N, SURKAN PJ, MELCHIOR M. The role of family history in mental health service utilization for major depression. J Affect Disorders 2013;**151**:461-6.

39. KENDLER KS, PEDERSEN N, JOHNSON L, NEALE MC, MATHE AA. A Pilot Swedish Twin Study of Affective-Illness, Including Hospital-Ascertained and Population-Ascertained Subsamples. Arch Gen Psychiat 1993;**50**:699-706.

40. NILSEN TS, BRANDT I, CZAJKOWSKI N, et al. The Norwegian twin registry. Nor Epidemiol 2016;**26**:19-27.

41. WITTCHEN H-U, PFISTER H. DIA-X interviews (M-CIDI). Frankfurt, Germany: Swets & Zeitlinger; 1997.

42. WITTCHEN HU. Reliability and Validity Studies of the Who Composite International Diagnostic Interview (CIDI): a critical review. J Psychiat Res 1994;**28**:57-84.

43. WITTCHEN HU, LACHNER G, WUNDERLICH U, PFISTER H. Test-retest reliability of the computerized DSM-IV version of the Munich Composite International Diagnostic Interview (M-CIDI). Soc Psych Psych Epid 1998;**33**:568-78.

44. WORLD ORGANIZATION OF NATIONAL COLLEGES ACADEMIES. ICPC-2-R: International classification of primary care. Oxford ; New York: Oxford University Press; 2005.

45. WORLD HEALTH ORGANIZATION. ICD-10: International statistical classification of diseases and related health problems. Geneva: World Health Organization; 1992.

46. MARTIN NG, EAVES LJ. The genetical analysis of covariance structure. Heredity 1977;**38**:79-95.

47. FALCONER DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. Ann Hum Genet 1965;**29**:51-76.

48. NEALE MC, ROYSAMB E, JACOBSON K. Multivariate genetic analysis of sex limitation and G x E interaction. Twin Res Hum Genet 2006;**9**:481-9.

49. AKAIKE H. Factor-analysis and AIC. Psychometrika 1987;**52**:317-32.

50. MAES HH, NEALE MC, OHLSSON H, et al. A bivariate genetic analysis of drug abuse ascertained through medical and criminal registries in Swedish twins, siblings and half-siblings. Behav Genet 2016;**46**:735-41.

51. WANG PS, ANGERMEYER M, BORGES G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007;**6**:177-85.

52. SAREEN J, JAGDEO A, COX BJ, et al. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. Psychiat Serv 2007;**58**:357-64.

53. ROSENSTROM T, JOKELA M. Reconsidering the definition of major depression based on collaborative psychiatric epidemiology surveys. J Affect Disorders 2017;**207**:38-46.

54. REGIER DA, KAELBER CT, RAE DS, et al. Limitations of diagnostic criteria and assessment instruments for mental disorders - Implications for research and policy. Arch Gen Psychiat 1998;**55**:109-15.

55. TEN HAVE M, OLDEHINKEL A, VOLLEBERGH W, ORMEL J. Does neuroticism explain variations in care service use for mental health problems in the general population? Soc Psych Psych Epid 2005;**40**:425-31.

56. TAMBS K, RONNING T, PRESCOTT CA, et al. The Norwegian Institute of Public Health twin study of mental health: examining recruitment and attrition bias. Twin Res Hum Genet 2009;**12**:158-68.

57. NILSEN RM, VOLLSET SE, GJESSING HK, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. Paediatr Perinat Epidemiol 2009;**23**:597-608.

58. YSTROM E, REICHBORN-KJENNERUD T, AGGEN SH, KENDLER KS. Alcohol dependence in men: reliability and heritability. Alcohol Clin Exp Res 2011;**35**:1716-22.

Table 1. Frequencies and cross-tabulated frequencies of three common mental disorders assessed at interview and the same and all mental disorders recorded by national primary and specialist care registries.

	Registered in full sample (%)	Registered in interviewed subsample (%)	Interview positive, registry positive	Interview negative, registry positive	Interview positive, registry negative	Interview negative, registry negative
Major depressive disorder (n=249; 11.0%)						
Primary, depression	815 (6.9%)	184 (8.1%)	89 (3.9%)	95 (4.2%)	160 (7.0%)	1927 (84.9%)
Specialist, depression	328 (2.8%)	63 (2.8%)	36 (1.6%)	27 (1.2%)	213 (9.4%)	1995 (87.8%)
Primary, all mental	2698 (23.0%)	578 (25.5%)	173 (7.6%)	405 (17.8%)	76 (3.3%)	1617 (71.2%)
Specialist, all mental	935 (8.0%)	186 (8.2%)	80 (3.5%)	106 (4.7%)	169 (7.4%)	1916 (84.4%)
Anxiety disorders (n=159; 7.0%)						
Primary, anxiety	359 (3.1%)	74 (3.3%)	34 (1.5%)	40 (1.8%)	125 (5.5%)	2072 (91.2%)
Specialist, anxiety	265 (2.3%)	56 (2.5%)	28 (1.2%)	28 (1.2%)	131 (5.8%)	2084 (91.8%)
Primary, all mental	2698 (23.0%)	578 (25.5%)	103 (4.5%)	475 (20.9%)	56 (2.5%)	1637 (72.1%)
Specialist, all mental	935 (8.0%)	186 (8.2%)	64 (2.8%)	122 (5.4%)	95 (4.2%)	1990 (87.6%)
Alcohol use disorder (n=90; 4.0%)						
Primary, alcohol	75 (0.6%)	9 (0.4%)	3 (0.1%)	6 (0.3%)	87 (3.8%)	2175 (95.8%)
Specialist, alcohol	61 (0.5%)	10 (0.4%)	6 (0.3%)	4 (0.2%)	84 (3.7%)	2177 (95.9%)
Primary, all mental	2698 (23.0%)	578 (25.5%)	36 (1.6%)	542 (23.9%)	54 (2.4%)	1639 (72.2%)
Specialist, all mental	935 (8.0%)	186 (8.2%)	17 (0.7%)	169 (7.4%)	73 (3.2%)	2012 (88.6%)

Notes: Data from 2008 until time of interview; average follow up time of 3.16 years; Primary, mental = P-chapter of ICPC-2; Specialist, mental = F-chapter of ICD-10; Primary, depression = P76 of ICPC-2; Specialist, depression = F32, F33 of ICD-10; Primary, anxiety = P74 and P79 of ICPC-2; Specialist, anxiety = F40, F41 of ICD-10; Primary, alcohol = P15 and P16 of ICPC-2; Specialist, alcohol = F10 of ICD-10; ICPC-2 = International Classification of Primary Care, second version; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, tenth version.

Table 2. Associations between three common mental disorders assessed at interview and the same and all mental disorders recorded by national primary and specialist care registries.

	Registration among interview positive	Registration among interview negative	Interview positive among registered	Interview positive among unregistered	r	OR	AOR
Major depressive disorder (n=249; 11.0%)							
Primary, depression	35.7%	4.7%	48.4%	7.7%	0.65 (0.58, 0.72)	11.3 (8.0, 15.8)	9.9 (7.0, 14.0)
Specialist, depression	14.5%	1.3%	57.1%	9.6%	0.58 (0.48, 0.69)	12.5 (7.4, 20.9)	11.2 (6.5, 19.0)
Primary, all mental	69.5%	20.0%	29.9%	4.5%	0.63 (0.58, 0.69)	9.1 (6.8, 12.1)	8.1 (6.1, 10.9)
Specialist, all mental	32.1%	5.2%	43.0%	8.1%	0.58 (0.51, 0.66)	8.6 (6.2, 11.7)	7.4 (5.4, 10.3)
Anxiety disorders (n=159; 7.0%)							
Primary, anxiety	21.4%	1.9%	45.9%	5.7%	0.63 (0.54, 0.73)	14.1 (8.5, 23.3)	12.1 (6.8, 21.6)
Specialist, anxiety	17.6%	1.3%	50.0%	5.9%	0.62 (0.52, 0.73)	15.9 (9.2, 27.6)	11.2 (6.2, 20.2)
Primary, all mental	64.8%	22.5%	17.8%	3.3%	0.54 (0.47, 0.61)	6.3 (4.5, 9.0)	5.2 (3.6, 7.4)
Specialist, all mental	40.3%	5.8%	34.4%	4.6%	0.63 (0.55, 0.70)	11.0 (7.5, 16.1)	8.9 (6.0, 13.2)
Alcohol use disorder (n=90; 4.0%)							
Primary, alcohol	3.3%	0.3%	33.3%	3.8%	0.47 (0.20, 0.74)	12.5 (3.1, 50.6)	-
Specialist, alcohol	6.7%	0.2%	60.0%	3.7%	0.69 (0.51, 0.87)	38.9 (12.3, 123.3)	-
Primary, all mental	40.0%	24.9%	6.2%	3.2%	0.21 (0.09, 0.32)	2.0 (1.3, 3.1)	2.7 (1.7, 4.2)
Specialist, all mental	18.9%	7.7%	9.1%	3.5%	0.26 (0.13, 0.40)	2.8 (1.6, 4.7)	3.6 (2.1, 6.4)

Notes: Data from 2008 until time of interview; average follow up time of 3.16 years; Primary, mental = P-chapter of ICPC-2; Specialist, mental = F-chapter of ICD-10; Primary, depression = P76 of ICPC-2; Specialist, depression = F32, F33 of ICD-10; Primary, anxiety = P74 and P79 of ICPC-2; Specialist, anxiety = F40, F41 of ICD-10; Primary, alcohol = P15 and P16 of ICPC-2; Specialist, alcohol = F10 of ICD-10; ICPC-2 = International Classification of Primary Care, second version; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, tenth version. r = polychoric correlation; OR = odds ratio; AOR is odds ratio adjusted for sex, age, educational attainment.

Table 3. Associations between major depressive disorder (MDD), anxiety disorders (AD), and alcohol use disorder (AUD) assessed at interview, primary care and specialist health care.

	r	OR	AOR
MDD and AD			
Interview	0.61 (0.53, 0.69)	9.94 (6.93, 14.27)	8.17 (5.58, 11.95)
Primary care	0.48 (0.42, 0.53)	7.82 (6.21, 9.85)	6.53 (5.12, 8.33)
Specialist care	0.60 (0.54, 0.66)	18.29 (13.65, 24.51)	14.49 (10.65, 19.72)
MDD and AUD			
Interview	0.26 (0.13, 0.39)	2.78 (1.71, 4.55)	4.12 (2.40, 7.06)
Primary care	0.40 (0.31, 0.50)	7.76 (4.81, 12.49)	7.85 (4.83, 12.75)
Specialist care	0.42 (0.30, 0.54)	10.77 (5.85, 19.82)	10.14 (5.13, 20.03)
AD and AUD			
Interview	0.18 (0.02, 0.33)	2.13 (1.14, 3.96)	2.83 (1.44, 5.56)
Primary care	0.35 (0.22, 0.47)	6.21 (3.31, 11.63)	5.27 (2.72, 10.22)
Specialist care	0.47 (0.35, 0.58)	16.30 (9.03, 29.42)	15.64 (8.02, 30.53)

Notes: Odds ratios (OR) and adjusted odds ratios (AOR) reflect risk of having the second disorder if the first disorder is present. AOR is adjusted for sex, age, educational attainment. r = polychoric correlations. P-values for differences between log odds is presented in Table S1.

Table 4. Heritability estimates, and the genetic (below diagonals), and environmental (above) correlations between the interview, primary and specialist care data.

Major depressive disorder	Heritability	1. Interview	2. Primary	3. Specialist
1. Interview	0.26 (0.10, 0.43)	-	0.55 (0.37, 0.72)	0.36 (0.16, 0.55)
2. Primary	0.49 (0.39, 0.59)	0.81 (0.50, 1.00)	-	0.66 (0.53, 0.78)
3. Specialist	0.46 (0.30, 0.59)	1.00 (0.60, 1.00)	0.80 (0.74, 0.94)	-
Anxiety disorders				
1. Interview	0.52 (0.40, 0.63)	-	0.51 (0.28, 0.71)	0.49 (0.23, 0.72)
2. Primary	0.45 (0.29, 0.60)	0.81 (0.72, 1.00)	-	0.57 (0.40, 0.75)
3. Specialist	0.36 (0.17, 0.56)	0.83 (0.78, 1.00)	0.93 (0.79, 1.00)	-
Alcohol use disorder				
1. Interview	0.45 (0.28, 0.60)	-	0.44 (-0.46, 0.96)	0.90 (0.20, 1.00)
2. Primary	0.40 (0.09, 0.70)	0.30 (-0.61, 1.00)	-	0.78 (0.52, 0.94)
3. Specialist	0.57 (0.24, 0.83)	0.30 (-0.31, 0.91)	1.00 (0.84, 1.00)	-

Notes: N=11,727 (2,271 with interview data). Results based on the best fitting model (AE model with sex differences only in mean).