28th June 2017
Abstract 249 words
Manuscript 5784 words
Tables = 1
Figures = 3
Supplementary material: Tables 5, Figures 19

Epidemiology of psychotic depression – systematic review and meta-analysis

E. Jääskeläinen[1,2,3]*ł, T. Juola[1], H. Korpela[1,2], H. Lehtiniemi[1], M. Nietola[4], J. Korkeila[5], J. Miettunen[1,2]ł

ł Equal contribution.

[1] Center for Life Course Health Research, University of Oulu, Finland; [2] Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Finland; [3] Department of Psychiatry, Oulu University Hospital, Finland; [4]Psychiatric Department, University of Turku and Turku University Hospital, Finland; [5] Psychiatric Department, University of Turku and Satakunta Hospital District, Finland.

*Address for the correspondence:

Erika Jääskeläinen

Center for Life Course Health Research

P.O. Box 5000

90014 University of Oulu, Finland

Fax. +358 8 5315037, Tel. +358 40 6742959

Email: erika.jaaskelainen@oulu.fi

Abstract

Large amount of data have been published on non-psychotic depression, schizophrenia and bipolar disorder, whilst psychotic depression as an own entity has received much smaller attention. We performed a systematic review and meta-analyses on epidemiology, especially incidence and prevalence, risk factors and outcomes of psychotic depression. A systematic search to identify potentially relevant studies was conducted using four electronic databases and a manual search. The search identified 1,764 unique potentially relevant articles, the final study included 99 articles. We found that the lifetime prevalence of psychotic depression varies between 0.35-1%, with higher rates in older age. Onset age of psychotic depression was earlier than that of non-psychotic depression in younger samples, but later in older samples. There were no differences in gender distribution in psychotic depression versus non-psychotic depression, but higher proportion of females was found in psychotic depression than in schizophrenia or in psychotic bipolar disorder. Risk factors have rarely been studied, the main finding being that family history of psychosis and bipolar disorder increases the risk of psychotic depression. Outcomes of psychotic depression were mostly worse when compared to nonpsychotic depression, but better compared to schizophrenia and schizoaffective disorder. The outcome compared to psychotic bipolar disorder was relatively similar, and somewhat varied depending on the measure of the outcome. Based on this review, the amount of research on psychotic depression is far from that of non-psychotic depression, schizophrenia and bipolar disorder. Based on our findings, psychotic depression seems distinguishable from related disorders and needs more scientific attention.

Keywords: psychotic depression, unipolar depression, major depressive disorder, epidemiology, risk, incidence, outcome, prognosis, psychotic disorders, psychosis

Introduction

Major depression with psychotic features (hereafter psychotic depression, PD) is a severe disorder with a high risk of recurrence and high mortality in both adult samples under 60 years (Lykouras & Gournellis, 2009) and older people samples (Gournellis *et al*. 2014). In spite of the severe course of illness there seems to be some difficulty identifying the disorder in clinical settings (Rothschild *et al.* 2008).

Originally, Kraepelin (Goodwin *et al. 2007*) considered PD as a type of manic-depressive illness. In the post-kraepelinian era, it has been classified among unipolar major depressive disorders. In ICD-10 PD is considered the most severe subtype of major depressive disorder (WHO, 1992), whereas in DSM-5 psychotic features are not an indicator of severity of major depression (APA, 2013). Due to a number of differences between PD and non-psychotic depression (hereafter NPD), it has for long been proposed that PD should be considered a distinct disease entity (Schatzberg & Rothschild, 1992; Keller *et al*. 2007).

The point prevalence of PD is estimated to be approximately 0.4%, with older adults being in the highest risk (Kivelä & Pahkala, 1989; Perälä *et al.* 2007). The prevalence of psychotic features in the adolescent outpatient major depression sample was 18% (Ryan *et al.* 1987) while the same figure was 45% in a hospitalized adolescent patient sample (Haley *et al.* 1988). There is lack of information concerning the risk factors for PD. Previous studies have often studied all affective psychosis, i.e. included bipolar disorder or studied PD as part of all major depressive disorders. Though there are marked similarities in PD and NPD risk factors, some differences are likely to exist. Also, there is considered a close link between PD and bipolar disorder (Keller *et al*. 2007; Østergaard *et al.* 2013). There are no previous systematic reviews on incidence, prevalence, or risk factors of PD.

Clinical course of illness in PD is more severe than in NPD. This applies especially in the short-term outcome but it has been suggested that in a longer follow-up the significance of psychotic features might fade (Keller *et al.* 2007; Lykouras & Gournellis, 2009). However, mortality is significantly higher in PD compared to NPD (Vythilingam *et al.* 2003), although there are conflicting findings (Suvisaari *et al.* 2013). The functional outcome has been suggested to be mostly better in PD compared to schizophrenia (SZ), and differences in outcomes between PD and psychotic bipolar disorder (PBD) have been unclear (Craig *et al.* 2000; Jarbin *et al.* 2003; Keller *et al*. 2007).

Huge amount of data and meta-analyses have been published on NPD, SZ and bipolar disorder, whilst PD as an own entity has received much smaller attention (Crebbin *et al.* 2008). Meanwhile, there has been a concern over the validity of PD diagnosis mainly due to diagnostic instability (Ruggero *et al*. 2011). There are some meta-analyses and reviews on pharmacological treatments (Wijkstra *et al*. 2015), cognition (Fleming *et al.* 2004; Zaninotto *et al.* 2015), genetics (Domschke, 2013) neuroimaging studies (Busatto, 2013), cortisol nonsuppression (Nelson & Davis, 1997) and PD in old age (Gournellis *et al.* 2014). Lykouras and Gournellis (2009) present a comprehensive review on neurobiology, treatments, epidemiology, course of illness, and outcomes of PD in comparison to NPD. However, they have not reported their results systematically, and some topics such as risk factors have not been studied. Earlier reviews presenting epidemiology of PD (Schatzberg & Rothschild, 1992; Gournellis & Lykouras, 2006; Lykouras & Gournellis, 2009) have not combined the data by meta-analytic means and they have compared their findings only to NPD.

Aims

Our aim was to perform a systematic review on epidemiology, especially incidence and prevalence, risk factors and outcomes of PD. We also aimed to do a meta-analysis on sex differences, onset age, and outcome of PD in comparison to NPD, SZ, PBD, and schizoaffective disorder (SZAFF).

Methods

Data Collection

We applied the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines for systematic reviews and meta-analyses (Moher *et al.* 2009).

In order to locate potentially suitable studies, we conducted database searches in May 2016 using four electronic databases: PubMed, Scopus, Web of Science, and CINAHL. The used search terms were the following: ("psychotic depression" OR "delusional depression" OR "depression with psychotic features") AND (epidemiology OR "risk factor" OR outcome OR employment OR occupational OR progression OR course OR stability OR relapse OR remission OR prevalence OR incidence OR "onset age" OR "diagnostic stability" OR mortality OR suicide OR physical OR somatic OR comorbidity OR "early intervention" OR prevention). No publication date exclusions were used in the search. Articles were also searched manually from the reference lists of the previous reviews.

All abstracts were independently analysed by two authors (HK and EJ). After the exclusion of irrelevant abstracts, all remaining articles were critically inspected by two authors (EJ or JM). For studies that met inclusion criteria, a third investigator (HL or TJ) independently extracted the data, and the collected data were checked by two authors (JM or EJ). When a disagreement occurred related to data extraction, this was resolved by consensus.

Study selection

Studies on prevalence and incidence were included if these were estimated from population surveys or used both inpatient and outpatient data to estimate prevalence or incidence.

Regarding studies on risk factors and outcomes in PD, to be included in the analyses, the studies had to be characterized by all of the following:

1. Original study included a sample of PD. Also studies including only delusional depression were included, as the early studies on the topic often used only this definition. The sample had to include at least 80% of PD. Studies focusing on psychotic depressive episode of SZ or PBD, or studies with postpartum psychotic depression were not included.
2. Diagnostic assessment and diagnostic criteria of PD was based on a commonly used diagnostic system or was otherwise adequately reported.
3. The sample size of PD was at least 15.
4. Studies presented risk or sociodemographic factors, or outcomes of PD.
5. Studies of risk factors and outcomes had to include a comparison group of NPD, PBD, SZ, SZAFF, or healthy controls without mental disorder. The size of the comparison group had to be at least 15 and the comparison group had to include at least 80% of NPD, PBD, SZ, or SZAFF.
6. The majority of subjects had onset age after 16 years.

Only observational (naturalistic) studies were included, whereas trials and intervention studies were excluded. While many intervention studies report clinical outcomes, the representativeness of these samples may vary widely according to the specific trial inclusion criteria. Thus, a large number of randomized controlled trials were excluded. To the current review, we finally included only studies published in English. In addition, studies analyzing neurobiological risk factors and correlates, treatments, mortality, suicides, and somatic comorbidities were excluded as being either out of the scope of this review (studies on treatments, neurobiology, somatic comorbidities) or being recently studied (mortality and suicides in PD; Lykouras & Gournellis, 2009; Rothschild, 2009; Zalpuri & Rothschild 2016).

Incidence, prevalence and risk factor studies

Studies on incidence and prevalence were reported with a systematic review. Regarding gender distribution and onset age we pooled studies using meta-analytic methods. Other risk factors were reported only narratively and in a literature table. The included risk factors encompassed both early risk factors and sociodemographic factors, such as marital status and education, collected at study entry.

Outcome studies

Of studies analyzing outcomes of PD, studies analyzing the severity of psychotic symptoms (positive, negative, total symptoms), severity of depression symptoms, number of hospitalizations during prospective follow-up, symptomatic remission, global clinical outcome, global outcome and occupational functioning were included. Global clinical outcome indicates outcome measured by the presence of clinical symptoms and severity of illness, without a specific instrument for the measurement. Global outcome indicates the outcome measured by Social and Occupational Functioning Assessment Scale (SOFAS), Global Assessment Scale (GAS) or Global Assessment of Functioning scale (GAF). Please see our earlier meta-analyses for the definitions of different outcome dimensions (Käkelä *et al.* 2014; Penttilä *et al.* 2014). Based on the number of studies (at least 3 studies from different samples per outcome), meta-analysis was performed on depression symptoms, total psychotic symptoms, positive and negative symptoms, global outcome, symptomatic remission and poor global clinical outcome. For the meta-analysis, we selected symptoms measured at the baseline of studies, since this was the most common time of assessment of symptoms. In meta-analysis, outcomes where compared between PD and NPD, SZ, and PBD when data was available. Systematic review (without meta-analysis) was done for hospitalizations and occupational functioning.

Statistical analyses

Random effects models were used in order to pool estimates of effect sizes between PD and

comparison groups in the meta-analyses of gender, onset age, symptoms, symptomatic remission, global clinical outcome, and hospitalizations, based on the expected heterogeneity of the associations. Meta-analyses were done if at least three studies investigated same outcome. In the random-effects analysis, each study was weighted by the inverse of its variance and the between-studies variance. In continuous variables (onset age and symptoms), the effect size of the standardised mean difference between groups was described with Hedges' g. Hedges' g values is comparable with Cohen's d but recommended with small sample sizes. It can be interpreted as small 0.20, moderate 0.50 and large 0.80 effect (Cohen, 1992). In categorical variables, pooled effect size was estimated using Relative Risk (RR) with 95% confidence interval (CI). When the number of studies allowed, we checked the results of meta-analyses in the subgroups of studies based on publication year (1973–1991, 1993–2003, and 2004–2016), mean study age (below 45, 45 to 55, and above 55 years), or mean age of illness onset (below 45 versus 45 or above). In addition, as a sensitivity analyses, we performed analyses in strata by sample size (studies under 50 cases vs. at least 50 cases with PD). In the current study, positive g values indicate that individuals with PD have more symptoms or later onset age than comparison group. Where multiple articles were available on the same or overlapping samples and presenting similar data, we selected one representative paper with the largest sample size or presenting outcomes measured by a more commonly used instrument for the meta-analysis. We assessed the heterogeneity of the studies using I^2 statistics, and the statistical significance in heterogeneity was tested using the chi-square test. Values of I^2 range from 0% to 100%, reflecting the proportion of total variation across studies beyond chance. A value of 25% describes low, 50% moderate, and 75% high heterogeneity (Higgins *et al*. 2003). An alpha level of 0.05 was used for all statistical tests. The *metan* command of the Stata version 13 (StataCorp 2013; Sterne 2009) was used in all analyses.

Results

Characteristics of the studies

Database searches identified 2,926 records, which reduced to 1,764 after the removal of duplicates. After analysing the abstracts, we were left with 279 articles that potentially fulfilled our inclusion criteria. The most common reason for exclusion during abstract screening was that the article did not present results separately to PD. Figure 1 shows the flow diagram that details the exclusion criteria after abstract reading. In total, 99 studies met all of our criteria regarding incidence/prevalence, risk factors or outcome, and were included in the systematic review. The studies included nine studies from manual search.

Insert Figure 1 here.

Incidence and prevalence

Studies reporting prevalence or incidence estimates in community samples and using estimates based on inpatient and outpatient data have been collected into the Supplement Table 1.

Only four studies estimated *prevalence* in community samples, trying also to detect cases not in treatment using different screening methods. In the nationally representative Finnish Health 2000 sample, Perälä *et al.* (2007) found a lifetime prevalence of 0.35% for DSM-IV PD. The prevalence was higher among those who were 65 years or more (0.43%) when compared with younger age groups, however differences were not statistically significant. In the US Epidemiological Catchment Area (ECA) study, Johnson *et al.* (1991) reported a lifetime prevalence of DSM-III PD to be 0.6%. In an older community based study of those with 60 years or more, the prevalence was 1.0% (0.6% for males, 1.2% for females) (Kivelä & Pahkala, 1989). In a large European telephone survey in five countries, an overall point prevalence for DSM-IV PD was 0.5%, and significantly higher rates were reported for females (0.6%) than males (0.3%) (Ohayon & Schatzberg, 2002).

Five studies used different in- and outpatient admission registers. Estimates for *annual incidence*

(per 100,000 persons) were reported in three studies. In a British study, Farquhar *et al.* (2007) reported an incidence of 3.4 in the year 1875–1924 and 3.0 in 1995–1999. In an Irish study, Baldwin *et al*. (2005) found an incidence of 6.4 (males 5.4 and females 7.4), similar estimates (6.0 for those with 16 years and over) were also in a British study by Reay *et al.* (2010). A Finnish study comparing two birth cohorts (from 1966 and 1986), found a substantial increase in cumulative incidence until age 27 years in the later cohort (0.02% vs. 0.21%) (Filatova *et al.* 2016).

Gender differences have been reported in some incidence and prevalence studies. In three studies, females had higher estimates for incidence (Kivelä & Pahkala, 1989; Ohayon & Schatzberg, 2002; Baldwin *et al.* 2005), however in one study higher lifetime prevalence was reported for males (0.41%) than females (0.29%) (Perälä *et al.* 2007).

Proportion of psychosis in depression

In studies (n=43) including both PD and NPD patients, the median proportion of PD patients was 28% of all depressive patients. Median proportion was lower in studies with mean age below 45 years (20%, n=16) than in the middle age samples (27%, n=11) or in older samples (34%, n=10). The median proportion of PD in depression patients was 29% among females and 26% among males. In the studies including only depressive inpatients (n=22), the median proportion of those with PD was 42% whereas in the studies including both in- and outpatients or only outpatients (n=21), corresponding proportion was 19%.

Gender differences in PD when compared with other patient samples

We compared gender distributions in the included studies on PD and patient control groups. In total 57 studies compared PD with other included patients samples. The median number of PD patients was 45 in these studies, whereas the total number of PD patients was 28,370. In total 43 studies compared gender distributions between PD and NPD, pooled RR being 1.03 (95% CI

0.97–1.08). Estimates of RR were relatively similar when studies were divided by mean study age or year of publication. The estimated RRs for PD for females are presented in Figure 2, for the total sample and by mean study age. Studies comparing PD and SZ and PBD found a higher proportion of females in PD than in SZ (14 studies; RR 1.40, 95% CI 1.20–1.71) or in PBD (3 studies; RR 1.36, 95% CI 1.01–1.83). The median percentage of females in PD was 65%, in NPD 65%, in SZ 37%, in SZAFF 57%, and in PBD 55%. Proportion of females in PD did not vary significantly when studies were divided by mean study age or year of publication. Proportions of females in different patient groups in the included studies are presented in the Supplement Table 2.

Insert Figure 2 here.

Onset age in PD when compared with other patient samples

Eighteen studies compared onset age between PD and NPD in different samples. Based on meta-analysis there was no significant difference between the groups (Hedges' g=0.08, p=0.44). However, when we divided the studies into three categories based on mean study age, we found conflicting results. In the studies of the youngest subjects (below 45 years, n=6), PD patients had earlier onset age (g=-0.39, p<0.001), whereas in the studies among the oldest (above 55 years, n=7) PD patients had later onset age (g=0.40, p<0.001) than NPD patients. The year of publication did not affect the results. A forest plot comparing onset age between PD and NPD by age groups and in the total sample is presented in Figure 3. In the six studies comparing mean onset age between PD and SZ, five found earlier onset age in SZ, and pooled meta-analysis found significant difference (g=0.53, p=0.013). When we compared PD and PBD, PD patients had non-significantly later onset age (g=0.34, p=0.069). Mean onset ages in different groups in the included studies are presented in the Supplement Table 3.

Insert Figure 3 here.

Regarding comparison in onset age between PD and NPD, only six studies had sample size of at least 50. There were only two studies from each three age groups, however all the statistical significant findings remained when compared to the original analyses.

Risk factors and sociodemographic factors

Studies on risk factors and sociodemographic factors in PD are summarized in the Supplement Table 4. In total 36 studies were found.

Studies on *early risk factors* are rare. The only study analyzing risk factors from birth (place of birth, gestational age, birth weight, small for gestational age, maternal and paternal age at birth) was a large Danish register study that did not find any significant differences in these factors between PD and healthy controls (HC) (Østergaard *et al.* 2013). Physical and sexual trauma was more likely in PD than in NPD in one study (84% vs. 64%, p=0.017) (Gaudiano & Zimmerman, 2010), but not in one (Gaudiano *et al.* 2016). Other premorbid factors linked with PD when compared with NPD were rural domicile (Ihezue, 1985), acute medical problems (Draper & Anstey, 1996), and poorer social competence score (Sands & Harrow, 1995). When compared with HC, PD patients differed in number of physical anomalies (Čulav-Sumić & Jukić, 2010) and also a loss of mother because of an unnatural cause after age 15 years associated with PD (Østergaard *et al.* 2013). *Ethnicity* was studies in eight articles. Individuals with PD were less likely to be Caucasian in five different studies (Johnson *et al.* 1991; Goldberg & Harrow, 2005; Gaudiano *et al.* 2009; Gaudiano & Zimmerman, 2010; Gaudiano *et al*. 2016). The British study by Heslin *et al*. (2016a) found that PD patients had less contact with friends, and they were more likely to have childhood adversity of neurological soft signs when compared with HC.

The family history of different psychiatric illnesses and suicides was analyzed in 14 articles. Most of the associations were non-significant. When PD patients were compared with NPD, they more often had a family history of psychosis (Buoli *et al.* 2013) and bipolar I disorder (Maj *et al.* 2007).

One study also found a higher likelihood of any mental illness in relatives (Okulate *et al.* 2001), whereas two other studies did not find differences (Frangos *et al.* 1983; Nakamura *et al.* 2015). Studies looking at the family history of affective or depressive disorders did not find differences between PD and NPD (Frangos *et al.* 1983; Parker *et al.* 1991; Simpson *et al.* 1999; Park *et al.* 2014). The large Danish register study by Østergaard *et al.* (2013) found several maternal, paternal and sibling psychiatric diagnoses to associate significantly with PD when compared with HC, the highest risk (Incidence Rate Ratio of 2.2) being in any maternal mental disorder. Also a recent study in the UK found the family history of any mental illness or psychosis to associate with PD (Heslin *et al*. 2016a).

Educational level or years of education between PD and other patient groups was compared in 18 studies. Differences were mainly non-significant. PD patients had less education when compared with PBD in one study (Breslau & Meltzer, 1988). When PD patients were compared with NPD, they had lower education in six studies (Ihezue, 1985; Karaaslan *et al*. 2003; Goldberg & Harrow, 2005; Gaudiano *et al.* 2009; Gaudiano & Zimmerman, 2010; Heslin *et al.* 2016b) but more years of education in one study (Park *et al.* 2014). In the ECA study PD patients had lower socioeconomic status when compared with NPD (Johnson *et al.* 1991). *Marital status* between PD and other patient groups was compared in 19 studies. Differences were mainly non-significant although two studies found PD patients to be more often single than NPD patients (Baldwin, 1995; Gaudiano *et al.* 2016) and in one study PD patients were less often single when compared to SZ patients (Heslin *et al*. 2016b).

Outcomes in psychotic depression

Study characteristic and quality. The studies included in outcome review are described in Supplement Table 5. We found altogether 44 articles presenting results from 37 separate studies. Several studies did not report characteristics of PD group in detail (Supplement Table 5). The sample sizes of PD varied between 16 and 190. 14 of the studies had sample size at least 50. In

25 of the studies there were more females than males. 25 studies included patients with onset age before 45 years of age (or if age of onset not reported, the sample was under 45 years at the study moment). Most of the studies (n=16) were cross-sectional, and 13 studies had over 5 years follow-up. Study populations were mostly mixed samples (n=25), with minority being first-episode (n=10) and consecutive samples (n=2). Outcomes were most frequently defined using validated scales, but in some studies the scale or its use were not clearly reported. Most commonly studied outcomes were different symptoms, remission, and global clinical outcome.

Outcome compared to NPD. Based on meta-analysis (Supplement Figures 1a-f), compared to NPD, the symptoms of depression were more severe in PD (Hedges' g=0.52, p<0.001). The difference in symptom severity was larger among three samples with onset age 45 years or older (g=0.84, p=0.004), but significant also in younger samples (g=0.40, p=0.005). Psychosis symptoms were more severe in PD (g=0.89, p=0.037). Symptomatic remission tended to be less common in PD though not statistically significantly (RR=0.82, p=0.052). There was no significant difference in the global clinical outcome or hospitalizations, though PD patients tended to have poorer outcomes. The global outcome (based on SOFAS, GAS, or GAF score) was somewhat worse in PD, but not statistically significantly (g=-0.43, p=0.065). Sensitivity analyses by sample size were performed for studies comparing depression symptoms, global outcome, symptomatic remission and poor global clinical outcome in PD versus NPD. The results were mixed. Regarding depression symptoms the difference in PD vs. NPD was not statistically significant in larger samples (50 cases or more), and the results of global outcome remained non-significant. Regarding symptomatic remission and poor clinical outcome, the difference between PD and NPD was statistically significant in large samples (Supplement Figures 4a-d).

Based on systematic review (Supplement Table 5), the rate of relapses was higher in PD compared to NPD (Baldwin 1988, Copeland 1983). In most of the studies analyzing occupational outcomes, individuals with PD had a somewhat poorer outcome compared to NPD (Coryell *et al*. 1984; Coryell & Tsuang, 1985). However, there were also studies indicating similar occupational

outcomes for PD and NPD (Jäger *et al*. 2005; Rush *et al*. 2006; Park *et al*. 2014). A good occupational outcome occurred in 60–79% of PD, and 68–78% on NPD (Coryell & Tsuang, 1985; Jäger *et al*. 2005; Park *et al*. 2014), and poor occupational outcome in 28% of PD and 19% of NPD (Coryell & Tsuang, 1985), and unemployment in 90% of PD and 81% of NPD (Rush *et al*. 2006). Based on only study analyzing full recovery (both symptomatic and functioning, Coryell *et al*. 1982), full recovery after 2–3 years of follow-up was more common in NPD (69%) than PD (40%).

Outcome compared to SZ. According to meta-analyses, when compared to SZ, there was no difference in severity of depression symptoms, but total psychosis symptoms (g=-0.77, p=0.000) and positive (g=-0.81, p=0.000) and negative symptoms (g=-0.89, p<0.001) were significantly less severe in PD. Global outcome was better in PD (g=0.80, p=0.001) (Supplement Figures 2a-e). All but one of the samples in the meta-analyses included patients with mean onset age below 40 years. The rate of relapses was lower in PD (Craig *et al*. 2000). Occupational functioning was better in PD, 60–79% of PD and 36–47% of SZ having a good occupational outcome, and 28–29% and 57–88% having poor, respectively (Coryell & Tsuang, 1985; Jarbin *et al*. 2003; Jäger *et al*. 2005). Full recovery (both symptomatic and functioning, Coryell *et al*. 1982) after 2–3 years of follow-up was more common in PD (40%) than SZ (7%) (Supplement Table 5).

Outcome compared to PBD. Symptoms of depression did not differ. Negative symptoms (g=0.65; p=0.001) were more severe in PD. However, PD had less severe positive symptoms (g=-0.44; p=0.046). There was no difference in global functioning between PD and PBD (Supplement Figures 3a-d). Rehospitalization rates were relatively similar in PD and PBD (Craig *et al*. 2000). The unemployment rate was similar in PD and PBD (63 vs. 53–69%) (Dell'Osso *et al*. 2002), as was functional recovery (32 vs. 37%, Tohen *et al*. 2000). Persons with PD were somewhat less often on a disability pension (29% vs. 33%), and they were less often unemployed (7% vs. 14%) (Supplement Table 5).

Outcome compared to SZAFF. Only a few studies comparing PD and SZAFF were found, and no

meta-analysis could be performed. Symptomatic remission (Coryell *et al*. 1990, Opjordsmoen *et al*. 1991), and employment (Opjordsmoen *et al*. 1991) were more common in PD, but there was no difference in number of relapses at follow-up (Opjordsmoen *et al*. 1991). In one study, there was no marked difference in syndromatic recovery between PD and SZAFF, but PD subjects had more often functional recovery (Tohen *et al*. 2000).

Discussion

Main results

Based on this systematic review, though not as common as for example schizophrenia, it seems that PD is relatively common, especially in older populations. However, this conclusion is based on a relatively few studies with varying methodology. Within depression, the onset age of PD was earlier than that of NPD in younger samples, but later in older samples. This may be due to PD at first episode being a marker of later bipolar disorder in younger samples. It seems that the proportion of PD is higher in inpatient samples. Based on this review, the median proportion of those with PD was 42% in inpatients, and 19% in outpatients. There was no difference in gender distribution in PD versus NPD, but higher proportion of females was found in PD than in SZ or in PBD. Risk factors have rarely been studied, and most of the findings were statistically nonsignificant. Family history of psychosis and bipolar I seems to increase the risk of PD.

To our knowledge, this is a first systematic review and meta-analysis comparing the outcomes of PD not only to NPD, but to also to SZ, SZAFF and PBD. Several outcomes of PD were mostly worse when compared to NPD, but better compared to SZ and SZAFF. The outcomes compared to PBD were relatively similar, though there were more negative and less positive symptoms in PD. The number of studies comparing PD to SZ and PBD, and especially SZAFF are very few.

Please see Table 1 for the summary of main results.

Clinical and public health implications

The number of studies on the epidemiology of PD are far from the large amount of studies on SZ (e.g. Matheson *et al*. (2014) or on unipolar depression in general (e.g. Kessler & Bromet, 2013; Hirschfeld, 2012) and on bipolar disorder (Benazzi, 2007; Esan & Esan, 2016; Sherazi *et al*. 2006). Many of the risk factors reviewed in reviews on NPD and SZ have not been studied on PD at all or only in a few small samples. Based on our review, there is lack of studies on epidemiology, especially risk factors, and longitudinal clinical and functional outcomes in PD. This is in line with the general notion of lack of clinical trials focusing on PD (Wijkstra *et al.* 2015). In addition, both treatment algorithms and clinical practice regarding PD are highly heterogeneous. This emphasizes the need for further studies also on the treatment of PD (Leadholm *et al.* 2013).

Our review supports the earlier conclusions about more severe depression symptoms in PD compared to NPD especially in older samples (Lykouras & Gournellis, 2009). Most of the studies included in the meta-analysis included patients with relatively young age at the study moment, and thus the other results on symptoms and global outcome can be generalized only to this age group.

Our review summarizes the outcomes of PD in comparison to SZ and PBD. After our database searches, very recently, an AESOP study was published, where 10-year outcomes in PD compared to SZ and PBD patients were investigated. The study found only minimal differences in the outcome between PD and PBD. Differences in clinical, social and service use outcomes between PD and SZ were more substantial with PD patients showing better outcomes on most variables. (Heslin *et al*. 2016a). These results of AESOP seems relatively similar to ours.

The burden of disease of mood disorders to society among EU nations is higher than in any other brain disorders, most of the costs resulting from disability (Olesen *et al*. 2012). There are not many studies on the disability due to PD. In PD, disability was found to be increased even when

compared to severe major depression in all functional dimensions of Short Form-36, there were, moreover, an increased number of absent days and days ill in bed (Kruijshaar *et al*. 2003). Severe forms of recurrent depressions, additionally, may have a scar effect in the form of an increase in disability (Ormel *et al*. 2004). Due to earlier age of onset and higher prevalence, the burden of disease on society is likely to be higher in SZ, although self-perceived suffering may be worse due to depression being a robust determinant of quality of life (Saarni *et al*. 2010).

Diagnostic instability has been a concern with PD (Bromet *et al*. 2011; Ruggero *et al.* 2011). In ten-year follow-up studies of relatively young patient samples the diagnosis of PD has remained in less than half of the cases (Bromet *et al*. 2011; Ruggero, 2011; Heslin *et al*. 2015) and Ruggero *et al*. (2011) have even suggested that PD diagnosis should be considered as a provisional diagnosis. However, in a two-year follow-up of slightly more aged sample the stability was 85% (Salvatore, 2011). The early onset of PD may well predict conversion to bipolar disorder (Østergaard *et al.* 2014). Additionally, changes in the symptom presentation seem to explain the instability (Bromet *et al*. 2011). A shift towards SZ has also been found during the course of a decade. Among these cases, poorer functioning and negative symptoms predicted the shift (Bromet *et al*. 2011). Altogether the stability of diagnosis in PD can be highly age-related as especially younger patients are more likely to develop PBD (Lykouras & Gournellis, 2009). The diagnosis might also be more stable in patients with medical co-morbidity (Tohen *et al*. 2012). Still, among mood disorders bipolar disorder has been found to best predict psychosis (Souery *et al*. 2011).

Considering the diagnostic validity of PD it is interesting that gender distribution in PD is similar to NPD while the proportion of females is lower in SZ and PBD. Meanwhile, the differences between PD and NPD are well documented (Keller et al. 2007) and our findings are in line with these. Also, the increasing prevalence, though not statistically significant, and proportion of psychosis in depression in older patient samples contradicts with the concept of psychotic illness, for example SZ, starting usually at early adulthood. In this systematic review onset age of PD was earlier than

that of NPD in younger samples, but later in older samples. It remains possible that there are two forms of PD. PD in young adulthood may be an etiologically and prognostically different illness than PD in late adulthood and in geriatric populations. Early onset form of PD may be more unstable, potentially an early expression for some patients of bipolar disorder, and for others perhaps other psychotic conditions. Among older onset cases it is possible that medical and neurological conditions partly explain the occurrence of PD. Future studies should address these questions and include also late-onset PD patients.

Strengths and limitations

There are several limitations related to this review. We included only articles published in English, meaning that especially older relevant articles on the topic may be missing. It should be acknowledged, that the oldest studies in this review were from 1980s. Although we consider our search criteria to be adequate, we may have missed some studies, especially older studies. Because of this we, have also done some manual work to locate these papers, e.g. using the reference lists of previous reviews. It should be noted that we excluded childhood onset samples. The included articles were quite mixed regarding methods, e.g. diagnostic criteria or other inclusion criteria. The original studies on incidence and prevalence rates were few, and they had very heterogeneus methodology. There were four population studies with different design and methods of ascertaining the PD cases, and five registry studies. These two sets of studies produced considerably different estimates of PD. It may be that the available data are too heterogeneous to make precise estimate if incidence and prevalence of PD.

The sample sizes were relatively small, e.g. in risk factors median sample size being 45, and mainly not based on population samples, but comparing clinical samples. Minority of the studies based on first-episode samples. Most of the studies on outcomes had sample size of PD under 50. Due to the low number of studies it is not possible have a clear picture on the effect of study quality (e.g. sample size) on the results. However, based on the study characteristic summarized in

Results, many of the original studies have important limitations (e.g. small sample size, short follow-ups, lack of long-term follow-ups in older populations). In outcome analyses, some of the definitions of outcomes were heterogeneous, e.g. definitions of symptomatic remission, global clinical outcome, global outcome varied.

The strength of this review was the comprehensive search strategy, as we searched four electronic databases. We read in detail studies analyzing depression in general, and whenever possible, extracted the data concerning PD as separate group. There was a relatively good amount of data on gender differences, differences in onset age, and differences in some of the outcome measures to also allow new conclusions on the epidemiology of PD.

Conclusions

To our knowledge, this is the first systematic review on different aspects of epidemiology of PD. Based on this review, the amount of research on PD is far from that of NPD, SZ and bipolar disorder. Based on differences in gender, onset age and outcomes in PD in comparison to other disorders, PD seems distinguishable from related disorders and needs more scientific attention.

Acknowledgements

This work was supported by grants from the NARSAD: Brain and Behavior Research Fund and the Academy of Finland (#268 336, #278 286). The funders had no role in the manuscript.

Declaration of Interest

None.

References

American Psychiatric Association (APA) (2013). *Diagnostic and statistical manual of mental disorders, 5th edition: DSM-5*. American Psychiatric Publishing: Arlington, VA.

Baldwin RC (1995). Delusional depression in elderly patients: characteristics and relationship to age at onset. *International Journal of Geriatric Psychiatry* **10**, 981-985.

Baldwin RC (1988). Delusional and non-delusional depression in late life. Evidence for distinct subtypes. *British Journal of Psychiatry* **152**, 39-44.

Baldwin P, Browne D, Scully PJ, Quinn JF, Morgan MG, Kinsella A, Owens JM, Russell V, O'Callaghan E, Waddington JL (2005). Epidemiology of first-episode psychosis: illustrating the challenges across diagnostic boundaries through the Cavan-Monaghan study at 8 years. *Schizophrenia Bulletin* **31**, 624-638.

Benazzi F (2007). Bipolar II disorder: epidemiology, diagnosis and management. *CNS Drugs* **21**, 727-740.

Breslau N, Meltzer HY (1988). Validity of subtyping psychotic depression: examination of phenomenology and demographic characteristics. *The American Journal of Psychiatry* **145**, 35-40.

Bromet EJ, Kotov R, Fochtmann LJ, Carlson GA, Tanenberg-Karant M, Ruggero C, Chang SW (2011). Diagnostic shifts during the decade following first admission for psychosis. *The American Journal of Psychiatry* **168**, 1186-1194.

Buoli M, Caldiroli A, Altamura AC (2013). Psychotic versus non-psychotic major depressive disorder: a comparative naturalistic study. *Asian Journal of Psychiatry* **6**, 333-337.

Busatto GF (2013). Structural and functional neuroimaging studies in major depressive disorder with psychotic features: a critical review. *Schizophrenia Bulletin* **39**, 776-786.

Cohen J (1992). A power primer. *Psychological Bulletin* **112**, 155-159.

Copeland JR (1983). Psychotic and neurotic depression: discriminant function analysis and five-year outcome. *Psychological Medicine* **13**, 373-383.

Coryell W, Keller M, Lavori P, Endicott J (1990). Affective syndromes, psychotic features, and prognosis. I. Depression. *Archives of General Psychiatry* **47**, 651-657.

Coryell W, Lavori P, Endicott J, Keller M, VanEerdewegh M (1984). Outcome in schizoaffective, psychotic, and nonpsychotic depression. Course during a six- to 24-month follow-up. *Archives of General Psychiatry* **41**, 787-791.

Coryell W, Tsuang MT (1985). Major depression with mood-congruent or mood-incongruent psychotic features: outcome after 40 years. *The American Journal of Psychiatry* **142**, 479-482.

Coryell W, Tsuang MT, McDaniel J (1982). Psychotic features in major depression. Is mood congruence important? *Journal of Affective Disorders* **4**, 227-236.

Craig TJ, Bromet EJ, Fennig S, Tanenberg-Karant M, Lavelle J & Galambos N (2000). Is there an association between duration of untreated psychosis and 24-month clinical outcome in a first-admission series? *American Journal of Psychiatry* **157**, 60-66.

Crebbin K, Mitford E, Paxton R, Turkington D (2008). Firstepisode psychosis: an epidemiological survey comparing psychotic depression with schizophrenia. *Journal of Affective Disorders* **105**, 117-124.

Čulav-Sumić J, Jukić V (2010). Minor physical anomalies in women with recurrent unipolar depression. *Psychiatry Research* **176**, 22-25.

Dell'Osso L, Pini S, Cassano GB, Mastrocinque C, Seckinger RA, Saettoni M, Papasogli A, Yale SA, Amador XF (2002). Insight into illness in patients with mania, mixed mania, bipolar depression and major depression with psychotic features. *Bipolar Disorders* **4**, 315-322.

Domschke K (2013). Clinical and molecular genetics of psychotic depression. *Schizophrenia Bulletin* **39**, 766-775.

Draper B, Anstey K (1996). Psychosocial stressors, physical illness and the spectrum of depression in elderly inpatients. *The Australian and New Zealand Journal of Psychiatry* **30**, 567-572.

Esan O, Esan A (2016). Epidemiology and burden of bipolar disorder in Africa: a systematic review of data from Africa. *Social Psychiatry and Psychiatric Epidemiology* **51**, 93-100.

Farquhar F, Le Noury J, Tschinkel S, Harris M, Kurien R, Healy D (2007). The incidence and prevalence of manic-melancholic syndromes in North West Wales: 1875–2005. *Acta Psychiatrica Scandinavica. Supplementum* **(433)**, 37-43.

Filatova S, Marttila R, Koivumaa-Honkanen H, Nordström T, Veijola J, Mäki P, Khandaker GM, Isohanni M, Jääskeläinen E, Moilanen K, Miettunen J (2016). A comparison of the cumulative incidence and early risk factors for psychotic disorder in young adults in the Northern Finland Birth Cohorts 1966 and 1986. *Epidemiology and Psychiatric Sciences* [Epub ahead of print].

Fleming SK, Blasey C, Schatzberg AF (2004). Neuropsychological correlates of psychotic features in major depressive disorders: a review and meta-analysis. *Journal of Psychiatric Research* **38,** 27-35.

Frangos E, Athanassenas G, Tsitourides S, Psilolignos P, Katsanou N (1983). Psychotic depressive disorder. A separate entity? *Journal of Affective Disorders* **5**, 259-265.

Gaudiano BA, Dalrymple KL, Zimmerman M (2009). Prevalence and clinical characteristics of psychotic versus nonpsychotic major depression in a general psychiatric outpatient clinic. *Depression and Anxiety* **26**, 54-64.

Gaudiano BA, Weinstock LM, Epstein-Lubow G, Uebelacker LA, Miller IW (2016). Clinical characteristics and medication use patterns among hospitalized patients admitted with psychotic vs nonpsychotic major depressive disorder. *Annals of Clinical Psychiatry* **28**, 56-63.

Gaudiano BA, Zimmerman M (2010). The relationship between childhood trauma history and the psychotic subtype of major depression. *Acta Psychiatrica Scandinavica* **121**, 462-470.

Goldberg JF, Harrow M (2005). Subjective life satisfaction and objective functional outcome in bipolar and unipolar mood disorders: a longitudinal analysis. *Journal of Affective Disorders* **89**, 79-89.

Goodwin FK, Jamison KR (2007). *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*. Oxford University Press: New York.

Gournellis R, Lykouras L (2006). Psychotic (delusional) major depression in the elderly: a review. *Current Psychiatry Reviews* **2**, 235-244.

Gournellis R, Oulis P, Howard R (2014). Psychotic major depression in older people: a systematic review. *Internatinal Journal of Geriatric Psychiatry* 29, 789-796.

Haley GM, Fine S, Marriage K (1988). Psychotic features in adolescents with major depression. *Journal of the American Academy of Child & Adolescent Psychiatry* **27**, 489-493.

Heslin M, Desai R, Lappin JM, Donoghue K, Lomas B, Reininghaus U, Onyejiaka A, Croudace T, Jones PB, Murray RM, Fearon P, Doody GA, Dazzan P, Fisher HL, Demjaha A, Craig T, Morgan C (2016a). Biological and psychosocial risk factors for psychotic major depression. Social Psychiatry and Psychiatric Epidemiology **51**, 233-245.

Heslin M, Lappin JM, Donoghue K, Lomas B, Reininghaus U, Onyejiaka A, Croudace T, Jones PB, Murray RM, Fearon P, Doody GA, Dazzan P, Craig TJ, Morgan C (2016b). Ten-year outcomes in first episode psychotic major depression patients compared with schizophrenia and bipolar patients. Schizophrenia Research **176**, 417-422.

Heslin M, Lomas B, Lappin JM, Donoghue K, Reininghaus U, Onyejiaka A, Croudace T, Jones PB, Murray RM, Fearon P, Dazzan P, Morgan C, Doody GA (2015). Diagnostic change 10 years after a first episode of psychosis. *Psychological Medicine* **45,** 2757-2769.

Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *British Medical Journal* **327**, 557-560.

Hirschfeld RM (2012). The epidemiology of depression and the evolution of treatment. *Journal of Clinical Psychiatry* **73**, Suppl 1, 5-9.

Ihezue UH (1985). Observations and comments on the psychosocial determinants of depressive illness among Nigerian adults. *Journal of the National Medical Association* **77**, 729-733.

Jäger M, Bottlender R, Strauss A, Möller HJ (2005). Fifteen-year follow-up of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition depressive disorders: the prognostic significance of psychotic features. *Comprehensive Psychiatry* **46**, 322-327.

Jarbin H, Ott Y, Von Knorring AL (2003). Adult outcome of social function in adolescent-onset schizophrenia and affective psychosis. *Journal of the American Academy of Child and Adolescent Psychiatry* **42**, 176-183.

Johnson J, Horwath E & Weissman MM (1991). The validity of major depression with psychotic features based on a community study. *Archives of General Psychiatry* **48**,1075-1081.

Käkelä J, Panula J, Oinas E, Hirvonen N, Jääskeläinen E, Miettunen J (2014). Family history of psychosis and social, occupational and global outcome in schizophrenia: a meta-analysis. *Acta Psychiatrica Scandinavica* **130**, 269-278.

Karaaslan F, Gonul AS, Oguz A, Erdinc E, Esel E (2003). P300 changes in major depressive disorders with and without psychotic features. *Journal of Affective Disorders* **73**, 283-287.

Keller J, Schatzberg AF, Maj M (2007). Current issues in the classification of psychotic major depression. *Schizophrenia Bulletin* **33**, 877-885.

Kessler RC, Bromet EJ (2013). The epidemiology of depression across cultures. *The Annual Review of Public Health* **34**, 119-138.

Kivelä SL, Pahkala K (1989). Delusional depression in the elderly: a community study. *Zeitschrift fur Gerontologie* **22**, 236-241.

Kruijshaar ME, Hoeymans N, Bijl RV, Spijker J, Essink-Bot ML (2003). Levels of disability in

major depression: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Journal of Affective Disorders* **77**, 53-64.

Leadholm AK, Rothschild AJ, Nolen WA, Bech P, Munk-Jørgensen P, Ostergaard SD (2013). The treatment of psychotic depression: is there consensus among guidelines and psychiatrists? *Journal of Affective Disorders* **145**, 214-220.

Lykouras L, Gournellis R (2009). Psychotic (delusional) major depression: new vistas. *Current Psychiatry Reviews* **5**, 1–28.

Maj M, Pirozzi R, Magliano L, Fiorillo A, Bartoli L (2007). Phenomenology and prognostic significance of delusions in major depressive disorder: a 10-year prospective follow-up study. *The Journal of Clinical Psychiatry* **68**, 1411-1417.

Matheson SL, Shepherd AM, Carr VJ (2014). How much do we know about schizophrenia and how well do we know it? Evidence from the Schizophrenia Library. *Psychological Medicine* **20**, 1–19.

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine* **6**:e1000097.

Nakamura K, Iga J, Matsumoto N, Ohmori T (2015). Risk of bipolar disorder and psychotic features in patients initially hospitalised with severe depression. *Acta Neuropsychiatrica* **27**, 113-118.

Nelson JC, Davis JM (1997). DST studies in psychotic depression: a meta-analysis. *The American Journal of Psychiatry* **154**, 1497-1503.

Ohayon MM, Schatzberg AF (2002). Prevalence of depressive episodes with psychotic features in the general population. *The American Journal of Psychiatry* **159**, 1855-1861.

Okulate GT, Oladapo HT, Osibogun A (2001). Comparison of three subtypes of depression. *The Nigerian Postgraduate Medical Journal* **8**, 41-45.

Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jönsson B; CDBE2010 study group; European Brain Council (2012). The economic cost of brain disorders in Europe. *European Journal of Neurology* **19**, 155-162.

Opjordsmoen S (1991). Unipolar delusional depression. Outcome in reactive versus endogenous types. *Psychopathology* **24**, 260-269.

Ormel J, Oldehinkel AJ, Nolen WA, Vollebergh W (2004). Psychosocial disability before, during, and after a major depressive episode: a 3-wave population-based study of state, scar, and trait effects. *Archives of General Psychiatry* **61**, 387-392.

Østergaard SD, Straszek S, Petrides G, Skadhede S, Jensen SO, Munk-Jørgensen P, Nielsen J (2014). Risk factors for conversion from unipolar psychotic depression to bipolar disorder. *Bipolar Disorders* **16**, 180-189.

Østergaard SD, Waltoft BL, Mortensen PB, Mors O (2013). Environmental and familial risk factors for psychotic and non-psychotic severe depression. *Journal of Affective Disorders* **147**, 232-240.

Park SC, Choi J, Kim JM, Jun TY, Lee MS, Kim JB, Yim HW, Park YC (2014). Is the Psychotic Depression Assessment Scale a useful diagnostic tool? The CRESCEND study. *Journal of*

Affective Disorders **166**, 79-85.

Parker G, Hadzi-Pavlovic D, Hickie I, Mitchell P, Wilhelm K, Brodaty H, Boyce P, Eyers K, Pedic F (1991). Psychotic depression: a review and clinical experience. *The Australian and New Zealand Journal of Psychiatry* **25**, 169-180.

Penttilä M, Jääskeläinen E, Hirvonen N, Isohanni M, Miettunen J (2014). Duration of untreated psychosis as predictor of long-term outcome in schizophrenia: systematic review and meta-analysis. *British Journal of Psychiatry* **205**, 88-94.

Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, Partonen T, Tuulio-Henriksson A, Hintikka J, Kieseppä T, Härkänen T, Koskinen S, Lönnqvist J (2007). Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Archives of General Psychiatry* **64**, 19-28.

Reay R, Mitford E, McCabe K, Paxton R, Turkington D (2019). Incidence and diagnostic diversity in first-episode psychosis. *Acta Psychiatrica Scandinavica* **121**, 315-319.

Rothschild AJ (2009). *Clinical Manual for Diagnosis and Treatment of Psychotic Depression*. American Psychiatric Publishing Inc.: Washington DC.

Rothschild AJ, Winer J, Flint AJ, Mulsant BH, Whyte EM, Heo M, Fratoni S, Gabriele M, Kasapinovic S, Meyers BS (2008). Study of Pharmacotherapy of Psychotic Depression (STOP-PD) Collaborative Study Group. Missed diagnosis of psychotic depression at 4 academic medical centers. *Journal of Clinical Psychiatry* **69**, 1293-1296.

Ruggero CJ, Kotov R, Carlson GA, Tanenberg-Karant M, Gonzalez DA, Bromet EJ (2011). Diagnostic consistency of major depression with psychosis across 10 years. *The Journal of Clinical Psychiatry* **72**, 1207-1213.

Rush AJ, Carmody TJ, Ibrahim HM, Trivedi MH, Biggs MM, Shores-Wilson K, Crismon ML, Toprac MG, Kashner TM (2006). Comparison of self-report and clinician ratings on two inventories of depressive symptomatology. *Psychiatric Services* **57**, 829-837.

Ryan ND, Puig-Antich J, Ambrosini P, Rabinovich H, Robinson D, Nelson B, Iyengar S, Twomey J (1987). The clinical picture of major depression in children and adolescents. *Archives of General Psychiatry* **44**, 854-861.

Saarni SI, Viertio S, Perälä J, Koskinen S, Lönnqvist J, Suvisaari J (2010). Quality of life of people with schizophrenia, bipolar disorder and other psychotic disorders. *The British Journal of Psychiatry* **197**, 386-394.

Salvatore P, Baldessarini RJ, Tohen M, Khalsa HM, Sanchez-Toledo JP, Zarate CA,Jr, Vieta E, Maggini C (2011). McLean-Harvard International First-Episode Project: two-year stability of ICD-10 diagnoses in 500 first-episode psychotic disorder patients. *The Journal of Clinical Psychiatry* **72**, 183-193.

Sands JR, Harrow M (1995). Vulnerability to psychosis in unipolar major depression: is premorbid functioning involved? *The American Journal of Psychiatry* **152**, 1009-1015.

Schatzberg AF, Rothschild AJ (1992). Psychotic (delusional) major depression: should it be included as a distinct syndrome in DSM-IV? *The American Journal of Psychiatry* **149**, 733-745.

Sherazi R, McKeon P, McDonough M, Daly I, Kennedy N (2006). What's new? The clinical epidemiology of bipolar I disorder. *Harvard Review of Psychiatry* **14**, 273-284.

Simpson S, Baldwin RC, Jackson A, Burns A (1999). The differentiation of DSM-III-R psychotic depression in later life from nonpsychotic depression: comparisons of brain changes measured by multispectral analysis of magnetic resonance brain images, neuropsychological findings, and clinical features. *Biological Psychiatry* **45**, 193-204.

Souery D, Zaninotto L, Calati R, Linotte S, Sentissi O, Amital D, Moser U, Kasper S, Zohar J, Mendlewicz J, Serretti A (2011). Phenomenology of psychotic mood disorders: lifetime and major depressive episode features. *Journal of Affective Disorders* **135**, 241-250.

StataCorp (2013). *Stata Statistical Software: Release 13*. College Station: StataCorp LP.

Sterne J (2009). *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*. College Station: Stata Press.

Suvisaari J, Partti K, Perälä J, Viertio S, Saarni SE, Lönnqvist J, Saarni SI, Härkänen T (2013). Mortality and its determinants in people with psychotic disorder. *Psychosomatic Medicine* **75**, 60-67.

Tohen M, Khalsa HM, Salvatore P, Vieta E, Ravichandran C, Baldessarini RJ (2012). Two-year outcomes in first-episode psychotic depression the McLean-Harvard First-Episode Project. *Journal of Affective Disorders* **136**, 1-8.

Tohen M, Strakowski SM, Zarate C,Jr, Hennen J, Stoll AL, Suppes T, Faedda GL, Cohen BM, Gebre-Medhin P, Baldessarini RJ (2000). The McLean-Harvard first-episode project: 6-month symptomatic and functional outcome in affective and nonaffective psychosis. *Biological Psychiatry* **48**, 467-476.

Vythilingam M, Chen J, Bremner JD, Mazure CM, Maciejewski PK, Nelson JC (2003). Psychotic depression and mortality. *The American Journal of Psychiatry* **160**, 574-576.

WHO (1992). *International Classification of Diseases and Related Health Problems. 10th revision.* World Health Organization: Geneva.

Wijkstra J, Lijmer J, Burger H, Cipriani A, Geddes J, Nolen WA (2015). Pharmacological treatment for psychotic depression. *The Cochrane Database of Systematic Reviews* **7**:CD004044. Doi: CD004044.

Zalpuri I, Rothschild AJ (2016). Does psychosis increase the risk of suicide in patients with major depression? A systematic review. *Journal of Affective Disorders* **198**, 23-31.

Zaninotto L, Guglielmo R, Calati R, Ioime L, Camardese G, Janiri L, Bria P, Serretti A (2015). Cognitive markers of psychotic unipolar depression: a meta-analytic study. *Journal of Affective Disorders* **174**, 580-588.

Table 1. Summary of the main results.

Topic	Main result
Occurrence	
Annual incidence	0.21-6.4/100 000 (higher in females).
Lifetime prevalence	0.35-1.0% (higher in older samples and females).
Point prevalence	0.5% (higher in females).
Proportion of all depressions	28% of all depressive patients, being higher in older samples and among inpatients.
Gender	
	Appr. 65% of the PD patients were females, this was comparable to NPD but higher than especially in SZ.
Onset age	No significant difference in onset age in PD versus NPD. Among youngest samples PD patients had earlier onset age, whereas in oldest samples PD patients had later onset age compared to NPD. SZ patients had younger age of illness onset than PD patients.
Risk factors	
	Lack of studies on early risk factors. Individuals with PD were less likely to be Caucasian and had more often family history of psychosis and bipolar I disorder when compared to NPD patients. Differences in educational level and marital status between PD and NPD were mostly non-significant.
Outcomes	
Depression symptoms	more severe in PD compared to NPD no difference in PD compared to SZ and PBD
Psychosis symptoms	more severe in PD compared to NPD less severe in PD compared to SZ
Positive symptoms	less severe in PD compared to SZ and PBD
Negative symptoms	less severe in PD compared to SZ more severe in PD compared to PBD
Symptomatic remission	somewhat less common in PD than NPD more common in PD compared to SZAFF
Clinical global outcome	somewhat poorer in PD than NPD
Relapses	higher in PD compared to NPD lower in PD compared to SZ relatively similar in PD and PBD
Global outcome	somewhat worse in PD compared to NPD better in PD compared to SZ no difference between PD and PBD
Occupational outcomes	somewhat poorer in PD, but in many studies also similar to NPD better in PD compared to SZ and SZAFF relatively similar in PD and PBD

Diagnoses: PD = psychotic depression, NPD = nonpsychotic depression, PD = bipolar disorder, PBD = psychotic bipolar disorder, SZ = schizophrenia, SZAFF = schizoaffective disorder

Fig. 1. Flow diagram of the selection of studies.

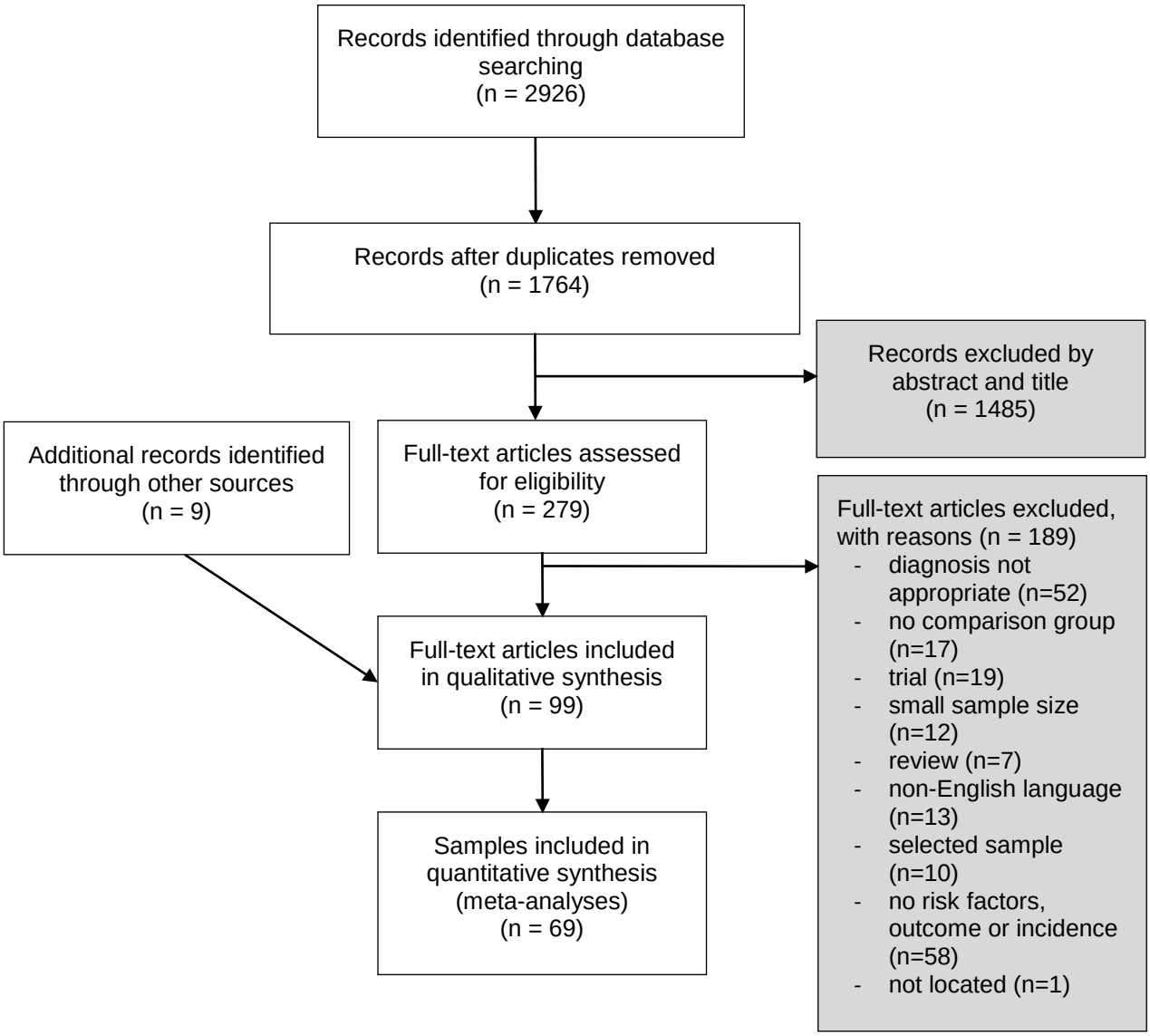


Fig. 2. Forest plot for estimated or relative rates (RR) psychotic depression among females in depression patients. Studies are grouped and analyzed separately by mean study age.

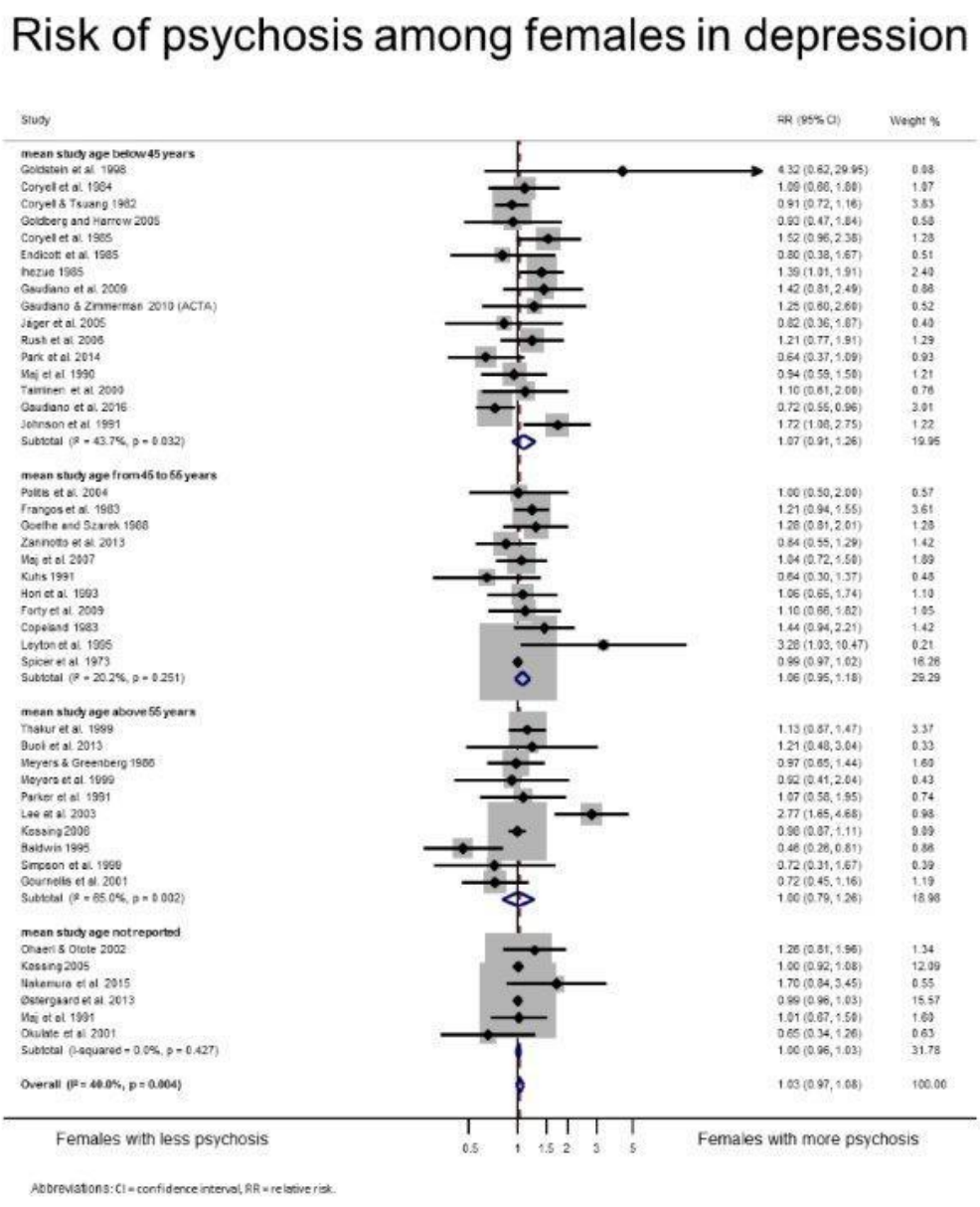
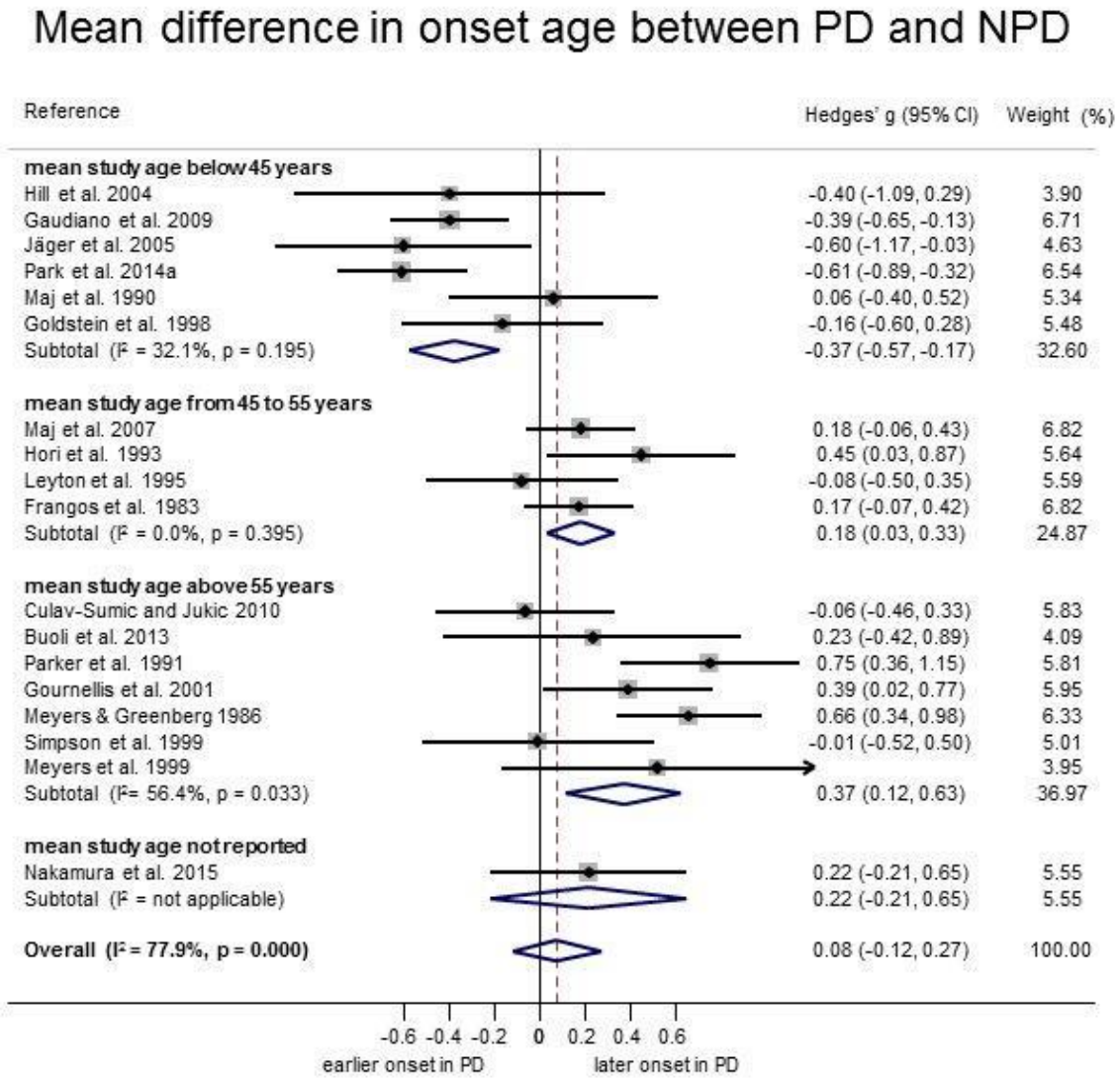


Fig. 3. Forest plot comparing mean onset age between psychotic depression and non-psychotic depression. Studies are grouped and analyzed separately by mean study age.



Abbreviations: PD = psychotic depression, NPD = non-psychotic depression, I^2= heterogeneity, CI = confidence interval.

Supplement Table 1. Studies on incidence or prevalence of psychotic depression.

Reference	Country	Diagnosis	Comments	Results
Community studies				
Johnson *et al*. 1991	USA (1980-1984)	DSM-III	Part of the Epidemiological Catchment Area (ECA) study (n=14212). Community sample of those with age of 18 or more living in five US cities (New Haven, St. Louis, Baltimore, Durham, Los Angeles). Diagnostic Interview Schedule was used as a screen, people were given a psychotic depression diagnosis if they had MDD in the depression module and endorsed psychotic symptoms in the psychosis module.	Lifetime prevalence rate for PD was in the total sample 0.6% (varying between 0.3 and 0.8 in different cities). Validity of the used psychotic depression diagnosis was poor.
Kivelä & Pahkala, 1989	Finland (1984-1986)	DSM-III	Community based study in Ähtäri, Finland, of those with 60 years or more (n=594). Zung Self-Rating Depression Scale was used as a screen.	Prevalence 1.0% for total sample (0.6% for males, 1.2% for females).
Ohayon & Schatzberg, 2002	Several European countries (1994-1999)	DSM-IV	Telephone survey of 18980 people between ages 15-100. Sleep-EVAL expert system covering DSM-IV was used as diagnostic instrument.	The overall point prevalence for PD was 0.5%. The point prevalence for PD was 0.5% in the United Kingdom, 0.4% in Germany, 0.2% in Italy, 0.5% in Portugal, and 0.2% in Spain. Point prevalence estimates by gender, age, marital status and occupation were also reported, significantly higher rates were reported for females (0.6%) than males (0.3%). Unemployed persons had prevalence of 1.2%, which differed from other occupation groups (e.g. 0.3% among daytime workers).
Perälä *et al*. 2007	Finland (2000-2001)	DSM-IV	Nationally representative Finnish Health 2000 sample (age 30 or older, n=8028). The study used information from Composite International Diagnostic Interview and from other sources (self-reported diagnoses, medical examination and national register) as a screen.	Lifetime prevalence 0.35% (males 0.41%, females 0.29%). By age: 30-44y (Total 0.30%, males 0.30%, females 0.30%), 45-54y (0.36%, males 0.52%, females 0.21%), 55-64y (0.31%, males 0.33%, females 0.30%), 65y or more (0.43%, males 0.54%, females 0.36%). Differences by age group or by sex within an age group were not statistically significant.
Studies based on in- and outpatient admissions				
Farquhar *et al*. 2007	UK (1875-1924,	ICD-10	Admissions in North West Wales, UK.	Annual incidence 3.4 per 100.000 persons in 1875-1924 and 3.0 per 100.000 in 1995-

	1995-2005)			1999.
Pederson *et al*. 1972	USA (1961-1962)	DSM-I	Psychiatric case register in Monroe County, New York, USA.	Overall age-adjusted yearly prevalence rate was 0.70/1000/year (men 0.53, women 0.87), whereas overall age-adjusted two-year incidence rate was 0.33/1000/year (men 0.27, women 0.37).
Filatova *et al*. 2016	Finland (1980-1993, 2000-2013)	ICD-9, ICD-10	Compares two Northern Finland Birth cohorts, born 1966 (NFBC1966) and 1986 (NFBC1986). Total sample size 12058 (NFBC1966) and 9432 (NFBC1986).	Cumulative incidence until age 27 was 0.02% in NFBC1966 and 0.21% in NFBC1986 ($p<0.001$).
Baldwin *et al*. 2005; Owoeye *et al*. 2013	Ireland (1995-2003)	DSM-IV	The Cavan-Monaghan First Episode Study. Population n=103054.	Annual incidence (per 100,000 population aged >15) was 6.4 (males 5.4 and females 7.4). In the follow-up study of 6 months, an incidence of 6.9 (males 6.5, females 7.4) per 100,000 of population was reported.
Reay *et al*. 2010; Crebbin *et al*. 2008	UK (1998-2005)	ICD-10	Used Census 2001 to estimate population. Used inpatient and outpatient admissions (age 16 or more) in Northumberland, UK.	Incidence per 100 000 for those with 16 years and over: was 6.0 (in age group 16-64 it was 5.4). PD was less common than NPD in younger people (under 36 years).

Abbreviations: DSM = Diagnostic and Statistical Manual of Mental disorders, ICD = International Classification of Diseases;, NFBC = Northern Finland Birth Cohort, PD = psychotic depression, NPD = nonpsychotic depression.

Supplement Table 2. Proportion of female patients in studies comparing psychotic depression with non-psychotic depression, schizophrenia or psychotic bipolar disorder.

Reference	Country	Diagnosis	Psychotic depression		Non-psychotic depression		Schizophrenia		Schizoaffective disorder		Psychotic bipolar disorder	
			N	Females	N	Females	N	Females	N	Females	N	Females
Baldwin 1995	UK	DSM-III-R	34	53%	100	77%						
Beiser *et al*. 1993	Canada	DSM-III	30	37%			72	22%			39	44%
Benazzi 1999	Italy	DSM-IV	40	55%							30	63%
Breslau & Meltzer, 1988	USA	RDC	39	64%					34	65%	38	68%
Buoli *et al*. 2013	Italy	DSM-IV-TR	18	83%	18	78%						
Copeland, 1983	UK	own system	55	75%	39	56%						
Coryell & Tsuang, 1982	USA	Feighner	122	53%	103	58%						
Coryell & Zimmerman, 1988	USA	RDC	29	62%			21	24%	47	62%		
Coryell *et al*. 1984b	USA	DSM-III	55	71%	180	67%						
Coryell *et al*. 1985[a]	USA	DSM-III	76	68%	473	57%						
Coryell *et al*. 1990[a]	USA	RDC	92	65%					42	64%		
Craig & Bromet, 2004	USA	DSM-IV	87	59%			227	35%			139	50%
Craig *et al*. 2000	USA	DSM-IV	75	60%			155	34%			119	52%
Crebbin *et al*. 2008	UK	ICD-10	105	54%			73	34%				
Dell'Osso *et al*. 2002	Italy	DSM-III-R	30	67%							147	55%
Endicott *et al*. 1985	USA	RDC	26	62%	178	67%					102	56%
Forty *et al*. 2009	UK	DSM-IV, ICD-10	64	70%	460	68%						
Frangos *et al*. 1983	Greece	RDC	145	70%	119	61%						
Gaudiano & Zimmerman, 2010	USA	DSM-IV	32	69%	591	63%						
Gaudiano *et al*. 2009	USA	DSM-IV	60	73%	1052	66%						
Gaudiano *et al*. 2016	USA	DSM-IV-TR	174	57%	1140	66%						
Goethe & Szarek, 1988	USA	DSM-III	77	71%	360	65%						
Goldberg & Harrow, 2005	USA	RDC	27	63%	95	65%						
Goldstein *et al*. 1998	USA	RDC	25	96%	93	82%						
Gournellis *et al*. 2001	Greece	DSM-IV	45	69%	73	79%						
Heslin *et al.* 2016	UK	ICD-10	72	50%			218	36%			70	53%
Hill *et al*. 2004	USA	DSM-III-R, DSM-IV	20	55%			86	38%				
Hori *et al*. 1993	Japan	DSM-III-R	38	55%	55	53%						

Ihezue, 1985	Nigeria	ICD-9	78	63%	64	45%						
Jarbin *et al*. 2003	Sweden	DSM-IV	17	53%			32	34%			25	64%
Johnson *et al*. 1991	USA	DSM-III	114	83%	662	73%						
Jäger *et al*. 2005	Germany	DSM-IV	20	80%	33	85%	64	56%				
Karaaslan *et al*. 2003	Turkey	DSM-IV	16	50%							20	40%
Kessing, 2005	Denmark	ICD-10	1497	66%	2962	62%						
Kessing, 2006	Denmark	ICD-10	705	66%	1044	66%						
Kuhs, 1991	Germany	ICD-9, DSM-III	23	48%	137	61%						
Lee *et al*. 2003	Taiwan	DSM-IV	48	69%	108	33%						
Leyton *et al*. 1995	Canada	DSM-III-R	25	88%	140	66%						
Maj *et al*. 1990	Italy	DSM-III	36	58%	36	61%						
Maj *et al*. 1991	Italy	DSM-III-R	41	63%	27	63%	28	54%	21	57%		
Maj *et al*. 2007	Italy	DSM-III	89	62%	240	60%						
Meyers & Greenberg, 1986	USA	DSM-III	72	76%	89	78%						
Meyers *et al*. 1999	USA	DSM-III-R	15	67%	20	70%						
Nakamura *et al*. 2015	Japan	ICD-10	33	79%	56	63%						
Ohaeri & Otote, 2002	ICD-10	Nigeria	51	73%	45	62%						
Okulate *et al*. 2001	ICD-10	Nigeria	31	39%	144	51%						
Opjordsmoen, 1991	Norway	DSM-III	50	60%					33	45%		
Østergaard *et al*. 2013a	Denmark	ICD-10	8260	65%	15913	65%						
Owoeye *et al*. 2013	Ireland	DSM-IV	77	53%			73	26%			54	52%
Park *et al*. 2014a	South Korea	DSM-IV	53	68%	441	78%						
Parker *et al*. 1991	Australia	DSM-III, RDC	35	66%	102	64%						
Politis *et al*. 2004	Greece	DSM-IV	16	50%	16	50%	20	40%				
Rush *et al*. 2006	USA	DSM-IV	106	82%	438	78%						
Rybakowski *et al*. 2007	Poland	DSM-IV	26	81%							92	60%
Simpson *et al*. 1999	UK	DSM-III-R	18	61%	81	70%						
Spicer *et al*. 1973	UK	ICD-9	14972	68%	13844	68%						
Taiminen *et al*. 2000	Finland	DSM-IV	23	57%	25	52%	17	71%				
Thakur *et al*. 1999	USA	DSM-III-R	189	67%	485	63%						
Zaninotto *et al*. 2013	European multicenter	DSM-IV	90	72%	609	76%						

[a] Overlapping samples, but different comparison group. Abbreviations: DSM = Diagnostic and Statistical Manual of Mental disorders, ICD = International Classification of Diseases, RDC = Research Diagnostic Criteria.

Supplement Table 3. Mean and Standard deviation (SD) of onset age of illness in studies comparing psychotic depression and non-psychotic depression, schizophrenia or psychotic bipolar disorder.

Reference	Country	Diagnosis	Study design	Psychotic depression			Non-psychotic depression			Schizophrenia / schizoaffective disorder[a]			Psychotic bipolar disorder		
				N	mean	SD	N	mean	SD	N	mean	SD	N	mean	SD
First-episode samples															
Baldwin *et al*. 2005	Ireland	DSM-IV	First-episode	39	45.6	22.3				66	31.3	16.6	32	34.8	16.2
Breslau & Meltzer, 1988[b]	USA	RDC	First-episode	22	26.3	11.2				22	24.8	9.7	22	23.6	7.4
Hill *et al*. 2004	USA	DSM-III-R. DSM-IV	First-episode	20	25.2	8.8	14	29.0	10.1	86	28.6	9.6			
Husted *et al*. 1995	Canada	DSM-III	First-episode	35	26.3	8.1				91	22.6	5.8	38	26.3	7.9
Nakamura *et al*. 2015	Japan	ICD-10	First-episode	33	51.5	16.3	56	48.1	14.9						
Owoeye *et al*. 2013	Ireland	DSM-IV	First-episode	77	51.2	22.0				73	30.9	14.5	54	31.9	13.9
Consecutive or mixed samples															
Benazzi, 1999	Italy	DSM-IV	Mixed	40	35.8	15.0							30	33.5	13.3
Buoli *et al*. 2013	Italy	DSM-IV-TR	Mixed	18	41.8	17.1	18	38.1	13.5						
Čulav-Sumić & Jukić, 2010	Croatia	ICD-10	Consecutive	50	39.3	13.4	50	40.1	11.5						
Dell'Osso *et al*. 2002	Italy	DSM-III-R	Mixed	30	36.2	12.7							147	36.5	12.3
Frangos *et al*. 1983	Greece	RDC	Mixed	145	44.2	1.2	119	41.7	14.2						
Gaudiano *et al*. 2009	USA	DSM-IV	Mixed	60	20.7	12.1	1052	26.1	13.8						
Goldstein *et al*. 1998	USA	RDC	Mixed	25	26.9	8.4	93	28.4	9.4						
Gournellis *et al*. 2001	Greece	DSM-IV	Consecutive	45	58.3	15.2	73	51.8	17.2						
Hori *et al*. 1993	Japan	DSM-III-R	Mixed	38	48.3	15.1	55	41.7	14.3						
Jäger *et al*. 2005	Germany	DSM-IV	Mixed	20	39.2	14.7	33	47.1	11.8	64	28.7	10.8			
Leyton *et al*. 1995	Canada	DSM-III-R	MIxed	25	38.4	2.3	140	39.4	13.0						
Maj *et al*. 1990	Italy	DSM-III	Mixed	36	32.0	3.0	36	31.8	3.5						
Maj *et al*. 2007	Italy	DSM-III	Mixed	89	33.6	6.4	240	32.5	5.8						
Meyers & Greenberg, 1986	USA	DSM-III	Mixed	72	62.4	15.5	89	51.5	17.2						
Meyers *et al*. 1999	USA	DSM-III-R	Mixed	15	67.5	16.2	20	58.0	19.1						
Park *et al*. 2014a	South Korea	DSM-IV	Mixed	53	28.6	14.1	441	38.3	16.2						
Parker *et al*. 1991	Australia	DSM-III. RDC	Mixed	35	48.1	16.5	101	36.2	15.5						
Simpson *et al*. 1999	UK	DSM-III-R	Mixed	18	63.0	14.1	81	63.1	11.0						

[a] Schizophrenia in other studies except schizoaffective disorder in Breslau & Meltzer (1988), [b] Estimated sample size. Abbreviations: DSM = Diagnostic and Statistical Manual of Mental disorders, ICD = International Classification of Diseases, RDC = Research Diagnostic Criteria, SD = standard deviation.

Supplement Table 4. Studies on sociodemographic and risk factors in psychotic depression and comparison groups.

Reference (city, country)	Diagnosis (setting)	Sample size (males/females)	Comparison groups	Age of the PD sample, mean (SD)	Sociodemographic or risk factors	Main results and comments
Baldwin, 1995 (Manchester, UK)	DSM-III-R (mood-incongruent delusional depression) (Not reported)	34 (16/18)	NPD (n=100)	median 75.5 [range 65-89]	marital status	Elderly sample. Individuals with PD were more often single (44%) than those with NPD (12%) (p=0.04).
Breslau & Meltzer, 1988 (Chicago, USA)	RDC (I)	39 (14/25)	PBD (n=38), SZAFF (n=34)	37.1 (13.2)	education, ethnicity	Those with PBD had more years of education (mean 13.3, SD 2.3) than those with PD (mean 11.6, SD 3.2) and SZAFF (mean 11.3, SD 2.8) (p<0.005). No difference in percentage of white (69% in PD, 61% in PBD and 62% in SZAFF).
Buoli *et al.* 2013 (Milan, Italy)	DSM-IV-TR (I)	18 (3/15)	NPD (n=18, matched for sample size)	59.7 (17.0)	family history for major psychoses	PD patients had more often family history for major psychoses than those with NPD (6% vs. 39%, p=0.016).
Coryell & Tsuang, 1982 (Iowa City, USA)	Feighner (delusional depression) (I)	122 (57/65)	nondelusional depression (122)	43.8 (Not reported)	marital status	Iowa 500 study. 16% of PD patients and 20% of the NPD patients had never been married. Difference in proportion was not significant.
Coryell & Zimmerman, 1988 (Iowa City, USA)	RDC (I)	29 (11/18)	SZ (n=21), SZAFF (n=47), HC (n=38)	42.8 (16.3)	marital status	Family study. Percentage of ever married was 66% in PD, 24% in SZ, 55% in SZAFF, and 63% among controls, differences in percentages were non-significant.
Coryell *et al.* 1984b (Iowa City, USA)	DSM-III (I)	55 (16/39)	NPD (n=180)	40.9 (16.0)	marital status	Individuals with PD were non-significantly more common single (28%) than those with NPD (21%).
Coryell *et al.* 1985 (five centers, USA)	DSM-III (I)	76 (24/52)	NPD (n=473)	39.0 (16.0)	marital status	Individuals with PD were non-significantly more common single (36%) than those with NPD (29%). Studied separately mood-congruent and mood-incongruent PD (here we have combined those).

Coryell *et al.* 1990 (five centers, USA)	RDC (I/O)	92 (32/60)	SZAFF depressed type (n=42)	39.1 (15.4)	education	Studied separately those who completed the follow-up and those with only baseline data (here we have combined those groups). Individuals with PD had non-significantly (p=0.20) higher level of education (mean 3.5, SD 1.5) than those with schizoaffective disorder (mean 3.2, SD 1.1).
Craig *et al.* 2000 [Craig *et al.* 1997] (Suffolk County, USA)	DSM-IV (I/O)	75 (30/45) [in 1997 study: 42 (24/18)]	SZ/SZAFF (n=155 [n=96]), PBD (n=119 [n=64])	28.0 (9.3)	ethnicity, marital status, social class, [education]	Data based on Suffolk County Mental Health Project, epidemiological study of new admissions. Groups did not differ significantly regarding proportion of African American (PD 7%, SZ/SZAFF 22%, PBD 6%); never married (55%, 79%, 61%; respectively); having high school education (87%, 77%, 87%; respectively); or being from lower social class (11%, 19%, 13%, respectively). Fennig *et al.* (1996) have published similar results from the same study, with smaller sample.
Čulav-Sumić & Jukić, 2010 (Zagreb, Croatia)	ICD-10 recurrent PD (I)	50 (0/50)	recurrent NPD (n=50), HC (n=50, hospital personal)	55.3 (10.2)	minor physical anomalies (51 different anomalies, also total sum score)	Sample included only women. Mean number of physical anomalies was highest in PD (mean 2.9, SD 1.4). Corresponding mean was 2.5 (SD 1.6) for NPD and 1.6 (SD 1.4) for HC. Difference between PD and HC was statistically significant (p<0.001), but not between PD and NPD (p=0.36).
Dell'Osso *et al.* 2002 (Pisa, Italy)	DSM-III-R (I)	30 (10/20)	PBD (n=147; most recent episode: pure mania, n=55; mixed mania, n=62; depressed, n=30)	36.2 (12.7)	education, marital status	PD did not differ significantly from PBD on having graduate/undergraduate education (57% vs. 54%) or in being unmarried (67% vs. 76%). Compared originally three PBD groups and PD, here PBD groups are pooled.
Draper & Anstey, 1996 (Sydney, Australia)	DSM-III, DSM-III-R, ICD-9 (I)	47 (Not reported)	NPD (148)	not reported for PD, mean age for the total sample	psychosocial and medical stressors at admission (eight categories)	Elderly sample. Compared originally to three comparison groups, here major and minor depression are pooled to NPD, whereas organic

				74.1 years		depression has been excluded. Acute medical problems were less common in PD (30%) than in NPD (51%) (p=0.012). No differences in psychosocial stressors.
Frangos *et al.* 1983 (Athens, Greece)	RDC (I)	145 (43/102)	NPD (n=119)	53.5 (1.1)	family history for mental illness and for affective disorders	Groups did not differ by family history (mental illness: PD 32%, NPD 30%; affective disorders: PD 19%, NPD 17%).
Gaudiano *et al*. 2009 [Gaudiano & Zimmerman, 2010] (Rhode Island, USA)	DSM-IV (O)	60 (16/44) [32 (10/22)]	NPD (n=1052 [591])	37.0 (11.7) [36.8 (13.1)]	education, ethnicity, marital status [childhood trauma]	PD patients were less often Caucasian than those with NPD (65% vs 86%, p<0.001). PD patients were also less often college graduates (13% vs. 34%, p=0.001). The difference between groups was not significant in amount of single patients (PD 22%, NPD 27%). Individuals with PD are more likely to have physical or sexual trauma than those with NPD (84% vs. 64%, p=0.017). Gaudiano *et al*. (2008) studied an overlapping sample.
Gaudiano *et al.* 2016 (Rhode Island, USA)	DSM-IV-TR (I/O)	174 (74/100)	NPD (n=1140)	44.9 (12.3)	ethnicity, race, marital status, history of physical, sexual or emotional abuse	PD patients were less often white (79% vs. 89%; p<0.001), more often Latino or Hispanic (26% vs. 6%, p<0.001), and less often married or with domestic partner (26% vs. 35%, p=0.017). There were no differences in history of abuse (PD 65% vs. NPD 58%, p=0.18).
Goldberg & Harrow, 2005 [Sands & Harrow, 1995] (Chicago, USA)	RDC (I)	27 (10/17) [22 (Not reported)]	NPD (n=95 [n=70])	21.9 (4.2) [23.3 (3.4]	education, marital status, ethnicity [premorbid social adjustment scale (SADS), Zigler-Phillips Social Competence Scale, Phillips Scale of Premorbid Adjustment (also subscales)]	Chicago Follow-up study. Individuals with PD were non-significantly less often Caucasian (67% vs. 83%; p=0.06) or married (11% vs. 20%; p=0.29) than those with NPD. Average years of education was 13.0 (SD 1.5) in PD and 13.8 (2.1) years in NPD (p=0.03). PD group had more patients with poor social adjustment score in SADS (47% vs. 15%, p=0.03) and with poor social competence score (23% vs. 13%,

						p=0.01). Groups did not differ significantly regarding proportion of poor premorbid adjustment in Phillips Scale (30% vs. 21%). Study included also a group of patients with any bipolar disorder (not shown here).
Gournellis *et al.* 2001 (Athens, Greece)	DSM-IV (I)	45 (14/31)	NPD (n=73)	69.6 (5.8)	education, marital status, family history of unipolar depression	PD and NPD group did not differ significantly regarding mean years of education (6.0 vs. 5.3), proportion of married (44% vs. 62%) or proportion of those with first degree relatives with unipolar depression (20% vs. 22%).
Heslin *et al*. 2016 (Nottingham and London, UK)	ICD-10 (I/O)	72 (36/36)	SZ (n=218), PBD (n=70), HC (n=391)	median 32.5 (interquartile range 25-41)	education, ethnicity, place of birth (UK vs. non-UK), current and lifetime relationship status, living alone, employment, contact with friends, contact with family, close confidants, severe life events, childhood adversity, family history of psychosis or any mental illness, neurological soft signs, minor physical abnormalities	First episode (incidence) study investigating several risk factors. Individuals with PD were statistically significantly more often living alone (OR 2.26), had basic level education (OR 2.89), were unemployed (OR 2.12), had less than monthly contact with friends (OR 4.24), no close confidants (OR 4.71), childhood adversity (OR 2.57), neurological soft signs (OR 1.15), and family history of any mental illness (OR 10.68) or psychosis (OR 12.85) when compared with controls. The study presented also ORs for SZ and PBD when compared to controls, the 95% confidence intervals did not overlap between SZ and PD in current relationship status were PD patients were less often single.
Hill *et al.* 2004 (Pittsburgh, USA)	DSM-III-R, DSM-IV (I)	20 (9/11)	SZ/SZAFF (n=86), HC (n=81)	25.2 (8.8)tot	education, ethnicity, own socioeconomic status (SES, Hollingshead Index of Social Position), parental SES	PD did not differ from other groups in mean years of education (PD 13.4, SZ/SZAFF 13.4, HC 14.5), in proportion of Caucasian (PD 65%, SZ/SZAFF 55%, HC 63%) or in own (mean scores: PD 2.9, SZ/SZAFF 3.3, HC 2.8) or parental SES (PD 2.7, SZ/SZAFF 2.9, HC 2.8). NPD group was small (n=14) and was not included here.
Ihezue, 1985	ICD-9 (I)	78 (29/49)	NPD (n=64)	~39 years	education, marital	First episode study. PD patients were

(Enugu, Nigeria)					status, urban/rural domicile, occupation	more often (62%) from rural domicile than patients with NPD (16%) (p<0.001). Proportion of single patients was 32% in PD and 53% in NPD (p=0.011). Regarding education, proportion of illiterate was 38% in PD and 13% in NPD (p=0.001). Amount of unskilled workers was 50% in PD and 34% in NPD (p=0.06).
Jarbin *et al.* 2003 (Lund, Sweden)	DSM-IV (I)	17 (8/9)	SZ (n=32), PBD (n=25)	median 16.2 [range 13.4-17.7]	heredity of bipolar and non-mood disorders.	First episode study, adolescent onset sample. Proportion of those with bipolar disorder heredity was 6% in PD, 30% in PBD, and 9% in SZ. Proportion of those with nonmood disorder heredity was 12% in PD, 17% in PBD, and 34% in SZ. Differences were nonsignificant. SZAFF group has been excluded here due to small (n=7) sample size.
Johnson *et al.* 1991 (New Haven, St. Louis, Baltimore, Durham, and Los Angeles, USA)	DSM-III (I/O)	114 (19/95)	NPD (n=662)	42.0 (16>.2)	ethnicity, marital status, socioeconomic status (SES)	Community based Epidemiological Catchment Area (ECA) study. Proportion of white was 61% in PD and 73% in NPD (p=0.005). Those with PD had lower SES than those with NPD (amount of those in lowest SES quartile was 21% in PD and 13% in NPD) (p=0.01). There were no differences in amount of married patients (38% in PD and 37% in NPD).
Karaaslan *et al.* 2003 (Turkey)	DSM-IV (I)	16 (8/8)	NPD (n=20), HC (n=20)	36.5 (8.7)	education	Mean years of education was lower (p<0.05) in PD (8.8; SD 1.16 years) than in NPD (10.4; 0.8) and in HC (10.02; 1.05).
Lee *et al.* 2003 (Taipei, Taiwan)	DSM-IV (I)	delusional depression 48 (15/33) [in family history n=37]	non-delusional depression (108) [in family history n=79]	74.7 (6.1)	education, family history of affective disorder	Mean years of education was 7.18 (SD 5.1) in delusional depression and 8.28 (4.9) in non-delusional depression. Proportion of those with family history of affective disorder was (5/37; 14%) in delusional depression and (8/79; 10%) in non-

						delusional depression. Differences were nonsignificant.
Maj *et al.* 1990 (Naples, Italy)	DSM-III (I/O)	36 (15/21)	NPD (n=36)	42.0 (6.3)	marital status, family history of affective disorders and schizophrenia	In PD 56% and in NPD 53% were married. Proportion of those with family history of affective disorders (PD 30% vs. NPD 25%) and schizophrenia (3% in both) did not differ between groups. NPD sample was restricted to same sample size than PD.
Maj *et al.* 2007 (Naples, Italy)	DSM-III (delusional depression) (I/O)	89 (34/55)	NPD (n=240)	45.2 (8.9)	family history of bipolar I disorder	Individuals with PD had more often family history of bipolar I disorder than those with NPD (16% vs. 7%, p=0.02).
Nakamura *et al.* 2015 (Tokushima, Japan)	ICD-10 (I)	33 (7/26)	NPD (n=56)	not reported for PD, mean age for the total sample ~62 years	family history of any psychiatric disease	Groups did not differ by family history of any psychiatric disease (PD: 41%, NPD: 46%).
Okulate *et al.* 2001 (Lagos, Nigeria)	ICD-10 (I/O)	31 (19/12)	NPD (n=144)	not reported for PD, mean age for the total sample 33.2 (9.7) years	marital status, religion, education, family history of mental illness, alcohol abuse, suicide or attempted suicide; precipitating factors of life events	Hospital with 40% patients being military personnel and 60% civilians. No significant difference between PD and NPD in being married (48% vs. 50%), being Christian (81% vs. 71%) or having secondary school or above education (68% vs. 73%). PD patients had more often family history of mental illness than those with NPD (45% vs. 15%, p<0.001), whereas there was no difference in family history of alcohol abuse (13% vs. 26%), family history of suicide (3% vs. 5%) or in precipitating life events (52% vs. 64%). Study included also a group of patients with any bipolar disorder (not shown here).
Østergaard *et al.* 2013b (whole Denmark)	ICD-8, ICD-10 (I/O)	2183 (Not reported)	HC (population sample of 2.4 Million)	not reported	history of mental disorders among mother, father or siblings, place of birth, born in city of Aarhus, gestational	Population-based historical prospective cohort study of all individuals born in Denmark between 1955 and 1990. Used in- and outpatient data from Danish Psychiatric Central Research

					age, birth weight, small for gestational age, maternal and paternal age at birth, loss of relative (mother, father or sibling; before and after age 15 years).	Register. Studied risk (incidence rate ratio, IRR) for PD (and also for NPD, data not shown here) in population using several register variables. Several maternal, paternal and sibling psychiatric diagnoses associated significantly with PD. Highest IRR (2.2, 95% CI 2.0-2.4; p<0.001) was in maternal any mental disorder. Loss of mother because of an unnatural cause after age 15 years associated with PD (IRR 1.7, 1.2-2.3; p=0.011). Other studied variables did not associate with PD.
Park *et al.* 2014a (18 centers, South Korea)	DSM-IV (I)	53 (17/36)	NPD (n=441)	40.7 (15.2)	education, marital status, religious affiliation, family history of depression and other mental disorders	Statistical tests are both unadjusted and adjusted with age and depression symptoms. Individuals with PD were had more education in years than those with NPD (12.2 vs. 10.3; p=0.003, adj. p=0.06). Differences were non-significant in being married (67% vs. 78%), having family history of depression (23% vs. 16%) or of other mental disorders (6% vs. 8%), or in having religious affiliation (62% vs. 65%).
Parker *et al.* 1991 (Sydney, Australia)	DSM-III, RDC (I/O)	35 (12/23)	NPD (n=102)	62.5 (12.1)	family history of depressive disorder, schizophrenia and alcoholism	PD did not differ significantly from those with NPD regarding proportion of those with family history of depressive disorder (31% vs. 39%), schizophrenia (9% vs. 5%) or alcoholism (11% vs. 14%). Studied also separately age- and sex-marched subsample of NPD (data not shown).
Politis *et al*. 2004 (Athens, Greece)	DSM-IV (I)	16 (8/8)	NPD (n=16), SZ (n=20), HC (n=20)	49 (15)	education	Mean (SD) education in years was 10 (4) in PD, 11 (3) in NPD, 11 (3) in SZ, and 12 (3) in HC. Groups did not differ significantly.
Rush *et al.* 2006 (Austin, USA)	DSM-IV (O)	106 (87/19)	NPD (n=438)	41.4 (10.7)	education, ethnicity, marital status	Proportion of white was lower in PD (38%) than in NPD (55%) (p<0.001). Proportion of married (27% in PD

						and 28% in NPD) and years of education (PD 11.1 (SD 2.6), NPD 11.2 (3.2)) did not differ between groups.
Simpson *et al.* 1999 (Manchester, UK)	DSM-III-R (I/O)	18 (7/11)	NPD (n=81)	75.2 (4.6)	education, premorbid IQ, family history of depression	Elderly population. PD and NPD did not differ by mean years of education (10.0 vs. 10.3), premorbid IQ (104.4 vs. 111.1) or by family history of depression (37% vs. 60%).
Thakur *et al.* 1999 (Durham, USA)	DSM-III-R (I/O)	189 (62/127)	NPD (n=485)	not reported (50% was 60 years or more)	family history of suicide	Family history of suicide was present in 32% among PD and in 24% among NPD patients (p=0.04).
Zaninotto *et al.* 2013 (multicenter: Brussels, Leuven, Milan, Paris, Sint-Truiden, Tel-Hashomer, Vienna)	DSM-IV (I/O)	90 (25/65)	NPD (n=609)	48.8 (14.6)	education, ethnicity, marital status, family history of major depressive disorder, bipolar disorder and suicide	Proportion of single patients was 12% in PD and 15% in NPD, PD group had 75% and NPD group 63% with secondary or higher education. Proportion of white was 97% in both PD and NPD. Regarding family history, proportion of depression was 62% in PD and 49% in NPD, of bipolar disorder 5% in PD and 7% in NPD, and of suicide 18% in PD and 15% in NPD. All differences were nonsignificant. NPD sample was originally in two groups (melancholic and non-melancholic).

Abbreviations: ICD: International Classification of Diseases; ICD-10-AM = ICD-10 Australian Modification; DSM: Diagnostic and Statistical Manual of Mental disorders; Odds Ratio = OR; RDC = Research Diagnostic Criteria; SD = standard deviation, SES = socioeconomic status.
Diagnoses: PD = psychotic depression, NPD = nonpsychotic depression, PBD = psychotic bipolar disorder, SZ = schizophrenia, SZAFF = schizoaffective disorder, HC = healthy controls.
Settings: I = inpatients, O = outpatients, I/O = both in- and outpatients.

Supplement Table 5. Studies comparing the outcome of psychotic depression to other disorders. For the results see also the Supplement Figures.

Reference (country)	Time of data collection City	Diagnosis Setting (I/O) Study design	Sample size of PD (males/females)	Comparison groups	Onset age, mean (SD) Mean age at study moment (SD)	Retrospective/prospective Follow-up time, mean (SD)	Outcome measures	Main results and comments
First-episode samples								
Baldwin, 1988 (UK)	1976–1981 Manchester	Feighner (delusional depression) (I) First-episode	24 (4/20)	non-delusional depression (n=24 matched by age, sex and follow-up length)	Not reported 74.2 [range 65–88] years	Retrospective 61-month follow-up [range 42-104]	Depressive symptoms (HAMD), course of depression symptoms Relapses requiring hospitalisation	Included in meta-analysis: Poor global clinical outcome (course of depression symptoms): Depressive invalidism: 21% vs. 33% Ill with depression throughout: 4% vs. 4% Not included in meta-analysis: Relapse rate: at 12 months: 13% vs. 13%; at 24 months: 33% vs. 17%; at 36 months: 33% vs. 21%; at 48 months: 46% vs. 25% Comment: Only cases older than 65 years included. See also Baldwin 1995 with partly overlapping sample (the results not presented here).
Breslau & Meltzer, 1988 (USA)	Not reported Chicago	RDC (I) First-episode	39 (14/25)	PBD (n=38), SZAFF (n=34)	26.3 (11.2) 37.1 (13.2)	Cross-sectional No follow-up	Depressive symptoms (SADS-C)	Included in meta-analysis: For depressive symptoms see Supplement Figures 2a and 3a.
Craig *et al*. 2000 (USA) (Fennig *et al*. 1996; Craig *et al*. 1997)	1989–1995 The Suffolk County Psychosis Project	DSM-IV) (I/O) First-episode	75 (30/45)	SZ/SZAFF (n=155), PBD (n=119)	Not reported At admission 31 (SD not reported)	Prospective 24-month follow-up	Rehospitalizations, remission, global outcome (GAF), and symptoms (SANS, SAPS,	Included in meta-analysis: For symptoms and global outcome see Supplement Figures 2a,c,d,e and 3a-c. Not included in meta-analysis: Rehospitalization: PD 29% vs. SZ/SZAFF 54% vs. PBD 33% Full remission: PD 60% vs. SZ/SZAFF 14% vs. PBD 78%

							BPRS, HAMD)	Partial or no remission: PD 33%, SZ/SZAFF 31%, PBD 19%
Hill *et al*. 2004 (USA)	Not reported Pittsburgh	DSM-III-R or DSM-IV (I) First-episode	20 (9/11)	SZ (n=86)	25.20 (8.84) Not reported	Prospective 2-year follow-up	Symptoms (HAMD, BPRS, SAPS, SANS) global outcome (GAF)	Included in meta-analysis: See Supplement Figures 2a-e for symptoms and GAF. Comment: All study groups were matched demographically on age, sex, and parental socioeconomic status. NPD comparison group not included here due to small sample size.
Husted *et al*. 1995 (Canada)	1982–1984 Vancouver (part of the Markers and Predictors Study (MAP))	DSM-III (I/O) First-episode	25 (16/9) (at follow-up)	PBD (n=27), SZ (n=66)	26.3 (8.09) Not reported	Prospective 18-month follow-up	Number of negative symptoms (not severity) based on DSM.	Included in meta-analysis: See Online Supplement figure 2d for mean number of negative symptoms.
Jarbin *et al*. 2003 (Sweden)	1982–1993 Lund	DSM-IV (I) First-episode	17 (8/9)	SZ (n=32), PBD (n=25)	Median 16.2 (range 13.4-17.7) Not reported	Prospective 10.5-year follow-up (range 5.1-18.2) after admission	GAF and employment combined to global outcome, symptoms	Not included in meta-analysis: Good outcome: PD 71%, SZ 6%, PBD 40%. Intermediate outcome: PD 18%, SZ 16%, PBD 24%. Poor outcome: PD 6%, SZ 12%, PBD 16%. Very poor outcome: PD 6%, SZ 66%, PBD 16%. Disability pension: PD 29%, SZ 88%, PBD 33%. No employment: PD 7%, SZ 52%, PBD 14%. Serious symptom severity: PD 7%, SZ 48%, PBD 15%. None symptoms: PD 87%, SZ 32%, PBD 6%. Mean GAF score (no SD given): PD 81,

								PBD 66, SZ 34. Comment: Adolescent onset psychoses, first episode sample. Diagnoses based on initial diagnosis at onset, SZAFF group not shown here due to small sample size.
Opjordsmoen 1991 (Norway)	1946–1948 (long-term group) and 1958-1961 (short–term group) Oslo	DSM-III (I) First-episode	50 (20/30)	SZAFF depressed type (n=33)	Not reported 40.8 (SD not reported)	Prospective mean 22.3 (range 3-39)	Global outcome (GAS) Employment Recovery Number of episodes	Not included in meta-analysis: Mean GAS score: PD 3.8 vs. SZAFF 3.4. Employed: PD 69% vs. SZAFF 46% Recovery: PD 66% vs. SZAFF 42% Only index episode: PD 46% vs. SZAFF 27% Mean number of episodes at follow-up: PD 3.1 vs. SZAFF 3.3 First admissions study.
Owoeye *et al*. 2013 (Ireland)	1995–2008 Cavan-Monaghan (CAMFEPS study)	DSM-IV (I/O) First-episode	77 (36/41)	PBD (n=54), SZ (n=73)	51.2 (22.0) Not reported	Prospective 6-month follow-up	Positive and negative symptoms (PANSS)	Included in meta-analysis: See Supplement Figure 2c,d and 3b,c for positive and negative symptoms.
Taiminen *et al*. 2000 (Finland)	1994–1998 Turku	DSM-IV (I/O) First-episode	23 (10/13)	NPD (n=25), SZ (n=17), HC (n=19)	Not reported 33.3 (9.4)	Cross-sectional	Symptoms (BPRS, HAMD)	Included in meta-analysis: See Supplement Figures 1a,b and 2b for symptoms.
Tohen *et al*. 2000 (USA)	1989-1995 Boston (McLean-Harvard study)	DSM-IV (I) First-episode	44 (Not reported)	PBD (n=128), SZAFF (n=20)	Not reported Age at admission for the whole study group 31.9 (5.0)	Prospective 6-month follow-up	Syndromal recovery (i.e. remission) Functional recovery	Included in meta-analysis: Syndromal recovery (i.e. remission): PD 75% vs. PBD 39%, see Online Supplement figure 1d. Not included in meta-analysis: Functional recovery: PD 32%, PBD 37%, SZAFF 0 Syndromal recovery (i.e. remission): PD 75% vs. SZAFF 70% Comment: See also Tohen *et al*. (1992) for partly

								overlapping sample (results not presented here)
Consecutive or mixed samples								
Angst, 1986 (Switzerland)	1959–1963 Zürich	DSM-III (I) Mixed	73 (Not reported)	NPD (n=100), PBD (n=143)	41.4 Not reported	Prospective Follow-up until 1980	Remission, chronic course (i.e. poor global clinical outcome), information from hospital notes	Included in meta-analysis: Remission: 26% in PD (40% mood congruent, 19% mood incongruent) and 33% in NPD 33% and 24% in PBD. Poor global clinical outcome: PD 24% mood congruent, 21% mood incongruent; NPD 7%; Bipolar disorder: mood congruent 5%, mood incongruent 20% Comment: Not exactly clear when outcomes were measured.
Benazzi, 1999 (Italy)	Not reported Not reported	DSM-IV (O) Mixed	40 (18/22)	PBD (n=30)	48.6 (14.9) Not reported	Not reported No follow-up	Course of illness (i.e. poor global clinical outcome: chronic and without full interepisode recovery) Depressive symptoms (MADRS) Global outcome (GAF)	Included in meta-analysis: For depressive symptoms and global outcome see Supplement Figures 3a and 3d. Not included in meta-analysis: Chronic and without full interepisode recovery of major depressive episode: PD 55% vs. BD: 43% (p=0.33)
Brockington *et al*. 1982 (UK)	1966–1968 London	DSM-III (I) Mixed	25 (Not reported)	NPD (n=45)	Not reported Not reported	Prospective 6-year follow-up	General outcome i.e. number of patients failed to recover from index episode	Not included in meta-analysis: General outcome: 36% in PD vs. 9% in NPD.
Buoli *et al*. 2013 (Italy)	Not reported	DSM-IV	18 (4/14)	NPD (n=18)	41.8 (17.1)	Cross sectional	Depression symptoms	Included in meta-analysis: Treatment remission: PD 5 (28%), NPD 11

	Milan	(I) Mixed			59.7 (17.0)	8-week follow-up including hospitalization	(HAMD) Treatment remission	(61%) For symptoms see Supplement Figure 1a.
Copeland, 1983 (UK)	Not reported Borough of Camberwell	Clinical diagnosis (I) Mixed	29 (Not reported) (at follow-up)	NPD (n=18) (at follow-up)	mean 41 years (baseline sample, n=55) mean 49 years (baseline sample, n=55)	Prospective 5-year follow-up	Relapsing at least once Mean number of readmissions Mean number of episodes	Not included in meta-analysis: Relapsing at least once: PD 69% vs. NPD 44% Mean number of readmission: PD 1.5 vs. NPD 0.2 (SD not reported) Mean number of episodes: PD 1.6 vs. NPD 0.6 (SD not reported)
Iowa study								
Coryell *et al*. 1986 (USA)	Iowa study	DSM-III (Not reported) Mixed	46 (Not reported)	NPD (n=159)	Not reported Not reported	Prospective 6-month follow-up	At baseline and 6 months follow up: depressive symptoms (HAMD) and global outcome (GAS) Recovery (return to premorbid level of functioning with absence of symptoms)	Included in meta-analysis: For symptoms and global outcome see Supplement Figures 1a and 1c. Not included in meta-analysis: Recovery, narrow: PD 37% vs. NPD 26% Recovery, broad: PD 52% vs. NPD 57%
Coryell *et al*. 1982 (USA)	1934–1940 (for SZ until 1945) Iowa study	DSM-III (I) Mixed	114 (Not reported)	NPD (n=78), SZ (n=171)	Not reported Median age at admission 42	Prospective Median follow-up 2–3 years	Clinical outcomes Recovery (return to premorbid level of functioning with absence of symptoms).	Included in meta-analysis: Unimproved (i.e. poor global clinical outcome): PD 35% vs. NPD 22% vs. SZ 77% Not included in meta-analysis: Recovered: PD 40% vs. NPD 69% vs. SZ 7% Improved (good global clinical outcome):

								PD 25% vs. NPD 10% vs. SZ 16% Comment: Schizophreniform disorder group (n=48) results not presented here.
Coryell & Tsuang, 1982 (USA)	1934–1940 (for SZ until 1945) Iowa study	DSM-III (I) Mixed	122 (51/71)	NPD (n=103)	Not reported 43.8 (SD not reported)	Not reported Mean follow up time 3.4–5.0 years	Clinical outcomes Recovery (return to premorbid level of functioning with absence of symptoms).	Not included in meta-analysis: Condition at discharge: Recovered: PD 28% (34/121) vs. NPD 46% (47/102) Improved: PD 19% (23/121) vs. NPD 18% (18/102) Unimproved: PD 53% (64/121) vs. NPD 36% (37/102) Follow-up: Recovered: PD 47% (51/108) vs. NPD 73% (69/95) Improved: PD 21% vs. NPD 7% Unimproved: PD 32% vs. NPD 20%
Coryell & Tsuang, 1985 (USA)	1934–1940 (for SZ until 1945) Iowa study	DSM-III delusional depression (I) Mixed	190 (Not reported)	NPD (n=106), SZ (n=219)	Not reported Not reported	Prospective 40-year follow-up or until the time of death	Occupational Mental	Not included in meta-analysis: Occupational outcome: Good: PD 60% vs. NPD 68% vs. SZ 36% Poor (%): PD 28% vs. NPD 18% vs. SZ 57% Mental outcome: Good: PD 55% vs. NPD 56% vs. SZ 22% Poor: PD 29% vs. NPD 18% vs. SZ 53% Comment: 43-78% of the sample were deceased and rated based on medical records. Schizophreniform disorder group (n=87), results not presented here.
NIMH study								
Coryell *et al*. 1984a (USA)	1978–1981 Five centers Boston,	RDC (I/O) Mixed	56 (19/37)	NPD (n=274)	Not reported At intake 40.7 (15.8)	Prospective 6-month follow-up	Global outcome (GAS) Recovery (i.e.	Included in meta-analysis: Symptomatic recovery (i.e. symptomatic remission) at any time: PD 46% vs. NPD 57% For global outcome see Supplement Figure

	Chicago, Iowa, New York, St Louis (NIMH study)						remission) at any time Mild, moderate or severe work impairment	1c. Not included in meta-analysis: Mild, moderate or severe work impairment: PD 71% (17/24) vs. NPD 34% (52/152) Comment: Schizoaffective disorder group (n=24), results not reported here.
Coryell *et al*. 1990 (USA)	1978–1981 Five centers Boston, Chicago, Iowa, New York, St Louis (NIMH study)	RDC (I/O) Mixed	73 (24/49) completers, 19 (8/11) noncompleters	SZAFF depressed type (completers n=30, noncompleters n=12)	Not reported At intake 39.2 (14.9) completers, 38.6 (16.6) noncompleters	Prospective 5-year follow-up	Global outcome (GAS) Recovery Relapses Rehospitalizations Symptoms	Not included in meta-analysis: Symptomatic recovery from index episode: PD 89% vs. SZAFF 73% Sustained psychotic features during final 6 months: PD 12% vs. SZAFF 30% Mean GAS at intake: PD completers 35.8, PD non completers 34.7, SZAFF completers 32.5, SZAFF non-completers 41.3. Mean GAS at last contact PD completers 58.0, PD non completers 46.5, SZAFF completers 55.9, SZAFF non-completers 32.6. Rehospitalized in the follow-up: PD 56%, SZAFF 77% Comment: Partly overlapping sample with Coryell et al. (1984a, 1987).
Coryell *et al*. 1987 (USA)	1978–1981 Five centers Boston, Chicago, Iowa, New York, St Louis (NIMH)	RDC (I/O) Mixed	55 (Not reported)	NPD (n=451)	32.5 (17.2) 40.8 (16.3)	Prospective 2-year follow-up	Global outcome (GAS) Recovery	Not included in meta-analysis: Mean GAS score: at 6 months PD 61 vs. NPD 55; at 24 months PD 63 vs. NPD 61. Symptomatic recovery from index episode by 6 months: PD 53% vs, NPD 58% Symptomatic recovery from index episode by 2 years: PD 75% vs. NPD 83% Comment: Partly overlapping sample with Coryell et al. (1984a, 1990).
Dell'Osso *et al*. 2002	1995–1997	DSM-III-R (I)	30 (10/20)	PBD (n=147)	29.6 (12.5) 36.2 (12.7)	Cross-sectional	Symptoms (BPRS),	Included in meta-analysis: For symptoms and global outcome see

(Italy)	Pisa	Mixed				No follow-up	global outcome (GAF), unemployment	Supplement Figures 3a-3d. Not included in meta-analysis: Unemployment rate: PD 63%, bipolar mania 60%, bipolar mixed episode 69%, depressed bipolar 53% Comment: PBD groups (pure mania (n=55), with mixed mania (n=62), with bipolar depression (n=30)) were studied separately.
Gaudiano *et al*. 2009 (USA)	Not reported Rhode Island (MIDAS project, data collected from the 1980's, still ongoing)	DSM-IV (O) Mixed	60 (16/44)	NPD (n=1052)	20.7 (12.1) 37.0 (11.7)	Cross-sectional No follow-up	Global outcome (GAF)	Included in meta-analysis: See Supplement Figure 1c for GAF. Comment: See also partly overlapping sample Gaudiano and Zimmerman, 2010 (USA) (results not presented here)
Gaudiano *et al*. 2016 (USA)	2013 Rhode Island	DSM-IV-TR (I/O) Mixed	174 (74/100)	NPD (n=1140)	Not reported 44.9 (12.3)	Retrospective No follow-up	Global outcome (GAF)	Included in meta-analysis: See Supplement Figure 1c for GAS.
Goethe & Szarek, 1988 (USA)	1981–1985 Admissions to private hospital	DSM-III (I) Mixed	77 (22/55)	NPD (n=360)	Not reported 52 (19.24)	Cross sectional No follow-up	Condition on discharge (symptomatic remission and poor clinical outcome)	Recovery (i.e. symptomatic remission): PD 4% vs. NPD 4% Unimproved (i.e. poor global clinical outcome): PD 8% vs. NPD 8% Comment: See also partly overlapping sample (Goethe *et al.*1988) (results not presented here)
Chicago follow-up								
Goldberg & Harrow, 2004	Not reported Chicago follow-up	RDC (I) Mixed	17 (Not reported)	NPD (n=72)	For the whole sample (n=123) mean age at index	Prospective Follow-ups at 2,	Global outcome (data from several	Not included in meta-analysis: Good overall outcome: At 2 years: PD 27% vs. NPD 40%

	study				hospitalization 23.2 (3.8)	4.5, 7.5 and 10 years	sources).	At 4.5 years: PD 41% vs. NPD 49% At 7.5 years: PD 35% vs. NPD 63% At 10 years: PD 47% vs. NPD 63%
Kettering *et al*. 1987 (USA)	Not reported Illinois, Chicago (subjects from Chicago follow-up study and from the Mental Health Clinical Research Center program)	RDC (I) Mixed	31 (Not reported)	NPD (n=28), SZ (n=51)	For the whole study group at follow-up 27.5 [range 19–56]	Prospective Average of 14-months follow-up [range 12–26]	Global outcome (GAS) Depressive symptoms (Katz adjustment scale, depressed mood)	Included in meta-analysis: See depressive symptoms and GAS in Online Supplement Figures 1a and 1c.
Sands & Harrow, 1994 (USA)	Not reported exactly Illinois, Chicago (subjects from Chicago follow-up study and from other clinical research centers)	RDC (I) Mixed	31 (Not reported)	NPD (n=63)	Age at index hospitalization for the whole the Chicago Follow-Up group 23.7 (3.6); for the patients from the mental health clinical research center 34.4 (13.2).	Prospective Mean 2.4 years, range 1-5 years	Clinical depression symptoms during last follow-up year (based on SADS)	Included in meta-analysis: Remission: PD 19% vs. NPD 29% Comment: Recruited during index hospitalization from two hospitals.
Gournellis *et al*. 2001 (Greece)	1997–2000 Athens	DSM-IV (I) Consecutive	45 (14/31)	NPD (n=73)	58.3 (15.2) 69.6 (5.8)	Prospective No follow-up	Depression symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for HAMD. Comment: Elderly patients.
Johnson *et al*. 1991 (USA)	1980–1985	DSM-III (I/O)	92 (Not reported)	NPD (n=532)	29.0 (13.9) 42.0 (16.2)	Cross-sectional	Currently depressed	Not included in meta-analysis: Currently depressed; PD 19% vs. NPD

	New Haven, St Louis, Baltimore, Durham, and Los Angeles (ECA study)	Mixed				1-year follow-up	(last month)	11%.
Jäger *et al*. 2005 (Germany)	1980–1982 Munich	ICD-9 and DSM-IV (I) Mixed	20 (4/16)	NPD (n=33), SZ (n=64)	39.2 (14.7) Not reported	Retrospective 15-year follow-up	Global outcome (GAS) Symptoms (PANSS, SANS, HAMD) Rehospitalizations Employment	Included in meta-analysis: For symptoms, global outcome and hospitalizations see Supplement Figures 1a-c and f, and 2a-e. Not included in meta-analysis: Regular employment: PD 79%, NPD 78%, SZ 47% Mean number of rehospitalizations: PD 2.8 vs. SZ 2.9 (p=0.045).
Karaaslan *et al*. 2003 (Turkey)	Not reported	DSM IV (I) Mixed	16 (8/8)	NPD (n=20) HC (n=20)	Not reported 36.45 (8.65)	Cross-sectional No follow-up	Depressive symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for HAMD. Comment: Only patients who had not taken any psychotropic medications during last 4 week were included.
Kuhs, 1991 (Germany)	1985–1987 Münster	ICD-9 and DSM-III (I) Mixed	23 (12/11)	NPD (n=137)	Not reported 54.7 (13.3)	Cross-sectional No follow-up	Depressive symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for HAMD.
Lee *et al*. 2003 (Taiwan)	Not reported Taipei	DSM-IV (I) Mixed	48 (33/5)	NPD (n=108)	67.9 (13.1) 74.7 (6.1)	Cross-sectional No follow-up	Depressive symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for HAMD. Comment: Elderly sample, hospital for veterans.
Lykouras *et al*. 1986 (Greece)	1982–1984 Athens	DSM-III (I) Mixed	22 (Not reported)	NPD (n=36)	Not reported Not reported	Cross-sectional No follow-up	Depressive symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for HAMD.

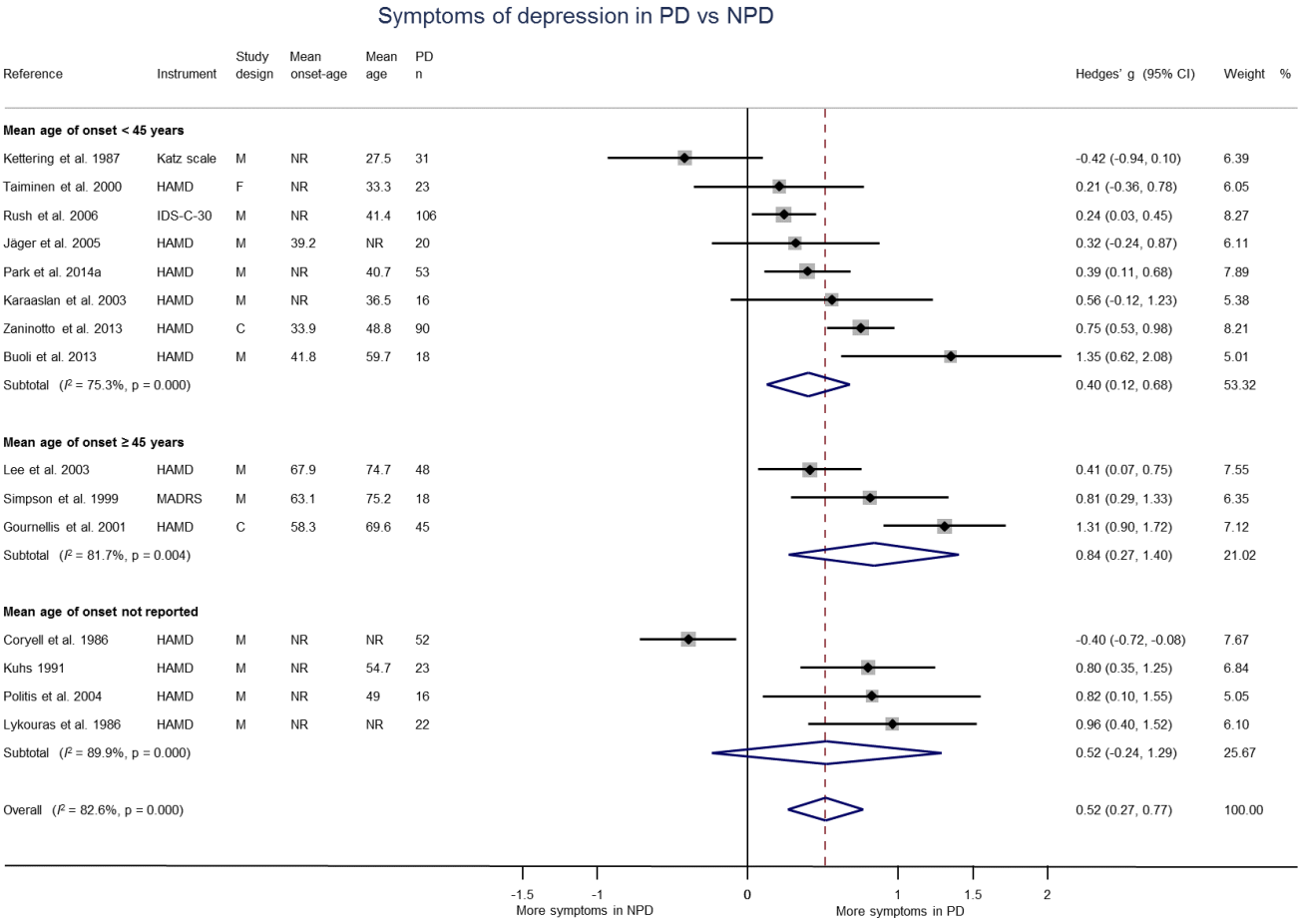
Maj *et al*. 1990 (Italy)	1980–1982 Naples	DSM-III (I) Mixed	36 (15/21)	NPD (n=36)	32.0 (3.0) 42.0 (6.3)	Prospective 7-year follow-up	Global outcome (GAS), psychosis symptoms (CPRS)	Included in meta-analysis: See Supplement Figure 1b and c for psychosis symptoms and global outcome.
Park *et al*. 2014a (South Korea)	1/2006–8/2008 18 hospitals (CRESCEND study)	DSM-IV (I/O) Mixed	53 (17/36)	NPD (n=441)	Not reported 40.7 (15.2)	Cross-sectional No follow-up	Depressive and total psychotic symptoms (HAMD-17, BPRS) Global outcome (SOFAS) Employment	Included in meta-analysis: For symptoms and global outcome see Supplement Figures 1a-c. Not included in meta-analysis: Employed: PD 73% vs. NPD 73% Comment: Patients in beginning of treatment for first episode or recurrent depression.
Park *et al*. 2014b (South Korea)	2006-2008 18 hospitals (CRESCEND study)	DSM-IV (I/O) Mixed	24 (6/18)	NPD (n= 942)	Not reported 44.7 (18.5)	Cross-sectional or 8 years (not clearly reported)	Symptoms (HAMD, BPRS) Global outcome (SOFAS) Employment	Not included in meta-analysis: Employed: PD 68% vs. NPD 75%. SOFAS mean score: PD 53.9 vs. NPD 57.4. HAMD mean score: PD 21.6 vs. NPD 20.2. BPRS mean score: PD 41.5 vs. NPD 20.9. Comment: Patients in beginning of treatment for first episode or recurrent depression. Results not in meta-analysis, due to overlap with Park *et al.* (2014a).
Politis *et al*. 2004 (Greece)	Not reported Athens	DSM-IV (I) Mixed	16 (8/8)	NPD (n=16)	Not reported 49 (15)	Cross-sectional No follow-up	Depressive symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for HAMD.
Robinson & Spiker, 1985 (USA)	1974–1982 Pittsburgh	RDC delusional depression (I) Mixed	52 (19/33)	NPD (n=52)	39.2 (13.4) 45.6 (14.2)	Prospective 1-year follow-up after discharge	Clinical status over first years Hospitalizations	Included in meta-analysis: Asymptomatic (i.e. remission): PD 42% vs. NPD 62% Chronically ill (i.e. poor global clinical outcome): PD 29% vs. NPD 6% Not included in meta-analysis: Total number of hospitalizations: PD 18 vs.

								NPD 10 (SD not reported) Comment: Controls matched for sex and age at first episode and index episode.
Rush *et al*. 2006 (USA)	1998–2000 Austin, Texas	DSM-IV (O) Mixed	106 (19/87)	NPD (n=438)	Not reported 41.4 (10.7)	Cross-sectional	Depressive symptoms (IDS-C-30) Employment	Included in meta-analysis: For depressive symptoms see Supplement Figure 1a. Not included in meta-analysis: Unemployed: PD 90% vs. NPD 81% Comment: Most of the sample were females.
Simpson *et al*. 1999 (UK)	Not reported Manchester	DSM-III-R (I/O) Mixed	18 (7/11)	NPD (n=81)	63.1 (11.0) 75.2 (4.6)	Cross-sectional No follow-up	Depressive symptoms (MADRS)	Included in meta-analysis: See Supplement Figure 1a for depression symptoms. Comment: Old-age study
Tsuang & Coryell, 1993 (USA)	1979-1982 Iowa	DSM-III-R (I) Mixed	32 (6/26) [two groups: with mood-congruent (n=17) and incongruent (n=15) features]	SZ (n=22)	Not reported Not reported	Prospective 8-year follow-up	Global outcome (GAS)	Included in meta-analysis: Please see Onluse supplement Figure 2e for GAS score. Not included in meta-analysis: Recovery from psychotic symptoms: PD 44% vs. SZ 0%
Zaninotto *et al*. 2013 (Several countries)	2000–2004 European multicenter project Brussels, Leuven, Milan, Paris, Sint-	DSM-IV (I/O) Consecutive	90 (25/65)	NPD (n=609) [two groups]	33.9 (14.7) 48.8. (14.6)	Cross-sectional	Depressive symptoms (HAMD)	Included in meta-analysis: See Supplement Figure 1a for depression symptoms.

	Truiden, Tel-Hashomer, Vienna

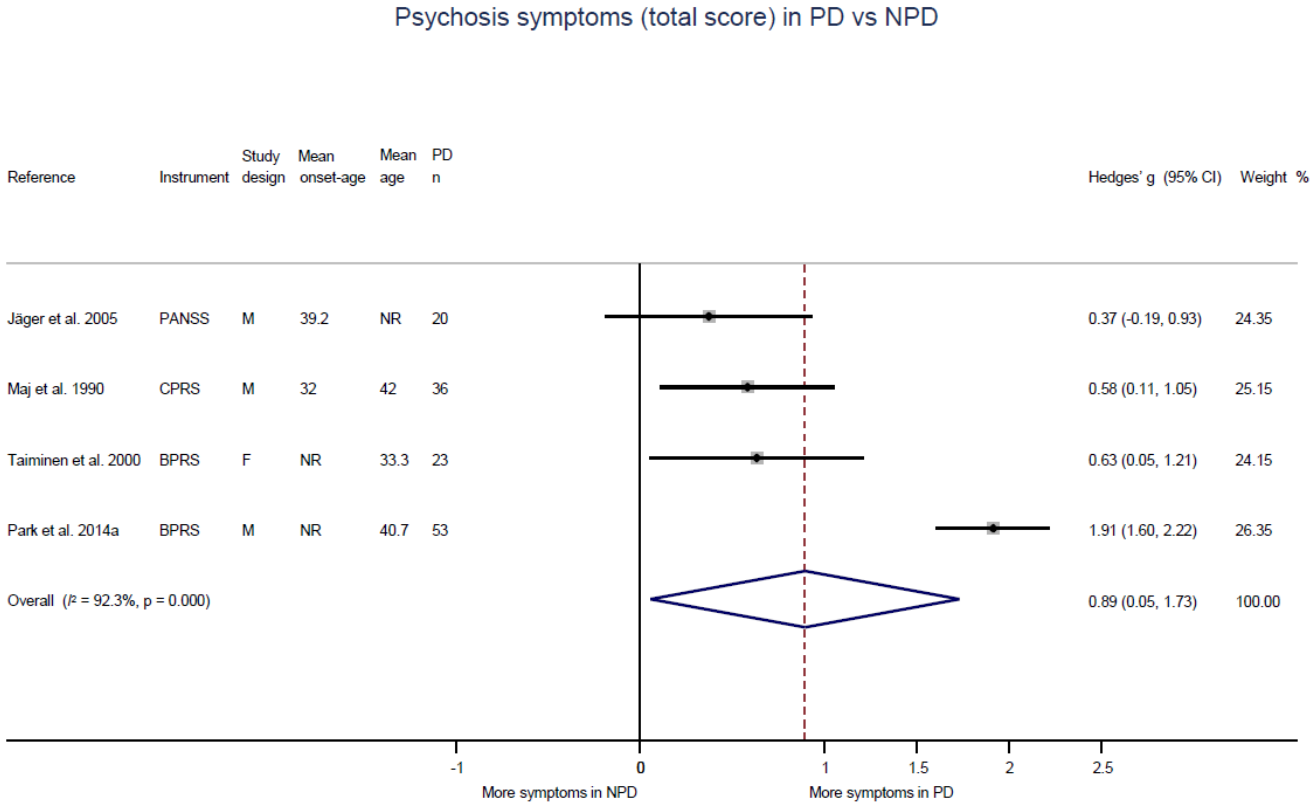
Abbreviations: BDI = Beck Depression Inventory; BPRS = Brief Psychiatric Rating Scale; CGI = Clinical Global Impression Scale; CI = Confidence Interval; CPRS = Comprehensive Psychopathology Rating Scale; DSM = Diagnostic and Statistical Manual of Mental disorders; ECA = Epidemiological Catchment Area, GAF = The Global Assessment of Functioning Scale; GAS = Global Assessment Scale; HAMD = Hamilton Rating Scale for Depression; ICD = International Classification of Diseases; IDS-C-30 = Inventory of Depressive Symptomatology (30 item); MADRS = Montgomery Asberg Depression Rating Scale; NIMH = National Institute of Mental Health; PANSS = Positive and Negative Syndrome Scale; PDAS = The Psychotic Depression Assessment Scale; PSE = Present State Examination; SADS = Schedule for Affective Disorders and Schizophrenia (-C = Change version); SANS = Scale for Assessment of Negative Symptoms; SAPS = Scale for Assessment of Positive Symptoms; SOFAS = Social and Occupational Functioning Assessment Scale.
Diagnoses: PD = psychotic depression, NPD = nonpsychotic depression, PD = bipolar disorder, PBD = psychotic bipolar disorder, SZ = schizophrenia, SZAFF = schizoaffective disorder, HC = healthy controls.
Settings: I = inpatients, O = outpatients, I/O = both in- and outpatients.

Supplement Figure 1a. Meta-analysis on symptoms of depression in psychotic depression (PD) and non-psychotic depression (NPD) in strata by age of illness onset.



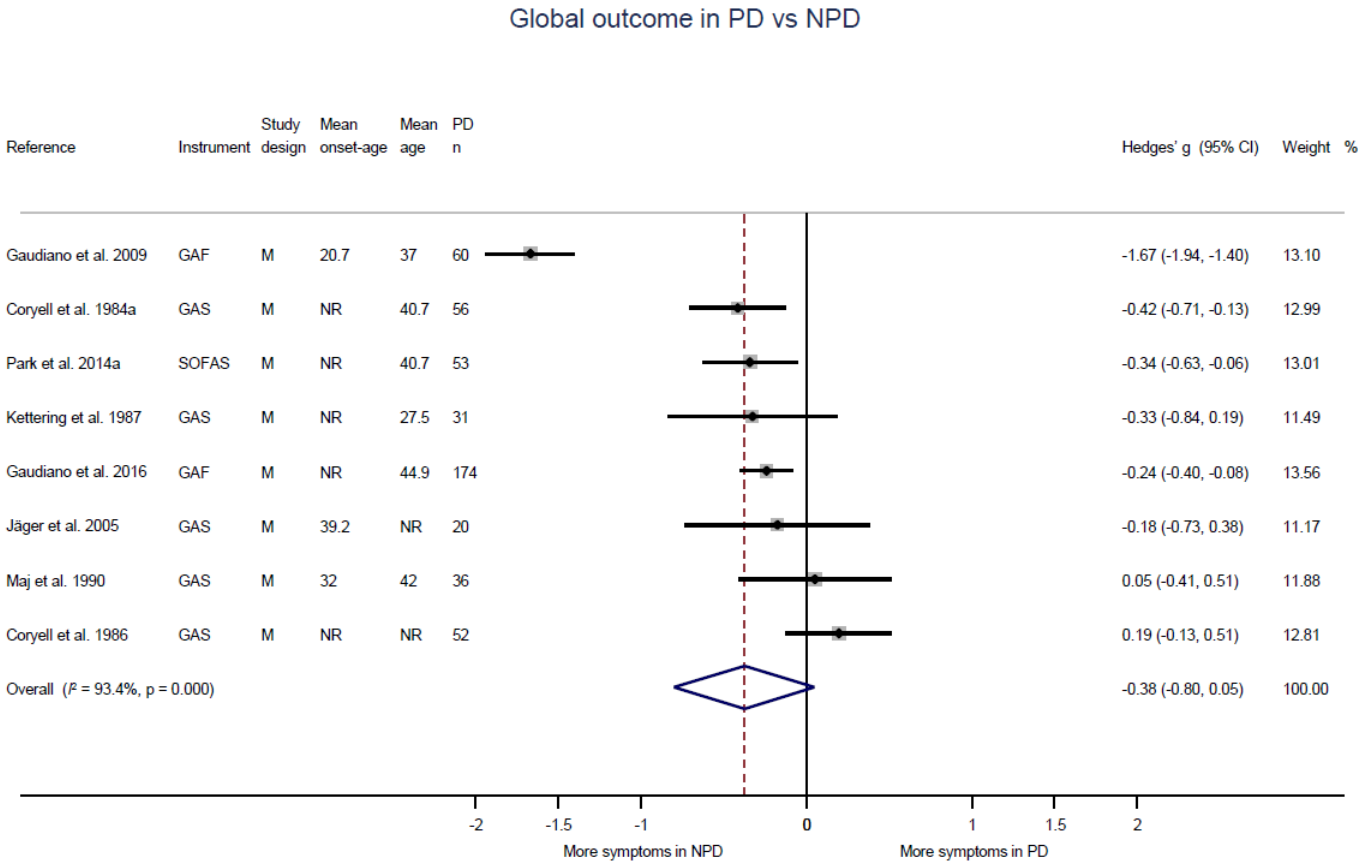
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, HAMD = Hamilton Rating Scale for Depression, IDS-C-30 = Inventory of Depressive Symptomatology (30 item), MADRS = Montgomery-Asberg Depression Rating Scale, M = mixed sample, F = first-episode sample, C = consecutive sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 1b. Meta-analysis on total psychotic symptoms in psychotic depression (PD) and non-psychotic depression (NPD).



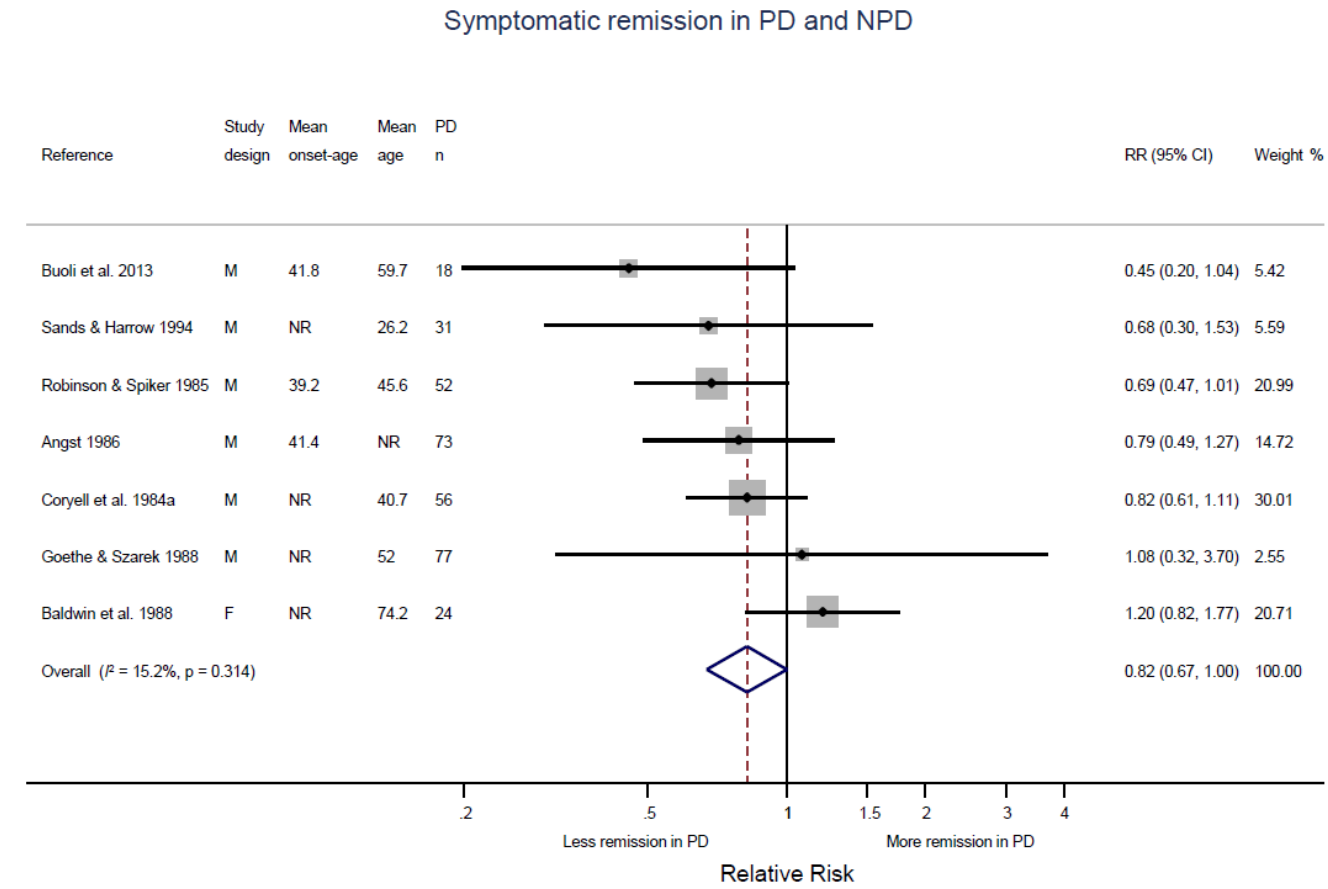
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, PANSS = Positive and Negative Syndrome Scale,CPRS = Comprehensive Psychopathology Rating Scale, BPRS = Brief Psychiatric Rating Scale, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 1c. Meta-analysis on global outcome in psychotic depression (PD) and non-psychotic depression (NPD).



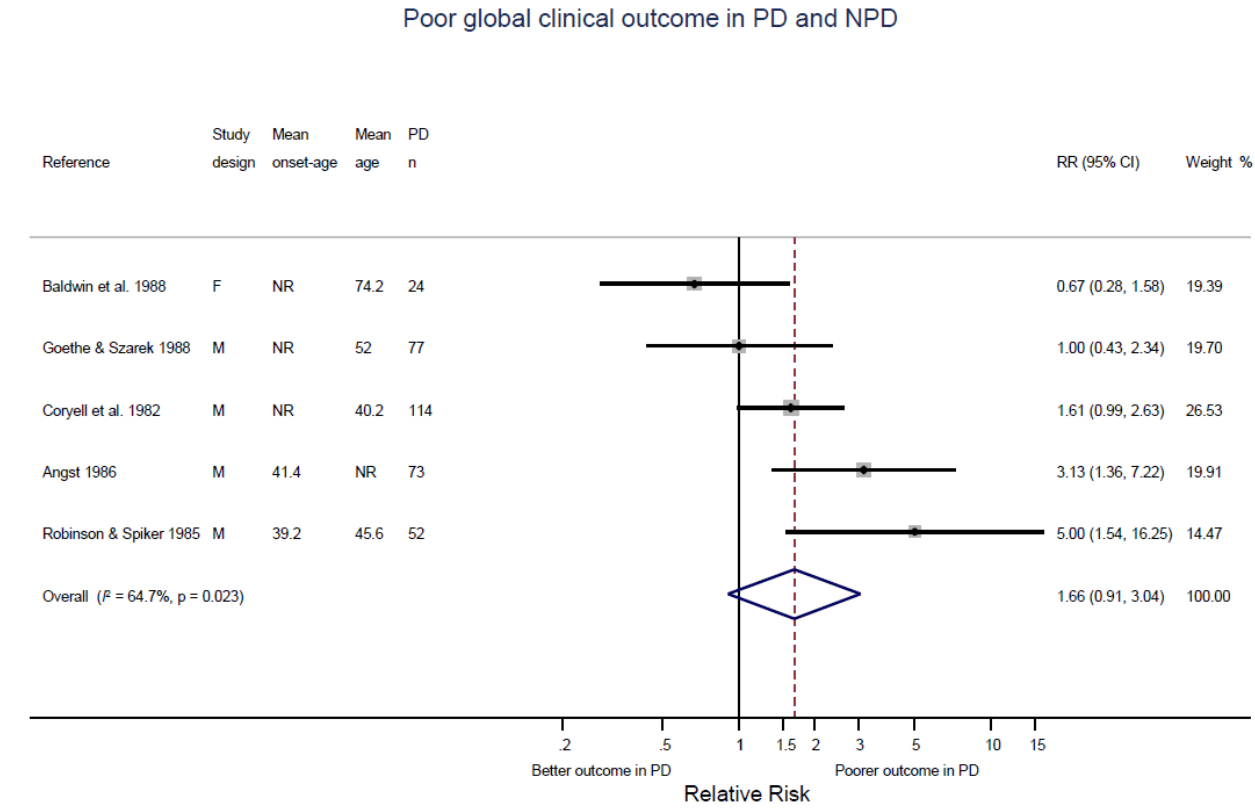
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, GAF = The Global Assessment of Functioning Scale, GAS = Global Assessment Scale, SOFAS = Social and Occupational Functioning Assessment Scale, M = mixed sample, NR = not reported, I^2 = heterogeneity, CI = confidence interval.

Supplement Figure 1d. Meta-analysis on remission in psychotic depression (PD) and non-psychotic depression (NPD).



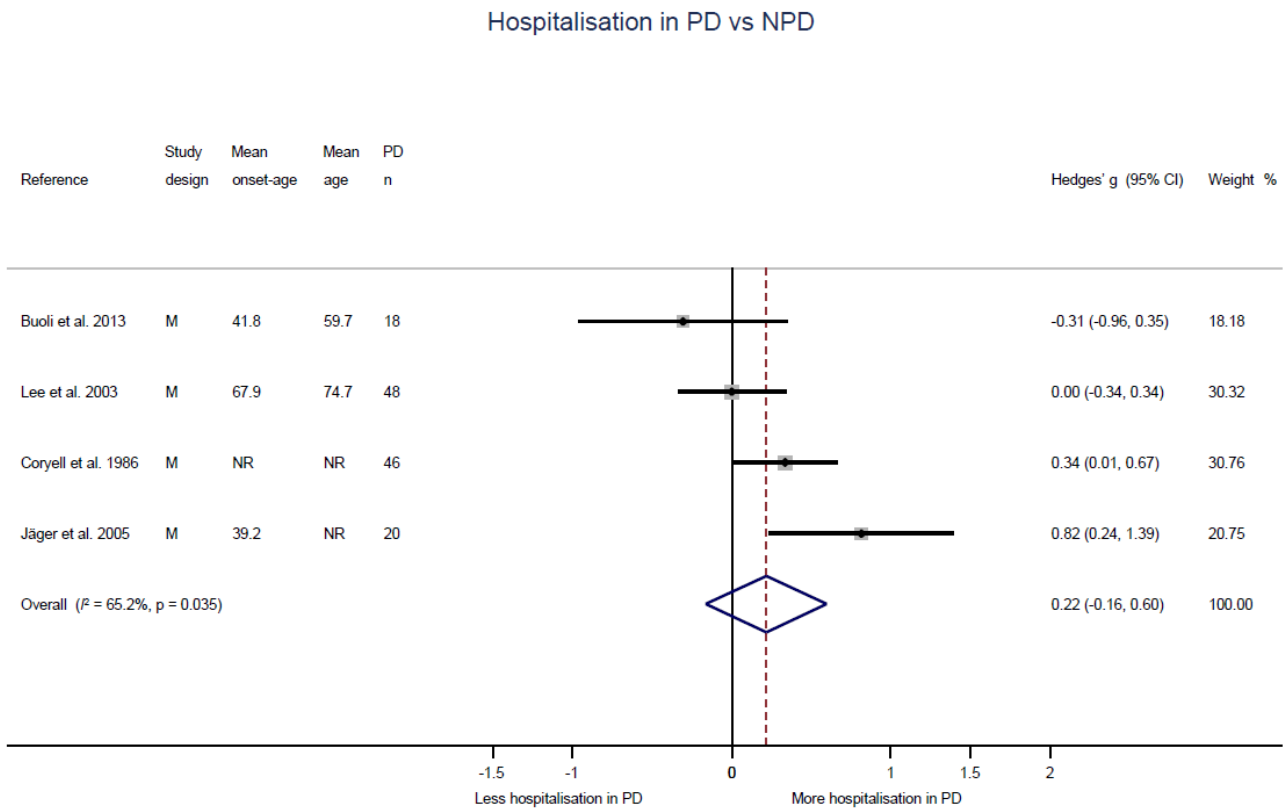
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, RR = Relative Risk, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 1e. Meta-analysis on poor global clinical outcome in psychotic depression (PD) and non-psychotic depression (NPD).



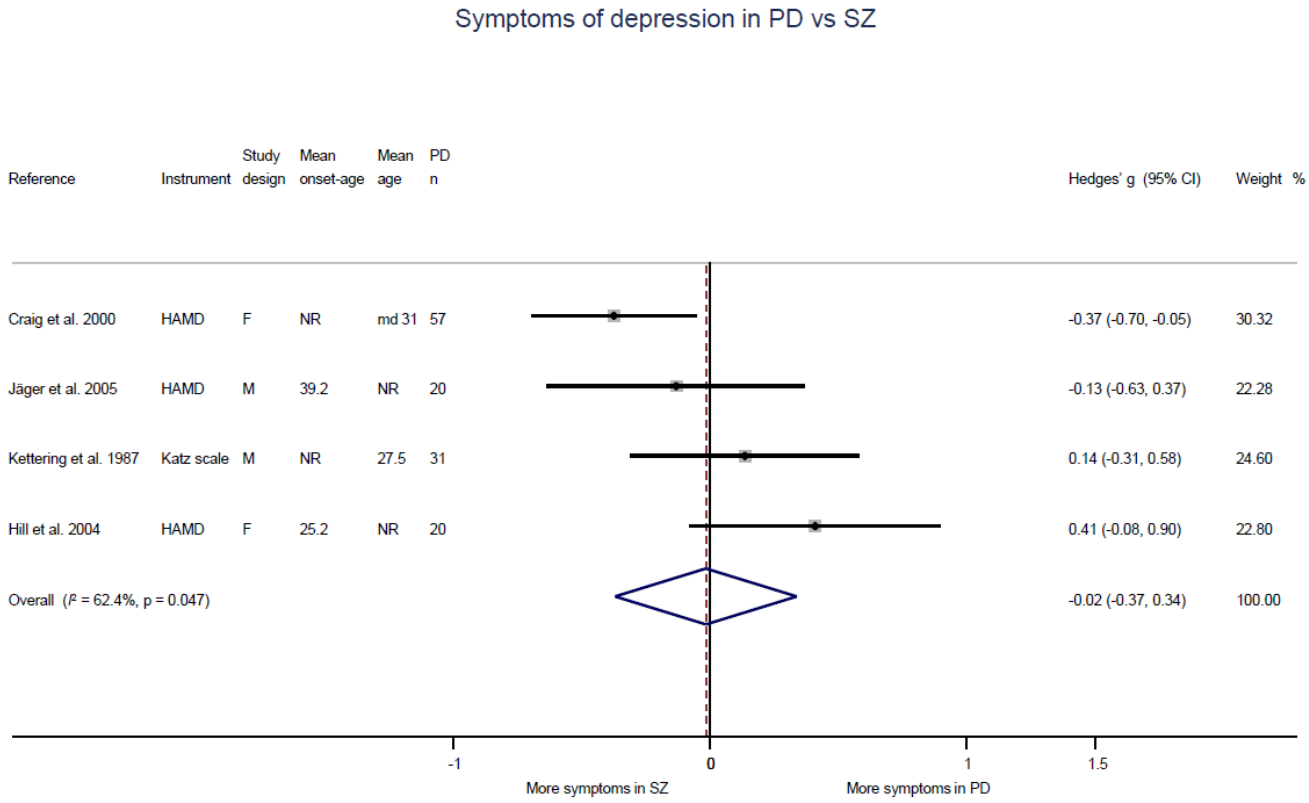
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, RR = Relative Risk, F = first-episode sample, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 1f. Meta-analysis on hospitalization in psychotic depression (PD) and non-psychotic depression (NPD).



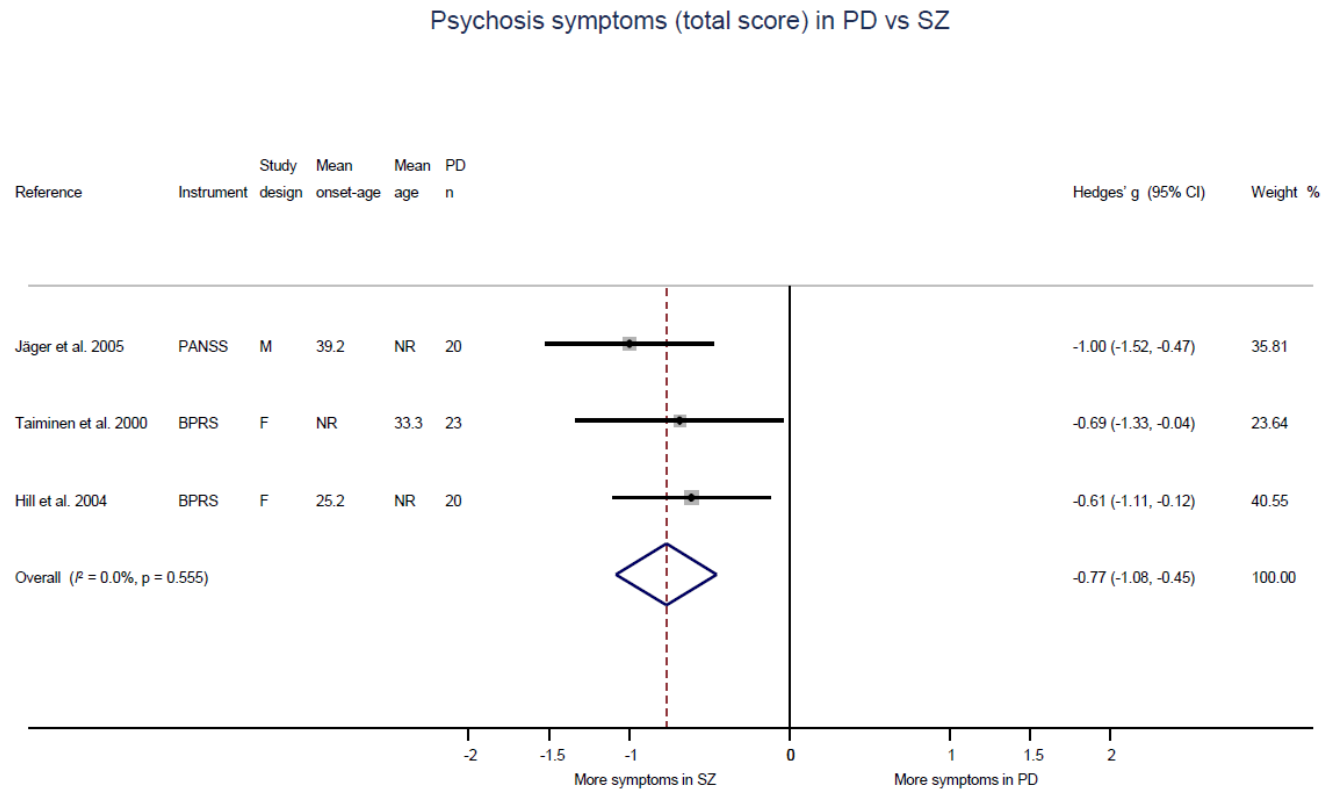
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 2a. Meta-analysis on symptoms of depression in psychotic depression (PD) and schizophrenia (SZ).



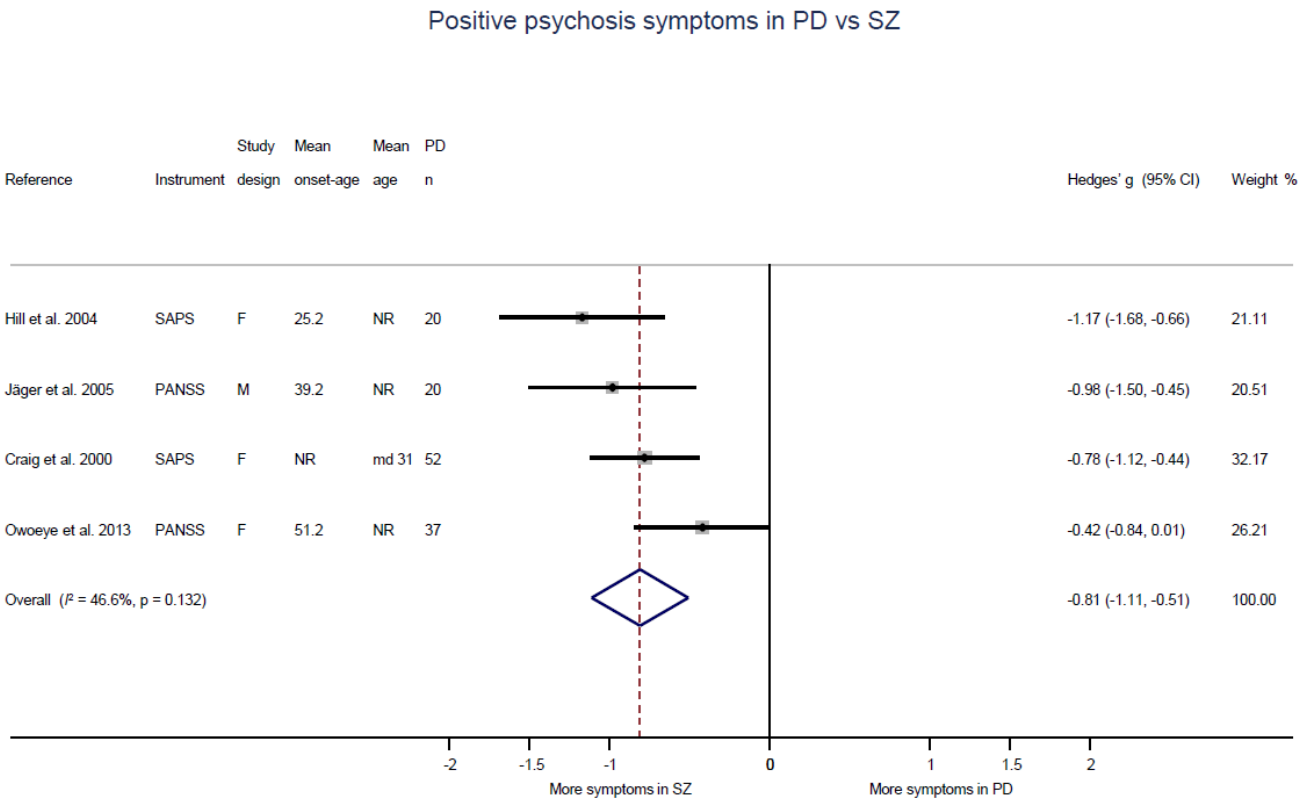
Abbreviations: PD = psychotic depression, SZ = schizophrenia, HAMD = Hamilton Rating Scale for Depression, F = first-episode sample, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 2b. Meta-analysis on total psychotic symptoms in psychotic depression (PD) and schizophrenia (SZ).



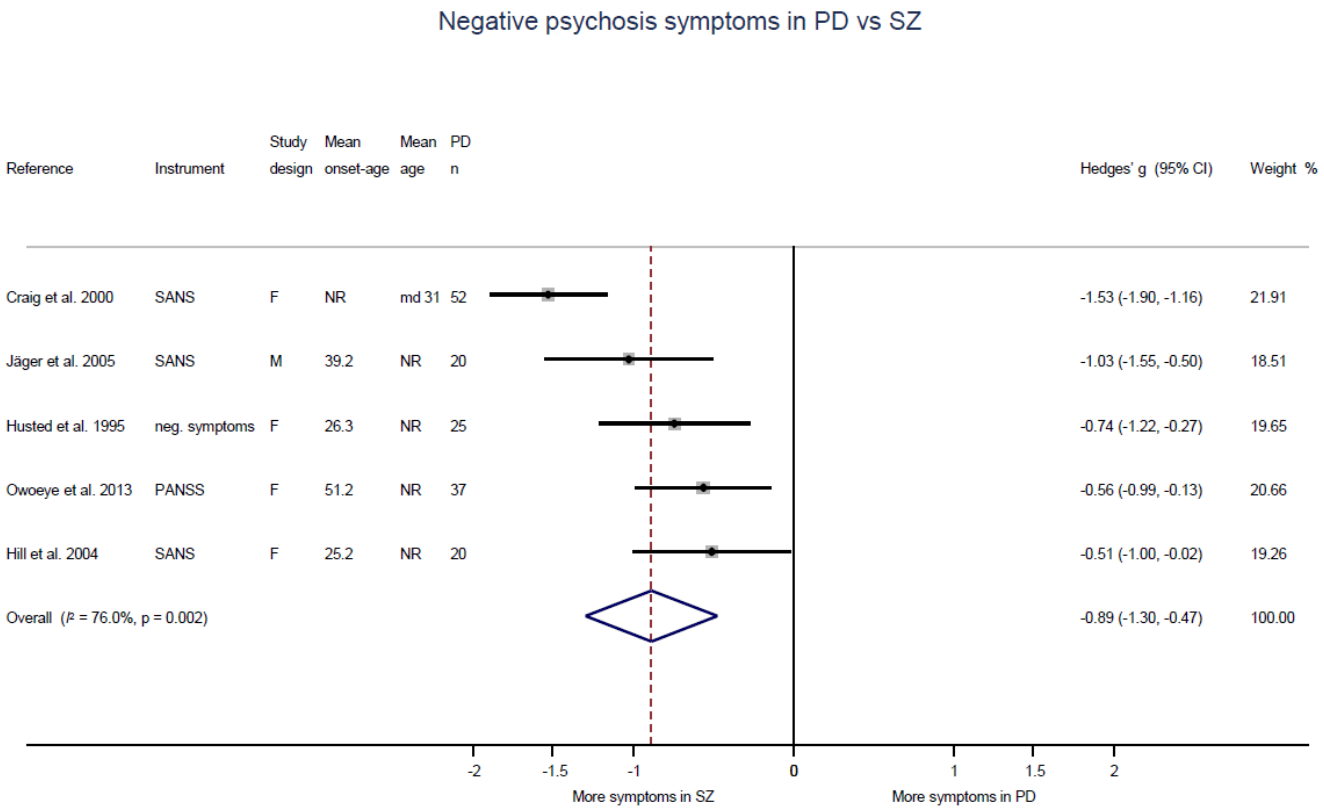
Abbreviations: PD = psychotic depression, SZ = schizophrenia, PANSS = Positive and Negative Syndrome Scale, BPRS = Brief Psychiatric Rating Scale, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval

Supplement Figure 2c. Meta-analysis on positive symptoms in psychotic depression (PD) and schizophrenia (SZ).



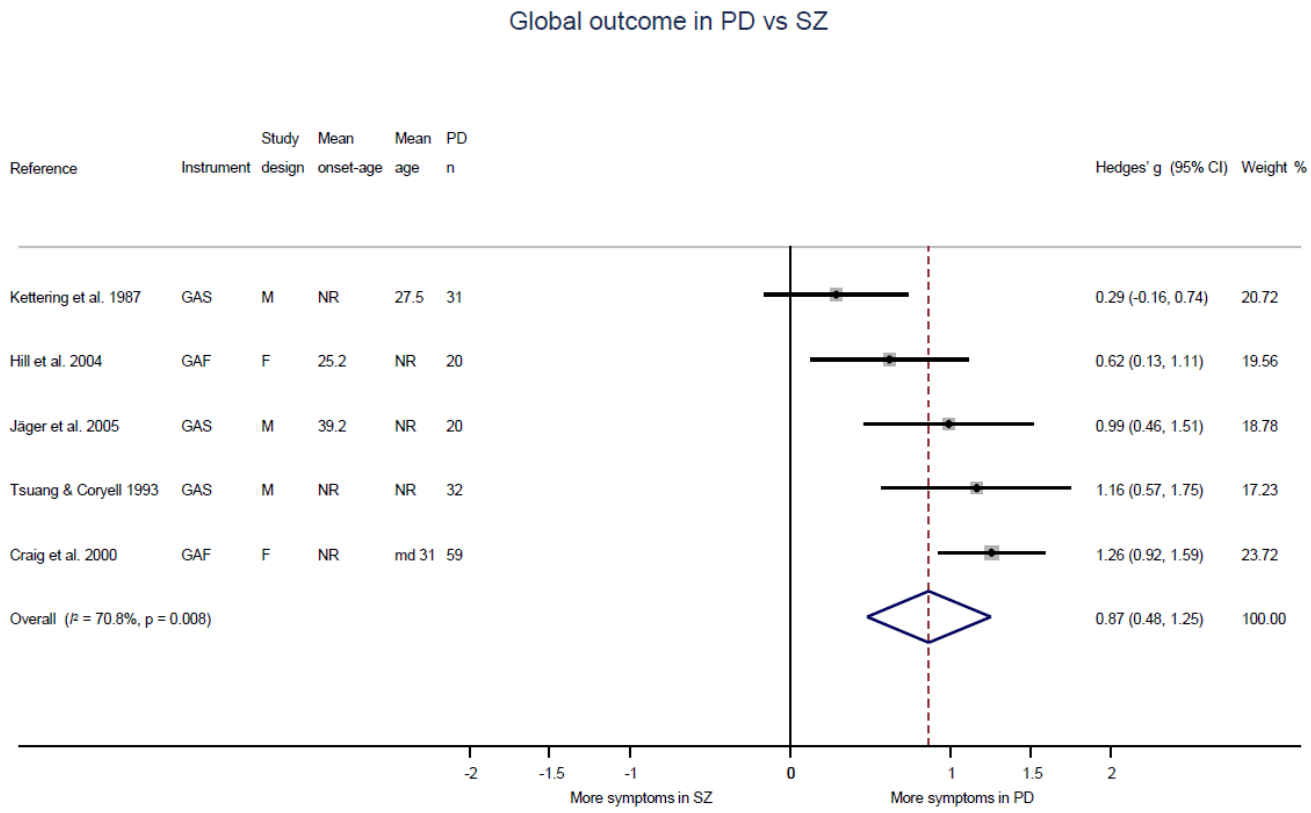
Abbreviations: PD = psychotic depression, SZ = schizophrenia, SAPS = Scale for Assessment of Positive Symptoms, PANSS = Positive and Negative Syndrome Scale, F = first-episode sample, M = mixed sample, NR = not reported, I^2 = heterogeneity, CI = confidence interval.

Supplement Figure 2d. Meta-analysis on negative symptoms in psychotic depression (PD) and schizophrenia (SZ).



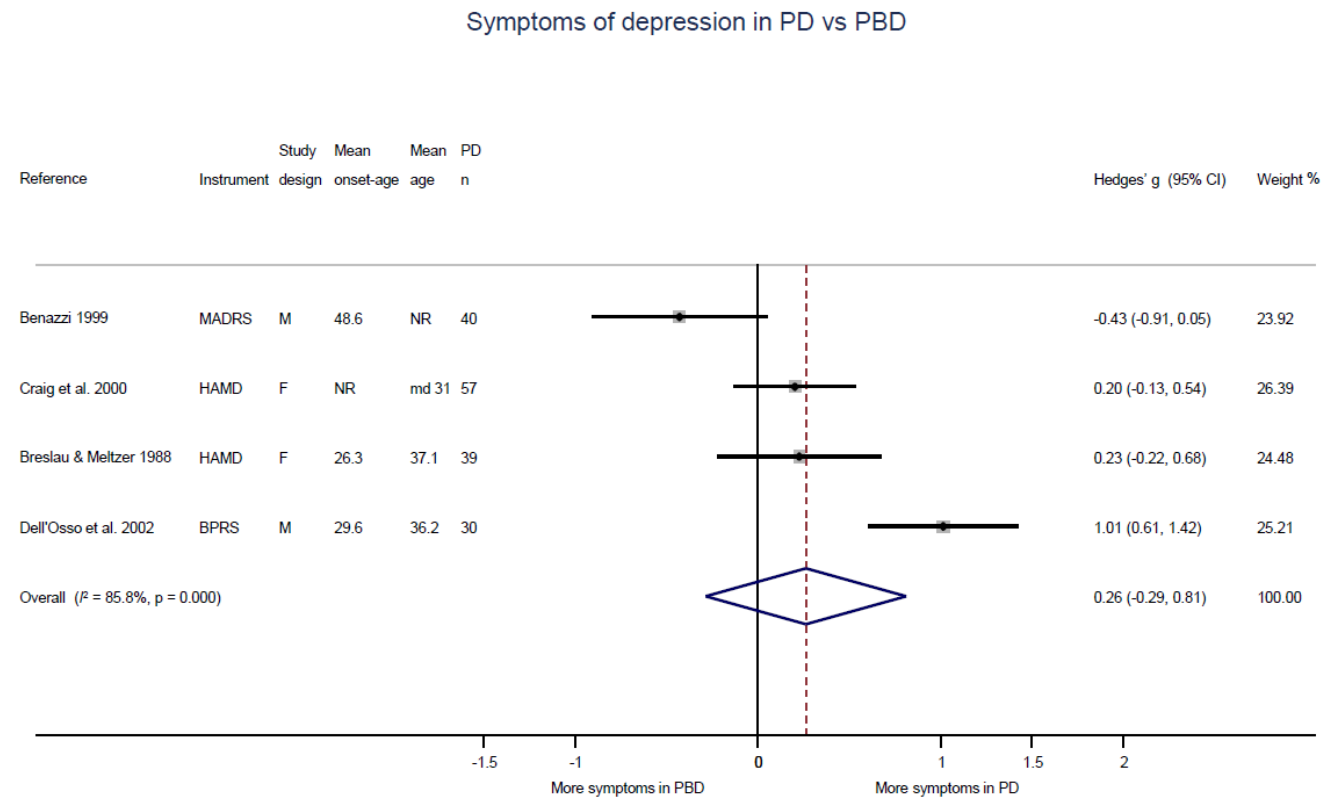
Abbreviations: PD = psychotic depression, SZ = schizophrenia, SANS = Scale for Assessment of Negative Symptoms, PANSS = Positive and Negative Syndrome Scale, F = first-episode sample, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 2e. Meta-analysis on global outcome in psychotic depression (PD) and schizophrenia (SZ).



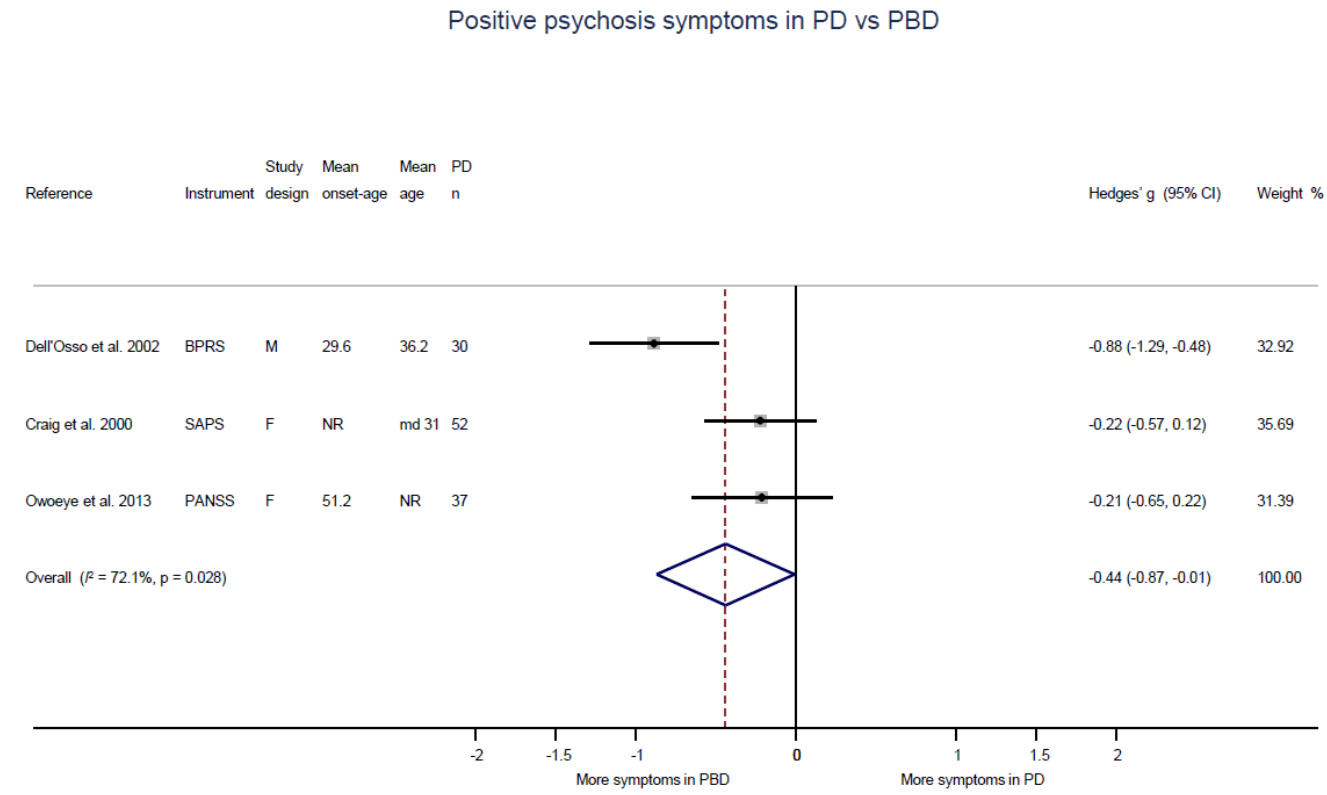
Abbreviations: PD = psychotic depression, SZ = schizophrenia, GAF = The Global Assessment of Functioning Scale, GAS = Global Assessment Scale, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 3a. Meta-analysis on symptoms of depression in psychotic depression (PD) and psychotic bipolar disorder (PBD).



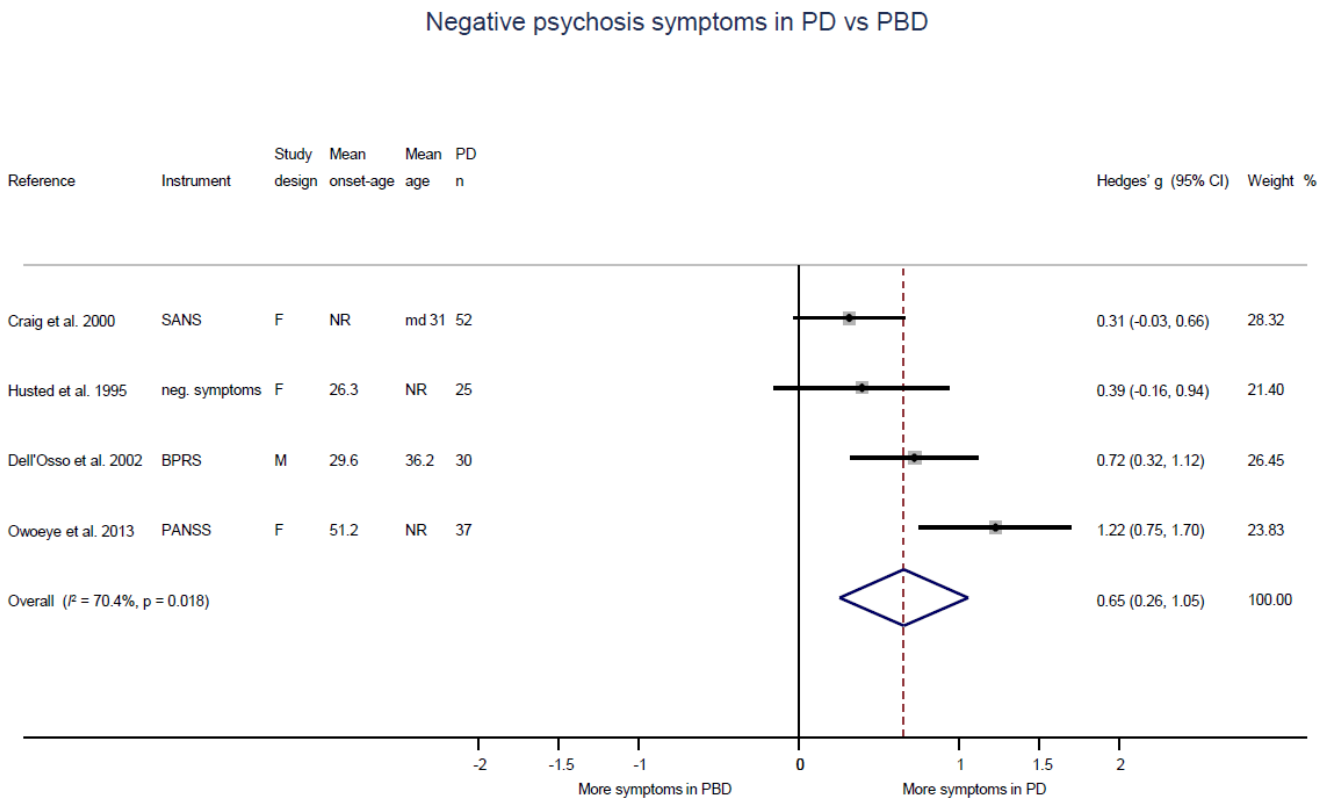
Abbreviations: PD = psychotic depression, PBD = psychotic bipolar disorder, HAMD = Hamilton Rating Scale for Depression, BPRS = Brief Psychiatric Rating Scale, MADRS = Montgomery-Asberg Depression Rating Scale, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 3b. Meta-analysis on positive symptoms in psychotic depression (PD) and psychotic bipolar disorder (PBD).



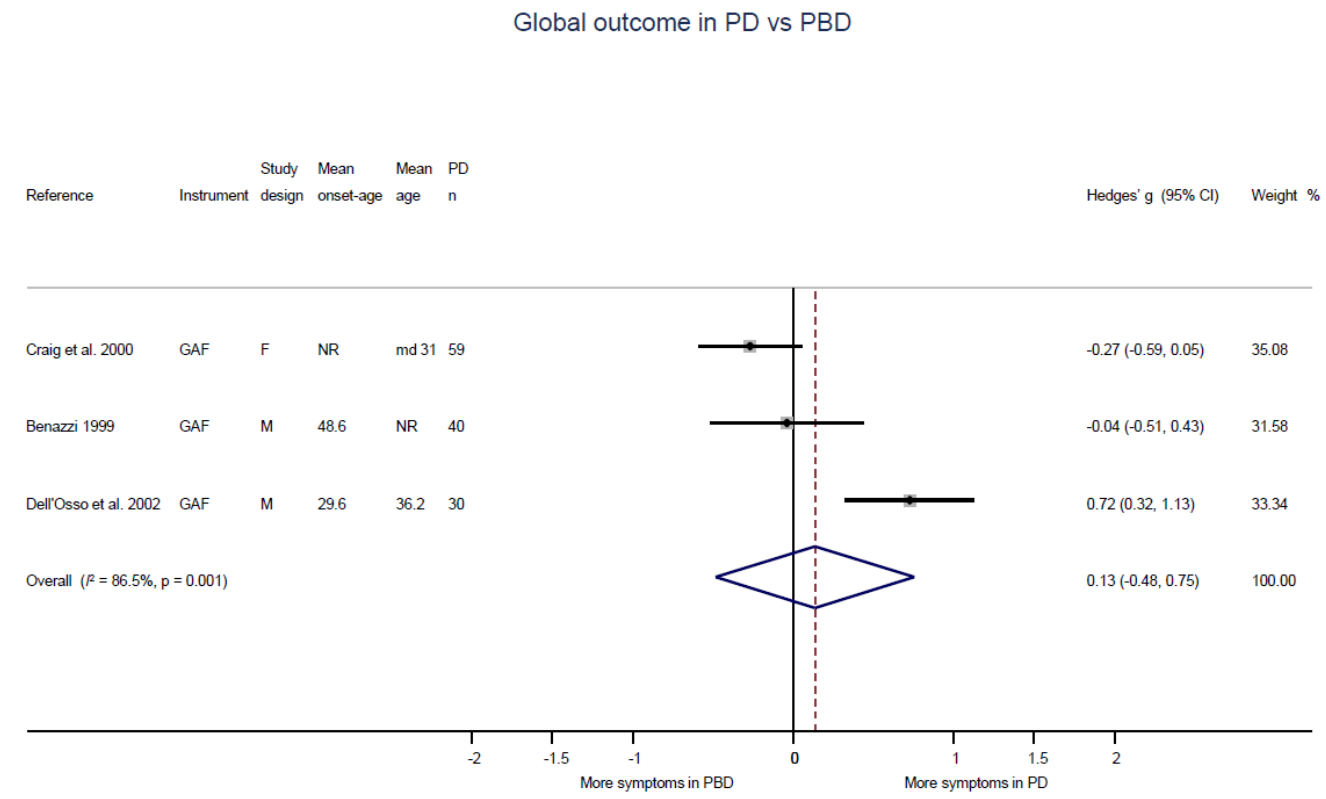
Abbreviations: PD = psychotic depression, PBD= psychotic bipolar disorder, SAPS = Scale for Assessment of Positive Symptoms, BPRS = Brief Psychiatric Rating Scale, PANSS = Positive and Negative Syndrome Scale, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 3c. Meta-analysis on negative symptoms in psychotic depression (PD) and psychotic bipolar disorder (PBD).



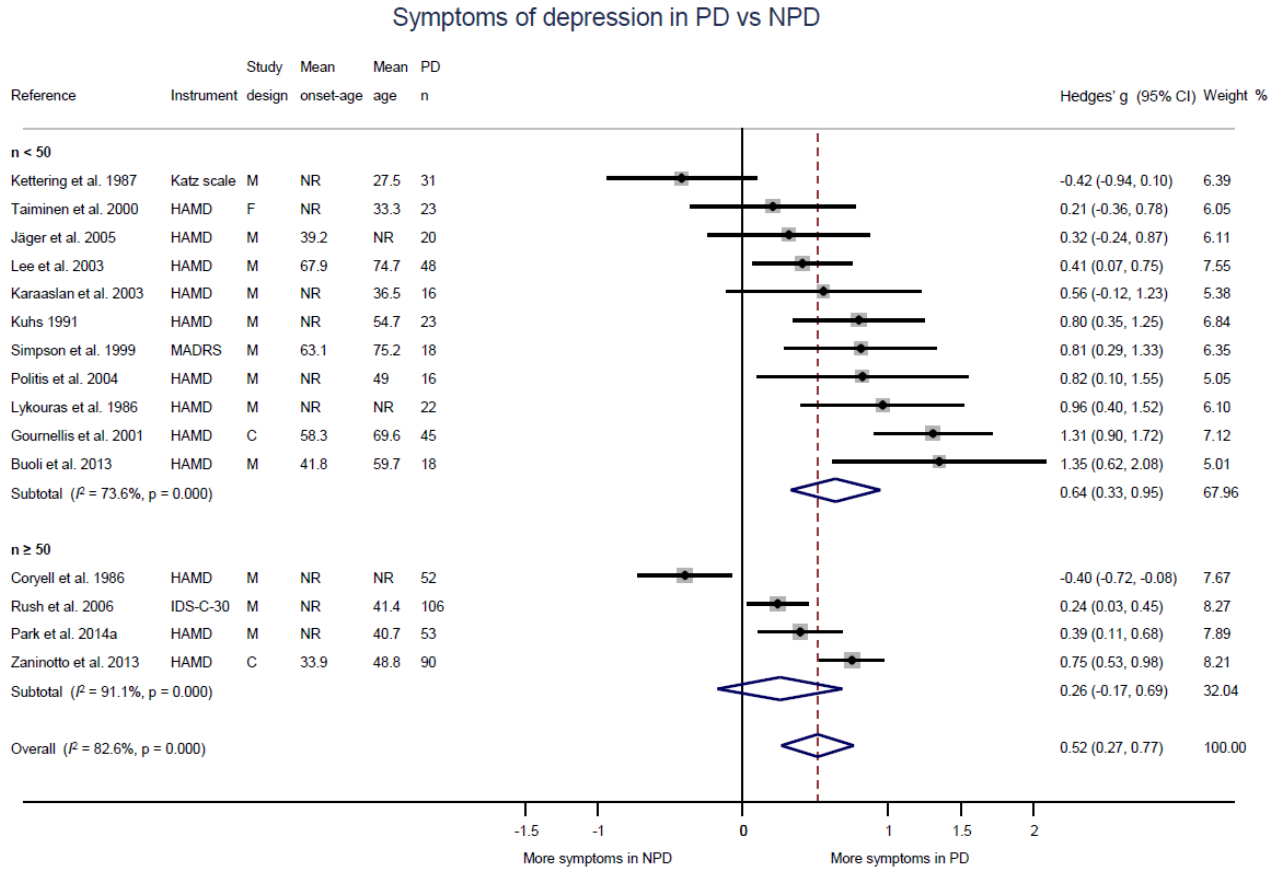
Abbreviations: PD = psychotic depression, PBD = psychotic bipolar disorder, SANS = Scale for Assessment of Negative Symptoms, BPRS = Brief Psychiatric Rating Scale, PANSS = Positive and Negative Syndrome Scale, F = first-episode sample, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 3d. Meta-analysis on global outcome in psychotic depression (PD) and psychotic bipolar disorder (PBD).



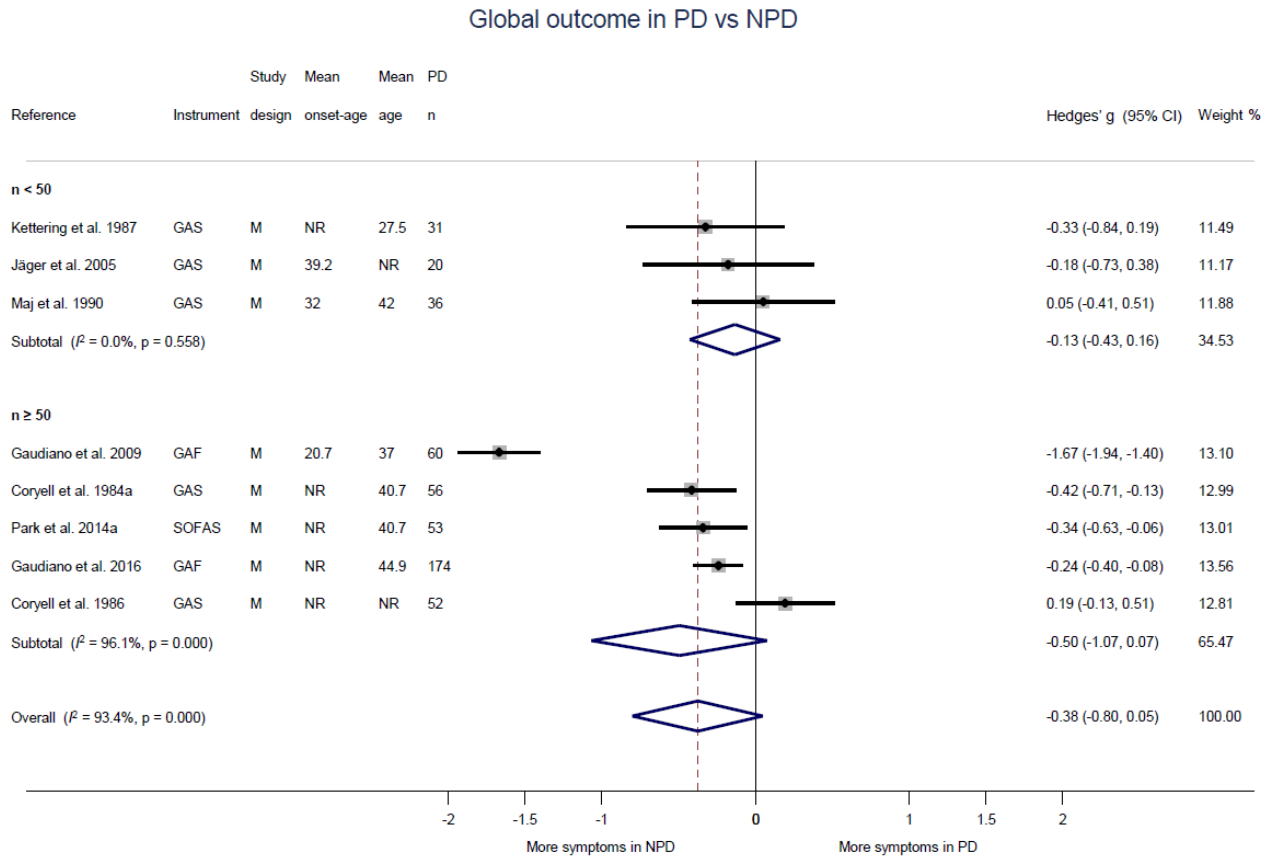
Abbreviations: PD = psychotic depression, PBD = psychotic bipolar disorder, GAF = The Global Assessment of Functioning Scale, F = first-episode sample, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 4a. Sensitivity analysis: Difference in depression symptoms in PD compared to NPD in strata by sample size.



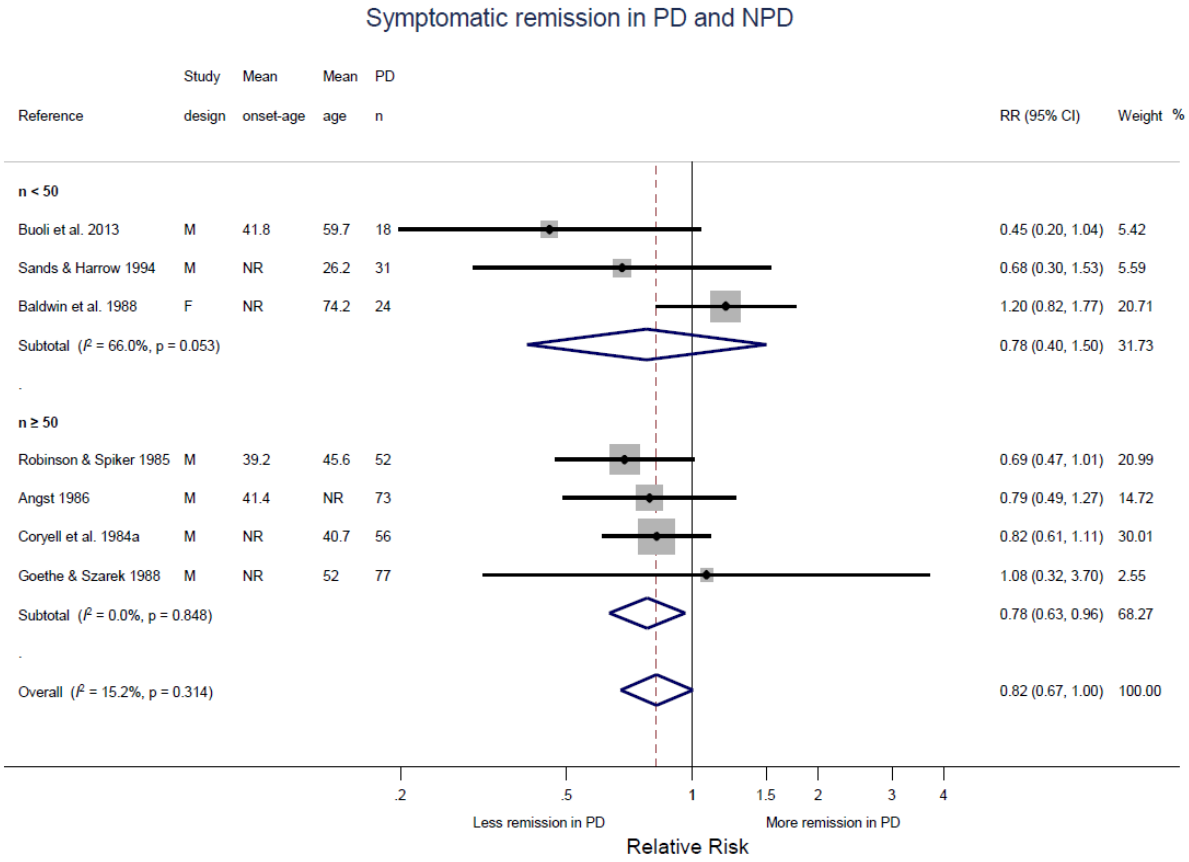
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, HAMD = Hamilton Rating Scale for Depression, IDS-C-30 = Inventory of Depressive Symptomatology (30 item), MADRS = Montgomery-Asberg Depression Rating Scale, M = mixed sample, F = first-episode sample, C = consecutive sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 4b. Sensitivity analysis: difference in global outcome in PD compared to NPD in strata by sample size.



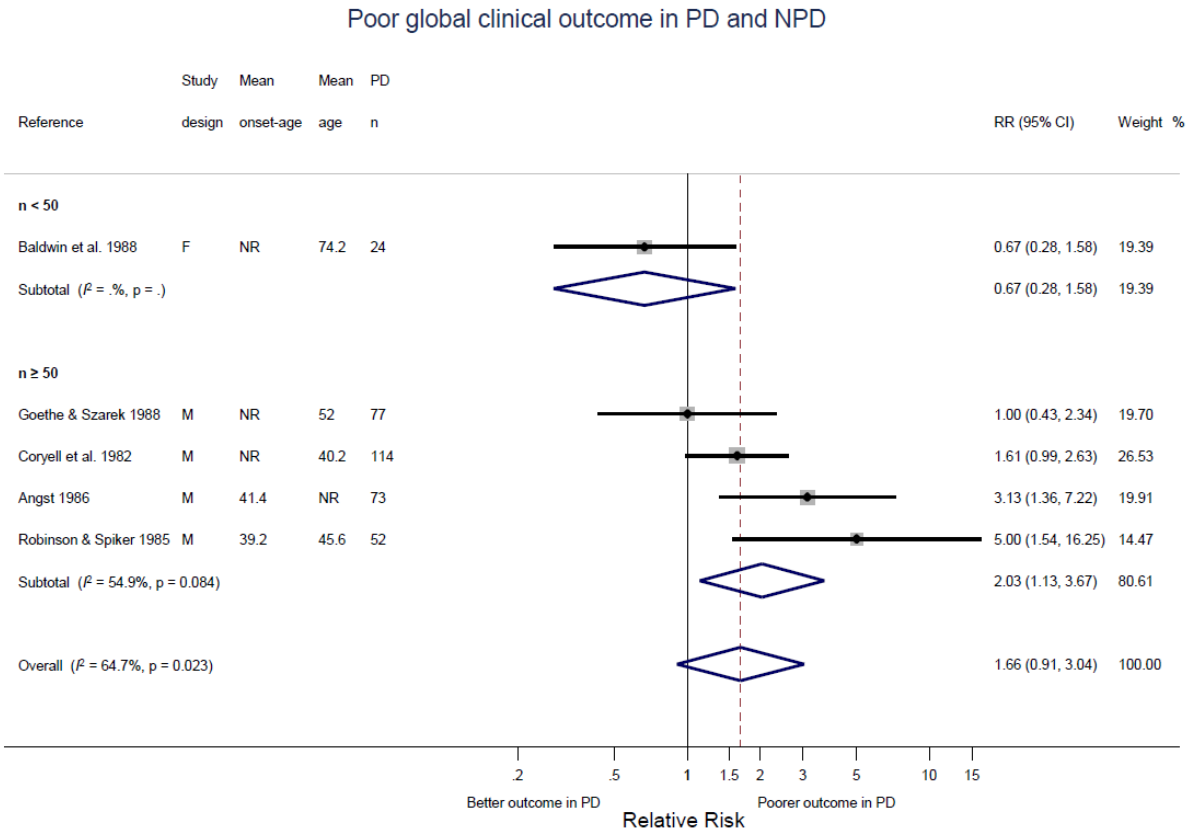
Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, GAF = The Global Assessment of Functioning Scale, GAS = Global Assessment Scale, SOFAS = Social and Occupational Functioning Assessment Scale, M = mixed sample, NR = not reported, I^2 = heterogeneity, CI = confidence interval.

Supplement Figure 4c. Sensitivity analysis: difference in symptomatic remission in PD compared to NPD in strata by sample size.



Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, RR = Relative Risk, M = mixed sample, F = first-episode sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement Figure 4d. Sensitivity analysis: difference in symptomatic remission in PD compared to NPD in strata by sample size.



Abbreviations: PD = psychotic depression, NPD = nonpsychotic depression, RR = Relative Risk, F = first-episode sample, M = mixed sample, NR = not reported, I^2= heterogeneity, CI = confidence interval.

Supplement References

Angst J (1986). The course of affective disorders. *Psychopathology* **19 Suppl 2**, 47-52.

Baldwin RC (1995). Delusional depression in elderly patients: characteristics and relationship to age at onset. *International Journal of Geriatric Psychiatry* **10**, 981-985.

Baldwin RC (1988). Delusional and non-delusional depression in late life. Evidence for distinct subtypes. *British Journal of Psychiatry* **152**, 39-44.

Baldwin P, Browne D, Scully PJ, Quinn JF, Morgan MG, Kinsella A, Owens JM, Russell V, O'Callaghan E, Waddington JL (2005). Epidemiology of first-episode psychosis: illustrating the challenges across diagnostic boundaries through the Cavan-Monaghan study at 8 years. *Schizophrenia Bulletin* **31**, 624-638.

Beiser M, Erickson D, Fleming JA, Iacono WG (1993). Establishing the onset of psychotic illness. *The American Journal of Psychiatry* **150**, 1349-1354.

Benazzi F (1999). Bipolar versus unipolar psychotic outpatient depression. *Journal of Affective Disorders* **55**, 63-66.

Breslau N, Meltzer HY (1988). Validity of subtyping psychotic depression: examination of phenomenology and demographic characteristics. *The American Journal of Psychiatry* **145**, 35-40.

Brockington IF, Helzer JE, Hillier VF, Francis AF (1982). Definitions of depression: concordance and prediction of outcome. *American Journal of Psychiatry* **139**, 1022-1027.

Buoli M, Caldiroli A, Altamura AC (2013). Psychotic versus non-psychotic major depressive disorder: a comparative naturalistic study. *Asian Journal of Psychiatry* **6**, 333-337.

Copeland JR (1983). Psychotic and neurotic depression: discriminant function analysis and five-year outcome. *Psychological Medicine* **13**, 373-383.

Coryell W, Endicott J, Keller M (1987). The importance of psychotic features to major depression: course and outcome during a 2-year follow-up. *Acta Psychiatrica Scandinavica* **75**, 78-85.

Coryell W, Endicott J, Keller M, Andreasen NC (1985). Phenomenology and family history in DSM-III psychotic depression. *Journal of Affective Disorders* **9**, 13-18.

Coryell W, Keller M, Lavori P, Endicott J (1990). Affective syndromes, psychotic features, and prognosis. I. Depression. *Archives of General Psychiatry* **47**, 651-657.

Coryell W, Lavori P, Endicott J, Keller M, VanEerdewegh M (1984a). Outcome in schizoaffective, psychotic, and nonpsychotic depression. Course during a six- to 24-month follow-up. *Archives of General Psychiatry* **41**, 787-791.

Coryell W, Tsuang MT (1982). Primary unipolar depression and the prognostic importance of delusions. *Archives of General Psychiatry* **39**, 1181-1184.

Coryell W, Tsuang MT (1985). Major depression with mood-congruent or mood-incongruent psychotic features: outcome after 40 years. *The American Journal of Psychiatry* **142**, 479-482.

Coryell W, Tsuang MT, McDaniel J (1982). Psychotic features in major depression. Is mood congruence important? *Journal of Affective Disorders* **4**, 227-236.

Coryell W, Pfohl B, Zimmerman M (1984b). The clinical and neuroendocrine features of psychotic depression. *Journal of Nervous and Mental Disease* **172**, 521-528.

Coryell W, Zimmerman M (1988). The Heritability of schizophrenia and schizoaffective disorder: a family study. *Archives of General Psychiatry* **45**, 323-327.

Coryell W, Zimmerman M, Pfohl B (1986). Outcome at discharge and six months in major depression the significance of psychotic features. *Journal of Nervous and Mental Disease* **174**, 92-96.

Craig T, Bromet EJ (2004). Parents with psychosis. *Annals of Clinical Psychiatry* **16**, 35-39.

Craig TJ, Bromet EJ, Fennig S, Tanenberg-Karant M, Lavelle J & Galambos N (2000). Is there an association between duration of untreated psychosis and 24-month clinical outcome in a first-admission series? *American Journal of Psychiatry* **157**, 60-66.

Craig TJ, Bromet EJ, Jandorf L, Fennig S, Tanenberg-Karant M, Ram R, Rosen B (1997). Diagnosis, treatment, and six-month outcome status in first-admission psychosis. *Annals of Clinical Psychiatry* **9**, 89-97.

Crebbin K, Mitford E, Paxton R, Turkington D (2008). First-episode psychosis: an epidemiological survey comparing psychotic depression with schizophrenia. *Journal of Affective Disorders* **105**, 117-124.

Čulav-Sumić J, Jukić V (2010). Minor physical anomalies in women with recurrent unipolar depression. *Psychiatry Research* **176**, 22-25.

Dell'Osso L, Pini S, Cassano GB, Mastrocinque C, Seckinger RA, Saettoni M, Papasogli A, Yale SA, Amador XF (2002). Insight into illness in patients with mania, mixed mania, bipolar depression and major depression with psychotic features. *Bipolar Disorders* **4**, 315-322.

Draper B, Anstey K (1996). Psychosocial stressors, physical illness and the spectrum of depression in elderly inpatients. *The Australian and New Zealand Journal of Psychiatry* **30**, 567-572.

Draper B, Low LF (2009). Patterns of hospitalisation for depressive and anxiety disorders across the lifespan in Australia. *Journal of Affective Disorders* **113**, 195-200.

Endicott J, Nee J, Andreasen N, Clayton P, Keller M, Coryell W (1985). Bipolar II. Combine or keep separate? *Journal of Affective Disorders* **8**, 17-28.

Farquhar F, Le Noury J, Tschinkel S, Harris M, Kurien R, Healy D (2007). The incidence and prevalence of manic-melancholic syndromes in North West Wales: 1875–2005. *Acta Psychiatrica Scandinavica. Supplementum* **433**, 37-43.

Fennig S, Bromet EJ, Galambos N, Putnam K (1996). Diagnosis and six-month stability of negative symptoms in psychotic disorders. *European Archives of Psychiatry and Clinical Neuroscience*

246, 63-70.

Fennig S, Bromet EJ, Karant MT, Ram R, Jandorf L (1996). Mood-congruent versus mood-incongruent psychotic symptoms in first-admission patients with affective disorder. *Journal of Affective Disorders* **37**, 23-29.

Filatova S, Marttila R, Koivumaa-Honkanen H, Nordström T, Veijola J, Mäki P, Khandaker GM, Isohanni M, Jääskeläinen E, Moilanen K, Miettunen J (2016). A comparison of the cumulative incidence and early risk factors for psychotic disorder in young adults in the Northern Finland Birth Cohorts 1966 and 1986. *Epidemiology and Psychiatric Sciences* [epub ahead of print].

Forty L, Jones L, Jones I, Cooper C, Russell E, Farmer A, McGuffin P, Craddock N (2009). Is depression severity the sole cause of psychotic symptoms during an episode of unipolar major depression? A study both between and within subjects. *Journal of Affective Disorders* **114**, 103-109.

Frances A, Brown RP, Kocsis JH & Mann JJ (1981). Psychotic depression: a separate entity? *The American Journal of Psychiatry* **138**, 831-833.

Frangos E, Athanassenas G, Tsitourides S, Psilolignos P, Katsanou N (1983). Psychotic depressive disorder. A separate entity? *Journal of Affective Disorders* **5**, 259-265.

Gaudiano BA, Dalrymple KL, Zimmerman M (2009). Prevalence and clinical characteristics of psychotic versus nonpsychotic major depression in a general psychiatric outpatient clinic. *Depression and Anxiety* **26**, 54-64.

Gaudiano BA, Weinstock LM, Epstein-Lubow G, Uebelacker LA, Miller IW (2016). Clinical characteristics and medication use patterns among hospitalized patients admitted with psychotic vs nonpsychotic major depressive disorder. *Annals of Clinical Psychiatry* **28**, 56-63.

Gaudiano BA, Young D, Chelminski I, Zimmerman M (2008). Depressive symptom profiles and severity patterns in outpatients with psychotic vs nonpsychotic major depression. C*omparative Psychiatry* **49**, 421-429.

Gaudiano BA, Zimmerman M (2010). The relationship between childhood trauma history and the psychotic subtype of major depression. *Acta Psychiatrica Scandinavica* **121**, 462-470.

Goethe JW, Szarek BL (1988). Do DSM-III criteria for major depression define distinct subtypes? *The Psychiatric hospital* **19**, 175-178.

Goethe JW, Szarek BL, Cook WL (1988). A comparison of adequately vs. inadequately treated depressed patients. *Journal of Nervous and Mental Disease* **176**, 465-470.

Goldberg JF, Harrow M (2005). Subjective life satisfaction and objective functional outcome in bipolar and unipolar mood disorders: a longitudinal analysis. *Journal of Affective Disorders* **89**, 79-89.

Goldberg JF, Harrow M (2004). Consistency of remission and outcome in bipolar and unipolar mood disorders: a 10-year prospective follow-up. *Journal of Affective Disorders* **81**, 123-131.

Goldstein RB, Horwath E, Wickramaratne PJ, Wolk SI, Warner V, Weissman MM (1998). Familial aggregation of delusional depression: re-examination in a recent family study. *Depression and*

Anxiety **8**, 160-165.

Gournellis R, Lykouras L, Fortos A, Oulis P, Roumbos V, Christodoulou GN (2001). Psychotic (delusional) major depression in late life: a clinical study. *International Journal of Geriatric Psychiatry* **16**, 1085-1091.

Heslin M, Desai R, Lappin JM, Donoghue K, Lomas B, Reininghaus U, Onyejiaka A, Croudace T, Jones PB, Murray RM, Fearon P, Doody GA, Dazzan P, Fisher HL, Demjaha A, Craig T, Morgan C (2016). Biological and psychosocial risk factors for psychotic major depression. *Social Psychiatry and Psychiatric Epidemiology* **51**, 233-245

Hill SK, Keshavan MS, Thase ME, Sweeney JA (2004). Neuropsychological dysfunction in antipsychotic-naive first-episode unipolar psychotic depression. *The American Journal of Psychiatry* **161**, 996-1003.

Hori M, Shiraishi H, Koizumi J (1993). Delusional depression and suicide. *The Japanese Journal of Psychiatry and Neurology* **47**, 811-817.

Husted JA, Beiser M, Iacono WG (1995). Negative symptoms in the course of first-episode affective psychosis. *Psychiatry Research* **56**, 145-154.

Ihezue UH (1985). Observations and comments on the psychosocial determinants of depressive illness among Nigerian adults. *Journal of the National Medical Association* **77**, 729-733.

Jäger M, Bottlender R, Strauss A, Möller HJ (2005). Fifteen-year follow-up of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition depressive disorders: the prognostic significance of psychotic features. *Comprehensive Psychiatry* **46**, 322-327.

Jarbin H, Ott Y, Von Knorring AL (2003). Adult outcome of social function in adolescent-onset schizophrenia and affective psychosis. *Journal of the American Academy of Child and Adolescent Psychiatry* **42**, 176-183.

Johnson J, Horwath E, Weissman MM (1991). The validity of major depression with psychotic features based on a community study. *Archives of General Psychiatry* **48**, 1075-1081.

Karaaslan F, Gonul AS, Oguz A, Erdinc E, Esel E (2003). P300 changes in major depressive disorders with and without psychotic features. *Journal of Affective Disorders* **73**, 283-287.

Kessing LV (2006). Gender differences in subtypes of late-onset depression and mania. *International Psychogeriatrics* **18**, 727-738.

Kessing LV (2005). Gender differences in patients presenting with a single depressive episode according to ICD-10. *Social Psychiatry and Psychiatric Epidemiology* **40**, 197-201.

Kettering RL, Harrow M, Grossman L, Meltzer HY (1987). The prognostic relevance of delusions in depression: a follow-up study. *American Journal of Psychiatry* **144**, 1154-1160.

Kivelä SL, Pahkala K (1989). Delusional depression in the elderly: a community study. *Zeitschrift fur Gerontologie* **22**, 236-241.

Kuhs H (1991). Depressive delusion. *Psychopathology* **24**, 106-114.

Lee TW, Tsai SJ, Yang CH, Hwang JP (2003). Clinical and phenomenological comparisons of delusional and non-delusional major depression in the Chinese elderly. *International Journal of Geriatric Psychiatry* **18**, 486-490.

Leyton M, Corin E, Martial J, Meaney M (1995). Psychotic symptoms and vulnerability to recurrent major depression. *Journal of Affective Disorders* **33**, 107-115.

Lykouras E, Malliaras D, Christodoulou GN, Papakostas Y, Voulgari A, Tzonou A, Stefanis C (1986). Delusional depression: phenomenology and response to treatment. A prospective study. *Acta Psychiatrica Scandinavica* **73**, 324-329.

Maj M, Pirozzi R, Di Caprio EL (1990). Major depression with mood-congruent psychotic features: a distinct diagnostic entity or a more severe subtype of depression? *Acta Psychiatrica Scandinavica* **82**, 439-444.

Maj M, Pirozzi R, Magliano L, Fiorillo A, Bartoli L (2007). Phenomenology and prognostic significance of delusions in major depressive disorder: a 10-year prospective follow-up study. *The Journal of Clinical Psychiatry* **68**, 1411-1417.

Maj M, Starace F, Pirozzi R (1991). A family study of DSM-III-R schizoaffective disorder, depressive type, compared with schizophrenia and psychotic and nonpsychotic major depression. *The American Journal of Psychiatry* **148**, 612-616.

Meyers BS, Alexopoulos GS, Kakuma T, Tirumalasetti F, Gabriele M, Alpert S, Bowden C, Meltzer HY (1999). Decreased dopamine beta-hydroxylase activity in unipolar geriatric delusional depression. *Biological psychiatry* **45**, 448-452.

Meyers BS, Greenberg R (1986). Late-life delusional depression. *Journal of Affective Disorders* **11**, 133-137.

Nakamura K, Iga J, Matsumoto N, Ohmori T (2015). Risk of bipolar disorder and psychotic features in patients initially hospitalised with severe depression. *Acta Neuropsychiatrica* **27**, 113-118.

Ohaeri JU, Otote DI (2002). Family history, life events and the factorial structure of depression in a Nigerian sample of inpatients. *Psychopathology* **35**, 210-219.

Ohayon MM, Schatzberg AF (2002). Prevalence of depressive episodes with psychotic features in the general population. *The American Journal of Psychiatry* **159**, 1855-1861.

Okulate GT, Oladapo HT, Osibogun A (2001). Comparison of three subtypes of depression. *The Nigerian Postgraduate Medical Journal* **8**, 41-45.

Opjordsmoen S (1991). Unipolar delusional depression. Outcome in reactive versus endogenous types. *Psychopathology* **24**, 260-269.

Østergaard SD, Petrides G, Dinesen PT, Skadhede S, Bech P, Munk-Jørgensen P, Nielsen J (2013a). The association between physical morbidity and subtypes of severe depression. *Psychotherapy and Psychosomatics* **82**, 45-52.

Østergaard SD, Waltoft BL, Mortensen PB, Mors O (2013b). Environmental and familial risk factors for psychotic and non-psychotic severe depression. *Journal of Affective Disorders* **147**, 232-240.

Owoeye O, Kingston T, Scully PJ, Baldwin P, Browne D, Kinsella A, Russell V, O'Callaghan E, Waddington JL (2013). Epidemiological and clinical characterization following a first psychotic episode in major depressive disorder: comparisons with schizophrenia and bipolar I disorder in the Cavan-Monaghan First Episode Psychosis Study (CAMFEPS). *Schizophrenia Bulletin* **39**, 756-765.

Park SC, Choi J, Kim JM, Jun TY, Lee MS, Kim JB, Yim HW, Park YC (2014a). Is the Psychotic Depression Assessment Scale a useful diagnostic tool? The CRESCEND study. *Journal of Affective Disorders* **166**, 79-85.

Park SC, Lee HY, Sakong JK, Jun TY, Lee MS, Kim JM, Kim JB, Yim HW, Park YC (2014b). Distinctive clinical correlates of psychotic major depression: The CRESCEND Study. *Psychiatry Investigation* **11**, 281-289.

Parker G, Hadzi-Pavlovic D, Hickie I, Mitchell P, Wilhelm K, Brodaty H, Boyce P, Eyers K, Pedic F (1991). Psychotic depression: a review and clinical experience. *The Australian and New Zealand Journal of Psychiatry* **25**, 169-180.

Pederson AM, Barry DJ, Babigian HM (1972). Epidemiological considerations of psychotic depression. *Archives of General Psychiatry* **27**, 193-197.

Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, Partonen T, Tuulio-Henriksson A, Hintikka J, Kieseppä T, Härkänen T, Koskinen S, Lönnqvist J (2007). Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Archives of General Psychiatry* **64**, 19-28.

Politis A, Lykouras L, Mourtzouchou P, Christodoulou GN (2004). Attentional disturbances in patients with unipolar psychotic depression: a selective and sustained attention study. *Comprehensive Psychiatry* **45**, 452-459.

Reay R, Mitford E, McCabe K, Paxton R, Turkington D (2010). Incidence and diagnostic diversity in first-episode psychosis. *Acta Psychiatrica Scandinavica* **121**, 315-319.

Robinson DG, Spiker DG (1985). Delusional depression. A one year follow-up. *Journal of Affective Disorders* **9**, 79-83.

Rush AJ, Carmody TJ, Ibrahim HM, Trivedi MH, Biggs MM, Shores-Wilson K, Crismon ML, Toprac MG, Kashner TM (2006). Comparison of self-report and clinician ratings on two inventories of depressive symptomatology. *Psychiatric Services* **57**, 829-837.

Rybakowski JK, Suwalska A, Lojko D, Rymaszewska J, Kiejna A (2007). Types of depression more frequent in bipolar than in unipolar affective illness: results of the Polish DEP-BI study. *Psychopathology* **40**, 153-158.

Sands JR, Harrow M (1994). Psychotic unipolar depression at follow-up: factors related to psychosis in the affective disorders. *The American Journal of Psychiatry* **151**, 995-1000.

Sands JR, Harrow M (1995). Vulnerability to psychosis in unipolar major depression: is premorbid functioning involved? *The American Journal of Psychiatry* **152**, 1009-1015.

Simpson S, Baldwin RC, Jackson A, Burns A (1999). The differentiation of DSM-III-R psychotic depression in later life from nonpsychotic depression: comparisons of brain changes measured by

multispectral analysis of magnetic resonance brain images, neuropsychological findings, and clinical features. *Biological Psychiatry* **45**, 193-204.

Spicer CC, Hare EH, Slater E (1973). Neurotic and psychotic forms of depressive illness. Evidence from age incidence in a national sample. *British Journal of Psychiatry* **123**, 535-541.

Taiminen T, Jääskeläinen S, Ilonen T, Meyer H, Karlsson H, Lauerma H, Leinonen KM, Wallenius E, Kaljonen A, Salokangas RK (2000). Habituation of the blink reflex in first-episode schizophrenia, psychotic depression and non-psychotic depression. *Schizophrenia Research* **44**, 69-79.

Thakur M, Hays J, Krishnan KR (1999). Clinical, demographic and social characteristics of psychotic depression. *Psychiatry Research* **86**, 99-106.

Tohen M, Strakowski SM, Zarate C,Jr, Hennen J, Stoll AL, Suppes T, Faedda GL, Cohen BM, Gebre-Medhin P, Baldessarini RJ (2000). The McLean-Harvard first-episode project: 6-month symptomatic and functional outcome in affective and nonaffective psychosis. *Biological Psychiatry* **48**, 467-476.

Tohen M, Stoll AL, Strakowski SM, Faedda GL, Mayer PV, Goodwin DC, Kolbrener ML, Madigan AM (1992). The McLean First-Episode Psychosis Project: six-month recovery and recurrence outcome. *Schizophrenia Bulletin* **18**, 273-282.

Tsuang D, Coryell W (1993). An 8-year follow-up of patients with DSM-III-R psychotic depression, schizoaffective disorder, and schizophrenia. *The American Journal of Psychiatry* **150**, 1182-1188.

Zaninotto L, Souery D, Calati R, Sentissi O, Kasper S, Akimova E, Zohar J, Montgomery S, Mendlewicz J, Serretti A (2013). Treatment resistance in severe unipolar depression: no association with psychotic or melancholic features. *Annals of Clinical Psychiatry* **25**, 97-106.