Depression during first episode psychosis and subsequent suicide risk: A systematic review and meta-analysis of longitudinal studies

Jessica McGinty\textsuperscript{a}, M. Sayeed Haque\textsuperscript{a,b} and Rachel Upthegrove*\textsuperscript{a,b,c},

\textsuperscript{a} University of Birmingham College of Medical and Dental Sciences
\textsuperscript{b} University of Birmingham, Institute of Mental Health
\textsuperscript{c} Forward Thinking Birmingham

* address for correspondence The Barberry, University of Birmingham, 25 Vincent Drive, Edgbaston Birmingham B152FG r.upthegrove@bham.ac.uk
Abstract

Background

Evidence suggests first episode psychosis (FEP) is associated with suicide, and the influence of depression on suicidal behaviour in cross sectional studies is clear. However the influence of depression during FEP on longer-term mortality is not certain. Existing evidence was synthesised to understand the influence of depressive symptoms during FEP on subsequent suicidal behaviour.

Methods

Medline, Embase, PsycINFO, Cochrane Library, Web of Science, OpenGrey, and NICE Evidence were searched from inception to Jan 25, 2017. Longitudinal observational studies assessing the relationship between depressive symptoms during FEP with a measure of suicidal behaviour at a specified follow-up time were included. Summary estimates were extracted. The Downs and Black Instrument was used to appraise study quality. Odds ratio (OR) of suicidal behaviour were calculated using random effects meta-analyses. The study protocol was registered with PROSPERO (CRD42017055581).

Results

Of 4210 articles found, 23 fulfilled eligibility criteria. 13 were included in meta-analysis (n=3002). 428 participants demonstrated suicidal behaviour in the study periods. Odds of suicidal behaviour during follow-up were significantly higher amongst patients with depressive symptoms during FEP compared to those without (OR=1.59, 95% CI 1.14–2.21; I²=50.0%, p=0.02). Meta-regression demonstrated no evidence of influence of length of follow-up on results.

Conclusions

Depressive symptoms during FEP are associated with increased longer-term risk of suicidal behaviour. This association should be acknowledged during early management planning. Large-scale clinical trials are needed to identify efficacious management of depression during FEP.

Key words:
Depression, First episode psychosis, Schizophrenia, Suicidal risk, Meta-analysis
1. Introduction

First episode psychosis (FEP) describes individuals in the early stages of psychotic illness or treatment. FEP is defined by first treatment contact, regardless of duration of untreated psychosis (Breitborde et al., 2009). FEP includes diagnoses such as schizophrenia, schizoaffective disorder and delusional disorder (WHO, 1992). Co-morbid depression is common in FEP (Upthegrove et al., 2010) and whilst historically there has been some thought that the presence of mood symptoms in ‘non-affective’ psychoses such as schizophrenia may be a good prognostic indicator, with patients appearing more on a ‘bipolar’ rather than deficit end of a psychosis continuum model (Craddock N, Craddock N, O'Donovan MC, Owen MJ. The genetics of schizophrenia and bipolar disorder: dissecting psychosis. J Med Genet. 2005 Mar;42(3):193-204.), more recent evidence suggests that depression in FEP is linked to poorer long term outcomes (Gardsjord et al., 2016; Upthegrove et al., 2010). (Conley et al., 2007) report that those with schizophrenia and depression are significantly more likely to relapse, to be a safety concern (violent, arrested, victimized), have greater substance-related problems and report poorer functioning, family relationships, and medication adherence.

Co-morbid depression is a significant factor in completed suicide in schizophrenia, more so than acting on command hallucinations (Crumlish et al., 2005; Dutta et al., 2011). Co-morbid depression is related to suicidal behavior in FEP, whether in the prodrome, acute or early post psychotic phases (Upthegrove et al., 2010). (Dutta et al., 2010) demonstrated that the first 12 months after FEP is a time of highest risk for completed suicide, however it is also clear that this risk extends to up to 5 years. In populations at high risk of developing psychosis, (Kelleher et al., 2013) show that attenuated psychotic experiences were relatively common among young people who had a diagnosis of moderate depressive disorder, and that the combination of experiences in this sample was significantly associated with suicidal behaviour.

In order to improve prognosis, optimal clinical management during FEP is key, as it is likely later trajectories are set during the early stages of psychosis (Harrison et al., 2001). Understanding the influence of depressive symptoms during FEP on the later risk of suicide could help inform early management, ultimately improving prognosis and care as co-morbid depressive symptoms present a modifiable target (Häfner et al., 2005). One existing review published in 2013, investigated multiple risk factors for deliberate self-harm after first episode psychosis, and found depression had a significant role (Challis et al., 2013). Another systematic review investigating risk factors associated with suicidal behaviour after FEP, concluded depressive symptoms were consistently affiliated (Coentre et al., 2017). However, meta-analysis was not performed. Further, the exclusion of studies focusing on depressive symptoms during acute psychosis meant the influence of depressive symptoms during FEP on long-term suicidal risk was not explored. Other reviews specifically investigating the longitudinal influence of risk factors (including depression) on
suicidal behaviour have been non-systematic or narrative in nature (Pompili et al., 2011; Ventriglio et al., 2016)

Therefore the aim of this study was to systematically synthesize the existing evidence of the influence of depressive symptoms during FEP on future, longer-term suicidal behaviors.

2. Methods

2.1. Search Strategy

This systematic review and meta-analysis has been completed in accordance with PRISMA (Moher et al., 2009) and MOOSE (Stroup et al., 2000) guidelines. Medline, Embase, PsycINFO (using ovid interface), Cochrane Library, Web of Science, OpenGrey, and NICE Evidence were searched from inception to Jan 25, 2017. The search term “[longitudinal or follow-up or prospective or retrospective or cohort or case-control] and [(initial or first or recent onset or early) adj3 (psychosis or psychotic or schizo*)]” was used for Medline. Search terms were adapted for differences in subject headings and proximity operators across databases. Searches were limited to humans. The Web of Science search was limited to conference papers and psychiatry. No time or language restrictions applied. Reference lists and citation histories of included studies were checked. Authors were asked for other recently published or ongoing studies.

Inclusion criteria were a) only longitudinal design (prospective or retrospective); b) participants experiencing FEP at baseline, in keeping with International Classification of Disease-10 codes F20-29, F30·2, F31·2, F31·5 or F32·3; c) mean age of participants between 13–45 years (to help ensure homogeneity of papers); d) a measure of depressive symptoms or level of depression during FEP (e.g. using a validated scale or diagnostic criteria); e) measure of suicidal behaviour (as defined by deliberate self-harm (DSH), suicide attempts, thoughts, plans and completed suicide) at specified follow-up time; and f) descriptive or summary statistics associating depressive symptoms or level of depression with suicidal behaviour. For meta-analysis eligibility, unadjusted summary estimates were required, including measures of variability. No minimum study duration was set.

Randomised controlled trials (RCTs) were excluded as they did not focus on the prognostic value of depressive symptoms, although published sub-studies of RCTs were eligible. Conference abstracts were excluded. A detailed list of eligibility criteria can be found in the protocol, preregistered with PROSPERO (CRD42017055881). One researcher (JM) undertook searches and screened titles, abstracts, and full texts. Another researcher (RU) blind checked the final included and excluded study list. Conflicts were discussed in reference to eligibility criteria and full text of articles.

2.2. Data Analysis
A standardised data extraction form was developed and piloted on five studies. One researcher (JM), not blind to study details, uploaded data into an excel spreadsheet (Microsoft 2016). In cases of duplicate data, the study with length of follow-up closest to the median follow-up time of included studies was chosen for meta-analysis, to help ensure homogeneity, and where follow-up was the same, the report with the most participants chosen. Data was extracted on general study details (setting, country), inclusion criteria (study design, participant demographics, measure of depression and outcome), outcome (summary estimate, variability), and details for risk of bias assessment (e.g. representative sample included, confounders).

Random effects models meta-analysis were used as inherent heterogeneity was anticipated. Odds ratio (OR) calculated show the odds of suicidal behaviour during follow-up among patients with depressive symptoms at inception of FEP over the odds among patients without depressive symptoms. Suicidal behaviour (defined as self-inflicted, potentially injurious behaviour, regardless of intent to die (Silverman et al., 2007a, b) was chosen as the outcome measure as this was the most frequently reported measure of suicidal behaviour, compared to suicidal related ideations and communications. As absolute number of events were not available in several studies, lnOR were used for calculations. One study reporting hazard ratios (Bakst et al., 2010) was included in analysis as hazard ratios were assumed similar to ORs as the frequency of the event (suicidal behaviour) is low. Standardised mean differences (SMD) were calculated for studies reporting depressive symptoms as a continuous variable and converted into lnOR under the guidance of the Cochrane handbook (Higgins and Green, 2011). Where studies reported median depression scores, means and standard deviations (SD) were estimated using the method recommended by (Hozo et al., 2005). I² was calculated to evaluate heterogeneity. Analyses were two-tailed, with a significance level of 0.05.

Small study effect was assessed through visual inspection of funnel plot asymmetry and by performing tests described by (Egger et al., 1997) and (Begg and Mazumdar, 1994). Funnel plots, with observed and imputed studies produced by the trim-and-fill method (Duval and Tweedie, 2000), are presented. Pre-specified sub-group analysis evaluated the impact of depressive symptom assessment on findings. Further post-hoc sensitivity analysis excluded studies where assumptions had been made, differed in study design, and differed in outcome measure. Meta-regression explored the influence of length of follow-up on OR. Review Manager (version 5.3) and Comprehensive Meta-analysis (version 3) were used for meta-analytic calculations.

The Downs and Black instrument (Downs and Black, 1998) was used to assess standardised risk of bias (Deeks et al., 2003). Six items specifically relating to RCTs were not used. Additionally, if characteristics of those lost to follow-up were not described and attrition rate was >20%, then high risk of attrition bias was assumed. Quality assessment was performed non-blind to study details. The study protocol was preregistered on PROSPERO (CRD42017055881).
3. Results

3.1. Study Retrieval and Selection

4210 unique papers were identified and 559 full text articles were assessed for eligibility (figure 1). 536 were excluded. A list of excluded studies with justification is available upon request. No suitable non-english articles were identified for inclusion in analysis. References of excluded non-English studies are provided (supplementary material 1). 23 articles, reporting on 3878 followed-up participants from 17 samples, were included in the qualitative review. Ten articles were excluded from meta-analysis as six were duplicate data (Klonsky et al., 2012; Madsen, 2016; Madsen, 2012; Nordentoft et al., 2002; Sanchez-Gistau, 2015; Yuen, 2014) and four had incomplete data for imputation (Ayese-Arriola, 2015; Hafner, 1999; Krausz, 1995).

The median number of participants recruited per sample was 180 (IQR 112–397). The mean age was 25.1 (SD 7.97) and 53.2% were male. Eight samples were with inpatients, four with inpatients and outpatients combined, four recruited from early intervention services, and in one, the setting was unclear. Articles were published between 1990–2016. A summary of each studies’ characteristics is presented in table 1.

3.2. Meta-analysis

Meta-analysis included 13 studies, with 3002 participants. The mean age was 23.6 (SD 6.97) and 59.2% were male. The median follow-up time was 24 months (IQR 12–60 months). Of 3002 participants, 428 were reported to have demonstrated suicidal behaviour during the study’s follow-up period.

Six used standardized measures (e.g. Classification Algorithm for the Determination of Suicide Attempt and Suicide, European Parasuide Study Interview Schedule, Clinical Global Impression of Severity of Suicidality Scale). Four studies used non-standardized measures: three used informal interviewing with medical records and one used medical records alone. For two studies, the method of assessing suicidal behaviour was unclear (Arrasate, 2016; Koreen et al., 1993). The final study measured suicide attempt and ideation using the suicidal behaviour score of the Calgary Depression Scale for Schizophrenia (CDSS) during the 2 weeks prior to follow-up (Barrett, 2015) and was excluded as part of a sensitivity analysis.

Measures of depressive symptoms or level of depression varied: eight used validated symptom scales (e.g. CDSS, Becks Depression Inventory), two used subscales of the Positive and Negative Syndrome Scale (PANSS), one used a subscale of the Brief Psychiatric Rating Scale (BPRS) and two used diagnostic
criteria. In one study, baseline depressive symptoms were assumed in one participant who attempted suicide as the timing of their depressive episode was unclear (Cohen et al., 1990). 11 studies used prospective cohort designs, one was a case-control study (Fedyszyn, 2012) and one was a sub-study of a RCT (Bertelsen, 2007). See Table 1.

The odds of suicidal behaviour during follow-up were significantly higher in patients with depressive symptoms during FEP compared to those without depressive symptoms (OR 1.59, 95% CI 1.14–2.21, p=0.006) (figure 2). The equivalent SMD of depressive symptoms between those who did and did not attempt suicide is 0.29, 95% CI 0.08–0.44. Significant moderate heterogeneity was present ($I^2$=50%, p=0.02).

Meta-regression suggests OR do not vary with different follow-up times (coefficient -0.08, 95%CI -0.23–0.07, p=0.309).

3.3. Study Heterogeneity

Visual inspection of funnel plots revealed no asymmetry (supplementary figure 1), and asymmetry was insignificant on Begg's (p=0.95) and Egger's test (p=0.88), suggesting no small study effect. The trim and fill method generated three missing studies. Effect size slightly increased and remained significant with these three studies imputed (OR 1.98, 95% CI 1.38–2.85).

3.4 Subset and sensitivity analysis

Findings did not substantially change when studies differing in study design or outcome measure were excluded during sensitivity analyses (see table 2). OR did not significantly differ between studies using different methods to measure depressive symptoms (validated scale vs subscale vs diagnostic criteria) (p=0.52).

However, in sub analysis, studies which dichotomised a continuous measure of depressive symptoms produced significantly larger effect sizes (OR 3.32, 95% CI 2.17–5.08) compared to studies using continuous measures (OR 1.18, 95% CI 0.91–1.54) (p<0.0001).

Including only studies using multiple methods to assess suicidal behaviour (Addington, 2004; Bakst et al., 2010; Fedyszyn, 2012; Gonzalez-Pinto, 2007; Sanchez-Gistau, 2013; Togay, 2015), which would have a low risk of information bias, increased the OR and this remained significant (OR 2.46, 95% CI 1.57–3.86, p=0.0001).

The relationship between depression and suicidal behaviour is dynamic. In order to address this, we performed a sub-group analysis to investigate if the timing of baseline depressive symptomology assessment influenced the overall effect size. No significant change in effect size was found between studies measuring depressive...
symptoms at different times (p=0.69) (within days of admission vs. at stabilisation vs. unclear timing).

3.5. Bias Assessment

Bias assessment for individual studies can be found in the supplementary material. The largest source of bias resulted from incomplete follow-up. 13 of 23 articles (56.5%) reported attrition rate greater than 20%, and of these only seven reported characteristics of those lost to follow-up. Of the remaining 10 articles where attrition was less than 20%, only six reported the characteristics of those lost to follow-up, making attrition bias difficult to assess in several papers.

Over 70% of studies (17 of 23) did not report whether outcomes were measured blind to depression status, a possible source of detection bias. Additionally, recall bias may be present in eight of 23 studies which assessed suicidal behavior through direct questioning alone, and record bias in two of 23 studies which used medical records alone. Eight of 23 studies assessing suicidal behaviour using mixed methods were deemed low risk of information bias. 18 of 23 articles (78.3%) approached a representative FEP sample, however for most (19 of 23 articles, 82.6%) it was unclear whether a representative FEP sample agreed to take part, possibly introducing non-response bias. In all studies, it is unlikely bias was introduced through inappropriate statistical analysis.

4. Discussion

To our knowledge, this is the first meta-analysis completed to date specifically investigating the association of depressive symptoms during FEP with long-term suicidal behaviour. The results suggest depressive symptoms during FEP are associated with greater odds of later suicidal behaviour, within a median follow-up time of 24 months. This review adds to the evidence that depressive symptoms during FEP carry poor longer-term prognostic significance. These findings are strengthened by sensitivity analyses, whereby findings remained significant.

Our findings add strength to the evidence of (Sönmez et al., 2016), who found persistent depression during the first year of treatment of FEP was associated with consistently higher levels of suicidal behaviour across a 10-year follow-up period, compared to persistent no depression during the first year. Further, our findings are consistent with those of a previous systematic meta-analysis which investigated several factors associated with deliberate self-harm (defined as self-injury regardless of lethal intent) in early psychosis, which demonstrated depression carries a similar risk (Challis et al., 2013). Our results specifically focus on depression during FEP, and the risk this may convey for subsequent suicidal behavior, suggesting the early focus of treatment is needed.
It is speculative to suggest how depression during FEP would convey a longer term risk of suicidal behaviour. It is possible that for some individuals, the presence of depression in FEP represents an enduring trait, or propensity to further depressive episodes and with these further suicidal behaviour. There is also some evidence that depression in FEP is associated with a range of poorer outcomes, relapse and repeated admission, it may be that the association between depression and future suicidal behavior is mediated through the cumulative burden of these additional negative events (Conley et al., 2007). We have also demonstrated that depression during FEP is associated with greater negative appraisals of loss, shame from the diagnosis of psychosis and feelings of entrapment (Upthegrove et al.). Recovery from these may not necessarily be co-linear with recovery from positive symptoms or depression: indeed some evidence suggests the opposite, that negative illness appraisals may need specific and targeted therapy or be subject to a lag in improvement (Brunet et al., 2012). These negative appraisals are significantly associated with risk of suicide and therefore may be the vehicle through which depression leads to longer term risk of suicidal behaviour (Upthegrove, 2016). This area warrants further targeted intervention trials to fully understand the direction of influence and potential improvement in outcome.

A clear strength of this review is the broad search strategy, devised with an aim to retrieve all longitudinal studies of FEP patients. The meta-regression calculation is likely to be adequately powered as the number of covariates to studies included is within the limits of expert recommendations (Baker et al., 2009). However, there are limitations. The review did not use individual participant data which would offer several advantages, such as performing consistent adjusted analyses (Abo-Zaid et al., 2012). Revised suicidal behaviour nomenclature ‘self-inflicted potentially injurious behaviour’ and stresses the need to clarify patient’s intent to die (no intent, undetermined intent, some intent), other research suggests that self-harm regardless of intent is significant in subsequent suicide (Hawton et al., 2015; Silverman et al., 2007a, b). However, studies included in this review used varying definitions of specific suicidal behaviour, such as suicide attempt or self-harm, and this may impact the reliability of combining studies. We were not able to adjust for potential confounders as crude estimate summaries were used and results for suicidal behavior may be weakened by sub-group analysis which found studies dichotomizing continuous measures of depressive symptoms produced significantly larger findings than those using continuous measures. Further, it was not possible to investigate the difference in effect size between affective and non-affective diagnoses, as studies included both diagnostic groups in their samples. It is possible the influence of depressive symptoms on longitudinal suicidal risk differs between affective and non-affective diagnoses, and is an area for further exploration. Future adequately designed longitudinal studies should also adjust for effects of confounders, such as previous suicide attempt and substance abuse.
Currently, there are no clear guidelines for the management of depression during FEP, although growing evidence demonstrates the safety and effectiveness of antidepressants in schizophrenia (Gregory A., 2017; Tiihonen et al., 2015). Cognitive behavioural therapy for psychosis has shown to be effective in reducing positive and negative symptoms, however depressive symptoms have not been considered as a primary outcome in trials (Jauhar et al., 2014). Suicide in psychotic disorders remains a significant concern, and the findings of this evidence synthesis adds to evidence for the need for effective identification and early treatment of depression during FEP to reduce the risk of subsequent suicidal behaviour.
References


Sanchez-Gistau, V.B., Inmaculada; Arango, Celso; Gonzalez-Pinto, Ana; de la Serna, Elena; Parellada, Mara; Graell, Motserrat; Paya, Beatriz; Llorente, Cloe; Castro-Fornieles, Josefina, 2013. Predictors of suicide attempt in early-onset, first episode psychoses: A longitudinal 24-month follow-up study. The Journal of Clinical Psychiatry 74(1), 59-66.


Upthegrove, R., Ross, K., Brunet, K., McCollum, R., Jones, L., Depression in First Episode Psychosis: The role of Subordination and Shame. Psychiatry Research(0).
Table 1: Characteristics of included studies

<table>
<thead>
<tr>
<th>Population</th>
<th>Study type</th>
<th>Dates</th>
<th>N recruit</th>
<th>N follow-up</th>
<th>Follow up period</th>
<th>Suicidality measure</th>
<th>N suicidal</th>
<th>Measure of depressive symptoms</th>
<th>Mean (SD) depression suicidal group</th>
<th>Mean (SD) depression non-suicidal group</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addington et al (2004) (ADDINGTON, 2004)</td>
<td>Early intervention service, inpatients and outpatients; Canada</td>
<td>Prospective cohort</td>
<td>Dec 96 – unknown</td>
<td>290</td>
<td>238</td>
<td>1 year</td>
<td>Suicidal behaviour (suicide or parasuicide - deliberate self-harm, regardless of suicidal intent); regular clinical practice, weekly case discussion, regional data from emergency department</td>
<td>7/238</td>
<td>CDSS</td>
<td>4·43 (4·16)</td>
<td>4·17 (3·85)</td>
</tr>
<tr>
<td>Arrasate et al (2016)</td>
<td>16–65 years, psychiatric inpatient unit; Spain</td>
<td>Prospective cohort</td>
<td>Jan 96 – Dec 97</td>
<td>112</td>
<td>82</td>
<td>5 years</td>
<td>Suicide attempts (suicide or a self-destructive act sufficient to require medical evaluation and carried out with probable suicidal intent); unclear how measured</td>
<td>19/82</td>
<td>HDRS</td>
<td>16·00 (7·57)</td>
<td>19·05 (7·77)</td>
</tr>
<tr>
<td>Barrett et al (2015)</td>
<td>18–65 years, inpatient and outpatient services; Norway</td>
<td>Prospective cohort</td>
<td>Oct 02 – Feb 11</td>
<td>207</td>
<td>146</td>
<td>1 year, (suicidality measured during past 2 weeks)</td>
<td>Suicidality defined as suicide ideation or attempts; CDSS &gt;7</td>
<td>29/146</td>
<td>G6 on PANSS</td>
<td>4* (range 1-6)</td>
<td>3* (range 1-6)</td>
</tr>
<tr>
<td>Canal-Rivero et al (2016)</td>
<td>14–54 years, public mental health service; Spain</td>
<td>Prospective cohort</td>
<td>2003 – 2005</td>
<td>65</td>
<td>65</td>
<td>0 – 6 mo.</td>
<td>Suicide attempts; SCAN interview</td>
<td>14/65</td>
<td>Depression factor of PANSS</td>
<td>2·12 (1·32)</td>
<td>2·02 (1·17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Age Range</td>
<td>Duration</td>
<td>Sample Size</td>
<td>Follow-Up</td>
<td>Description</td>
<td>Completed Suicide</td>
<td>Methodology</td>
<td>Results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang et al (2015)</td>
<td>15–25 years, early intervention program; China</td>
<td>Jul 01 – Aug 03</td>
<td>700</td>
<td>3 years</td>
<td>Suicide behaviour (suicide attempts and completed suicide); systematic medical file review, regularly updated by standardised clinical assessment forms</td>
<td>70/700</td>
<td>Systematic medical file review, regularly updated by standardised clinical assessment forms</td>
<td>CGI-BP depressive score: 2·30 (1·20) 2·30 (1·30) 0·83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohen et al (1990)</td>
<td>18–30 years; USA</td>
<td>Sub-study of a randomised controlled trial</td>
<td>1978 – 1986</td>
<td>112</td>
<td>Completed suicide; legal authorities</td>
<td>8/77</td>
<td>Completed suicide; systematic medical file review, regularly updated by standardised clinical assessment forms</td>
<td>Symptom Checklist 90 Revised Depression Score: 5·89* 3·52* &lt;0·05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>González Pinto et al (2007)</td>
<td>15–65 years, inpatient psychiatric ward; Spain</td>
<td>Prospective cohort</td>
<td>Feb 97 – Jan 99</td>
<td>112</td>
<td>Suicidal behaviour (self-destructive act sufficient to require medical evaluation and carried out with suicidal intent, completed suicide); direct interview, Spanish National Statistical Register, medical and forensic records</td>
<td>12/83</td>
<td>Suicidal behaviour (self-destructive act sufficient to require medical evaluation and carried out with suicidal intent, completed suicide); direct interview, Spanish National Statistical Register, medical and forensic records</td>
<td>HDRS: 18·42 (8·68) 17·93 (7·98) 0·85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Häfner et al (1999)</td>
<td>12–59 years, inpatient facilities; Germany</td>
<td>Prospective cohort</td>
<td>1987 – 1989</td>
<td>232</td>
<td>Suicide ideation or fairly harmless suicidal acts; PSE assessment</td>
<td>37/115</td>
<td>Suicide ideation or fairly harmless suicidal acts; PSE assessment</td>
<td>Mannheim Version of Disability Assessment Schedule, depressive symptom score: 2·11 1·90 (total sample) NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klonsky et al (2012)</td>
<td>15–60 years, inpatient facilities; USA*</td>
<td>Prospective cohort</td>
<td>Sep 89 – Dec 95</td>
<td>628</td>
<td>Suicide attempt; DSM-III-R SCID, suicide item on HDRS, reason for hospitalisation</td>
<td>44/231</td>
<td>Suicide attempt; DSM-III-R SCID, suicide item on HDRS, reason for hospitalisation</td>
<td>HDRS: 18·3 (7·69) 15·2 (7·85) NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robinson et al (2010)</td>
<td>15–30 years, early</td>
<td>Prospective</td>
<td>Apr 93 –</td>
<td>413</td>
<td>Suicide attempts and completed</td>
<td>61/282</td>
<td>Suicide attempts and completed</td>
<td>BDI: 10·3 (6·9) 8·5 (7·3) 0·068</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies measuring suicidality, with depressive symptoms as a categorical variable</td>
<td>Odds ratio</td>
<td>95% CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ayesa-Arriola et al (2015)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–60 years, inpatient and outpatient facilities; Spain</td>
<td>Retrospective cohort</td>
<td>2001–2010</td>
<td>397</td>
<td>397</td>
<td>1 mo. before admis. – 2 mo. after</td>
<td>Suicide behaviour and attempts; medical records</td>
<td>35/397</td>
<td>CDSS, dichotomised &gt;6</td>
<td>4·41†</td>
<td>1·60-12·18†</td>
<td>0·004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 mo. – 3 years</td>
<td></td>
<td></td>
<td></td>
<td>2·98†</td>
<td>1·09-8·11†</td>
<td>0·033</td>
</tr>
<tr>
<td><strong>Bakst et al (2010)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–60 years, inpatient facilities; USAa</td>
<td>Prospective cohort</td>
<td>Sep 89–Dec 95</td>
<td>628</td>
<td>529</td>
<td>2 years (at least 6 mo.)</td>
<td>Suicide attempts; SCID, suicide item on HDRS, specifically asking of suicide attempts</td>
<td>NA</td>
<td>HDRS, dichotomised &gt;18</td>
<td>2·56</td>
<td>1·38-4·77</td>
<td>&lt;0·01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 years (at least 6 mo.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2·18</td>
<td>1·57-3·04</td>
</tr>
<tr>
<td>Study</td>
<td>Age Range</td>
<td>Setting</td>
<td>Study Design</td>
<td>Outcomes</td>
<td>Measured Criteria</td>
<td>Effect Size</td>
<td>CI</td>
<td>p Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bertelsen et al (2007)</td>
<td>18–45 years, inpatient and outpatient facilities; Denmark</td>
<td>Sub study of a randomised controlled trial</td>
<td>Jan 98 – Dec 00</td>
<td>547</td>
<td>419</td>
<td>1 year</td>
<td>Suicide plans; selected questions from EPSIS II</td>
<td>NA</td>
<td>Fulfil ICD-10 criteria for depression</td>
<td>1.16</td>
<td>0.74 – 1.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Suicide attempts; selected questions from EPSIS II</td>
<td>NA</td>
<td>1.54</td>
<td>0.78 – 3.02</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fedyszyn et al (2012)</td>
<td>15–24 years, early intervention centre; Australia</td>
<td>Case control</td>
<td>Dec 02 – Nov 05</td>
<td>180</td>
<td>174</td>
<td>18 mo.</td>
<td>Suicide attempts; medical files, suicide risk assessment checklist (completed by clinical staff during inpatient admission), data from emergency department, incident cases rated using CAD-SAS</td>
<td>72/180</td>
<td>Depressed mood and at least one other (DSM-IV-TR) symptom of major depressive episode</td>
<td>4.23</td>
<td>2.21–8.09</td>
</tr>
<tr>
<td>Koreen et al (1993)</td>
<td>14–40 years, inpatient facilities; USA</td>
<td>Prospective cohort</td>
<td>1987 – 1991</td>
<td>70</td>
<td>64</td>
<td>5 years</td>
<td>Suicide ideation; unclear how measured</td>
<td>24/64</td>
<td>HDRS, dichotomised &gt;14, or ‘syndromal’ definition of depression as defined by RDC</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Krausz et al (1995)</td>
<td>14–18 years, inpatient facilities; Germany</td>
<td>Prospective cohort</td>
<td>1972 – 1978</td>
<td>61</td>
<td>NA</td>
<td>Average 30-5 years (at least 11 years)</td>
<td>Completed suicide; unclear how measured</td>
<td>8</td>
<td>Depressive traits; unclear how measured</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Madsen et al (2012)</td>
<td>18–45 years, inpatient and outpatient facilities; Denmark</td>
<td>Sub study of a randomised controlled trial</td>
<td>Jan 98 – Dec 00</td>
<td>491</td>
<td>386</td>
<td>1 year</td>
<td>Increasing suicidal tendency; selected questions from EPSIS II</td>
<td>NA</td>
<td>Fulfil ICD-10 criteria for depression</td>
<td>NA</td>
<td>OR given for increasing suicidality.</td>
</tr>
<tr>
<td>Study</td>
<td>Age Range</td>
<td>Setting</td>
<td>Study Design</td>
<td>Start - End</td>
<td>Sample Size</td>
<td>Duration</td>
<td>Suicide Trajectory</td>
<td>Suicidal Thoughts</td>
<td>Suicide Plans</td>
<td>Suicide Attempts</td>
<td>Depression Criteria</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>---------------</td>
<td>-------------</td>
<td>----------</td>
<td>--------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Madsen et al (2016)</td>
<td>18–45</td>
<td>Inpatient and outpatient</td>
<td>Sub study of a randomised controlled trial</td>
<td>Jan 98 – Dec 00</td>
<td>547</td>
<td>10 years</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Fulfil ICD-10 criteria for depression</td>
</tr>
<tr>
<td>Nordentoft et al (2002)</td>
<td>18–45</td>
<td>Inpatient and outpatient</td>
<td>Sub study of a randomised controlled trial</td>
<td>Jan 98 – Sept 99</td>
<td>321</td>
<td>1 year</td>
<td>Suicide thoughts</td>
<td>Suicide thoughts</td>
<td>Suicide plans</td>
<td>Suicide attempts</td>
<td>Fulfil ICD-10 criteria for depression</td>
</tr>
<tr>
<td>Sanchez-Gistau et al (2013)</td>
<td>9–17</td>
<td>Inpatient and outpatient</td>
<td>Prospective cohort</td>
<td>Mar 03 – Nov 05</td>
<td>110</td>
<td>2 years</td>
<td>Suicide attempts</td>
<td>Suicide attempts</td>
<td>Suicide attempts</td>
<td>Suicide attempts</td>
<td>HDRS, dichotomised &gt;19</td>
</tr>
</tbody>
</table>

- Refers to studies reporting on the same sample
- *Shows medians not means
- +Odds ratios show odds of suicidality measure among those with depressive symptoms during FEP over those without depressive symptoms during FEP
- †Only multivariate data available
- ‡Assumes both participants who attempted suicide had depressive symptoms during FEP

Abbreviations: admis, admission; BDI, Beck Depression Inventory; BPRS, Brief Psychiatric Rating scale; CAD-SAS, Classification Algorithm for the Determination of Suicide Attempt and Suicide; CDSS, Calgary Depression Scale for Schizophrenia; CGI-BP, Clinical Global Impression Scale for Bipolar Disorder; CGI-SS, Clinical Global Impression of Severity of Suicidality Scale; DSM, Diagnostic and Statistical Manual of Mental Disorders; EPSIS, European Parasuicide Study Interview Schedule; HDRS, Hamilton Depression Rating Scale; ICD, International Classification of Diseases; MANSA, Manchester Short Assessment of Quality of Life; mo., months; NA, not applicable; NS, not significant (when no P value is reported); OR, Odds Ratio; PANSS, Positive and Negative Syndrome Scale; RDC, Research Diagnostic Criteria; SCAN, Schedule for Clinical Assessment in Neuropsychiatry; SCID, Structured Clinical Interview for DSM-V; WHO, World Health Organisation.
Table 2: Summary of Sensitivity Analyse

<table>
<thead>
<tr>
<th></th>
<th>N studies</th>
<th>N Participants</th>
<th>Odds ratio (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effect model</td>
<td>13</td>
<td>3,002</td>
<td>1·57 (1·27 – 1·94)</td>
<td>&lt;0·0001</td>
</tr>
<tr>
<td>Exclude study reporting suicidal ideation with attempts (Barrett, 2015)</td>
<td>12</td>
<td>2,856</td>
<td>1·59 (1·11 – 2·29)</td>
<td>0·01</td>
</tr>
<tr>
<td>Exclude study where baseline depressive symptoms assumed (Cohen et al., 1990)</td>
<td>12</td>
<td>2,938</td>
<td>1·58 (1·13 – 2·22)</td>
<td>0·008</td>
</tr>
<tr>
<td>Exclude study reporting hazard ratios (Bakst et al., 2010)</td>
<td>12</td>
<td>2,473</td>
<td>1·50 (1·06 – 2·12)</td>
<td>0·02</td>
</tr>
<tr>
<td>Exclude three studies mentioned above</td>
<td>10</td>
<td>2,263</td>
<td>1·49 (1·00 – 2·23)</td>
<td>0·05</td>
</tr>
<tr>
<td>Include only prospective cohort studies</td>
<td>11</td>
<td>2,409</td>
<td>1·40 (1·03 – 1·90)</td>
<td>0·03</td>
</tr>
</tbody>
</table>
Figure 1: Study Selection Process

Records identified through database searching
n = 6,675
- Medline = 3,139
- PsycINFO = 562
- Embase = 2,452
- Cochrane Library = 144
- OpenGrey = 169
- NICE Evidence = 100
- Web of Science = 109

Additional records identified through other sources
n = 151
- Hand searching reference lists = 103
- Citation Search = 46
- Authors = 2

Records after duplicates removed
(n = 4,210)

Abstracts screened
(n = 4,210)

Records excluded
(n = 3,651)

Full-text articles excluded, with reasons
(n = 536)
- 15 Book/thesis/letters
- 12 Reviews
- 3 Protocols
- 2 Qualitative Studies
- 36 Not longitudinal
- 5 Ambiguous follow-up time
- 37 Do not focus on FEP patients
- 118 No suicidality outcome
- 255 No specific measure of depression at baseline
- 53 Baseline depression is not associated with outcome

Full-text articles assessed for eligibility
(n = 559)

Studies included in qualitative synthesis
(n = 23)

Studies included in quantitative synthesis (meta-analysis)
(n = 13)

Excluded (n=10)
- 6 duplicate data
- 2 no estimate of sample variance
- 1 only multivariate information available
- 1 incomplete data set
Figure 2: Odds ratio of subsequent suicidal behaviour among those with vs. without depressive symptoms during FEP

<table>
<thead>
<tr>
<th>Study</th>
<th>Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrasate et al (2016)</td>
<td>0.49 (0.19-1.27)</td>
</tr>
<tr>
<td>Chang et al (2015)</td>
<td>1.00 (0.64-1.57)</td>
</tr>
<tr>
<td>Gonzalez-Pinto et al (2007)</td>
<td>1.12 (0.37-3.37)</td>
</tr>
<tr>
<td>Addington et al (2004)</td>
<td>1.14 (0.29-4.43)</td>
</tr>
<tr>
<td>Canal-Rivero et al (2016)</td>
<td>1.16 (0.40-3.37)</td>
</tr>
<tr>
<td>Barrett et al (2015)</td>
<td>1.52 (0.72-3.19)</td>
</tr>
<tr>
<td>Bertelsen et al (2007)</td>
<td>1.54 (0.78-3.01)</td>
</tr>
<tr>
<td>Robinson et al (2010)</td>
<td>1.57 (0.94-2.64)</td>
</tr>
<tr>
<td>Togay et al (2015)</td>
<td>1.85 (0.61-5.65)</td>
</tr>
<tr>
<td>Koreen et al (1993)</td>
<td>1.92 (0.09-41.52)</td>
</tr>
<tr>
<td>Bakst et al (2010)</td>
<td>2.56 (1.38-4.76)</td>
</tr>
<tr>
<td>Fedyszyn et al (2012)</td>
<td>4.22 (2.21-8.08)</td>
</tr>
<tr>
<td>Sanchez-Gistau et al (2013)</td>
<td>4.66 (1.10-19.70)</td>
</tr>
<tr>
<td>Overall</td>
<td>3,002</td>
</tr>
</tbody>
</table>

No depressive symptoms during FEP  Depressive symptoms during FEP