Role of age, gender and marital status in prognosis for adults with depression: An individual patient data meta-analysis

J. E. J. Buckman1,2, R. Saunders1, J. Stott3, L.-L. Arundell1, C. O’Driscoll1, M. R. Davies3, T. C. Eley3, S. D. Hollon4, T. Kendrick5, G. Ambler4, Z. D. Cohen7, E. Watkins8, S. Gilbody9, N. Wiles10, D. Kessler11, D. Richards12,13, S. Brabyn9, E. Littlewood9, R. J. DeRubeis14, G. Lewis15 and S. Pilling1,16

1Research Department of Clinical, Educational & Health Psychology, Centre for Outcomes Research and Effectiveness (CORE), University College London, 1-19 Torrington Place, London WC1E 7HB, UK; 2Icope – Camden & Islington NHS Foundation Trust, St Pancras Hospital, London NW1 0PE, UK; 3Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London SE8 8AF, UK; 4Department of Psychology, Vanderbilt University, Nashville, TN 37240, USA; 5Faculty of Medicine, Primary Care, Population Sciences and Medical Education, University of Southampton, Southampton SO16 5ST, UK; 6Statistical Science, University College London, London WC1E 7HB, UK; 7Department of Psychiatry, University of California, Los Angeles, Los Angeles, CA, 90095, USA; 8Department of Psychology, University of Exeter, Exeter EX4 4QG, UK; 9Department of Health Sciences, University of York, York YO10 5DD, UK; 10Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield House, Bristol BS8 2BN, UK; 11Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, Bristol, UK; 12Institute of Health Research, University of Exeter College of Medicine and Health, Exeter EX1 2LU, UK; 13Department of Health and Caring Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway; 14Department of Psychology, School of Arts and Sciences, 425 S. University Avenue, Philadelphia PA, 19104-60185, USA; 15Division of Psychiatry, University College London, London W1T 7NF, UK and 16Camden & Islington NHS Foundation Trust, 4 St Pancras Way, London NW1 0PE, UK

Abstract

Aims. To determine whether age, gender and marital status are associated with prognosis for adults with depression who sought treatment in primary care.

Methods. Medline, Embase, PsycINFO and Cochrane Central were searched from inception to 1st December 2020 for randomised controlled trials (RCTs) of adults seeking treatment for depression from their general practitioners, that used the Revised Clinical Interview Schedule so that there was uniformity in the measurement of clinical prognostic factors, and that reported on age, gender and marital status. Individual participant data were gathered from all nine eligible RCTs (N = 4864). Two-stage random-effects meta-analyses were conducted to ascertain the independent association between: (i) age, (ii) gender and (iii) marital status, and depressive symptoms at 3–4, 6–8, 12 months post-baseline, and remission at 3−4 months. Risk of bias was evaluated using QUIPS and quality was assessed using GRADE. PROSPERO registration: CRD42019129512. Pre-registered protocol https://osf.io/e5zup/.

Results. There was no evidence of an association between age and prognosis before or after adjusting for depressive ‘disorder characteristics’ that are associated with prognosis (symptom severity, duration of depression and anxiety, comorbid panic disorder and a history of antidepressant treatment). Difference in mean depressive symptom score at 3–4 months post-baseline per−5-year increase in age = 0.02 to 0.02). There was no evidence for a difference in prognoses for men and women at 3–4 months or 9–12 months post-baseline, but men had worse prognoses at 6−8 months (percentage difference in depressive symptoms for men compared to women: 15.08% (95% CI: 4.82 to 26.35)). However, this was largely driven by a single study that contributed data at 6–8 months and not the other time points. Further, there was little evidence for an association after adjusting for depressive ‘disorder characteristics’ and employment status (12.23% (~1.69 to 28.71%). Participants that were either single (percentage difference in depressive symptoms for single participants: 9.25% (95% CI: 2.78 to 16.13) or no longer married (8.02% (95% CI: 1.31 to 15.18)) had worse prognoses than those that were married, even after adjusting for depressive ‘disorder characteristics’ and all available confounders.

Conclusion. Clinicians and researchers will continue to routinely record age and gender, but despite their importance for incidence and prevalence of depression, they appear to offer little information regarding prognosis. Patients that are single or no longer married may be expected to have slightly worse prognoses than those that are married. Ensuring this is recorded routinely alongside depressive ‘disorder characteristics’ in clinic may be important.
Introduction

Patients and clinicians can benefit from prognostic information (Hayden et al., 2013). Such knowledge can inform routine clinical assessments and triages of patients seeking treatment for depression and might guide the collaborative treatment decision-making process. This information does not help determine which treatment may be most likely to benefit any individual patient, but instead, informs considerations of the intensity, dose or duration of treatment, the regularity of progress reviews and referrals to specialist treatment centres (Buckman et al., 2021). This would be particularly true if the differences in prognoses for patients with one compared to another pre-treatment characteristic were of a clinically important magnitude (Button et al., 2015). Prognostic information is often wanted by patients and clinicians (Trusheim et al., 2007). However, prognostication for adults with depression is challenging as most prior studies have considered prognosis following one particular type of treatment, or in samples where either no treatment was given or details regarding any treatment(s) are unknown (Buckman et al., 2020; Buckman et al., 2021). As there is an approximate equivalence (on average) in the effectiveness of the most common types of treatment for depression (Menchetti et al., 2010; Cuijpers et al., 2013; Weitz et al., 2017; Cipriani et al., 2018), and a number of prognostic factors have been found to be associated with outcomes from one type of treatment but not others (Dodd et al., 2014; Chekroud et al., 2016; Nakabayashi et al., 2018; Buckman et al., 2021), there is a need to study prognosis regardless of the type of treatment received. Doing so would identify factors that are associated with prognosis in general, which would be informative for clinicians assessing and treating patients with depression in a breadth of settings. We call this prognosis independent of treatment. This form of prognosis has been largely overlooked in studies of depression (Buckman et al., 2020; Buckman et al., 2021). We recently found that a number of clinical factors were associated with prognosis independent of treatment, and that each contributed to explaining prognosis over and above depressive symptom severity (Buckman et al., 2021). We termed these depressive ‘disorder characteristics’. These were: the durations of depression and anxiety problems, comorbid panic disorder and a history of antidepressant medication, each should be routinely assessed in clinic alongside depressive symptom severity. Whether other factors are similarly associated with prognosis and add incremental value to prognostic assessments is uncertain.

For decades it has been shown that age, gender and marital status are associated with the prevalence of depression (Ensel, 1982; Kessler and Essex, 1982; Bebbington, 1987; Faravelli et al., 2013). Previous reviews of the relationship between these factors and post-treatment prognosis have found inconsistent results. For example, three reviews reported that older age was associated with worse prognosis (Cuijpers et al., 2018; Nakabayashi et al., 2018; Sockol, 2018); two found no such association (Johnsen and Friberg, 2015; Karyotaki et al., 2017); and one found that prognosis improved with increasing age (Noma et al., 2019). Similarly, four reviews reported that women had poorer prognoses than men (Carter et al., 2012; Dodd et al., 2014; Johnsen and Friberg, 2015; Noma et al., 2019), while four reported no such association (Karyotaki et al., 2017; Cuijpers et al., 2018; Nakabayashi et al., 2018; Schoemaker et al., 2018). Two reviews reported better prognoses for married individuals (Carter et al., 2012; Sockol, 2018) while a third found no evidence for this association (Karyotaki et al., 2017).

There are clear limitations, most of these past studies are not directly comparable as the reviews examined specific treatments only (e.g., duloxetine or Interpersonal Psychotherapy), and the sample settings were unclear or unstated. None of the past reviews presented results adjusted for the effects of multiple known clinical prognostic factors (Buckman et al., 2021). To do so accurately, for factors that differ at the individual patient level requires individual patient data (IPD) (Rothwell, 2005; Fisher, 2015) and only four of the past reviews had IPD. Two were not systematic reviews and made no adjustment for any clinical factors (Nakabayashi, Hara and Minami, 2018; Noma et al., 2019). Two others adjusted for baseline depressive symptom severity (Dodd et al., 2014; Karyotaki et al., 2017). The first (Dodd et al., 2014) made no adjustments for between-study effects and no assessments of either attrition or heterogeneity, limiting the robustness and interpretability of the findings. The second was an investigation of potential moderators of treatment effects and found high levels of heterogeneity that could not be explained (Karyotaki et al., 2017). Clarity on the association of these factors with prognosis is needed, particularly for patients initially seeking treatment in primary care as large proportions of adults are either initially screened and assessed or are wholly treated for depression in primary care settings (McManus et al., 2016; Olsson et al., 2016; Thornicroft et al., 2017).

Given these contrasting results, we aimed to determine whether age, gender and marital status are associated with prognosis for adults with depression who sought treatment in primary care. We investigated these associations after accounting for the effect of known clinical prognostic factors, the depressive ‘disorder characteristics’ we identified previously (Buckman et al., 2021), and regardless of the type of treatment. In so doing, we aimed to determine whether these factors add incrementally to knowledge of prognosis at the point patients initially seek treatment.

Methods

This systematic review with IPD meta-analysis is reported in accordance with the PRISMA-IPD statement (Stewart et al., 2015) (checklist in online Supplementary materials).

Identification and selection of studies

A protocol for this study was pre-registered (https://osf.io/e5zup/). A general protocol for forming the Depression in General Practice (Dep-GP) IPD dataset and the plan for analysing those data (Buckman et al., 2020) and pre-registered methods for identifying studies are also available (PROSPERO: CRD42019129512 (01/04/2019)), and were reported in accordance with the PRISMA-P statement (Shamseer et al., 2015). Searches have been reported in accordance with PRISMA-S (Rethlefsen et al., 2021), with a brief description below and more details provided in online Supplementary Materials. Searches were conducted on Medline, Embase, International Pharmaceutical Abstracts, PsyInfo and Cochrane Central (searched from inception to 1st December 2020), we also hand-searched reference lists, and contacted experts for unpublished or missed studies. Search terms included variations of phrases such as ‘depression’ or ‘major depression’, ‘RCT’ or ‘Randomised Controlled Trial’ and ‘CIS-R’ or ‘Clinical Interview Schedule’. See online Supplementary Table 1 for a full list and results of the searches.
A single reviewer (JB) screened titles and abstracts of potentially eligible studies, these were then read in full and judged against inclusion/exclusion criteria by two reviewers (JB and GL) with consultation with a third (SP) to resolve any uncertainties by consensus.

**Inclusion and exclusion criteria**

Studies were included if they were RCTs of adults (aged 16 or over) with unipolar depression, depressive symptoms significant enough to seek treatment, or a Revised Clinical Interview Schedule (CIS-R) (Lewis et al., 1992) score of ≥12 (the cut-off for common mental disorder); recruited from primary care; had at least one active treatment arm; used the CIS-R at baseline, and assessed marital status, age and gender at baseline.

Studies were excluded if they: included patients with depression secondary to a diagnosis of personality disorder, psychotic conditions or neurological conditions; were studies of adults with bi-polar or psychotic depressions; were studies of children or adolescents; were feasibility studies; were studies of just one gender, one particular age group or one marital status group only.

**Data extraction**

The included studies are detailed in Table 1. Data were extracted for each study participant on all measures included in Table 2 and additional socio-demographics listed in Table 3 by the chief investigators or data managers of each study. Data were independently cleaned by two independent reviewers (JB and RS). Issues were resolved by consensus between four reviewers (JB, RS, GL and SP).

**Data integrity checks**

Integrity of all baseline and endpoint data for each study was checked with the study team and against publications from each study. We removed two participants from IPCRESS and one from PANDA due to them missing data on over 75% of baseline variables. There were also data available on 36 more participants in ITAS than reported in the publication about that study.

**Ethical considerations**

All studies were granted NHS Research ethical approvals and all participants gave informed consent, see online Supplementary Table 2. No additional ethical approval was required for this study: HRA reference 712/86/32/81.

**Confounders and covariates**

Data on some potential confounders were available in all studies whereas data on others were not (i.e., systematically missing). Following our protocol, we adjusted for the following in separate models of each outcome:

1) Depressive ‘disorder characteristics’: depressive symptom severity, the duration of the current depressive episode, mean duration of anxiety problems, comorbid panic disorder and a history of antidepressant treatment.
2) Age, gender and marital status (excluding whichever of these was already included in the model as the potential prognostic factor).
3) Employment status.

In sensitivity analyses we then added each of the systematically missing variables one at a time:

1) Housing status (available in eight studies).
2) Long-term physical health condition status (yes/no; available in eight studies).

Exponentiation of the regression coefficient provides an estimate of the percentage difference in symptoms.

It was expected that these methods would give similar results but the log outcome might have greater utility as percentage differences are more easily understood and do not require division by standard deviation estimates.

**Secondary outcomes**

1) Remission on the primary depression measure in each study at 3–4 months post-baseline (a score of <10 on the BDI-II or PHQ-9, whichever was the primary measure collected at that time point in each study).
2) Depressive symptoms at 6–8 months post-baseline, as above but also including the GHQ-12, captured with (i) the z-score calculated using the mean and standard deviation for the scores at 3–4 months and (ii) the logarithm of scores at 6–8 months.
3) Depressive symptoms at 9–12 months post-baseline, as above, captured with (i) the z-score calculated using the mean and standard deviation for the scores at 3–4 months and (ii) the logarithm of scores at 9–12 months.

**Prognostic indicators under consideration**

1) Age was considered both as a continuous variable per 5-year increase, and in age groups: (i) 16–29 (ii) 30–39; (ii) 40–49; (iii) 50–59; (iv) ≥60 years old.
2) Gender was taken from a question asking participants for their self-reported gender at baseline, in all included studies response options were binary. This was commonplace in studies at the time, it does not capture the appropriate range of gender identities, neither does it capture differences between gender and sex assigned at birth.
3) Marital status was initially recorded in five categories: (1) married or cohabiting (living with partner); (2) single (never married), (3) separated, (4) divorced, (5) widowed. However, some of the studies had very few participants that were divorced (n = 0 in ITAS) and widowed (n = 4 in GENPOD, n = 1 in IPCRESS and n = 6 in TREAD). Therefore, marital status was recoded into three categories to meet power requirements by combining categories 3–5 above in a new category ‘no longer married’.

**Residuals as outcome**

There were also data available on 36 more participants in ITAS than reported in the publication about that study.

**Data analysis plan**

The pre-registered analysis plan (https://osf.io/e5zup/) is outlined below. It builds on the plan outlined in our general protocol (Buckman et al., 2020).

**Primary outcomes**

Depressive symptoms at 3–4 months post-baseline were captured in two ways: (1) z-score (standardised mean) of the scores on the depressive symptom measures (BDI-II and PHQ-9) used at 3–4 months post-baseline within each study (see Table 1). The score at 3–4 months was divided by the standard deviation for that measure calculated at 3–4 months. (2) The logarithm of depression scale scores irrespective of the measure used.

Exponentiation of the regression coefficient provides an estimate of the percentage difference in symptoms.

It was expected that these methods would give similar results but the log outcome might have greater utility as percentage differences are more easily understood and do not require division by standard deviation estimates.

**Secondary outcomes**

1) Remission on the primary depression measure in each study at 3–4 months post-baseline (a score of <10 on the BDI-II or PHQ-9, whichever was the primary measure collected at that time point in each study).
2) Depressive symptoms at 6–8 months post-baseline, as above but also including the GHQ-12, captured with (i) the z-score calculated using the mean and standard deviation for the scores at 3–4 months and (ii) the logarithm of scores at 6–8 months.
3) Depressive symptoms at 9–12 months post-baseline, as above, captured with (i) the z-score calculated using the mean and standard deviation for the scores at 3–4 months and (ii) the logarithm of scores at 9–12 months.

**Prognostic indicators under consideration**

1) Age was considered both as a continuous variable per 5-year increase, and in age groups: (i) 16–29 (ii) 30–39; (ii) 40–49; (iii) 50–59; (iv) ≥60 years old.
2) Gender was taken from a question asking participants for their self-reported gender at baseline, in all included studies response options were binary. This was commonplace in studies at the time, it does not capture the appropriate range of gender identities, neither does it capture differences between gender and sex assigned at birth.
3) Marital status was initially recorded in five categories: (1) married or cohabiting (living with partner); (2) single (never married), (3) separated, (4) divorced, (5) widowed. However, some of the studies had very few participants that were divorced (n = 0 in ITAS) and widowed (n = 4 in GENPOD, n = 1 in IPCRESS and n = 6 in TREAD). Therefore, marital status was recoded into three categories to meet power requirements by combining categories 3–5 above in a new category ‘no longer married’.

**Residuals as outcome**

There were also data available on 36 more participants in ITAS than reported in the publication about that study.

**Data analysis plan**

The pre-registered analysis plan (https://osf.io/e5zup/) is outlined below. It builds on the plan outlined in our general protocol (Buckman et al., 2020).

**Primary outcomes**

Depressive symptoms at 3–4 months post-baseline were captured in two ways: (1) z-score (standardised mean) of the scores on the depressive symptom measures (BDI-II and PHQ-9) used at 3–4 months post-baseline within each study (see Table 1). The score at 3–4 months was divided by the standard deviation for that measure calculated at 3–4 months. (2) The logarithm of depression scale scores irrespective of the measure used.

Exponentiation of the regression coefficient provides an estimate of the percentage difference in symptoms.

It was expected that these methods would give similar results but the log outcome might have greater utility as percentage differences are more easily understood and do not require division by standard deviation estimates.
| Study          | N    | Inclusion criteria                                                                 | Age Mean (S.D) | Gender % Female | Marital status % married, single, no longer married | To depressive severity mean (S.D) | Remission % at 3–4 months | Interventions                          | Outcome measure               |
|---------------|------|-------------------------------------------------------------------------------------|----------------|----------------|---------------------------------------------------|---------------------------------|--------------------------|-------------------------------------|-------------------------------|
| CADET         | 527  | Adults ≥18, ICD-10 Depressive Episode                                               | 44.4(13.2)     | 72%            | 41.4% married; 30.2% single; 28.5% no longer married | PHQ-9 = 17.7(5.1)               | 41%                      | Collaborative Care vs TAU          | PHQ-9                        |
| COBALT        | 469  | Adults 18–75 with treatment resistant depression, scoring ≥14 BDI-II               | 49.6(11.7)     | 72%            | 52.9% married; 19.0% single; 28.1% no longer married | BDI-II = 31.8(10.7)              | 34%                      | CBT + TAU vs TAU                   | BDI-II (PHQ-9)                |
| GENPOD        | 601  | Adults 18–74 with depressive episode                                               | 38.8(12.4)     | 68%            | 52.6% married; 29.1% single; 18.3% no longer married | BDI-II = 33.7(9.7)               | 41%                      | Citalopram vs Reboxetine           | BDI-II                       |
| IPCRESS       | 295  | Adults scoring ≥14 BDI-II and GP confirmed diagnosis of depression                 | 34.9(11.6)     | 68%            | 36.6% married; 47.8% single; 15.6% no longer married | BDI-II = 33.2(8.8)               | 34%                      | iCBT + TAU vs TAU + waiting list for iCBT | BDI-II                       |
| ITAS          | 798  | Adults ≥16, scored ≥12 on CIS-R                                                    | 43.2(14.8)     | 68%            | 58.7% married; 22.4% single; 18.9% no longer married | GHQ = 7.7(3.2)                  | N/A; at 6-8 months 46%          | Recommendation + TAU vs TAU       | GHQ-12                        |
| MIR           | 480  | Adults ≥18 taking SSRIs or SNRIs at adequate dose for ≥6 weeks, and scored ≥14 on BDI-II | 50.7(13.2)     | 69%            | 59.4% married; 19.6% single; 21.0% no longer married | BDI-II = 31.1(9.9)               | 30%                      | Mirtazapine vs Placebo             | BDI-II (PHQ-9)                |
| PANDA         | 652  | Adults presenting with low mood or depression to GP in last 2 years, free of ADM for 8 weeks up to baseline | 39.7(15.0)     | 59%            | 39.1% married; 45.4% single; 15.5% no longer married | BDI-II = 23.9(10.3)              | 69%                      | Sertraline vs Placebo              | PHQ-9 (BDI-II)                |
| REEACT        | 685  | Adults with PHQ-9 ≥10 presenting to GP with depression                             | 39.9(12.7)     | 67%            | 50.9% married; 33.5% single; 15.6% no longer married | PHQ-9 = 16.7(4.3)                | 53%                      | Moodgym vs Beating the Blues vs TAU | PHQ-9                        |
| TREAD         | 361  | Adults 18-69 who met diagnostic criteria for MDD and scored ≥14 on BDI-II          | 39.9(12.6)     | 66%            | 46.3% married; 32.1% single; 21.6% no longer married | BDI-II = 32.1(9.2)               | 35%                      | Physical Activity + TAU vs TAU     | BDI-II                       |

Abbreviations: ADM, antidepressant medication; BDI-II, Beck Depression Inventory; EPDS, Edinburgh Postnatal Depression Scale; GHQ-12, General Health Questionnaire 12 item version; HADS-D, Hospital Anxiety and Depression Scale, depression subscale; iCBT (internet-based therapist delivered cognitive behavioural therapy); MDD, Major Depressive Disorder; T0 - Baseline; TAU, treatment as usual; TCA, tricyclic antidepressant.
study-level. meant that all models accounted for clustering effects at the adjusted for the effects of the randomised treatments. It also able in all models. This meant that associations between the ent category in a single categorical variable and including this vari-

5) Social support (available in six studies).

4) Financial wellbeing (available in six studies).

3) Highest level of educational attainment (available in seven studies).

In addition, all models were adjusted for treatment allocation by assigning each of the randomised groups across the studies a different category in a single categorical variable and including this variable in all models. This meant that associations between the prognostic factors and outcomes could be investigated having adjusted for the effects of the randomised treatments. It also meant that all models accounted for clustering effects at the study-level.

**Data handling and data management**

Details of the pre-processing stages and handling of missing data are given in the original protocol (Buckman et al., 2020). Multiple imputation with chained equations was used with baseline and outcome variables to generate 50 imputed datasets for each study. Systematically missing variables were not imputed.

**Primary analyses**

Two-stage DerSimonian and Laird random effects meta-analyses were conducted for each prognostic factor (each fitted into separate models) for each outcome variable listed above, using the 'admetan' package (Fisher, 2015) in Stata 16 (StataCorp, 2019). Linear models were fitted for the z-score and log outcomes and

---

**Table 2. Measures used across the studies of the Dep-GP IPD dataset meeting inclusion criteria for the present study**

| Measure | Details | Scores and Cut-offs for Remission |
|---------|---------|----------------------------------|
| The CIS-R (Lewis et al., 1992) | Consists of 14 symptom subsections scored 0–4 covering core features of depression, depressive thoughts (scored 0–5), fatigue, concentration/forgetfulness, and sleep, generalised anxiety, worry, irritability, obsessions, compulsions, health anxiety, somatic concerns, phobic anxiety (split into agoraphobia, social phobia, and specific phobia), and panic. A final section measures general health, impairment and weight change. | The total score ranges from 0-57 with a cut-off of ≥12 used to indicate likely common mental disorder, primary and secondary diagnoses using ICD-10 criteria are given as are binary indicators of diagnosis for all the disorders assessed. The duration of each type of problem is also assessed for the present episode (or sub syndromal episode) up to the point of completing the CIS-R. Duration items are measured in five categories: (1) less than two weeks; (2) between two weeks and six months; (3) between six months and one year; (4) between one and two years; and (5) more than two years. |
| Beck Depression Inventory 2nd Edition (BDI-II) (Beck, Steer and Brown, 1996) | Consists of 21 items to assess depressive symptoms, each item is scored −3. | There is a maximum score obtainable of 63, and a cut-off of ≥10 is used to indicate significant symptoms of depression, scores of <10 are therefore used to indicate remission in those that were previously depressed/scored ≥10. |
| Patient Health Questionnaire 9-item version (PHQ-9) (Kroenke, Spitzer and Williams, 2001) | This is a depression screening measure, with respondents asked to rate how often they have been bothered by each of the nine symptom items over the preceding two weeks. Each item is scored 0–3. | There is a maximum score of 27 with a cut-off of ≥10 is used to indicate ‘caseness’ for depression, a score of 9 or below for those that were previously depressed is therefore considered to indicate remission |
| General Health Questionnaire (12-item version) (GHQ-12) (Goldberg, 1992) | Consists of 12 items related to present and recent health over the ‘few weeks’ prior to completion. Each item is related to depression or generalised anxiety, they are scored 0-0-1-1 for the four response options. | A cut-off of ≥2 is used to indicate the likely presence of common mental disorder, and so scores of <2 for those formally scoring above this would be considered to indicate remission. |
| Social Support Scale - adapted by authors of RCTs (Kessler et al., 2009) included in this IPD by adding one item to the Health and Lifestyles Survey Social Support Measure (Cox et al., 1987). | An 8-item instrument (the first seven of which are from the Health and Lifestyles Survey) assessing the degree to which participants rated the social support of their friends and family in each of the following domains: (1) being accepted for who one is; (2) feeling cared about; (3) feeling loved; (4) feeling important to them; (5) being able to rely on them; (6) feeling well supported and encouraged by them; (7) being made to feel happy by them; and (8) feeling able to talk to them whenever one might like. Items are scored 1-3, with total scores ranging from 8 to 24; higher scores indicate higher levels of perceived social support. The authors of the Health and Lifestyles Survey suggested the maximum score for social support (which was 21 on that scale) indicated ‘no lack of social support’; scores between 18 and 20 indicated a ‘moderate lack of social support’, and scores of 17 or below indicated a ‘severe lack of social support’. | N/A |

---

CIS-R was used in all studies n = 4864. BDI-II was used in six studies (COBALT, GENPOD, IPCRESS, MIR, PANDA and TREAD), n = 2858; PHQ-9 was used in five studies (CADET, COBALT, MIR, PANDA and REEACT) n = 2807; GHQ was used in ITAS only n = 796; and the Social Support Scale was used in six studies (COBALT, GENPOD, IPCRESS, MIR, PANDA and TREAD) n = 2858.
logistic models were fitted for remission. The degree of heterogeneity was assessed using prediction intervals and its impact assessed using the $I^2$ statistic (Higgins et al., 2003).

### Table 3. Baseline characteristics of study sample using observed data only

| Self-reported baseline characteristics | Factor | N(%) or N and Mean(s.d.) |
|---------------------------------------|--------|--------------------------|
| Total sample size ($N$)               | 4864   |                          |
| Age in years                          | N: 4864, Mean(s.d.): 42.45(13.99) |
| Age 16–29 years old                   | 1079(22.18%) |
| Age 30–39 years old                   | 1043(21.44%) |
| Age 40–49 years old                   | 1208(24.84%) |
| Age 50–59 years old                   | 919(18.89%)  |
| Age 60 + years old                    | 615(12.64%)  |
| Gender                                | Female | 3279(67.41%)              |
|                                       | Male   | 1583(32.55%)              |
|                                       | Missing| 2(0.04%)                  |
| Marital status                        | Married/cohabiting | 2412(49.59%) |
|                                       | Single | 1477(30.37%)              |
|                                       | No longer married | 975(20.05%)  |
| Employment status                     | Employed | 2713(55.81%)          |
|                                       | Not seeking employment | 1199(24.67%) |
|                                       | Unemployed | 949(19.52%)             |
|                                       | Missing | 3(0.06%)                  |
| Housing status                        | Home owner | 2148(48.92%)           |
|                                       | Tenant  | 1677(38.19%)             |
|                                       | Other   | 566(12.89%)               |
|                                       | Missing | 473(9.72%)                |
| Financial wellbeing                   | Doing OK financially | 1537(42.14%)         |
|                                       | Just about getting by | 1171(32.11%)          |
|                                       | Struggling financially | 939(25.75%)           |
|                                       | Missing | 1217(25.02%)             |
| Highest level of educational attainment | Degree or above | 959(28.02%)          |
|                                       | A-level of diplomas | 905(26.44%)          |
|                                       | GCSE    | 1016(29.68%)             |
|                                       | None or other | 543(15.86%)             |
|                                       | Missing | 1441(39.63%)             |
| Long-term physical health condition   | No      | 3244(73.81%)             |
|                                       | Yes     | 1151(26.19%)             |
|                                       | Missing | 469(9.64%)                |
| Social support scale score           | N: 2858, Mean(s.d.): 20.25(3.85) |
| Past antidepressant use              | No      | 1241(25.53%)             |
|                                       | Yes     | 3620(74.47%)             |
| CIS-R durations                      | Depression | N: 4864, Mean(s.d.): 3.32(1.37) |
|                                       | Average anxiety duration | N: 4813, Mean(s.d.): 2.05(0.98) |

(Continued)

| Self-reported baseline characteristics | Factor | N(%) or N and Mean(s.d.) |
|---------------------------------------|--------|--------------------------|
| Comorbid panic disorder               | No     | 4439(91.26%)             |
|                                       | Yes    | 425(8.74%)               |
| Baseline BDI-II score                 | N: 2858, Mean(s.d.): 30.44(10.53) |
| Baseline PHQ-9 score                  | N: 2812, Mean(s.d.): 15.71(5.65) |
| Baseline GHQ-12 score                 | N: 795, Mean(s.d.): 7.69(3.22)  |
| Attrition at 3–4 months               | No     | 3411(70.13%)             |
|                                       | Yes    | 658(13.53%)              |
|                                       | Not applicable | 795(16.34%)          |
| BDI-II score 3–4 months               | N:1918, Mean(s.d.): 16.07(11.99) |
| PHQ-9 score 3–4 months                | N:2393, Mean(s.d.): 10.28(6.65) |
| Remission 3–4 months                  | No     | 1928(56.57%)             |
|                                       | Yes    | 1480(43.43%)             |
| BDI-II score 6–8 months               | N:1236, Mean(s.d.): 18.64(11.44) |
| PHQ-9 score 6–8 months                | N:814, Mean(s.d.): 10.33(6.78)  |
| GHQ-12 score 6–8 months               | N:585, Mean(s.d.): 3.80(4.10)   |
| Attrition at 6–8 months               | No     | 1236(25.41%)             |
|                                       | Yes    | 369(7.59%)               |
|                                       | Not applicable | 3259(67.00%)          |
| BDI-II score 9–12 months              | N:1028, Mean(s.d.): 16.78(12.87) |
| PHQ-9 score 9–12 months               | N:1764, Mean(s.d.): 9.51(6.71)  |
| Attrition at 9–12 months              | No     | 2005(41.22%)             |
|                                       | Yes    | 516(10.61%)              |
|                                       | Not applicable | 2343(48.17%)          |

Risk of bias

Two reviewers (JB & RS) conducted independent risk of bias assessments using the Quality in Prognosis Studies (QUIPS) (Hayden et al., 2013) and the quality of evidence for each prognostic indicator was assessed using the GRADE framework (Guyatt et al., 2008). Disagreements were resolved by consensus with two additional reviewers (GL and SP).

Sensitivity analyses

In addition to those listed above sensitivity analyses were conducted if there was considerable heterogeneity (Higgins et al., 2003).
Green, 2011) either from inspection of the forest plots or if $I^2$ was 75% or above, or if either study quality was low or risk of bias was high, removing the study contributing most to the heterogeneity/low quality/high risk of bias.

**Results**

**Characteristics of the included studies**

In total nine RCTs met the inclusion criteria. IPD from all 4864 participants formed the present dataset, see Fig. 1.

**Quality assessments and risk of bias**

Overall, the risk of bias was low, and quality was high in all studies, so no sensitivity analyses were required in relation to these, although half of the studies had a moderate risk of bias related to attrition, see online Supplementary Tables 3 and 4. There was near perfect agreement between the reviewers, with interrater reliability (Cohen’s Kappa) $k = 0.96$ for QUIPS and $k = 1.00$ for GRADE.

**Descriptive statistics**

Across all nine studies, age ranged between 16 and 84 years old at baseline with a mean age of approximately 42 years, 80% of the sample were aged between 23 and 61 years old. Approximately 67% of the participants were women, half of participants were married or cohabiting, approximately 30% were single and 20% no longer married (Table 3).

**Association between age and prognosis independent of treatment**

Overall, there was no evidence of an association between age and prognosis at 3–4 months post-baseline. This was the case when age was modelled as a continuous variable, an ordinal variable of age groups, comparing each age group to those aged 16–29 years old (see Table 4), and with age modelled with a quadratic term ($p = 0.233$). There was also no evidence of an association between age and prognosis at 6–8 months (online Supplementary Tables 7 and 8) or 9–12 months (online Supplementary Tables 9 and 10).

**Associations between gender and prognosis independent of treatment**

There was no evidence for a difference in prognosis between men and women at 3–4 months (Tables 4 and 5) or 9–12 months (online Supplementary Tables 9 and 10) post-baseline, but there was evidence that men had worse prognoses at 6–8 months (percentage difference in depressive symptoms for men compared to women: 15.08% (95% CI: 4.82 to 26.35), see online Supplementary Table 8). However, this was attenuated when adjusting for socio-demographic factors: age, marital status and employment, 12.23% (−1.69 to 28.12).

**Associations between marital status and prognosis independent of treatment**

There was some evidence of an association between marital status and prognosis at 3–4 months (Tables 4 and 5). Participants who were single (difference in mean depressive symptom score at 3–4 months (0.22(95% CI: 0.15 to 0.30)) and those who were no longer married (0.28(95% CI: 0.20 to 0.37)), had more depressive symptoms at 3–4 months post-baseline on average, compared to married participants. After adjusting for the ‘disorder characteristics’ and socio-demographic factors there was still evidence that single or no longer married participants had worse prognoses than those that were married, but the strength of the associations was somewhat attenuated (see Tables 4 and 5). There were similar effects at 6–8 months post-baseline although confidence intervals included zero when controlling for employment status in addition to all other confounders with the log outcome (online Supplementary Tables 7 and 8). Further, when comparing just those participants who reported being single to those who reported being married, there was a lack of evidence for an association with prognosis at 6–8 months post-baseline after adjusting for depressive ‘disorder characteristics’. At 9–12 months there was no evidence for an association between marital status and prognosis after adjusting for the depressive ‘disorder characteristics’ (online Supplementary Tables 9 and 10).

**Sensitivity analyses**

Sensitivity analyses including systematically missing confounding factors did not substantively change our conclusions (see online Supplementary tables 12–18), nor did analyses that removed studies due to high levels of heterogeneity (online Supplementary Table 19).

**Discussion**

This study investigated the associations between age, gender and marital status, with the prognosis of depression for adults that sought treatment in primary care, regardless of the type of treatment they were given. We found no evidence for an association between age and prognosis at 3–4, 6–8 or 9–12 months post-baseline and no evidence that gender was associated with prognosis after adjusting for clinical prognostic factors and employment status. There was evidence that single, and either separated, divorced or widowed patients (those no longer married) had worse prognoses than married patients at 3–4 months. However, the effect was weaker after adjusting for clinical prognostic factors and employment status.

Our recent review of systematic reviews and meta-analyses found inconsistent and at times contradictory findings regarding the associations between age, gender, marital status and prognosis for depressed patients (Buckman et al., 2021). Our findings suggest that the pattern of associations found in those past reviews may have been due to chance, or the specific features of each review. With greater power to detect effects due to having a large IPD sample here, and additional robustness by adjusting for a range of treatments and for important clinical and demographic prognostic factors, we found no overall association between either age or gender and prognosis. Although we found the that men appeared to have worse prognoses than women at 6–8 months this was attributable to a single large study that contributed data at the 6–8 months endpoint and not at 3–4 or 9–12 months, and the association was minimal after adjusting for clinical and demographic covariates. There had been limited evidence for an association between marital status and prognosis in prior reviews, here we did find evidence of such an association in our data (married patients did better with treatment). On the basis of our findings here it is unlikely that knowledge of patients’ self-reported marital statuses will inform clinically meaningful differences in prognosis alone, particularly after accounting for depressive ‘disorder characteristics’.

https://doi.org/10.1017/S2045796021000342 Published online by Cambridge University Press
and employment status (Button et al., 2015). However, it may be important to consider in conjunction with those other factors. The strength of the association with prognosis likely lies somewhere between the prognostic effects found for a history of antidepressant medication and comorbid panic disorder (Buckman et al., 2021).

Strengths and limitations
To our knowledge, this is the first study to use a large individual patient dataset formed from a number of RCTs to consider the prognostic status of age, gender, and marital status across different types of treatment, bringing greater precision to estimates of these associations than in past studies. In contrast to past reviews, we selected studies that included treatment seeking adults with depression recruited in primary care settings, as this is a very common route into treatment (McManus et al., 2016; Thornicroft et al., 2017). This may have limited the number of studies found to meet inclusion criteria for this review, but has the advantage of ensuring there is a minimum population for whom our findings may be generalisable. Further, as many of

Fig. 1. Flow of studies through selection process for IPD meta-analysis.
Table 4. Difference in Z-score of depressive symptoms (‘mean difference’) at 3–4 months post-baseline per unit increase in baseline prognostic indicator

| Baseline variable | Adjusted for treatment | Additionally adjusted for depressive severity and ‘disorder characteristics’ | Additionally adjusted for age, gender, and marital status | Additionally adjusted for age, gender, marital status, and employment status |
|-------------------|------------------------|---------------------------------|---------------------------------|---------------------------------|
|                   | Z-score of depressive symptoms | Number of Studies | Heterogeneity | Z-score of depressive symptoms | Number of Studies | Heterogeneity | Z-score of depressive symptoms | Number of Studies | Heterogeneity | Z-score of depressive symptoms | Number of Studies | Heterogeneity |
| Age (per 5 year increase) | 0.01 (−0.04 to 0.01) | 8 | 73 | 0 (−0.02 to 0.02) | 8 | 62 | −0.06 (−0.16 to 0.04) | 8 | 0 | −0.06 (−0.16 to 0.04) | 8 | 0 |
| Age group | 0.03 (−0.08 to 0.02) | 8 | 74 | 0.01 (−0.03 to 0.05) | 8 | 65 | −0.04 (−0.13 to 0.05) | 8 | 0 | −0.05 (−0.14 to 0.04) | 8 | 0 |
| Age (16–29 – reference) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Age (30–39) | 0.07 (−0.17 to 0.03) | 8 | 0 | −0.03 (−0.13 to 0.07) | 8 | 13 | 0.00 (−0.11 to 0.11) | 8 | 7 | −0.01 (−0.11 to 0.09) | 8 | 0 |
| Age (40–49) | −0.00 (−0.21 to 0.21) | 8 | 78 | 0.01 (−0.13 to 0.16) | 8 | 55 | 0.04 (−0.10 to 0.19) | 8 | 47 | 0.03 (−0.12 to 0.18) | 8 | 49 |
| Age (50–59) | 0.03 (−0.15 to 0.16) | 8 | 45 | 0.07 (−0.08 to 0.21) | 8 | 49 | 0.08 (−0.06 to 0.23) | 8 | 33 | 0.06 (−0.08 to 0.19) | 8 | 25 |
| Age (60 + ) | −0.14 (−0.35 to 0.06) | 8 | 63 | 0.05 (−0.11 to 0.20) | 8 | 39 | 0.06 (−0.11 to 0.24) | 8 | 40 | 0.05 (−0.13 to 0.23) | 8 | 46 |
| Gender (women – reference) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gender (men) | 0.01 (−0.06 to 0.09) | 8 | 7 | 0.05 (−0.02 to 0.11) | 8 | 0 | 0.05 (−0.02 to 0.11) | 8 | 0 | 0.03 (−0.03 to 0.09) | 8 | 0 |
| Marital status | 0.15 (0.11 to 0.2) | 8 | 0 | 0.1 (0.06 to 0.14) | 8 | 0 | 0.1 (0.06 to 0.14) | 8 | 0 | 0.08 (0.04 to 0.12) | 8 | 0 |
| Married (reference) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Single | 0.22 (0.15 to 0.30) | 8 | 0 | 0.14 (0.07 to 0.21) | 8 | 2 | 0.15 (0.08 to 0.23) | 8 | 0 | 0.10 (0.03 to 0.18) | 8 | 0 |
| No longer married | 0.28 (0.20 to 0.37) | 8 | 0 | 0.18 (0.10 to 0.26) | 8 | 0 | 0.18 (0.10 to 0.26) | 8 | 0 | 0.13 (0.05 to 0.21) | 8 | 0 |

Note: Association for ordinal variables is per category increase from first category shown below the variable down to the last (i.e. married to single, single to no longer married). 'Disorder characteristics' adjusted for are: baseline BDI-II score, average anxiety duration, depression duration, comorbid panic disorder, and history of antidepressant treatment.
Table 5. Percentage difference (‘% difference’) in depressive symptom scale scores at 3–4 months post-baseline per unit increase in baseline prognostic indicator

| Baseline variable | Adjusted for treatment |  |  | Additionally adjusted for depressive severity and 'disorder characteristics' |  |  | Additionally adjusted for age, gender, and marital status |  |  | Additionally adjusted for age, gender, marital status, and employment status |  |  |
|------------------|------------------------|---|---|----------------------------------|---|---|----------------------------------|---|---|----------------------------------|---|---|
|                  | Percentage difference in depressive symptoms | Number of studies | Heterogeneity | Percentage difference in depressive symptoms | Number of studies | Heterogeneity | Percentage difference in depressive symptoms | Number of studies | Heterogeneity | Percentage difference in depressive symptoms | Number of Studies | Heterogeneity |
| Age (per 5 year increase) | −0.75 (−2.61 to 1.14) | 8 | 73 | 0.45 (−1.27 to 2.20) | 8 | 68 | −2.10 (−9.68 to 6.11) | 8 | 0 | −2.47 (−9.97 to 5.66) | 8 | 0 |
| Age group | −1.45 (−5.26 to 2.51) | 8 | 73 | 0.97 (−2.7 to 4.78) | 8 | 69 | −3.23 (−10.19 to 4.27) | 8 | 0 | −4.03 (−10.87 to 3.32) | 8 | 0 |
| Age (16–29 – reference) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Age (30–39) | −8.41 (−16.62 to 0.60) | 8 | 19 | −6.06 (−14.13 to 2.78) | 8 | 18 | −3.4% (−11.46 to 5.27) | 8 | 6 | −3.80 (−11.60 to 4.70) | 8 | 0 |
| Age (40–49) | −3.08 (−16.93 to 13.08) | 8 | 73 | −1.13 (−10.69 to 9.46) | 8 | 40 | 2.29 (−7.49 to 13.11) | 8 | 28 | 2.09 (−7.82 to 13.06) | 8 | 29 |
| Age (50–59) | −0.58 (−11.03 to 11.08) | 8 | 40 | 3.57 (−9.17 to 18.11) | 8 | 57 | 6.24 (−6.25 to 20.38) | 8 | 43 | 4.50 (−7.40 to 17.94) | 8 | 38 |
| Age (60+) | −8.92 (−24.30 to 9.59) | 8 | 71 | 3.47 (−11.23 to 20.62) | 8 | 58 | 6.47 (−9.63 to 25.43) | 8 | 56 | 4.04 (−12.67 to 23.94) | 8 | 60 |
| Gender (women – reference) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gender (men) | 2.9% (−2.53 to 8.63) | 8 | 0 | 4.13 (−1.1 to 9.63) | 8 | 0 | 4.33 (−1.03 to 9.99) | 8 | 4 | 2.61 (−3.02 to 8.56) | 8 | 15 |
| Marital status | 10.07 (6.61 to 13.64) | 8 | 0 | 6.82 (3.64 to 10.1) | 8 | 0 | 6.81 (3.61 to 10.12) | 8 | 0 | 4.85 (1.61 to 8.19) | 8 | 3 |
| Married (reference) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Single | 18.12 (11.34 to 25.31) | 8 | 0 | 12.02 (5.95 to 18.44) | 8 | 0 | 13.76 (7.10 to 20.83) | 8 | 0 | 9.25 (2.78 to 16.13) | 8 | 0 |
| No longer married | 19.17 (11.64 to 27.21) | 8 | 0 | 12.72 (5.92 to 19.94) | 8 | 0 | 11.97 (5.10 to 19.29) | 8 | 0 | 8.02 (1.31 to 15.18) | 8 | 0 |

Note: Association for ordinal variables is per category increase from first category shown below the variable down to the last (i.e. married to single, single to no longer married). ‘Disorder characteristics’ adjusted for are: baseline BDI-II score, average anxiety duration, depression duration, comorbid panic disorder and history of antidepressant treatment.
the randomised treatments here are commonly used in other settings, our findings may be informative for clinicians assessing depressed patients in other settings where there are multiple treatment options.

In addition, our inclusion criteria specified such that all studies used the same measure to determine diagnosis, assess baseline symptoms and the depressive ‘disorder characteristics’ confounders, minimising bias in harmonising the data across studies. This reduced the potential pool of studies that might have been included here. However, as IPD data were available for all studies that met these criteria we did not introduce a common source of selection bias that can occur when only a subset of trials provide IPD. It is noteworthy that from our preliminary searches only two other studies (Hegerl et al., 2010; Perroud et al., 2012) (totaling n = 612 patients) might potentially have met all other inclusion criteria and used other comprehensive measures of depressive and anxiety symptoms and disorders at baseline. An alternative approach might have been taken by not specifying which measure was used at baseline but instead only specifying that all of the ‘disorder characteristics’ noted to be associated with prognosis in our previous study should have been assessed. It is noteworthy though that many studies that were excluded contained no comprehensive measure of anxiety disorder symptoms or diagnoses and so most would not have assessed for panic disorder, thus would have failed to meet inclusion criteria in this way too.

We used robust methods whereby data were extracted, cleaned and checked by multiple reviewers (Buscemi et al., 2006), but only one reviewer screened the initial titles and abstracts of the studies returned in the searches which may have introduced additional bias. Although adjustments were made for a number of potential confounders, including clinical, demographic and socio-economic variables, we cannot rule out residual confounding. Further, it is possible that adjusting for baseline depressive severity may have led to underestimating the effects of marital status on prognosis, as higher levels of baseline severity might be associated with a greater likelihood of separation, divorce or of being single. In addition, using a standardised outcome is a method that has been criticised but the results using the z-score outcome were similar to those with the log outcome and the secondary and sensitivity outcomes, suggesting the use of the standardised outcome metric did not unduly affect the results. The included studies were conducted at such a time when only two categories of gender were routinely assessed and did not give participants the option to identify with non-binary gender identities. Two participants had no gender data recorded; it is not clear whether they would have chosen another gender identity were other categories available to them.

**Implications and conclusions**

Clinicians and researchers will continue to routinely assess for age and gender however the value of including information on these factors when considering prognosis is likely to be very limited. On the basis of these findings we might conclude that although age and gender are important for the incidence and prevalence of depression (Kessler et al., 2005), they appear to not be important for prognosis with treatment. Patients that are married appear to have better prognoses than patients that are single or no longer married. On its own, marital status is unlikely to inform a clinically important difference in prognosis after accounting for the severity of depressive illness, captured by the depressive ‘disorder characteristics’. However, as the magnitude of effects was similar to those of a number of other clinical prognostic factors, it may be equally important to assess for marital status in clinic and to include it in predictive models of prognosis. We were unable to investigate why married people have better prognoses. It is noteworthy that when adjusting for housing status, financial wellbeing and social support (in sensitivity analyses) the associations were weaker, so married patients might differ from those that are not married in these social domain, and these in turn may be associated with prognosis.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S2045796021000342

**Data availability.** Requests for sharing of the IPD used in this study can be made to the corresponding author, any sharing of data will be subject to obtaining appropriate agreements from the chief investigators or data custodians for each individual trial dataset used here.

**Acknowledgements.** The authors would like to thank Professor Ian R. White for his contribution to determining the methods for this study and writing the protocol paper that this study followed.

**Authors contributions.** JB, GL and SP conceived of the original project, all three along with SDH, RJD, SG, TK, EW and GA applied for and received funding to support this work. All nine of the above, RS and ZDC, wrote the initial protocol document and plan for the current study. JB and RS co-reviewed the studies with consultation from GL and SP. ZDC, NW, DK, GA and GL provided data and liaison to resolve issues and discrepancies between received datasets and publications about those studies. JB, RS, GL and SP were responsible for the screening of studies, data extraction, additional data cleaning. JB was responsible for the data analysis with support from GA, GL, CO’D, L-LA, JS and RS in particular, and consultation from all other authors. JB wrote the original manuscript with support from RS, L-LA, JS, CO’D, GA, TE, MD, SDH, SB, EL, ZDC, GL and SP. All authors contributed to consecutive drafts and approved the final manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting criteria have been omitted.

**Financial support.** This work was supported by the Wellcome Trust through a Clinical Research Fellowship to JB (201192/Z/16/Z), MQ Foundation (for ZC: MQDS16/72), the Higher Education Funding Council for England (RS, L-LA, CO’D and SP), the National Institute of Health Research (NIHR), NIHR University College London Hospitals Biomedical Research Centre (RS, L-LA and SP), University College London (GA, GL), Vanderbilt University (SDH), University of Southampton (TK), University of Exeter (EW) and University of York (SG). NIHR Biomedical Research Centre at the University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol (NW: The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care), Alzheimer’s Society (grant code: 457 (AS-PG-18-013) for JS). National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London (TE and MD: The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health).

The studies that IPD for this study were funded by:

1. **CADET:** UK Medical Research Council (MRC; reference G0701031), managed by the National Institute for Health Research (NIHR) on behalf of the MRC-NIHR partnership
2. **COBALT:** This research was funded by the National Institute for Health Research Health Technology Assessment (NIHR HTA) programme (project number 06/404/02).
3. **GENPOD:** Medical Research Council and supported by the Mental Health Research Network.
4. **IPCRESS:** BUPA Foundation
5. **ITAS:** Primary Secondary Care Interface initiative of the United Kingdom National Health Service Research and Development Programme (PSI 2-58).

https://doi.org/10.1017/S2045796021000342 Published online by Cambridge University Press
Conflict of interest. All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf. Declaration of Interest: None.

References

Bebbington P (1987) Marital status and depression: a study of English
ure form atwww.icmje.org/coi_disclosure.pdf. Declaration of Interest: None.
References

9. TREAD: National Institute for Health Research (NIHR) Health Technology
8. REEACT: UK National Institute for Health Research (NIHR) Health
7. MIR: National Institute for Health Research (NIHR) Health Technology
6. J. E. J. Buckman
5. Buscemi N, Hartling L, Vandermeer B, Tjosvold L and Klassen TP
4. Button KS, Kounali D, Thomas L, Wiles NJ, Peters TJ, Welton NJ, Ades AE
3. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi
2. White IR, Lewis G and Pilling S
1. All authors have completed the ICMJE uniform disclosur
Kroenke K, Spitzer RL and Williams JBW (2001) The PHQ-9. *Journal of General Internal Medicine* 16, 606–613.

Lewin GS, Pelosi AJ, Araya R and Dunn G (1992) Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychological Medicine* 22, 465–486.

Lewis G, Duffy L, Ades A, Amos R, Araya R, Brabyn S, Button KS, Churchill R, Derrick C, Dowrick C, Gilbody S, Fawsitt C, Hollingworth W, Jones V, Kendrick T, Kessler D, Kounali D, Khan N, Lanham P, Pervin J, Peters TJ, Rizozie D, Salaminios G, Thomas L, Welton NJ, Wiles N, Woodhouse R and Lewis G (2019) The clinical effectiveness of sertraline in primary care and the role of depression severity and duration (PANDA): a pragmatic, double-blind, placebo-controlled randomised trial. *The Lancet Psychiatry* 6, 903–914.

McManus S, Bebbington P, Jenkins R and Brugha TS (2016) Mental Health and Wellbeing in England: Adult Psychiatric Morbidity Survey 2014. Leeds, UK: NHS Digital.

Menchetti M, Bortolotti B, Rucci P, Scocco P, Bombi A and Berardi D (2010) Depression in primary care: interpersonal counseling vs selective serotonin reuptake inhibitors. The DEPICS study. A multicenter randomized controlled trial. *Rationale and design. BMC Psychiatry* 10, 1-9. Available at: https://bmcpsychiatry.biomedcentral.com/articles/10.1186/1471-244X-10-97 (Accessed 4 May 2021).

Nakabayashi T, Har A and Minami H (2018) Impact of demographic factors on the antidepressant effect: a patient-level data analysis from depression trials submitted to the pharmaceuticals and medical devices agency in Japan. *Journal of Psychiatric Research* 98, 116–123.

Noma H, Furukawa TA, Maruo K, Imai H, Shinohara K, Tanaka S, Ikeda K, Yamawaki S and Cipriani A (2019) Exploratory analyses of effect modifiers in the antidepressant treatment of major depression: individual-participant data meta-analysis of 2803 participants in seven placebo-controlled randomized trials. *Journal of Affective Disorders* 250, 419–424.

Olff M, Blanco C and Marcus SC (2016) Treatment of adult depression in the United States. *JAMA Internal Medicine* 176, 1482–1491.

Perroud N, Uher R, Ng MYM, Guipponi M, Hauser J, Henigsberg N, Maier W, Morris O, Gennarelli M, Rietschel M, Souery D, Dermouzis P, Stamp AS, Lathrop M, Farmer A, Breen G, Atchison KJ, Lewis CM, Craig IW and McGuffin P (2012) Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *The Pharmacogenomics Journal* 12, 68–77.

Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ and Koffel JB (2021) PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews* 10, 1-19. Available at: https://systematicerviewsjournal.biomedcentral.com/articles/10.1186/s13643-020-01542-z (Accessed 4 May 2021).

Richards DA, Hill JJ, Gask I, Lovell K, Chew-Graham C, Bower P, Cape J, Pilling S, Araya R, Kessler D, Bland JM, Green C, Gilbody S, Lewis G, Manning C, Hughes-Morley A and Barkham M (2013) Clinical effectiveness of collaborative care for depression in UK primary care (CADET): cluster randomised controlled trial. *British Medical Journal* 347, f4913–f4913.

Rothwell PM (2005) Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *The Lancet* 365, 176–186.

Schoemaker JH, Killan S, Emsley R and Vingerhoets AJIM (2018) Factors associated with placebo response in depression trials: a systematic review of published meta-analyses (1990–2017). *Neurology Psychiatry and Brain Research* 30, 12–21.

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA and the PRISMA-P Group (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *British Medical Journal* 349, 1–25.

Sockel LE (2018) A systematic review and meta-analysis of interpersonal psychotherapy for perinatal women. *Journal of Affective Disorders* 232, 316–328.

StataCorp (2019) Stata Statistical Software: Release 16. *Stata Statistical Software*. College Station, TX: StataCorp LLC.

Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G and Tierney JF (2015) Preferred reporting items for a systematic review and meta-analysis of individual participant data. *Journal of the American Medical Association* 313, 1657–1665.

Thomas HV, Lewis G, Watson M, Thomas HV, Lewis G, Watson, M, Bell T, Lyons I, Lloyd K, Weich S and Sharp D (2004) Computerised patient-specific guidelines for management of common mental disorders in primary care: a randomised controlled trial. *British Journal of General Practice* 54, 832–837.

Thornicroft G, Chatterji S, Evans-Lacko S, Gruber M, Sampson N, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, Andrade L, Borges G, Bruffaerts R, Bunting B, De Almeida JMC, Florescu S, De Girolamo G, Gureje O, Haro JM, He Y, Hinkov H, Karam E, Kawakami N, Lee S, Navarro-Mateu F, Piazza M, Posada-Villa J, De Galvis Y and Kessler RC (2017) Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry* 210, 119–124.

Trusheim MR, Berndt ER and Douglas FL (2007) Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nature Reviews Neuroscience* 8, 287–293.

Weitz E, Kleiboer A, Van Straten A, Hollon SD and Cuijpers P (2017) Individual patient data meta-analysis of combined treatments versus psychotherapy (with or without pill placebo), pharmacotherapy or pill placebo for adult depression: a protocol. *BMJ Open* 7, 1–8.

Wiles NJ, Mulligan J, Peters TJ, Cowen PJ, Mason V, Nutt D, Sharp D, Tallon D, Thomas L, O’Donovan MC and Lewis G (2012) Severity of depression and response to antidepressants: GENPOD randomised controlled trial. *British Journal of Psychiatry* 200, 130–136.

Wiles N, Thomas L, Abel A, Ridgway N, Turner N, Campbell J, Garland A, Hollinghurst S, Jerrom B, Kessler D, Kuyken W, Morrison J, Turner K, Williams C, Peters T and Lewis G (2013) Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the CoBaTR randomised controlled trial. *The Lancet* 381, 375–384.

https://doi.org/10.1017/S2045796021000342 Published online by Cambridge University Press