Real-world effect of antidepressants for depressive disorder in primary care: protocol of a population-based cohort study

Franco De Crescenzo [1,2] Cesar Garriga [3] Anneka Tomlinson,[1] Carol Coupland,[4] Orestis Efthimiou [5] Seena Fazel,[1] Julia Hippisley-Cox,[3] Andrea Cipriani [1,2]

ABSTRACT

Introduction Clinical guidelines recommend antidepressants as the first line of treatment for adults with moderate-to-severe depression. Randomised trials provide the best evidence on the comparative effectiveness of antidepressants for depression, but are limited by a short follow-up and a highly selected population. We aim to conduct a cohort study on a large database to assess acceptability, efficacy, safety and tolerability of antidepressant monotherapy in people with depressive disorder in primary care.

Methods and analysis This is a protocol for a cohort study using data from the QResearch primary care research database, which is the largest general practice research database in the UK. We will include patients registered for at least 1 year from 1 January 1998, diagnosed with a new episode of depression and on antidepressant and a comparison group not on antidepressant. The exposure of interest will be treatment with antidepressant medications. Our outcomes will be acceptability (treatment discontinuation due to any cause), efficacy (clinical response and remission); safety (adverse events (AEs) and all-cause mortality); and tolerability (dropouts due to any AE) measured at 2 months, 6 months and 1 year. For each outcome, we will estimate the absolute risks for all antidepressants, and relative effects between antidepressants using Cox's proportion hazards models. We will calculate HRs and 99.9% CIs for each outcome of interest.

Discussion The main limitation is the observational nature of our study, while the major strengths include the large representative population contained in QResearch and the possibly high generalisability.

INTRODUCTION

Depression is a very common major health problem; 350 million people in the world are affected.[1] There are several effective treatments for depression including drug treatment and talking therapies,[2] but clinical guidelines often recommend antidepressant medications as the first method of treatment for adults with moderate-to-severe depression.[3–5] Antidepressants are very commonly used (in England alone, 64.7m prescriptions for antidepressants were dispensed in 2016). After 2 months of treatment, an average 26.4% of patients discontinue antidepressants due to any cause, while an average of 10.4% of patients discontinue antidepressants for side-effects that cannot be tolerated.[6] This might happen because antidepressants are prescribed without a clear understanding of which drug is the most appropriate medication for each patient taking into account their characteristics, so people often stop the antidepressant early because they are prescribed a drug which might work for the 'average person' but has not been tailored to them individually.[7] Clinical trials provide the best evidence regarding average comparative efficacy of drugs, but they are usually not designed to assess AEs, especially if they are rare or less common. In addition, their external validity is limited because they focus on treatments over short periods of time (usually 8–12 weeks) in highly selected patients rather than in more complicated, real-world cases, such as patients with multiple comorbidities and concurrent long-term medication that are seen in routine practice by general practitioners. To understand the effects of antidepressants in real-world conditions, we aim to conduct a cohort study using a large and representative sample from primary care in England, to assess acceptability, efficacy, safety and tolerability of antidepressants in adults with depression.

METHODS

Setting

We will undertake the study using data from the QResearch primary care research registry (www.qresearch.org). This registry is the largest general practice research database in the UK, and it contains in total the anonymised electronic healthcare records of over 35 million patients ever registered with 7500 general practices throughout UK.[8] Consent to provide data for QResearch was sought from all UK practices using the Egton Medical Information Systems (EMIS) medical records system. EMIS is the major supplier of primary care computer systems in the UK and is in use in two-thirds of all UK general practices.
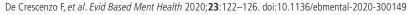
The information recorded on the QResearch database includes patient demographic data (year of birth, gender, socioeconomic data), characteristics (height, weight, smoking status), symptoms, clinical diagnoses, consultations, referrals, prescribed medications and results of investigations. The latest version of the QResearch database, which is updated quarterly, will be used for the analysis.

Detailed analyses have compared QResearch practices with all UK practices and found that practices contributing to QResearch are somewhat larger than UK practices overall but are very similar

in other respects (eg, age–sex distribution and prevalence of chronic diseases).[8] Within QResearch, we will use data coming from English general practices.

Population

We will initially identify an open cohort of patients aged 18–100 years at the study entry date drawn from patients registered with eligible practices since 1 January 1998. Patients will be included if they have been registered with the practice for at least 12 months. Patients will be followed up for 12 months.

We will use Read codes to identify patients with a diagnosis of depression, using as a starting point, case definitions that have been used in previous studies (see online supplementary file 1).[9 10] We will exclude:

► Patients with a recorded diagnosis of bipolar disorder or schizophrenia spectrum disorder made at any point during their lifetime.
► Patients with a diagnosis of post partum depression made within 180 days before or up to 180 days after the first diagnosis of depression.
► Patients prescribed more than one antidepressant at baseline.
► Patients prescribed antipsychotics or mood stabilisers.
► Patients if they had received prescriptions for an antidepressant or they had a diagnosis of depression before their entry date in the cohort.
► Patients with a diagnosis of depression made 2 months before or 2 months after starting an antidepressant.

Our final cohort for analysis will include exposed groups of patients on each antidepressant monotherapy, and an unexposed group of patients not on antidepressants.

Drug exposure

The primary exposure of interest is the use of licensed antidepressants, which will be considered individually and also grouped according to the four main classes in the British National Formulary (BNF; https://bnf.nice.org.uk). The groups will consist of tricyclics (TCAs; ie, amitriptyline, amoxapine, butriptyline, clomipramine, desipramine, dosulepin, doxepin, imipramine, iprindole, lofepramine, maprotiline, mianserin, nortriptyline, protriptyline, trimipramine, viloxazine), selective serotonin reuptake inhibitors (SSRIs; ie, citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline), monoamine oxidase inhibitors (MAOIs; ie, isocarboxazid, moclobemide, phenelzine, tranylcypromine) and other antidepressants (ie, agomelatine, duloxetine, mirtazapine, nefazodone, reboxetine, tryptophan, trazodone, venlafaxine, vortioxetine).[11]

Information will be extracted from all prescriptions for antidepressants issued during 12-month follow-up. We will calculate the duration of each prescription in days by dividing the number of tablets prescribed by the number of tablets to be taken each day. If the information on tablets per day is missing or not sufficiently detailed (expected to be <5% of total prescriptions), we will estimate the duration of the prescription based on the number of tablets prescribed, as in previous studies.[10] Patients will be classified as continually exposed to an antidepressant during periods where there were no gaps of more than 30 days between the end of one prescription and the start of the next (most antidepressants at the beginning of a treatment are prescribed for not more than 28–30 days). Patients will be also classified as exposed for the first 30 days after the estimated date of stopping an antidepressant in order to account for any delays in starting the prescription or accumulation of tablets as well as to attribute the outcomes occurring during withdrawal periods to the antidepressant, as done in previous studies.[10] The daily dose of each prescription will be calculated by multiplying the number of tablets to be taken each day by the dose of each tablet, and then converted to a defined daily dose using values assigned by the WHO's Collaborating Centre for Drug Statistics Methodology (www.whocc.no/atc_ddd_index).

Outcomes

We will measure the following outcomes up to 2 months, 6 months and 1 year from the initial prescription of antidepressants or from the date of the first episode of depression for patients not on antidepressants. We will consider extending the follow-up to 2 years if data are available.

Primary outcomes

1. Acceptability of treatment will be measured as all-cause treatment dropout rate.
   Treatment dropout will be computed if:
   – A patient has a gap of more than 30 days between the end of a prescription of an antidepressant and the start of the next prescription, or
   – a patient switches to another antidepressant, or
   – a patient is prescribed an additional antidepressant, a mood stabiliser, or an antipsychotic (augmentation).
2. Tolerability of treatment, measured as treatment dropout (measured as described above) after any AE occurred (ie, no subsequent prescriptions or implementation of a switch/combination/augmentation strategy after the occurring of an AE). Treatment dropout would have to occur within 30 days of an AE to be assumed to be a consequence of the AE. The list of AEs is described below.

Secondary outcomes

3. Safety, measured as:
   – number of patients with at least one AE among the ones specified below. We will use the Read codes for AEs that have been shown in randomised trials as the frequent and severe AEs of antidepressant. These include: nausea, headache, dry mouth, insomnia, dizziness, sedation/somnolence, diarrhoea, constipation, sexual dysfunction, fatigue, rhinitis/nasopharyngitis, hyperhidrosis, respiratory disorder (infection, cough), anxiety, decreased appetite, increased appetite, tremor, pain, vomiting, abdominal pain/discomfort, dyspepsia, agitation, visual impairment, ejaculation disorder/erectile dysfunction, weight increased, weight decreased, arrhythmia/heart rate disorder, abnormal dreams, infection, blood pressure increased, blood pressure decreased, extrapyramidal disorders, suicidal ideation, suicide behaviour or self-harm, hot flush, dysuria, skin disorder, flatulence, urinary disorders, injury, yawning, eye disorders, paraesthesia, nervous system symptoms, feeling cold, menstrual disorder, chest pain, disturbance in attention, libido increased, psychiatric symptoms, fall, confusional state, salivary hypersecretion, accidental overdose, cardiovascular symptoms (eg, angina), sleep disturbance, oedema, aggression, completed suicide, affect lability, fever, euphoric mood, hypersomnia, memory impairment, muscular skeletal problems, serotonin syndrome, withdrawal syndrome, fractures, upper gastrointestinal bleeding, bleeding at any site, epilepsy/seizures.

The classification of the above-mentioned side effects is performed based on the use of preferred terms from

MedDRA (https://www.meddra.org/) to categorise each AE into categories. MedDRA has been developed by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use to provide a single standardised international medical terminology which can be used for regulatory communication and evaluation of data pertaining to medicinal products for human use. As a result, MedDRA is designed for use in the registration, documentation and safety monitoring of medicinal products through all phases of the development cycle (ie, from clinical trials to postmarketing surveillance). The side effects proposed in our protocol are similar to those considered in the study by .Tomlinson *et al*[12] as we are focusing on clinically relevant outcomes which have been evaluated also in randomised controlled trials (RCTs).

– All-cause mortality. Patients who have died will be identified using death data recorded on the patients' general practice record as in previous studies.[10 13]

4. Efficacy, measured at the last observation before 2, 6 and 12 months, as:

– Clinical remission, measured as scoring less than a pre-specified threshold on a standardised rating scale,[6] such as the Hamilton Depression Rating Scale (HDRS), the Montgomery and Asberg Depression Rating Scale, the Beck Depression Inventory, the Patient Health Questionnaire, the Geriatric Depression Scale (eg, a score <7 for the HDRS would be considered as remission).

– Clinical response to antidepressant treatment, measured as 50% reduction on a standardised rating scale, compared with the most recent value recorded in the 6 months before the first antidepressant prescription.

Confounder variables

Confounders will include baseline variables that we consider to be potential risk factors for the outcomes and are also potentially associated with the likelihood of receiving a particular antidepressant treatment, based on previous studies of antidepressants.[9] These include:

▶ age at study entry, in years;
▶ sex;
▶ Body mass index;
▶ year of diagnosis of depression;
▶ type of diagnosis (major depressive disorders, minor depression, other);
▶ severity of index diagnosis of depression (categorised as mild, moderate or severe, using the Read codes published by Martinez *et al*[14];
▶ deprivation (Townsend deprivation score corresponding to the patients postcode, in fifths)[15];
▶ smoking status (non-smoker, ex-smoker, light smoker: 1–9 cigarettes/day, moderate smoker: 10–19 cigarettes/day, heavy smoker: ≥20 cigarettes/day, not recorded);
▶ alcohol intake (none, trivial: <1 unit/day, light: 1–2 units/day, medium: 3–6 units/day, heavy: 7–9 units/day, very heavy: >9 units/day, not recorded);
▶ ethnic group (categorised as either white/not recorded or non-white (Indian, Pakistani, Bangladeshi, other Asian, black African, black Caribbean, Chinese, other including mixed));
▶ comorbidities at baseline (binary variables for each of coronary heart disease, stroke/transient ischaemic attack, diabetes, hypertension, cancer, epilepsy/seizures, hypothyroidism, osteoarthritis, rheumatoid arthritis, suicidal ideation/behaviour or self-harm, obesity, asthma/chronic obstructive airways disease, osteoporosis, liver disease, renal disease, obsessive–compulsive disorder); and
▶ use of other drugs at baseline (binary variables for each of anticonvulsants, hypnotics/anxiolytics, antihypertensive drugs, aspirin, statins, anticoagulants, non-steroidal anti-inflammatory drugs, bisphosphonates, oral contraceptives, hormone replacement therapy).

Handling of missing data

Excluding subjects with missing values would reduce the sample study size, thus decreasing precision and power.[16] We will use multiple imputation by chained equations to impute values when actual values are not available. We will first analyse the pattern of distribution of missing values in order to assess if the data are suitable for multiple imputation using chained equations. For each imputation, we will generate 10 imputed datasets and we will combine coefficient estimates across these using Rubin's rules.[17] We will include all of the confounding variables in the multiple imputation process, along with the outcome variable as it carries information about predictors' missing values.

Statistical analysis

We will use Stata MP V.16.0 to conduct statistical analysis. We will follow the 'REporting of studies Conducted using Observational Routinely-collected health Data' Statement.[18]

Our primary analysis will be time to event, while self-controlled case-series and propensity score will be secondary analysis.

Descriptive statistics

We will split the population into antidepressant users and non-users.

In the study cohort of patients with a diagnosis of depression, we will summarise baseline characteristics with descriptive statistics and will describe patterns of antidepressant use according to type of antidepressant prescribed, duration of use and dose. We will describe the severity of depression (classified as mild, moderate or severe) in the study cohort, overall and by age and gender. We will describe patterns of antidepressant use according to severity of depression.

Time to event analysis

We will determine absolute and relative risks according to type of antidepressant prescribed for dropouts due to any cause (acceptability), dropouts due to any AE (tolerability), response and remission (efficacy), number of patients experiencing at least one AE and all-cause mortality (safety).

Incidence rates of all the outcomes above described will be calculated in the study cohort of people with a new diagnosis of depression. The statistical analysis will comprise a series of survival analyses to assess the relationship between exposure to each antidepressant monotherapy (and antidepressants grouped according to the four main classes in the BNF) and acceptability, efficacy, safety and tolerability. Dose of antidepressant drugs will be examined in the analyses. The date of entry into the survival analyses will be the date of starting an antidepressant for patients on antidepressants and the date of the first episode of depression for patients not on antidepressants. The right censor date will be the earliest of the following: date of dropout, date of switch to another antidepressant, date of a new antidepressant added, date of a mood stabiliser added, date of an antipsychotic added, date of the outcome of interest, date of death, date of leaving the practice or the study end date. Death will be a date of censorship

for all of the AEs, except for the analysis of all-cause mortality where date of death is the outcome date.

We will explore non-linear relationships between continuous variables and the outcome using fractional polynomials.[19]

Cox's proportion hazards models accounting for practice clustering will be used for all time-to-event analyses. The analyses will calculate HRs and 99.9% CIs comparing:

► The risk of acceptability, efficacy, safety and tolerability for each antidepressant will be directly compared with each other.

► The risk of efficacy and safety in patients on any type of antidepressant compared with patients with no antidepressant treatment.

► Each separate class of antidepressants (SSRIs, TCAs, MAOIs and other antidepressants) will be directly compared with each other for acceptability, efficacy, safety and tolerability and compared with no treatment for efficacy and safety.

► Analyses will also calculate HRs according to prescribed dose of antidepressant; where numbers are sufficient, individual antidepressants within each class will be examined.

► Analyses of interaction will be carried out to examine the extent to which patient's characteristics (age, gender), use of other medications and comorbidities modify the relationship between antidepressant use and outcomes.

Adjustment will be made for potential confounders as listed above. The assumptions of the Cox proportional hazards model will be checked. In case these assumptions are implausible, we will instead employ alternative parametric survival models.

For acceptability (discontinuation due to any cause), according to results from RCTs, a mean of 26.4% of patients discontinued the treatment for any cause after the first 2 months of treatment.[6] Considering an alpha of 0.01 and a beta of 0.2, to detect a HR of 2, 1.5 and 1.2, we require a sample size of at least 369, 1077 and 5324 patients per treatment, respectively.

Self-controlled case-series analysis

The self-controlled case-series analysis only uses the patients in the cohort who have the outcomes of interest.[20] Cases with each type of AE will be identified; these will be cases with a diagnosis of the AE since 1 January 1998, who had a new diagnosis of depression (see above for details). Information on prescriptions for antidepressants in these cases will be extracted and the analysis (such as conditional fixed-effects Poisson regression model) will compare rates of the outcomes of interest in periods following a first prescription for an antidepressant compared with a baseline period free of antidepressant treatment for the same patient.[21] We will account for multiple periods of exposure in the analysis, defining a period of antidepressant treatment as one without gaps of more than 30 days between the end of a prescription and the start of the next prescription. A prescription after more than 30 days will count as a new treatment episode. The 28 days before the first prescription in each treatment period will be considered separately, as occurrence of the outcome of interest in this period could affect the probability of an antidepressant prescription. All other time periods outside these specified risk periods will contribute to the baseline person time, that is, the unexposed periods.

Propensity score

A propensity score for antidepressant use will be calculated using a logistic regression model,[22 23] which will initially include all confounders above mentioned. We will be using the whole sample for the analysis. Antidepressant users will be matched to a common comparator (eg, fluoxetine users) and to non-antidepressant users on the logit of the propensity score using a calliper width of 0.2 of the SD of the logit of the propensity score. A matching ratio of 1:1 will be used.[24] We will assess standardised differences between the two groups with 10% or more considered as suggestive of imbalance.[25] We will then perform a multivariable cox regression to estimate the effect of antidepressant on outcome occurrence.

DISCUSSION

This protocol describes a large cohort study which aims to assess the acceptability, efficacy, safety and tolerability of antidepressant monotherapy in adults with depressive disorder in primary care.

This study is part of the 'Personalise antidepressant treatment for unipolar depression combining individual choices, risks and big data (PETRUSHKA)' project. PETRUSHKA seeks to ultimately develop and test a precision medicine approach to the pharmacological treatment of major depressive disorder by synthesising data coming from RCTs and data coming from observational studies and patient registries.[12]

This study builds on a solid foundation of research performed on QResearch on the safety of antidepressants use in people aged 20–64 years,[10] and in older people[9] and contributes to the field by focusing on clinically relevant outcomes which have been evaluated also in RCTs.

At present, the best evidence in terms of comparative efficacy and acceptability of antidepressants comes from the 'Group of Researchers Investigating Specific Efficacy of individuaL Drugs for Acute depression (GRISELDA)' study, which was a network meta-analysis including only RCTs.[6] Indeed, RCTs are the most reliable source of information on relative treatment effects. However, RCTs have strictly experimental settings and employ inclusion criteria which might limit their ability to predict results in real-world clinical practice.[26] Large observational studies may complement evidence provided by RCTs and potentially address some of their limitations. Recently, Efthimiou et al[27] developed statistical methods for combining randomised and non-randomised evidence in a network meta-analysis.[27]

Our study aims to investigate the effectiveness of antidepressant monotherapy in a real-world setting.[28] The study design is a cohort study and thus observational in nature.[29] Compared with the GRISELDA study, the outcomes will include an in-depth evaluation of side effects, with a longer follow-up (1 year). Moreover, in our cohort study, we will include a representative population of people with depression in England, also including people with multiple comorbidities, who would normally be excluded in an RCT. This is because people with multiple comorbidities represent the added value of a real-world study, whose results aim to be generalised to a wide population.

The major strengths of our study include the large population contained in QResearch and the potentially high generalisability (external validity). Our study suffers from a number of limitations. First, the main limitation is the weaker internal validity of our study when compared with RCTs due to potential indication bias and residual confounding. A recent umbrella review showed that most of available evidence from observational studies on antidepressants is actually flawed by confounding by indication.[30] Although we used a large number of confounders and propensity score matching to control confounding, we are aware that RCTs are generally considered to be at a lower risk of bias compared with observational studies, when aiming to estimate causal effects of interventions.

Second, some information before date of entry in QResearch database may have not been recorded. For example, we will exclude patients if they had received prescriptions for an antidepressant or they had a diagnosis of depression before their entry date in the cohort, but this may have not been recorded in all cases.

Third, we will be not adjusting antidepressants dose for concomitant drugs which may increase or decrease antidepressants' metabolism (eg, carbamazepine).

Fourth, time to event analyses in patients on any type of antidepressant compared with patients with no antidepressant treatment may suffer from survival bias as the start of follow-up for exposed individuals is the date of antidepressant prescription, while for unexposed individuals it is the date of depression diagnosis. However, this will not affect comparisons between antidepressants which is the main focus of our study.

In the future, we aim to predict the effects of different antidepressants given information on patient-level characteristics, by using both statistical and machine learning tools (such as artificial neural networks and support vector machines). The present study will identify a range of possible prognostic factors and effect modifiers which will be used to inform the predictive model. The predictive model will then be used in PETRUSHKA to develop a web-based treatment algorithm to help clinicians, patients and carers to personalise the choice of antidepressant in primary care.[12]

ORCID iDs
Franco De Crescenzo http://orcid.org/0000-0002-2478-7763
Cesar Garriga http://orcid.org/0000-0001-7073-3611
Orestis Efthimiou http://orcid.org/0000-0002-0955-7572
Andrea Cipriani http://orcid.org/0000-0001-5179-8321

REFERENCES

1 James SL, Abate D, Abate KH, *et al*. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1789–858.

2 Cuijpers P, Noma H, Karyotaki E, *et al*. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry* 2020;19:92–107.

3 Kennedy SH, Lam RW, McIntyre RS, *et al*. Canadian network for mood and anxiety treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 3. pharmacological treatments. *Can J Psychiatry* 2016;61:540–60.

4 Malhi GS, Bassett D, Boyce P, *et al*. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Aust N Z J Psychiatry* 2015;49:1087–206.

5 National Institute for Health and Care Excellence. Depression in adults: recognition and management. London, UK: NICE, 2009. Available: https://www.nice.org.uk/guidance/cg90 [Accessed 2 Feb 2020].

6 Cipriani A, Furukawa TA, Salanti G, *et al*. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;391:1357–66.

7 Kernot C, Tomlinson A, Chevance A, *et al*. One step closer to personalised prescribing of antidepressants: using real-world data together with patients and clinicians' preferences. *Evid Based Ment Health* 2019;22:91–2.

8 Hippisley-Cox J, Vinogradova Y, Coupland C, *et al*. *Comparison of key practice characteristics between general practices in England and Wales and general practices in the QResearch database*. Nottingham, UK: University of Nottingham, 2005.

9 Coupland C, Dhiman P, Morriss R, *et al*. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. *BMJ* 2011;343:d4551.

10 Coupland C, Hill T, Morriss R, *et al*. Antidepressant use and risk of adverse outcomes in people aged 20-64 years: cohort study using a primary care database. *BMC Med* 2018;16:36.

11 Faria R, Hernandez Alava M, Manca A, *et al*. The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. NICE dsu technical support document 17, 2015. Available: http://www.nicedsu.org.uk/Observational-data-TSD(2973296).htm [Accessed 2 Feb 2020].

12 Tomlinson A, Furukawa TA, Efthimiou O, *et al*. Personalise antidepressant treatment for unipolar depression combining individual choices, risks and big data (PETRUSHKA): rationale and protocol. *Evid Based Ment Health* 2020;23:52–6.

13 Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. *BMJ* 2013;346:f2573.

14 Martinez C, Rietbrock S, Wise L, *et al*. Antidepressant treatment and the risk of fatal and non-fatal self harm in first episode depression: nested case-control study. *BMJ* 2005;330:389.

15 Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality and the North*. London, UK: Routledge, 1998.

16 Sterne JAC, White IR, Carlin JB, *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.

17 Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996;91:473–89.

18 Benchimol EI, Smeeth L, Guttmann A, *et al*. The reporting of studies conducted using observational Routinely-collected health data (record) statement. *PLoS Med* 2015;12:e1001885.

19 Harrell FE. *Regression modeling strategies with applications to linear models, logistic regression, and survival analysis*. New York, NY: Springer, 2001.

20 Whitaker HJ, Farrington CP, Spiessens B, *et al*. Tutorial in biostatistics: the self-controlled case series method. *Stat Med* 2006;25:1768–97.

21 Nagra NS, Robinson DE, Douglas I, *et al*. Antibiotic treatment and flares of rheumatoid arthritis: a self-controlled case series study analysis using CPRD gold. *Sci Rep* 2019;9:8941.

22 Austin PC. A tutorial and case study in propensity score analysis: an application to estimating the effect of in-hospital smoking cessation counseling on mortality. *Multivariate Behav Res* 2011;46:119–51.

23 Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Stat Med* 2014;33:1242–58.

24 Austin PC. Comparing paired vs non-paired statistical methods of analyses when making inferences about absolute risk reductions in propensity-score matched samples. *Stat Med* 2011;30:1292–1.

25 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083–107.

26 Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365:82–93.

27 Efthimiou O, Mavridis D, Debray TPA, *et al*. Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med* 2017;36:1210–26.

28 Cipriani A, Geddes JR, Furukawa TA, *et al*. Metareview on short-term effectiveness and safety of antidepressants for depression: an evidence-based approach to inform clinical practice. *Can J Psychiatry* 2007;52:553–62.

29 Barbui C, Cipriani A, Brambilla P, *et al*. "Wish bias" in antidepressant drug trials? *J Clin Psychopharmacol* 2004;24:126–30.

30 Dragioti E, Solmi M, Favaro A, *et al*. Association of antidepressant use with adverse health outcomes: a systematic umbrella review. *JAMA Psychiatry* 2019. doi:10.1001/jamapsychiatry.2019.2859