BJPsych | The British Journal of Psychiatry (2015) 206, 360–370. doi: 10.1192/bjp.bp.114.154377

Review article

Quetiapine immediate release *v.* placebo for schizophrenia: systematic review, meta-analysis and reappraisal

Paul Hutton, Peter J. Taylor, Lee Mulligan, Sarah Tully and Joanna Moncrieff

Background
Immediate-release (IR) quetiapine has been used to treat schizophrenia since 1997, although all the principal placebo-controlled trials have >50% missing outcome data. New studies with relatively lower rates of participant withdrawal have since been published.

Aims
To assess the efficacy and adverse effects of quetiapine IR for schizophrenia, with consideration of outcome quality and clinical meaningfulness of results, and to examine the potential impact of missing data on the main efficacy findings.

Method
We conducted a systematic review and meta-analysis of randomised controlled trials comparing quetiapine IR and placebo (or subtherapeutic dose in relapse prevention trials) for the treatment of schizophrenia (PROSPERO registration CRD4201100165). Primary outcomes were change in overall symptoms and response rates. We also examined whether high rates of participant withdrawal (⩾50%) attenuated effect sizes, and assessed the impact of making different assumptions about these people's outcomes.

Results
We identified 15 relevant trials (including 2 unpublished), providing the first 12-week data for this drug and the first data on self-reported quality of life. We found quetiapine IR to have a weighted mean difference (WMD) of 6.5 points (95% CI −8.9 to −4) on Positive and Negative Syndrome Scale (PANSS) total scores, which corresponds to a standardised mean difference (SMD) of −0.33 (95% CI −0.46 to −0.21). Longer trials reported larger mean differences favouring quetiapine IR, but the overall estimate was smaller if more conservative assumptions about the outcomes of people who left the trial early were made. Approximately 21 people needed to take quetiapine IR for 1 person to experience at least a 50% improvement in PANSS score. No difference in quality of life was observed (two RCTs), although small to moderate improvements in social functioning were found (three RCTs). Quetiapine IR caused sedation and increased rates of clinically significant weight gain, but no extrapyramidal effects were observed.

Conclusions
Quetiapine IR has a small beneficial effect on overall psychotic symptoms over 2–12 weeks, but also leads to weight gain and sedation.

Declaration of interest
None.

Quetiapine is a widely prescribed antipsychotic.[1] First manufactured by AstraZeneca, it was initially introduced for the treatment of schizophrenia and non-affective psychosis in the late 1990s, but is now licensed in several countries for the treatment of bipolar disorder and other conditions. The original immediate-release (IR) version was the third most frequently prescribed antipsychotic in the UK from 2004 to 2007, and by 2008 had been taken by over 25 million people worldwide.[2] The patent for quetiapine IR expired in March 2012 and the generic version is now comparable in cost to haloperidol, leading to considerable cost savings. Although older reviews of comparative and placebo-controlled trials concluded it was an effective treatment for schizophrenia and non-affective psychosis,[3,4] these were based on a limited number of trials which suffered from severe attrition. The Cochrane review, for example, noted that most of the original placebo-controlled studies were severely compromised by missing outcome data, and that their results were therefore 'impossible to interpret with confidence'.[3] Three of the four included studies were missing more than half of their 6-week outcome data, and the remaining study had only 12 participants. In each case the outcomes of those leaving early were estimated by the method of carrying their last available observation forward, an imputation strategy now regarded as unreliable.[5] In their 2009 review Leucht *et al* found that, overall, second-generation antipsychotics had a moderate effect on symptoms (Hedges' $g = 0.51$).[4] However, they suggested that the high withdrawal rate in these studies might have attenuated the drug–placebo difference. Indeed, in most placebo-controlled trials more than a quarter of participants leave the study, and in a significant number of trials more than half do so.[6] Such high rates of missing data cannot be safely ignored. A recent survey by the Cochrane Schizophrenia Group found consultant psychiatrists, patients, carers and Cochrane researchers in agreement that trials with over 25% missing data lack credibility,[7] and there is largely a consensus that no statistical approach can produce reliable results when assumptions about the outcomes of participants carry more weight than actual observations.[7,8] Understanding the impact of missing data is particularly challenging if it is missing for non-random reasons that are related to outcome, as may be the case for antipsychotic trials.[9] The development of a sustained release version of quetiapine (quetiapine XR) has led to new randomised controlled trials (RCTs) comparing the immediate release version with placebo, some of which had relatively low rates of missing data. Owing to the uncertainty introduced by high attrition in the older studies of quetiapine IR, we set out to perform a new systematic review and meta-analysis.

We had two main objectives. The first was to provide a comprehensive assessment of the efficacy and adverse effects of quetiapine IR for schizophrenia when compared with placebo, with consideration of both outcome quality and the clinical meaningfulness of the results, as informed by recent advances in our understanding of what constitutes a minimum clinically important difference in Positive and Negative Syndrome Scale

(PANSS) total scores.[10–12] Our second objective was to examine the potential impact of missing data on the primary outcomes. More specifically, we examined whether trials with high rates of missing data had smaller effect sizes,[4] and we used a recently published approach to examine the impact on our efficacy estimates of changing assumptions about the likely outcomes of the large numbers of people who leave these trials early.[13]

Method

Our search strategy and protocol detailing our inclusion and exclusion criteria are provided in an online data supplement. Two researchers independently searched publication databases, clinical trial registries and previous reviews for randomised controlled trials in which participants with a diagnosis of schizophrenia or early psychosis were randomly allocated to receive double-blind treatment with either placebo or quetiapine IR. No pre-specified limits were placed on study duration.

Data extraction and outcomes

Two reviewers independently extracted data from each study using data extraction forms. We attempted to trace missing summary data by contacting first authors or the study sponsor. Our primary outcomes were the average reduction at study end-point in total score on the PANSS or, if that were not available, on the Brief Psychiatric Rating Scale (BPRS), and the numbers of people achieving an important clinical response. We defined the latter as a greater than 50% reduction in PANSS or BPRS score.[14] When these were not reported or provided, we imputed them from means and standard deviations using the validated method of Furukawa *et al*.[15] (See 'Changes from protocol' section in the online supplement.) Our secondary efficacy outcomes included relapse, positive symptoms, negative symptoms, depression, quality of life and need for additional antipsychotic medication or sedatives. We also examined the numbers of participants leaving the study early for any reason, need for hospital care and functioning. For adverse effects, we looked at use of anti-Parkinsonian medication, extrapyramidal side-effects, withdrawal due to adverse events, sedation, total number of drug-attributable adverse events, insomnia, weight gain and weight loss.

We used a strict intention-to-treat (ITT) analysis for dichotomous outcomes, using the total numbers randomised to each group as the denominator in each case. Where possible, we assumed those leaving early or otherwise unaccounted for had an unchanged outcome from randomisation, but carried out sensitivity analyses to test this. Data incorporating last observation carried forward (LOCF) assumptions were used only when there was no alternative. We also wished to use a strict ITT analysis for continuous data, but expected to be limited to summary data derived from smaller samples excluding participants leaving the study early or those without at least one post-baseline assessment. For all outcomes we intended to use summary data based on the mixed-model repeated measures (MMRM) imputation method, followed by LOCF or observed case data if not available. Missing standard deviations were, where possible, calculated from *t*-test values, *P*-values, standard errors or confidence intervals.[16] If no variance parameters were reported for a particular study, we imputed standard deviations using the medians of the other studies. Similarly to previous studies,[4,17] we planned to use data from study arms where participants received an optimal drug dose of more than 250 mg. However, we carried out a sensitivity analysis excluding doses of more than 400 mg, as per the recent International Consensus on Antipsychotic Dosing and recent Leucht group analysis.[18,19]

Meta-analytic calculations

For continuous data we calculated the Hedges' *g* standardised mean difference (SMD) using Comprehensive Meta-Analysis version 2 for Windows 7. For the primary analysis of the 2–12 week study end-point data we converted BPRS scores (mean and s.d.) to PANSS scores using recently published conversion charts (PANSS total score = 1.538 × BPRS total score),[20] thus allowing us to present also the unstandardised weighted mean difference (WMD) in PANSS total scores for all the studies combined. When a trial had two or more relevant arms we combined the data following procedures in the Cochrane *Handbook*.[16] For binary data we calculated the relative risk (RR) of the unfavourable outcome, together with 95% confidence intervals, as well as the absolute risk difference and numbers needed to treat (NNT) or harm (NNH). If a trial had eligible binary data from two or more active treatment arms, we combined these into one. We used a random effects analysis for all outcomes. For the primary outcomes we also performed a sensitivity analysis using fixed effects, but not if heterogeneity was moderate or more, defined as an I^2 statistic of $\geqslant 40\%$.[16]

Impact of missing data

We tested the hypothesis that trials with severe rates of missing data ($\geqslant 50\%$ at end-point) had smaller drug–placebo differences on our primary outcomes than trials with less severe rates ($<50\%$). The 50% cut-off was chosen because it marks the point at which estimated data carry more weight than actual observations, and because the National Institute for Health and Care Excellence (NICE), the Cochrane Schizophrenia Group and others often exclude trials with this degree of missing data from their reviews.[6,21,22] We also wished to compare studies with $<25\%$ and $\geqslant 25\%$ attrition at end-point,[7] but were unable to do so because no study of 6–12 weeks' duration had less than 25% attrition. When observed case data were available we were also able to examine the impact of missing data on the primary outcome by imputing values for those who left the trial early using new guidelines provided by Ebrahim *et al*.[13] Their method involves testing whether the overall treatment effect is robust under four increasingly more conservative strategies – two of which we applied here. Strategy 1 is non-extreme and involves replacing missing data in both arms of each trial with the observed case mean of the control arm. Strategy 2 is more conservative yet plausible, and uses the highest observed control arm mean to replace missing control arm data, and the lowest observed intervention arm mean to replace missing intervention arm data. For both approaches we imputed the missing data treatment and placebo standard deviations with the medians of the control arms of all the included trials, as recommended.[13]

Analysis of clinical significance

The minimum clinically important difference (MCID) has been defined by Jaeschke *et al* as 'the smallest difference in a score in the domain of interest which patients [or providers] perceive as beneficial and which would mandate in the absence of troublesome side-effects and excessive cost, a meaningful change in the patient's management'.[11,23] An analysis of data from 14 antipsychotic trials ($n = 5970$) found a rater-determined MCID on the PANSS of roughly 15 points,[10] a criterion that has since been replicated by separate analyses of two large non-industry effectiveness trials ($n = 1650$).[11,12] Data from a large naturalistic study ($n = 398$) suggested a lower criterion of 10 points,[24] which is similar to the patient-rated MCID of 11 points derived from the Clinical Antipsychotic Trials of Intervention Effectiveness

(CATIE).[12] We tested the validity of these definitions by comparing them with the median of mean changes that the included trials were designed to detect. We assumed that trial sponsors had provided enough resources to detect with adequate power what they regarded *a priori* as the smallest difference between the groups that was important to detect.[25]

Risk of bias and study quality

Two raters independently assessed both study-level risk of bias with the Cochrane Collaboration risk of bias tool,[16] and outcome quality using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.[26] Further details on method and ratings are provided in the online supplement. We tested for publication bias using funnel plots for the PANSS/BPRS total score effect sizes (Hedges' *g*) of all studies. Ratings of bias and quality were used to inform interpretation of reliability and magnitude of effects.

Registration of review protocol and subsequent changes

The review protocol was registered in advance with the International Prospective Register of Systematic Reviews (PROSPERO), protocol CRD4201100165. Subsequent changes, in addition to those outlined above, are detailed in the online supplement. We abandoned the use of the response rate hierarchy used by Leucht *et al*,[4] given their recently expressed concerns that response rate estimates are particularly vulnerable to selective reporting bias,[19] and used only the top of the hierarchy instead (50% or more reduction in PANSS/BPRS score). This criterion is now recommended for use in studies of acutely ill patients with non-refractory illness,[27] and we used the method of Furukawa *et al* to impute this data when not reported or provided.[15] This method has been recently validated using individual patient-level data from 16 antipsychotic trials,[28] and has the additional advantage of allowing for the use of adjusted PANSS and BPRS total scores when calculating percentage change, thus avoiding underestimation of response.[27] Additional changes included using meta-regression to assess the association between study duration (measured in weeks) and year of publication on total symptoms and clinically significant improvement. These were conducted in Stata version 9 using the *Metareg* command and Knapp–Hartung variance estimator.[29]

Results

The process of selecting studies is detailed in Fig. 1. We identified 15 relevant trials, 11 of which assessed short-term efficacy ($n = 2259$). Lundbeck provided us with summary reports for two unpublished 12-week placebo-controlled studies,[30,31] both of which were terminated early owing to the inefficacy of the investigational drug (bifeprunox; quetiapine IR was an active comparator in these trials). AstraZeneca, the makers of quetiapine, provided us with a considerable amount of additional unpublished data in relation to many of their trials. They decided not to provide us with the report for one unpublished long-term trial comparing therapeutic and subtherapeutic doses of quetiapine,[32] arguing that the lack of a placebo control meant it did not meet our original inclusion criteria. However, we managed to acquire an extract detailing the main results, and other published summaries allowed us partly to assess risk of bias. We therefore included data from a total of 15 studies. An overview of included studies is provided in Table 1, and excluded studies

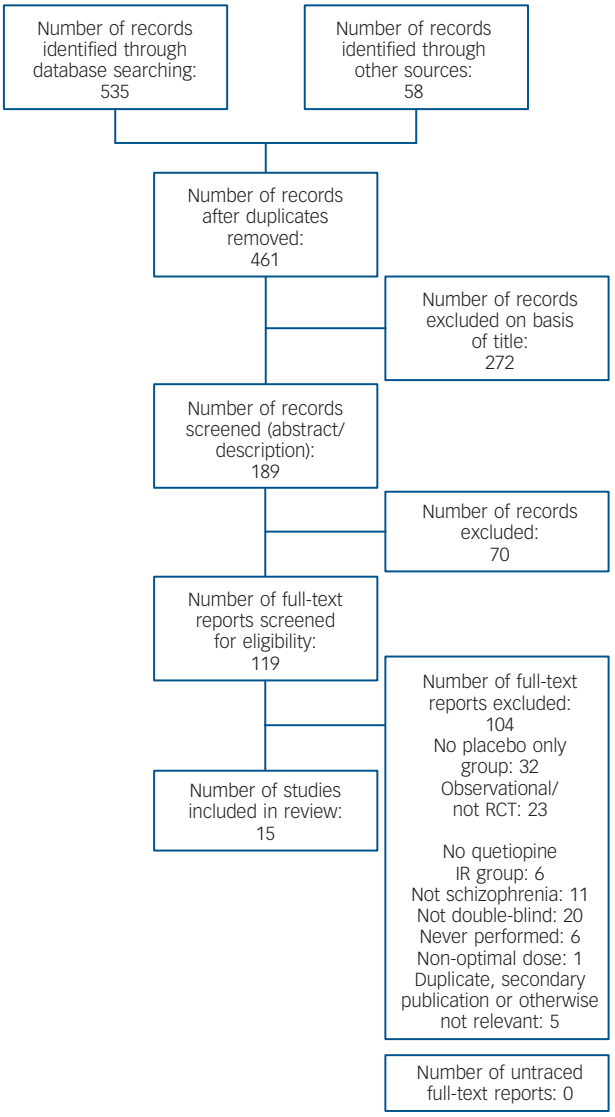

Fig. 1 Study search procedure. IR, immediate release; RCT, randomised controlled trial.

are listed in the online supplement, together with a table of trial characteristics and baseline demographic data.

Risk of bias and quality ratings

Table 1 provides the main risk of bias ratings, and the right-hand columns of Tables 2 and 3 provide the outcome quality ratings for the main primary and safety outcomes. Ratings for secondary outcomes and additional safety outcomes are provided in the online supplement. Our rationale for the ratings is also provided online, alongside ratings produced by other research groups (where available). In our judgement the main problem with these trials is a somewhat high risk of selective reporting bias in relation to secondary outcomes and adverse effects, coupled with a very high risk of attrition bias for most outcomes. We also judge unblinding due to sedative effects to be likely,[33] and the double-blind design might not protect against the risk of researchers adopting a high threshold for recording effects (e.g. adverse effects) where the desired outcome is 'no difference'.[34] There is also evidence from documents released through legal proceedings in the USA that AstraZeneca have historically not published all

Table 1 Included studies and Cochrane risk of bias ratings

	Year published/ completed	Primary publication available?	Clinical study report synopsis or extract available?	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Performance bias (masking of participants and personnel)	Detection bias (masking of assessments)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias[a]
Arvanitis & Miller	1997	Yes[38]	Yes[b]	Unclear	Unclear	Unclear	Unclear	High	Unclear	High
Small *et al*	1997	Yes[41]	Yes[b]	Low	Unclear	Unclear	Unclear	High	Unclear	High
Borison *et al*	1996	Yes[42]	Yes[b]	Unclear	Unclear	Unclear	Unclear	High	High	High
Kahn *et al*	2007	Yes[43]	Yes	Unclear	Unclear	Unclear	Unclear	High	High	Unclear
Canuso *et al*	2009	Yes[37]	Yes[c]	Low	Low	Unclear	Unclear	Low	Unclear	Unclear
Potkin *et al*	2006	Yes[36]	No	Low	Low	Unclear	Unclear	Low	High	Unclear
Chen *et al*	2010	Yes[45]	No	Low	Low	Unclear	Unclear	High	High	High
Lindenmayer *et al*	2008	Yes[39]	Yes	Unclear	Unclear	Unclear	Unclear	High	High	Unclear
Cutler *et al*	2010	Yes[44]	Yes	Unclear	Unclear	Unclear	Unclear	High	High	Unclear
Hough *et al*	2011	Yes[51]	Yes	Low	Low	Unclear	High	Low	High	Unclear
Chapel *et al*	2009	Yes[52]	No	Unclear	Unclear	Unclear	High	Unclear	High	Unclear
Findling *et al*	2012	Yes[40]	Yes	Low	Low	Unclear	Unclear	High	Low	Low
Study 11915A	2009	No	Yes[30]	Unclear	Unclear	Unclear	Unclear	High	High	High
Study 11916A	2009	No	Yes[31]	Unclear	Unclear	Unclear	Unclear	High	High	High
Arvanitis & Scott (Study 15)	1995	No	Partially[32]	Unclear	Unclear	Low	Low	High	High	High

a. Not including financial conflict of interest of sponsor or researcher.
b. AstraZeneca supplied extract.
c. Pfizer supplied extract.

active-comparator quetiapine trials, or have not reported all outcomes.[35]

Validation of MCID criterion

Researchers and trial sponsors designed their trials to detect with adequate power a mean change in PANSS total score or equivalent of approximately 12 points (range 9.0–15.5), which corresponds to an SMD of 0.55, and is similar in magnitude to the empirically derived estimate of MCID of 11–15 points.[10–12]

Primary efficacy outcomes

Moderate- to high-quality evidence suggested that quetiapine IR was statistically superior to placebo from 2 weeks to 12 weeks in terms of reducing overall symptoms, but the effect was small (WMD = −6.5 points, 95% CI −8.89 to −4.00; SMD = −0.33, 95% CI −0.46 to −0.21) and the 95% confidence intervals excluded the MCID of 11–15 points (Table 2, Figs 2 and 3). Low- to moderate-quality evidence suggested the NNT for much improvement was 21 (95% CI 13 to 63).

Sensitivity analyses and meta-regression

We identified a significant effect of study duration (weeks) on the effect size for total PANSS score (B = −0.04, 95% CI −0.08 to −0.01; $P = 0.02$), with a more treatment-favourable outcome associated with longer duration. Treatment duration did not significantly moderate the effect for treatment response (B < −0.01, RR = 1.00, 95% CI 0.98 to 1.01; $P = 0.70$). Excluding the two 2-week studies (Potkin *et al*, Canuso *et al*) was associated with a marginal increase in average PANSS change (WMD = −7.7 points, 95% CI −10.0 to −5.3; SMD = −0.38, 95% CI −0.50 to −0.26) and response rates (NNT = 19, 95% CI 11 to 59).[36,37] Year of publication did not significantly predict outcomes for total PANSS (B = 0.01, 95% CI −0.01 to 0.04; $P = 0.28$), but there was a small association for treatment response (B = 0.01, RR = 1.01, 95% CI 1.00 to 1.02; $P = 0.01$), with a less treatment-favourable outcome associated with a more recent publication date. Meta-regression bubble plots are provided in the online supplement.

Removing data from arms employing doses smaller than 400 mg reduced the contribution made by two trials,[38,39] but this had little effect on the overall estimate of average change (WMD = −6.3, 95% CI −8.7 to −3.8; SMD = −0.32, 95% CI −0.44 to −0.20) or response rates (NNT = 20, 95% CI 13 to 53). Excluding the study with an adolescent sample (Findling *et al*) also had little effect on estimates (WMD = −6.3, 95% CI −8.9 to −3.6; SMD = −0.32, 95% CI −0.45 to −0.19; NNT = 22, 95% CI 13 to 83).[40]

Impact of missing data

Overall, the seven 2–12 week trials with less than 50% attrition had a mean PANSS advantage of −5.4 points (95% CI −8.0 to −2.9; SMD = −0.29, 95% CI −0.44 to −0.15) and an NNT of 42 (19, 250H; H = harm), whereas the four 6-week studies with 50% or more attrition had a mean PANSS advantage of 9.2 points (95% CI −15 to −3.4; SMD = −0.39, 95% CI −0.62 to −0.17) and an NNT of 13 (95% CI 7 to 250). The three 6-week studies with less than 50% attrition had a mean advantage of 6.1 points (95% CI −9.9 to −2.3; SMD = −0.27, 95% CI −0.43 to −0.12) and a non-significant NNT of 35 (95% CI 12 to 42H).

Strategy 1 of the Ebrahim approach involved testing whether the overall results would be different if we assumed that participants who withdrew early from both groups had the same degree of change as participants in the control group who stayed until the end.[13] To illustrate this, consider the study by Small *et al*.[41] Here the mean change for the 49 quetiapine and 39 placebo group participants who completed the trial was, after conversion of BPRS to PANSS scores, −23.3 points (s.d. = 17.7) and −14.9 points (s.d. = 17.7) respectively – a between-group difference of around 8.4 points. Carrying forward the last available scores of the 104 people who did not complete this trial reduced the quetiapine estimate to −13.5 points (s.d. = 24.5; $n = 94$) and the placebo estimate to −1.5 points (s.d. = 24.0; $n = 94$), and increased the

Table 2 Primary efficacy outcomes

Outcome	Time (weeks)	No. of included studies	Quetiapine *n* events/*n*	Placebo *n* events/*n*	Hedges' *g* (95% CI)	Difference Mean (95% CI)	Risk ratio (95% CI)	Absolute difference (95% CI)	NNTB/H (95% CI)	Heterogeneity for *g* or RR	Quality (GRADE)
Overall symptoms (mean change in PANSS total score) based on LOCF or MMRM	2–12	11	1346	912	−0.33 (−0.44, −0.21)*	−6.44 (−8.89, −4.00)*				$I^2 = 47\%$; $\chi^2 = 18.9$ ($P = 0.040$)	Moderate to high
Overall symptoms (mean change in PANSS total score) using strategy 1 imputations	2–12	11	1373	931	−0.23 (−0.35, −0.11)*	−4.25 (−6.46, −2.04)*				$I^2 = 52\%$; $\chi^2 = 20.7$ ($P = 0.023$)	
Overall symptoms (mean change in PANSS total score) using strategy 2 imputations	2–12	11	1373	931	−0.15 (−0.30, 0.01)	−2.66 (−5.46, 0.15)				$I^2 = 70\%$; $\chi^2 = 32.9$ ($P < 0.001$)	
Significant improvement (⩾50% reduction in PANSS/BPRS score) based on LOCF	2–12	11	1126/1375	816/933			0.95 (0.91, 0.98)*	−0.047 (−0.016, −0.016)*	21B (13B, 63B)*	$I^2 = 43\%$; $\chi^2 = 17.5$ ($P = 0.070$)	Low to moderate

BPRS, Brief Psychiatric Rating Scale; LOCF, last observation carried forward; MMRM, mixed-models repeated measures; NNTB/H, number needed to treat (benefit/harm); PANSS, Positive and Negative Syndrome Scale; RR, relative risk.
*$P < 0.05$.

Table 3 Main safety outcomes

Outcome (definition, imputation strategy)	Time (weeks)	No. of included studies	Quetiapine *n* events/*n*	Placebo *n* events/*n*	Hedges' *g* (95% CI)	Difference Mean (95% CI)	Risk ratio (95% CI)	Absolute difference (95% CI)	NNTB/H (95% CI)	Heterogeneity for *g* or RR	Quality (GRADE)
Serious adverse event	2–12	8	50/851	45/658			0.94 (0.64, 1.39)	−0.001 (−0.023, 0.021)	1000B (44B, 48H)	$I^2 = 47\%$; $\chi^2 = 3.3$ ($P = 0.885$)	Very low
Any adverse event	2–12	9	754/1112	438/756			1.14 (1.06, 1.22)*	0.089 (0.045, 0.134)*	11H (22H, 8H)*	$I^2 = 0\%$; $\chi^2 = 6.6$ ($P = 0.583$)	Low
Simpson–Angus Scale (worsening)	6	7	128/869	83/596			0.97 (0.73, 1.29)	0.007 (−0.033, 0.047)	143H (30B, 21H)	$I^2 = 15\%$; $\chi^2 = 7$ ($P = 0.317$)	Low
Abnormal Involuntary Movement Scale (worsening)	6	4	88/534	56/265			0.694 (0.521, 0.924)*	−0.047 (−0.130, 0.037)	21B (8B, 27H)	$I^2 = 0\%$; $\chi^2 = 2.6$ ($P = 0.460$)	Low
Barnes Akathisia Rating Scale (worsening)	2–6	7	67/973	49/643			0.866 (0.609, 1.234)	−0.005 (−0.030, 0.020)	200B (33B, 50H)	$I^2 = 0\%$; $\chi^2 = 3.5$ ($P = 0.745$)	Low
Needing medication for extrapyramidal side-effects	2–6	9	97/1071	65/698			0.838 (0.597, 1.176)	0.004 (−0.025, 0.032)	250H (40B, 31H)	$I^2 = 12\%$; $\chi^2 = 9.1$ ($P = 0.334$)	Moderate
Mean weight change, kg	2–12	12	1410	948	0.640 (0.428, 0.852)*	1.753 (1.104, 2.402)*				$I^2 = 83\%$; $\chi^2 = 65.4$ ($P < 0.001$)	Very low
Significant weight-gain (⩾7% or recorded as adverse effect)	2–12	10	140/1220	32/863			2.988 (2.048, 4.362)*	0.076 (0.044, 0.109)*	13H (23H, 9H)*	$I^2 = 0\%$; $\chi^2 = 6.9$ ($P = 0.648$)	Moderate
Sedation or somnolence	2–12	12	247/1419	57/958			2.818 (1.963, 4.047)*	0.115 (0.078, 0.151)*	9H (7H, 13H)*	$I^2 = 31\%$; $\chi^2 = 16.1$ ($P = 0.138$)	Moderate
Leaving early owing to adverse effects	2–12	11	97/1263	74/885			1.009 (0.753, 1.351)	0.010 (-0.010, 0.031)	100H (100B, 32H)	$I^2 = 0\%$; $\chi^2 = 9.4$ ($P = 0.495$)	Moderate

GRADE, Grading of Recommendations Assessment, Development and Evaluation; NNTB/H, number needed to treat (benefit/harm); RR, risk ratio.
*$P < 0.05$.

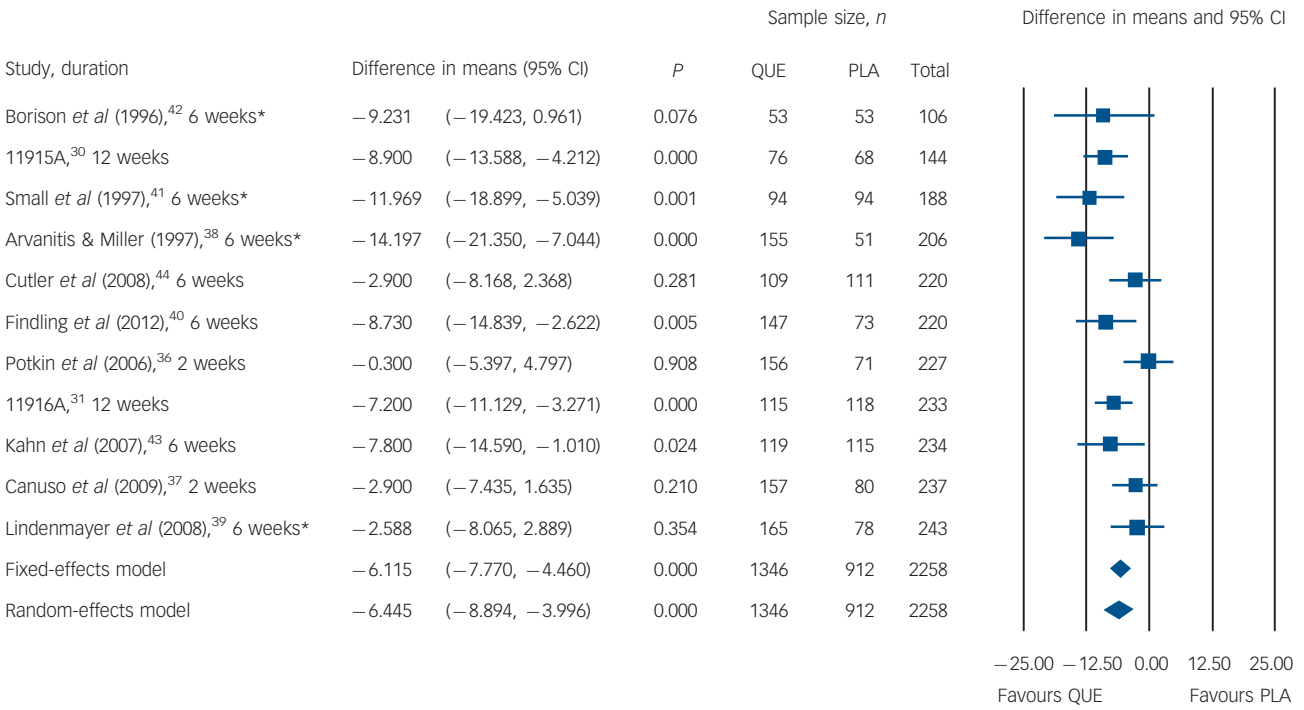

Fig. 2 Mean change in Positive and Negative Syndrome Scale (PANSS) total scores or equivalent, using mostly original last observation carried forward estimates to impute missing data. PLA, placebo; QUE, quetiapine. *Severe attrition (⩾50%).

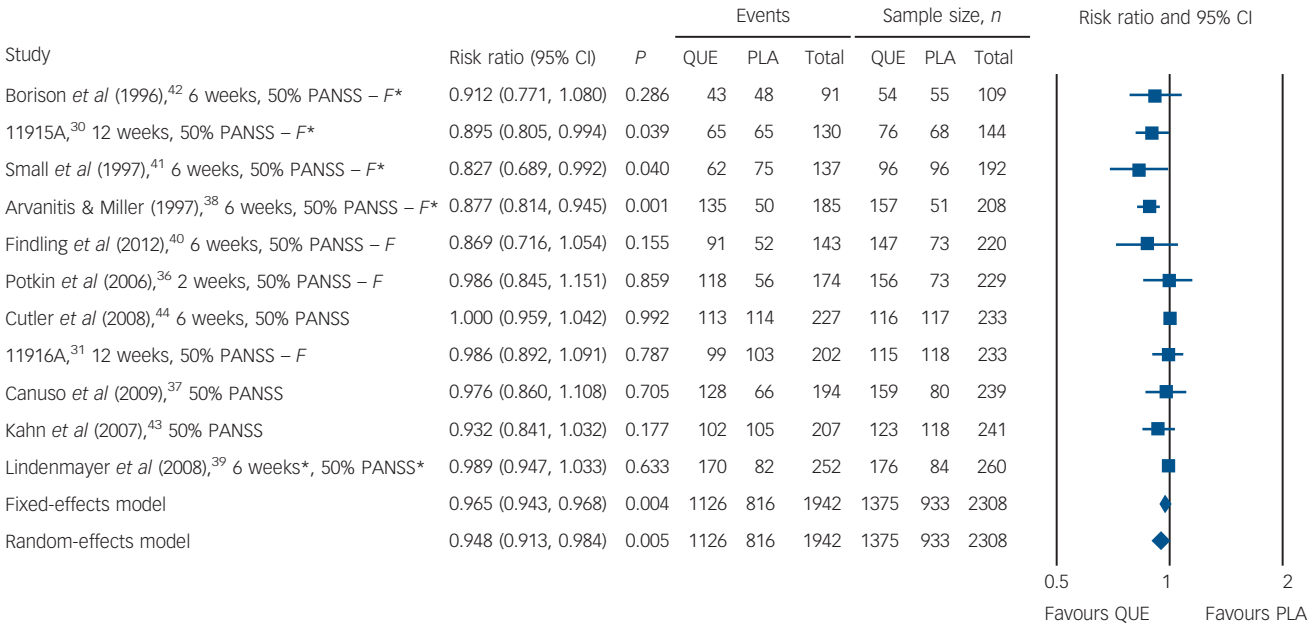

Fig. 3 Relative risk of not achieving at least 50% reduction in Positive and Negative Syndrome Scale (PANSS) total scores or equivalent, based mostly on last observation carried forward data; *F* estimated using Furukawa method. PLA, placebo; QUE, quetiapine. *Severe attrition (⩾50%).

between-group difference to around 12 points. These are the figures we used in the main analysis. Introducing the strategy 1 assumption that those who did not complete the trial had a similar outcome to those in the placebo group who did complete it (−14.9 points) reduced the overall estimate for the quetiapine group to −19.1 points (s.d. = 18.2) and reduced the advantage over placebo to 4.2 points. We repeated this procedure for the other five trials for which we had completer data and where no usable MMRM estimate was provided,[38,39,42–44] and entered the revised estimates into the overall meta-analysis. Table 2 shows that the overall advantage for quetiapine over placebo fell to 4.3 points (95% CI −6.5 to −2.0; SMD = −0.23, 95% CI −0.35 to −0.11).

Strategy 2 of the Ebrahim approach involved testing whether the overall results were robust to assuming, first, that participants in the quetiapine non-completers group had the smallest treatment response observed, and second, that those in the placebo non-completers group had the largest placebo response observed. In the study by Small *et al* this involved assuming the 47 people in the quetiapine non-completers group had the same degree of response as the quetiapine completers group in the study by Lindenmayer *et al* (−17.4 points),[39] and that the 57 people in the placebo non-completers group had the same degree of response as the placebo completers group in the 2007 study by Kahn *et al* (−23.1 points).[43] The revised quetiapine and placebo estimates were −20.4 (s.d. = 18.2) and −19.8 points (s.d. = 18.3) respectively, leading to a between-group difference of 0.6 points. As shown in Table 2, applying strategy 2 to the six trials for which we had completer data reduced the overall advantage for quetiapine to 2.7 points (95% CI −5.5 to 0.2; SMD = −0.15, 95% CI −0.30 to 0.01). Revised forest plots for strategies 1 and 2 are provided in the online supplement.

Publication bias

We detected some asymmetry in the funnel plot of clinically significant change, but not in relation to mean change in overall symptoms or most other outcomes. Funnel plots for the primary outcomes are provided in the online supplement.

Secondary efficacy outcomes

Full details concerning secondary efficacy outcomes are given in sections H and K of the online supplement.

Relapse, exacerbation and need for hospital care

Evidence from one study indicated that quetiapine IR was effective for prevention of symptom exacerbation in people with early psychosis who had responded to quetiapine,[45] but an unpublished study suggested there was no effect of therapeutic dose (300–600 mg) over a subtherapeutic dose (75 mg) in relapse prevention in chronic schizophrenia.[32] The combined estimate was therefore heterogeneous ($I^2 = 87\%$) and not significant (NNT = 5, 95% CI 2 to 13H). Quetiapine IR was associated with a marginally reduced need for hospital care after 2–6 weeks in three RCTs (NNT = 19, 95% CI 10 to 143).[36,37,44] One trial suggested quetiapine IR had a small effect over 52 weeks in relation to reducing readmission to hospital due to relapse (NNT = 11, 95% CI 6 to 143),[45] but the results were not robust to changing assumptions about the outcome of those leaving early. Overall the relapse and readmission data were very low to low in quality.

Other outcomes

There was a small effect on positive symptoms (SMD = −0.32, 95% CI −0.44 to -0.20; moderate-quality evidence) and a marginal to small effect on negative symptoms (SMD = −0.21, 95% CI −0.32 to −0.10; moderate-quality evidence) over 2–12 weeks, and a marginal effect on depression over 2–6 weeks (SMD = −0.13, 95% CI −0.23 to −0.02; low-quality evidence). Forest plots are provided in the online supplement. We did not investigate whether these estimates were robust to missing data, but the sensitivity analyses for the primary outcome of total symptoms suggest that this is unlikely. Those taking quetiapine IR had a marginally reduced need for additional sedative medication after 2–6 weeks (NNT = 34, 95% CI 13 to 53H; six RCTs), but we judged the evidence as low quality because of selective reporting and missing data, and no reduced need for antipsychotic medication was observed in the two 6-week RCTs where additional medication was not restricted (NNT = 24, 95% CI 7 to 19H; moderate-quality evidence).[36,37]

Pooled self-report end-point data from the two 12-week trials did not indicate any benefit of quetiapine IR on quality of life,[30,31] as measured by the Schizophrenia Quality of Life (S-QoL) scale (SMD = 0.11, 95% CI −0.15 to 0.36), but we judged the evidence to be very low in quality owing to early termination of the trials, missing data and possible selective reporting from the other trials. No significant effect was observed on any of the subscales, including psychological well-being (SMD = −0.02, 95% CI −0.28 to 0.24) or family relationships (SMD = 0.01, 95% CI −0.25 to 0.28). Since only observed case S-QoL data were reported, we imputed missing data using strategy 1 from Ebrahim *et al*.[13] This reduced the overall effect from 0.11 to 0.06 (95% CI −0.14 to 0.27). Long-term quality of life data from two RCTs remain unpublished.[32,45]

An analysis of data from three RCTs (one studying adolescents, two with adult samples) covering a period of 6–12 weeks found quetiapine IR had a small to moderate benefit on functioning,[30,31,40] as assessed by a combination of Children's Global Assessment Scale data and Personal and Social Performance (PSP) data (SMD = 0.39, 95% CI 0.18 to 0.60). Global Assessment of Functioning data were also reported, but unlike the PSP this measure assesses symptom severity as well as functioning. After imputing missing PSP data using strategy 1, the effect size was small (SMD = 0.28, 95% CI 0.09 to 0.46). One study found no benefit of 12 months of quetiapine IR maintenance treatment over placebo in relation to employment status.[45] Overall, the functioning and employment data were very low in quality owing to selective reporting, early termination of studies, imprecision and missing data. High-quality evidence from 11 trials suggested quetiapine IR had a marginal effect on rates of early discontinuation over a period of 2–6 weeks (NNT = 21, 95% CI 10 to 333H).

Safety outcomes

Safety outcomes are detailed in Table 3 and Figs 4–6. There was low-quality evidence that quetiapine IR was associated with a small to moderately increased risk of non-serious adverse effects over the short term (NNH = 11, 95% CI 8 to 22). There was no evidence of extrapyramidal side-effects and no evidence of an increased risk of serious adverse events. Moderate-quality evidence suggested no need for additional anti-Parkinsonian medication in quetiapine-treated participants in the short term, but longer-term data were not reported. Data from 12 trials suggested quetiapine IR had a moderate to large effect on weight gain over 2–12 weeks (SMD = 0.64, 95% CI 0.43 to 0.85). Participants gained an extra 1.75 kg (95% CI 1.10 to 2.40) on average, but we rated the evidence as very low quality because of non-reporting of variance parameters in 7 out of 12 studies, high rates of withdrawal and high heterogeneity. Moderate-quality evidence suggested that around 12% of participants treated with quetiapine experienced a clinically significant increase in weight over 2–12 weeks, compared with 4% of those taking placebo (NNH = 13, 95% CI 9 to 23), and 35% reported sedation or somnolence as an adverse effect compared with 6% of those taking placebo (NNH = 9, 95% CI 7 to 13). Details on additional safety outcomes and forest plots are reported in sections I and K of the online supplement.

Discussion

Using published and unpublished data, we found the average change in PANSS total score attributable to quetiapine IR over 2–12 weeks to be small. Although the 95% confidence intervals excluded the minimum clinically important difference of 11–15

Study	Difference in means (95% CI)		*P*	Sample size, *n* QUE	PLA	Total
Hough *et al.* (2011),[51] 2-wk s.d. imputed	1.400	(0.056, 2.744)	0.041	43	22	65
Borison *et al.* (1996),[42] 6-wk* s.d. imputed	5.000	(3.999, 6.001)	0.000	54	55	109
11915A,[30] 12-wk s.d. imputed	1.600	(0.731, 2.469)	0.000	76	68	144
Small *et al.* (1997),[41] 6-wk* s.d. imputed	1.900	(1.146, 2.654)	0.000	96	96	192
Arvanitis & Miller (1997),[38] 6-wk* s.d. imputed	3.100	(2.282, 3.918)	0.000	157	51	208
Cutler *et al.* (2010),[44] 6-wk	1.080	(−0.353, 2.513)	0.140	108	107	215
Findling *et al.* (2012),[40] 6-wk	1.800	(1.067, 2.533)	0.000	147	75	222
Potkin *et al.* (2006),[36] 2-wk	0.500	(0.019, 0.981)	0.042	156	73	229
11916A,[31] 12-wk, s.d. imputed	1.900	(1.217, 2.583)	0.000	115	119	234
Canuso *et al.* (2009),[37] 2-wk	0.600	(0.208, 0.992)	0.003	159	80	239
Kahn *et al.* (2007),[43] 6-wk	0.920	(0.194, 1.646)	0.013	123	118	241
Lindenmayer *et al.* (2006),[39] 6-wk*	1.490	(0.601, 2.379)	0.001	176	84	260
Fixed-effects model	1.361	(1.160, 1.563)	0.000	1410	948	2358
Random-effects model	1.753	(1.104, 2.402)	0.000	1410	948	2358

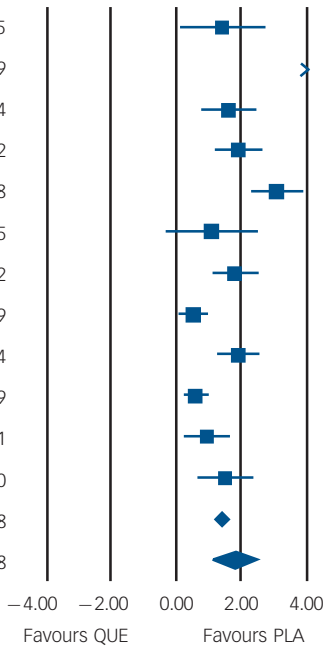

Fig. 4 Mean weight gain, kg. PLA, placebo; QUE, quetiapine. *Severe attrition (⩾50%).

Study	Risk ratio (95% CI)		*P*	Events QUE	PLA	Total	Sample size, *n* QUE	PLA	Total
Borison *et al.* (1996),[42] 6-wk*	7.130	(1.701, 28.889)	0.007	14	2	16	54	55	109
11915A,[30] 12-wk	11.649	(0.668, 203.016)	0.092	6	0	6	76	68	144
Small *et al.* (1997),[41] 6-wk*	4.800	(1.911, 12.057)	0.001	24	5	29	96	96	192
Arvanitis & Miller (1997),[38] 6-wk*	2.166	(0.671, 9.989)	0.196	20	3	23	157	51	208
Findling *et al.* (2012),[40] 6-wk	3.912	(1.213, 12.610)	0.022	23	3	26	147	75	222
Cutler *et al.* (2010),[44] 6-wk	2.185	(0.860, 5.553)	0.100	13	6	19	116	117	233
11916A,[31] 12-wk	8.207	(1.043, 64.589)	0.046	8	1	9	116	119	235
Canuso *et al.* (2009),[37] 2-wk	2.516	(0.299, 21.173)	0.396	5	1	6	159	80	239
Kahn *et al.* (2007),[43] 6-wk	1.812	(0.841, 3.904)	0.129	17	9	26	123	118	241
Lindenmayer *et al.* (2006),[39] 6-wk*	2.386	(0.535, 10.650)	0.254	10	2	12	176	84	260
Fixed-effects model	2.988	(2.048, 4.362)	0.000	140	32	172	1220	863	2083
Random-effects model	2.988	(2.048, 4.362)	0.000	140	32	172	1220	863	2083

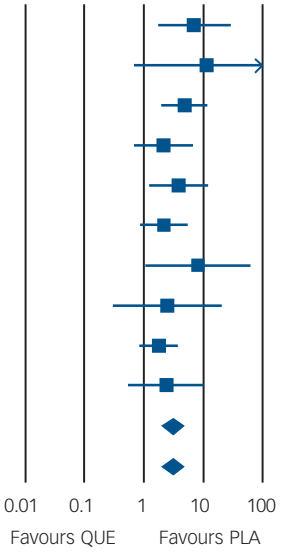

Fig. 5 Relative risk of clinically significant weight gain (normally ⩾7% increase). PLA, placebo; QUE, quetiapine. *Severe attrition (⩾50%).

points, it should be noted that few treatments for psychosis reach this threshold.[19] Furthermore, as study duration increased, so did the effect size. Marginal advantages were observed at 2 weeks, whereas moderate effects that approached the threshold for change of at least minimal clinical importance were observed in two as yet unpublished 12-week trials. On the other hand, the overall effect size was smaller if we assumed that the substantial number of participants who left early would have had the same degree of change as placebo-treated participants who stayed. There was no evidence to suggest that drug-attributable benefits had been underestimated because of the severe rates of withdrawal in the older studies. Approximately 21 people needed to take quetiapine IR for 1 person to experience much improvement, defined in accordance with recent recommendations as a 50% or greater reduction in PANSS total score.[27]

Null to small effects were observed for depression and negative symptoms respectively. Although moderate effects on positive symptoms were observed in the two unpublished 12-week trials, the pooled effect over 2–12 weeks was small. The two 12-week trials also reported small to moderate effects on functioning but

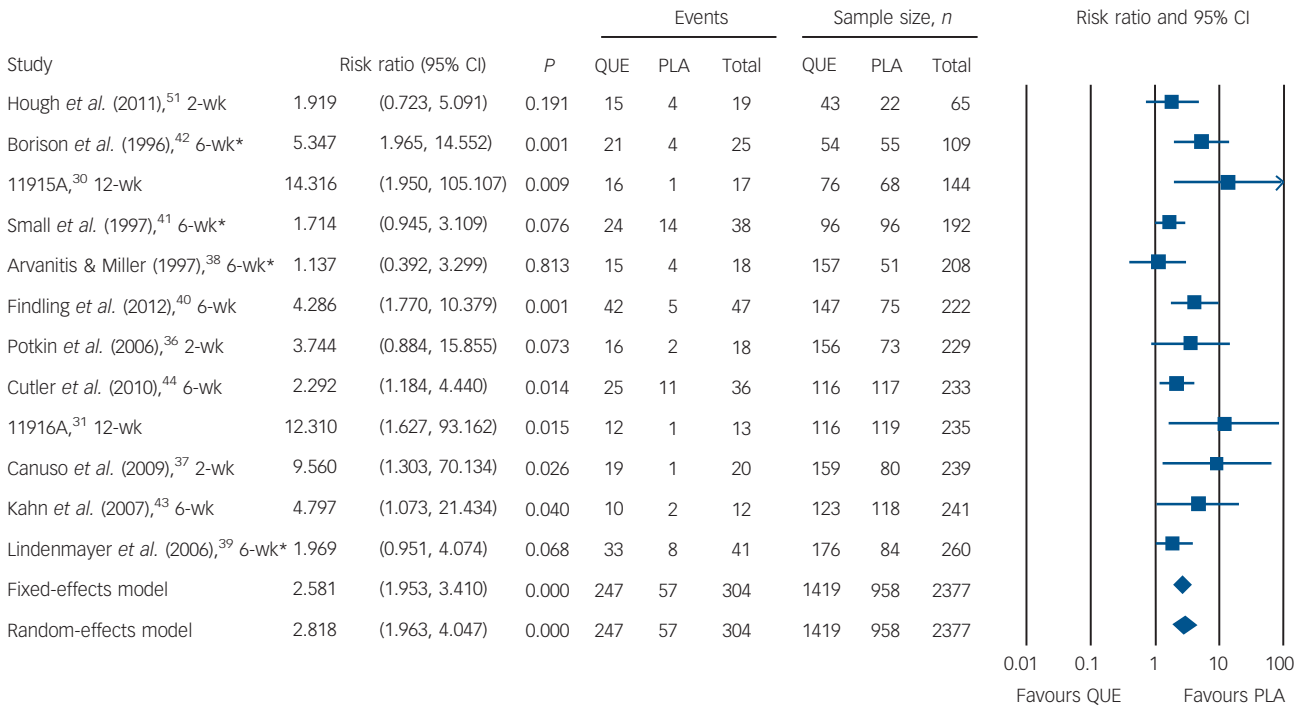

Fig. 6 Relative risk of somnolence or sedation. PLA, placebo; QUE, quetiapine. *Severe attrition (⩾50%).

found no difference between quetiapine IR and placebo on participant-reported quality of life. Quetiapine IR caused weight gain and sedation, but did not lead to extrapyramidal side-effects. Although there was no evidence of increased serious adverse effects, the evidence was very low quality owing to imprecision and incomplete reporting.

We found some evidence that estimates of clinically significant response derived from more recent trials were lower than in older trials, which is consistent with results from other antipsychotic meta-analyses,[4,46] although no relationship between publication year and total symptoms was observed. It remains unclear whether reduced antipsychotic response in recent, large multisite trials with multiple treatment arms reflects a change in the characteristics of participants taking part, improvements in study quality or reporting, increased variability due to an increasing number of sites,[47] or better masking of treatment allocation due to reduced expectation of receiving placebo in such trials.[48]

Older reviews of quetiapine IR have reported effect sizes of around 0.4 when compared with placebo,[3,4,19] and NNTs of around 10 or 11.[3,4] However, most of the trials available at the time had high rates of participant withdrawal and examined quetiapine IR as a target drug for regulatory approval rather than as a control for other preparations. Previous reviews have not been able to account for selective reporting in relation to response rates, examine the impact of changing assumptions about missing data on estimates, or consider the clinical significance of the change attributable to quetiapine IR. A more recent review pooled 4–12 week outcome data for quetiapine IR with outcome data for the more recent extended release version of quetiapine (quetiapine XR) and reported an overall moderate effect size of 0.44.[19] Since quetiapine XR was judged by its manufacturer to be sufficiently novel to warrant a separate patent application and significantly greater cost, our *a priori* view was that pooling the data for the two formulations would give an inaccurate appraisal of both. In relation to duration, we planned to include 2-week quetiapine IR data in our review because this was the approach favoured by preceding reviews that were available at the time of protocol writing.[4] We adhered to this decision because several prescribing guidelines recommend a minimum 2-week trial, and evidence on prescribing practices suggests psychiatrists normally wait only 3–3.5 weeks before switching to another antipsychotic because of non-response.[49] Nonetheless, it is important to consider that overall efficacy was positively associated with trial duration in our review, and might have been larger still had we included quetiapine XR data. Our data may help explain why a recent meta-analysis found that quetiapine IR was significantly less effective at reducing positive symptoms than first-generation antipsychotics (nine RCTs).[17] In an unpublished study therapeutic dose quetiapine IR was significantly less effective than haloperidol in preventing relapse.[32]

Missing data

Levels of missing data were high in the included trials. In order to reduce this, trial researchers should continue to assess participants who stop treatment early, as this will inform realistic estimates of likely outcome had they stayed, both in relation to efficacy and adverse effects. Since many early studies of other second-generation antipsychotics also suffered from severe attrition,[6] the robustness of their effects to changing assumptions about missing data may also need to be examined. Although meta-regression has been used to examine the relationship between withdrawals and effect size,[19] such analyses are inevitably limited by the fact that few trials have low rates of missing data.[6] Application of the Ebrahim approach would help prescribers and patients appreciate the extent of uncertainty in estimates of antipsychotic benefits and costs.[13] In addition to attrition bias, the proper assessment of both drug and non-drug treatments for psychosis continues to be limited by incomplete and selective reporting of outcomes, low external validity and non-publication of negative trials.[8] Indeed, a recent meta-analysis found the median effect of currently available antipsychotics over placebo fell from moderate to small after adjusting for the tendency for small studies to report larger effects,[19] and selective reporting bias

is a particular concern when, as documented by Spielmans & Parry and others,[35] it biases our understanding of the severity of adverse effects.

Study limitations

We took advantage of several important developments in methodology which were published after we registered our protocol.[13,20,28] Changes *post hoc* do raise a risk of bias, but we had to balance this against taking the opportunity to increase the quality, robustness and usefulness of our estimates, and we hope we have provided enough information for readers to judge the merit of these decisions. Our claim that a score of 11–15 points is required for minimal clinical improvement might be controversial, not least because few treatments achieve such change in psychosis.[19] However, we note the evidence supporting this minimum threshold is now quite consistent across different populations,[10–12] and we demonstrated that quetiapine trial researchers designed their trials to be able to detect with adequate power only differences of approximately 12 points. Although it has been argued that small benefits might have value at a public health level,[4] there is clearly a need for further debate on this issue. As with non-inferiority and equivalence trials,[34] researchers planning superiority trials might consider stating in advance what they believe constitutes a minimum important difference on continuous outcomes. Although this can be inferred from power calculations, it needs to be stated explicitly.

We were unable to access the full clinical study reports for each trial, which is problematic given a recent study found a much better quality of reporting in these documents when compared with registry reports or peer-reviewed publications.[50] Although we have acquired a large amount of previously unpublished data, access to the reports would have raised the quality of many of the outcomes, in particular the assessments of mean weight gain and response rates. Acquiring unpublished data was challenging, and we doubt we would have been successful had a public debate on publication bias in clinical trials not been taking place at the time. This is an unsatisfactory and unsustainable situation, and a change in the law is clearly required to ensure that all trials past and present are registered, and their full methods and summary results reported, as advocated by the Alltrials campaign (www.alltrials.net).

Paul Hutton, DClinPsy, Department of Clinical Psychology, School of Health in Social Science, University of Edinburgh; **Peter J. Taylor**, PhD, DClinPsy, Department of Clinical Psychology, Institute of Psychology, Health and Society, University of Liverpool; **Lee Mulligan**, BSc, Department of Clinical Psychology, School of Psychological Sciences, University of Manchester; **Sarah Tully**, BSc, Psychosis Research Unit, Greater Manchester West Mental Health Foundation National Health Service Trust, Manchester; **Joanna Moncrieff**, MD, Mental Health Sciences Unit, School of Life and Medical Sciences, University College London, UK

Correspondence: Dr Paul Hutton, Department of Clinical Psychology, Doorway 6, Medical Building, Teviot Place, Edinburgh EH8 9AG, UK. Email: paul.hutton.cf@ed.ac.uk.

First received 7 Dec 2013, final revision 8 Jul 2014, accepted 30 Sep 2014

Funding

This work was not funded. P.H. has been a co-investigator on a National Institute of Health Research (NIHR) funded pilot trial of cognitive therapy for people with psychosis who refuse antipsychotics, and is a co-investigator on another NIHR-funded pilot trial of cognitive therapy *v.* antipsychotics in early psychosis; J.M. is co-chairperson of the Critical Psychiatry Network.

Acknowledgements

We thank the following individuals and their organisations for providing us with unpublished data from various clinical trials: Dr Carla Canuso from Johnson & Johnson, Rakesh Kantaria, John Ramsey, Jasmine Lichfield and Craig Shering from AstraZeneca and Mads Kronborg and Dr Andrew Roberts from Lundbeck. We also thank Giovana Pezzi from AstraZeneca, Jan Yonge and Dr Chris Bushe from Eli Lilly, Sanofi (Medical Information), the UK Medicines and Healthcare Regulatory Authority, the Danish Medicines Agency, Dr Ben Goldacre, Edd Howard and Kerry Roberts.

References

1 National Prescribing Centre. *Schizophrenia Data Focussed Commentary*. NPC, 2010 (http://www.npc.nhs.uk/therapeutics/cns/schizophrenia/resources/dfc_schizophrenia.pdf).

2 Medicines and Healthcare Products Regulatory Authority. *UKPAR Seroquel XL 50mg, 200mg, 300mg & 400mg prolonged-release tablets*. MHRA, 2008.

3 Srisurapanont M, Maneeton B, Maneeton N. Quetiapine for schizophrenia. *Cochrane Database Syst Rev* 2004; **2**: CD000967.

4 Leucht S, Arbter D, Engel RR, Kissling W, Davis JM. How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *Mol Psychiatry* 2009; **14**: 429–47.

5 Hamer RM, Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *Am J Psychiatry* 2009; **166**: 639–41.

6 Hutton P, Morrison AP, Yung AR, Taylor PJ, French P, Dunn G. Effects of drop-out on efficacy estimates in five Cochrane reviews of popular antipsychotics for schizophrenia. *Acta Psychiatr Scand* 2012; **126**: 1–11.

7 Xia J, Adams C, Bhagat N, Bhagat V, Bhoopathi P, El-Sayeh H, et al. Losing participants before the end of the trial erodes credibility of findings. *Psychiatr Bull* 2009; **33**: 254–7.

8 Leucht S, Heres S, Hamann J, Kane JM. Methodological issues in current antipsychotic drug trials. *Schizophr Bull* 2008; **34**: 275–85.

9 Rabinowitz J, Davidov O. The association of dropout and outcome in trials of antipsychotic medication and its implications for dealing with missing data. *Schizophr Bull* 2008; **34**: 286–91.

10 Leucht S, Kane JM, Etschel E, Kissling W, Hamann J, Engel RR. Linking the PANSS, BPRS, and CGI: clinical implications. *Neuropsychopharmacology* 2006; **31**: 2318–25.

11 Thwin SS, Hermes E, Lew R, Barnett P, Liang M, Valley D, et al. Assessment of the minimum clinically important difference in quality of life in schizophrenia measured by the Quality of Well-Being Scale and disease-specific measures. *Psychiatry Res* 2013; **209**: 291–6.

12 Hermes ED, Sokoloff D, Stroup TS, Rosenheck RA. Minimum clinically important difference in the Positive and Negative Syndrome Scale with data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). *J Clin Psychiatry* 2012; **73**: 526–32.

13 Ebrahim S, Akl EA, Mustafa RA, Sun X, Walter SD, Heels-Ansdell D, et al. Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers. *J Clin Epidemiol* 2013; **66**: 1014–21, e1.

14 Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophr Res* 2005; **79**: 231–8.

15 Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* 2005; **20**: 49–52.

16 Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Cochrane Collaboration, 2011.

17 Leucht S, Corves C, Arbter D, Engel RR, Li C, Davis JM. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet* 2009; **373**: 31–41.

18 Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ. International consensus study of antipsychotic dosing. *Am J Psychiatry* 2010; **167**: 686–93.

19 Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 2013; **382**: 951–62.

20 Leucht S, Rothe P, Davis JM, Engel RR. Equipercentile linking of the BPRS and the PANSS. *Eur Neuropsychopharmacol* 2013; **23**: 956–9.

21 National Institute for Health and Clinical Excellence. *Schizophrenia: Core Interventions in the Treatment and Management of Schizophrenia in Adults in Primary Care and Secondary Care (Update).* NICE, 2009.

22 Bagnall AM, Jones L, Ginnelly L, Lewis R, Glanville J, Gilbody S, et al. A systematic review of atypical antipsychotic drugs in schizophrenia. *Health Technol Assess* 2003; **7**: 1–193.

23 Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989; **10**: 407–15.

24 Schennach-Wolff R, Obermeier M, Seemuller F, Jager M, Schmauss M, Laux G, et al. Does clinical judgment of baseline severity and changes in psychopathology depend on the patient population? Results of a CGI and PANSS linking analysis in a naturalistic study. *J Clin Psychopharmacol* 2010; **30**: 726–31.

25 Man-Son-Hing M, Laupacis A, O'Rourke K, Molnar FJ, Mahon J, Chan KB, et al. Determination of the clinical importance of study results. *J Gen Intern Med* 2002; **17**: 469–76.

26 Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336**: 924–6.

27 Leucht S, Davis JM, Engel RR, Kissling W, Kane JM. Definitions of response and remission in schizophrenia: recommendations for their use and their presentation. *Acta Psychiatr Scand Suppl* 2009; **438**: 7–14.

28 Samara MT, Spineli LM, Furukawa TA, Engel RR, Davis JM, Salanti G, et al. Imputation of response rates from means and standard deviations in schizophrenia. *Schizophr Res* 2013; **151**: 209–14.

29 Harbord RM, Higgins JPT. Meta-regression in STATA. *Stata J* 2008; **8**: 493–519.

30 Lundbeck. Synopsis – Study 11915A: a one-year multi-national, multi-centre, randomised, double-blind, parallel-group, fixed-dose bifeprunox study combining a 12-week placebo-controlled, quetiapine-referenced phase with a 12-month quetiapine-controlled phase in patients with schizophrenia. Lundbeck, 2010.

31 Lundbeck. Synopsis – Study 11916A: a one-year multi-national, multi-centre, randomised, double-blind, parallel-group, fixed-dose bifeprunox study combining a 12-week placebo-controlled, quetiapine-referenced phase with a 12-month quetiapine-controlled phase in patients with schizophrenia. Lundbeck, 2010.

32 Arvanitis L, Scott M. A multicenter, double blind, randomized, controlled, multiple fixed dose and dose regimen comparison of SEROQUELTM (ICI 204,636) and haloperidol in the prevention of psychotic relapse in outpatients with chronic or subchronic schizophrenia (5077IL/0015). AstraZeneca,1996.

33 Perlis RH, Ostacher M, Fava M, Nierenberg AA, Sachs GS, Rosenbaum JF. Assuring that double-blind is blind. *Am J Psychiatry* 2010; **167**: 250–2.

34 Treadwell JR, Uhl S, Tipton K, Shamliyan T, Viswanathan M, Berkman ND, et al. Assessing equivalence and noninferiority. *J Clin Epidemiol* 2012; **65**: 1144–9.

35 Spielmans GI, Parry PI. From evidence-based medicine to marketing-based medicine: evidence from internal industry documents. *J Bioeth Inq* 2010; **7**: 13–29.

36 Potkin SG, Gharabawi GM, Greenspan AJ, Mahmoud R, Kosik-Gonzalez C, Rupnow MF, et al. A double-blind comparison of risperidone, quetiapine and placebo in patients with schizophrenia experiencing an acute exacerbation requiring hospitalization. *Schizophr Res* 2006; **85**: 254–65.

37 Canuso CM, Dirks B, Carothers J, Kosik-Gonzalez C, Bossie CA, Zhu Y, et al. Randomized, double-blind, placebo-controlled study of paliperidone extended-release and quetiapine in inpatients with recently exacerbated schizophrenia. *Am J Psychiatry* 2009; **166**: 691–701.

38 Arvanitis LA, Miller BG. Multiple fixed doses of 'Seroquel' (quetiapine) in patients with acute exacerbation of schizophrenia: a comparison with haloperidol and placebo. The Seroquel Trial 13 Study Group. *Biol Psychiatry* 1997; **42**: 233–46.

39 Lindenmayer JP, Brown D, Liu S, Brecher M, Meulien D. The efficacy and tolerability of once-daily extended release quetiapine fumarate in hospitalized patients with acute schizophrenia: a 6-week randomized, double-blind, placebo-controlled study. *Psychopharmacol Bull* 2008; **41**: 11–35.

40 Findling RL, McKenna K, Earley WR, Stankowski J, Pathak S. Efficacy and safety of quetiapine in adolescents with schizophrenia investigated in a 6-week, double-blind, placebo-controlled trial. *J Child Adolesc Psychopharmacol* 2012; **22**: 327–42.

41 Small JG, Hirsch SR, Arvanitis LA, Miller BG, Link CG. Quetiapine in patients with schizophrenia. A high- and low-dose double-blind comparison with placebo. Seroquel Study Group. *Arch Gen Psychiatry* 1997; **54**: 549–57.

42 Borison RL, Arvanitis LA, Miller BG. ICI 204,636, an atypical antipsychotic: efficacy and safety in a multicenter, placebo-controlled trial in patients with schizophrenia. U.S. SEROQUEL Study Group. *J Clin Psychopharmacol* 1996; **16**: 158–69.

43 Kahn RS, Schulz SC, Palazov VD, Reyes EB, Brecher M, Svensson O, et al. Efficacy and tolerability of once-daily extended release quetiapine fumarate in acute schizophrenia: a randomized, double-blind, placebo-controlled study. *J Clin Psychiatry* 2007; **68**: 832–42.

44 Cutler AJ, Tran-Johnson T, Kalali A, Astrom M, Brecher M, Meulien D. A failed 6-week, randomized, double-blind, placebo-controlled study of once-daily extended release quetiapine fumarate in patients with acute schizophrenia: lessons learned. *Psychopharmacol Bull* 2010; **43**: 37–69.

45 Chen EY, Hui CL, Lam MM, Chiu CP, Law CW, Chung DW, et al. Maintenance treatment with quetiapine versus discontinuation after one year of treatment in patients with remitted first episode psychosis: randomised controlled trial. *BMJ* 2010; **341**: c4024.

46 Agid O, Siu CO, Potkin SG, Kapur S, Watsky E, Vanderburg D, et al. Meta-regression analysis of placebo response in antipsychotic trials, 1970–2010. *Am J Psychiatry* 2013; **170**: 1335–44.

47 Leucht S, Heres S, Davis JM. Increasing placebo response in antipsychotic drug trials: let's stop the vicious circle. *Am J Psychiatry* 2013; **170**: 1232–4.

48 Mallinckrodt CH, Zhang L, Prucka WR, Millen BA. Signal detection and placebo response in schizophrenia: parallels with depression. *Psychopharmacol Bull* 2010; **43**: 53–72.

49 Hamann J, Kissling W, Leucht S. How long do psychiatrists wait for response before they switch to another antipsychotic? *Psychopharmacol Bull* 2007; **40**: 149–54.

50 Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. *BMJ* 2012; **344**: d8141.

51 Hough DW, Natarajan J, Vandebosch A, Rossenu S, Kramer M, Eerdekens M. Evaluation of the effect of paliperidone extended release and quetiapine on corrected QT intervals: a randomized, double-blind, placebo-controlled study. *Int Clin Psychopharmacol* 2011; **26**: 25–34.

52 Chapel S, Hutmacher MM, Haig G, Bockbrader H, de Greef R, Preskorn SH, et al. Exposure-response analysis in patients with schizophrenia to assess the effect of asenapine on QTc prolongation. *J Clin Pharmacol* 2009; **49**: 1297–308.

EXTRA CONTENT ONLINE