Early Improvement As a Predictor of Later Response to Antipsychotics in Schizophrenia: A Diagnostic Test Review

Myrto T. Samara, M.D., Claudia Leucht, M.D., Mariska M. Leeflang, Ph.D., Ion-George Anghelescu, M.D., Young-Chul Chung, Ph.D., M.D., Benedicto Crespo-Facorro, Ph.D., M.D., Helio Elkis, Ph.D., M.D., Kotaro Hatta, Ph.D., M.D., Ina Giegling, Ph.D., John M. Kane, M.D., Monica Kayo, M.D., Martin Lambert, M.D., Ching-Hua Lin, Ph.D., M.D., Hans-Jürgen Möller, Ph.D., M.D., José María Pelayo-Terán, Ph.D., M.D., Michael Riedel, M.D., Dan Rujescu, Ph.D., M.D., Benno G. Schimmelmann, M.D., Alessandro Serretti, Ph.D., M.D., Christoph U. Correll, M.D., Stefan Leucht, M.D.

Objective: How long clinicians should wait before considering an antipsychotic ineffective and changing treatment in schizophrenia is an unresolved clinical question. Guidelines differ substantially in this regard. The authors conducted a diagnostic test meta-analysis using mostly individual patient data to assess whether lack of improvement at week 2 predicts later nonresponse.

Method: The search included EMBASE, MEDLINE, BIOSIS, PsycINFO, Cochrane Library, CINAHL, and reference lists of relevant articles, supplemented by requests to authors of all relevant studies. The main outcome was prediction of nonresponse, defined as <50% reduction in total score on either the Positive and Negative Syndrome Scale (PANSS) or Brief Psychiatric Rating Scale (BPRS) (corresponding to at least much improved) from baseline to endpoint (4–12 weeks), by <20% PANSS or BPRS improvement (corresponding to less than minimally improved) at week 2. Secondary outcomes were absent cross-sectional symptomatic remission and <20% PANSS or BPRS reduction at endpoint. Potential moderator variables were examined by meta-regression.

Results: In 34 studies (N=9,460) a <20% PANSS or BPRS reduction at week 2 predicted nonresponse at endpoint with a specificity of 86% and a positive predictive value (PPV) of 90%. Using data for observed cases (specificity=86%, PPV=85%) or lack of remission (specificity=77%, PPV=88%) yielded similar results. Conversely, using the definition of <20% reduction at endpoint yielded worse results (specificity=70%, PPV=55%). The test specificity was significantly moderated by a trial duration of <6 weeks, higher baseline illness severity, and shorter illness duration.

Conclusions: Patients not even minimally improved by week 2 of antipsychotic treatment are unlikely to respond later and may benefit from a treatment change.

Am J Psychiatry 2015; 172:617–629; doi: 10.1176/appi.ajp.2015.14101329

When to change the treatment of patients who do not respond to a recently initiated antipsychotic drug is an unresolved clinical question. For decades the dogma of a delayed onset of antipsychotic drug action determined clinical decisions and guidelines in this regard (1–6). In 2003, a meta-analysis by Agid et al. (7) challenged that theory by demonstrating that the greatest symptom reduction occurred during the first weeks of treatment. This "early onset of antipsychotic drug action hypothesis" was corroborated by a subsequent analysis using longer-term, individual patient data (8). As a consequence, numerous studies have since examined whether the degree of early improvement could predict later response (9–27). Most studies showed such associations, but the lack of consensus about the definitions of early improvement and later response made uniform guideline recommendations impossible. For instance, some studies defined early improvement and/or later response as ≥20% reduction in the total score on the Positive and Negative Syndrome Scale (PANSS), whereas others used a ≥30%, ≥40%, or ≥50% score reduction.

Therefore, the statements in treatment guidelines have remained inconsistent and are often not based on evidence. For example, the American Psychiatric Association (APA) (28) suggests, "Patients may take between 2 and 4 weeks to show an initial response" on the basis of a small initial study from Correll et al. (29). The guidelines from the Schizophrenia Patient Outcomes Research Team (PORT) (30, 31) and the World Federation of Societies of Biological Psychiatry (32, 33) recommend waiting for at least 2 weeks before switching medication, but again no solid evidence is provided. The guidelines from the British Association of Psychopharmacology (34) and from the National Institute of Clinical Excellence (NICE) (35) recommend trying an antipsychotic at the optimum dose for 4–6 weeks before switching, also without providing firm evidence supporting this recommendation.

Given the uncertainty about these questions, we examined whether lack of improvement at week 2 can predict later nonresponse by a diagnostic test review. Diagnostic test reviews are novel meta-analytic techniques that allow

See related feature: **Clinical Guidance** *(Table of Contents)*

researchers to synthesize the results of studies on diagnostic tests and obtain their overall test parameters, such as sensitivity, specificity, and positive and negative predictive value (36). As diagnostic meta-analyses are far more challenging than conventional reviews of interventions, we collaborated with the authors of the included studies, most of whom provided us with individual patient data. Based on prior individual study evidence (12, 15, 22, 23, 29), we hypothesized that a lack of significant improvement at week 2 would predict ultimate treatment failure.

METHOD

In studies that examine the accuracy of diagnostic tests, the results of a new test ("index test") are compared with a "reference standard." Then a diagnostic test review synthesizes all such studies to provide the overall accuracy of the new test. This concept can be applied to our research question, where the index test is a predefined degree of nonimprovement at week 2 and the reference standard is nonresponse at a later stage. We decided to analyze nonimprovement and nonresponse to antipsychotic treatment, instead of improvement and response, because it is the nonresponders that we wish to identify and change treatment for as early as possible. In other words, predicting response is worthy, but responders do not need to have their treatment changed, whereas it is the nonresponders who require a change of treatment. We followed the general principles described by the Cochrane diagnostic test review group (37). An a priori written protocol was published in the PROSPERO database (International Prospective Register of Systematic Reviews; registration number CRD42012002905), and further protocol details are given in the online data supplement accompanying the online version of this article.

Inclusion Criteria

We included all studies that examined the identification of responders to an antipsychotic at follow-up by the degree of improvement in overall symptoms of schizophrenia at 2 weeks as measured by the PANSS total score (38) or the total score on the Brief Psychiatric Rating Scale (BPRS) (39), irrespective of the design (randomized controlled trial or naturalistic study), the setting (inpatient or outpatient), and the blinding procedure. The follow-up time to assess response was preferably 6 weeks, but it could vary from a minimum of 4 to a maximum of 12 weeks (28). We applied no language restriction, to avoid the problem of language bias (40). Any antipsychotic marketed in at least one country was included. The antipsychotic had to be given orally, since intramuscular formulations are usually given either for short-term treatment of acute agitation or as depot injections for long-term relapse prevention. We excluded fixed doses below the lower bounds of the target dosage ranges suggested by the International Consensus Study of Antipsychotic Dosing (41). In flexible-dosage studies, the upper limit had to at least include the lower bound of the target dosage range as suggested by the International Consensus Study of Antipsychotic Dosing (41). Lower dosages were acceptable for studies in first-episode patients or adolescents (41). We did not exclude studies on the basis of the speed of dosage titration, but in a post hoc sensitivity analysis we tested whether the exclusion of studies that applied a slow titration changed the overall results. Other concomitant medications, such as benzodiazepines, antidepressants, or anticholinergics, were allowed since these are usually given in routine clinical care, thus enhancing generalizability of the findings. We included people with acute exacerbations of schizophrenia (no restriction in age, setting, gender, ethnicity) or schizophrenia-like psychoses (schizophreniform and schizoaffective disorders), irrespective of the diagnostic criteria used.

Outcome Variables

Index test: lack of improvement at 2 weeks. Early nonimprovement of symptoms was primarily defined as less than 20% reduction of the total PANSS score from baseline to 2 weeks because a number of individual studies had previously identified this cutoff as predictive (12, 15, 29, 42). Moreover, studies have shown that this cutoff roughly means less than minimal improvement according to ratings on the Clinical Global Impressions scale (CGI) (43) made by clinicians (44, 45). Studies that used the BPRS (39) instead of the PANSS were also included, because the two scales are highly correlated (46). In secondary analyses, we examined the results of other index cutoffs of the total PANSS or BPRS score reduction, such as less than 0%, 10%, 15%, 25%, 30%, 40%, and 50%, for the primary definition of the reference standard. Furthermore, in an attempt to examine whether the degree of improvement in positive symptoms alone, instead of total score, could better predict nonresponse at endpoint, we did a post hoc assessment of several index cutoffs of the reduction in PANSS or BPRS positive symptom subscore at 2 weeks (less than 0%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, and 50%) for the primary definition of the reference standard. The PANSS positive symptom subscore was based on items 1 to 7, and the BPRS positive symptom subscore was based on items 4, 11, 12, and 15 (47).

Reference standard: nonresponse at endpoint/follow-up. Nonresponse was primarily defined as less than 50% reduction of the PANSS or BPRS total score from baseline to endpoint. Several studies have shown that this cutoff is clinically meaningful, roughly equal to a CGI rating of much improved (43) according to the equipercentile linking method (46, 48). Secondary definitions were the absence of a cross-sectional symptomatic remission as defined by Andreasen et al. (49) and less than 20% reduction of PANSS or BPRS total score.

Search

We searched electronic databases (EMBASE, MEDLINE, BIOSIS, PsycINFO, Cochrane Library, and CINAHL) by combining terms for multiple antipsychotic drugs and

schizophrenia with terms for prediction of response to treatment: (schizophrenia* or schizo* or psychotic*) AND (antipsychoti* or neurolept* or amisulpride or asenapine or benperidol or chlorpromazine or chlorprothixene or clopenthixol or clozapine or cyamemazine or droperidol or fluphenazine or flupenthixol or fluphenazine or haloperidol or iloperidone or levomepromazine or loxapine or lurasidone or mesoridazine or molindone or olanzapine or paliperidone or pericyazine or perphenazine or pimozide or pipothiazine or prochlorperazine or promazine or promethazine or quetiapine or risperidone or sertindole or stelazine or sulpiride or thioridazine or thiothixene or trifluoperazine or triflupromazine or ziprasidone or zotepine or zuclopenthixol) AND (early* or predict* or improvement* or nonrespon* or respon*). Moreover, we searched ClinicalTrials.gov and the International Clinical Trials Registry Platform for further relevant trials. We inspected the reference lists of all included studies and previous narrative reviews on the topic. We contacted the first or corresponding authors of all relevant studies to ask for individual patient data, missing information, and information about further studies.

FIGURE 1. Explanation of Sensitivity, Specificity, and Positive and Negative Predictive Values in a Diagnostic Test Meta-Analysis of Studies on Lack of Early Improvement As a Predictor of Nonresponse to Antipsychotics

Reference Standard (endpoint)	Index Test (2 weeks)	
	Negative Index Test: Patient Shows Improvement at 2 Weeks (percentage score reduction higher than defined cutoff)	Positive Index Test: Patient Does Not Show Improvement at 2 Weeks (percentage score reduction lower than defined cutoff)
Negative reference standard: patient is a responder	*a*	*b*
Positive reference standard: patient is a nonresponder	*c*	*d*

Sensitivity, also called True Positive Rate

$d/(c+d)$

The probability that a patient having the target condition will be correctly identified as such by the index test; in this analysis, the probability that *a nonresponder* will also be rated as *not improved at 2 weeks*

Specificity, also called True Negative Rate

$a/(a+b)$

The probability that a patient not having the target condition will be correctly identified as such by the index test; in this analysis, the probability that *a responder* will also be rated as *improved at 2 weeks*

Positive Predictive Value

$d/(b+d)$

The probability that a patient with a positive index test result will indeed have the target condition; in this analysis, the probability that a patient rated as *not improved* at 2 weeks will be *a nonresponder* at endpoint

Negative Predictive Value

$a/(a+c)$

The probability that a patient with a negative index test result will not have the target condition; in this analysis, the probability that a patient rated as *improved at 2 weeks* will be *a responder* at endpoint

Note that, mathematically, positive and negative predictive values directly depend on prevalence:

PPV = (sens * prev) / [sens * prev + (1 − spec) * (1 − prev)]

NPV = [spec * (1 − prev)] / {[(1 − sens) * prev] + [spec * (1 − prev)]}

PPV: positive predictive value; NPV: negative predictive value; sens: sensitivity; prev: prevalence; spec: specificity.

Data Extraction

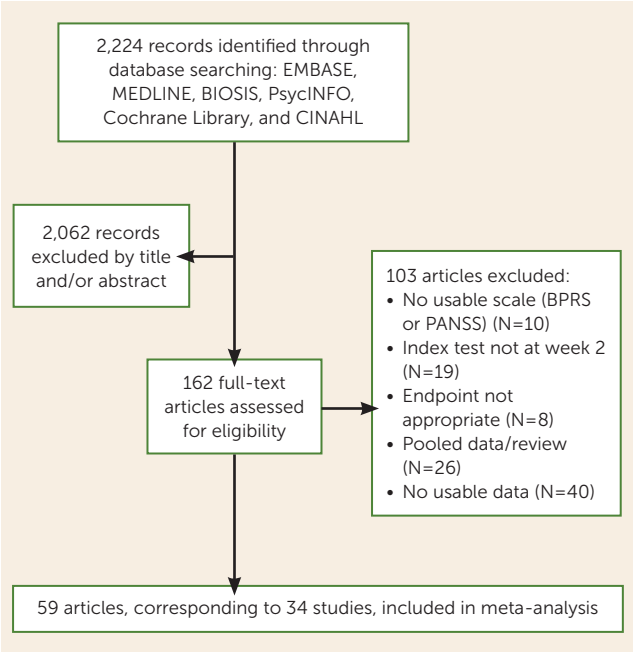
Two reviewers (C.L., M.T.S.) independently selected the studies and extracted the data. We reconstructed the diagnostic two-by-two tables (true positive, false positive, true negative, and false negative index test results) for each study on the basis of individual patient data. For this purpose, the original study authors had sent us either their entire data sets or the exact required numbers. Intention-to-treat data sets were used in the primary analysis. In a sensitivity analysis of the primary outcome, we examined whether the use of data sets of observed cases (in other words, completer data) would lead to substantial differences. In addition to the values needed for the two-by-two tables, our standardized extraction sheets included characteristics of participants (diagnosis, sex, age, baseline severity, and duration of illness), index tests and reference standards, antipsychotic drug and dosage used, and study methods.

Methodological quality was assessed by using the Quality Assessment of Diagnostic Accuracy Studies–2 (QUADAS-2) (50). QUADAS-2 consists of four key domains that refer to patient selection, index test, reference standard, and flow of patients through the study as well as timing of the index tests and reference standard ("flow and timing"). Each domain is assessed for the risk of bias, and the first three domains are also assessed in terms of applicability. We classified each of the seven items as having "low" (adequately addressed), "high" (inadequately addressed), or "unclear" risk of bias or concern for applicability. Disagreements were resolved by discussion and, if necessary, by consulting a third review author (S.L.). We did not exclude studies on the basis of this assessment, but we did evaluate the overall quality of the available data.

Meta-Analytic Calculations

A bivariate logitnormal random-effects meta-analysis, using a nonlinear mixed model approach, was performed to calculate summary estimates of sensitivity and specificity. The bivariate method models the logits of sensitivity and specificity in one model and allows for correlation between the two (36). The logit is the natural logarithm of sensitivity (or specificity) divided by 1 minus sensitivity (or specificity).

FIGURE 2. PRISMA Diagram of Process for Selecting Studies on Lack of Early Improvement As a Predictor of Nonresponse to Antipsychotics[a]



[a] PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (54).

Study estimates and a summary point, with its 95% confidence region, were plotted in a summary receiver-operating characteristic (SROC) plot. The bivariate approach was used to calculate summary estimates of positive and negative predictive values (51). Since predictive values depend highly on the prevalence of the condition examined (in our case, nonresponse), all estimates referred to the mean prevalence of the studied population. A diagram relating predictive values with prevalence was also created.

Most variation in sensitivity would be explained by variation in specificity, and vice versa (37). Therefore, in diagnostic accuracy meta-analyses, statistical tests and I-squares, as used in meta-analyses of interventions, are not helpful to test for heterogeneity. We assumed heterogeneity to be present in our data, and we dealt with it by using random effects models and by investigating potential sources of heterogeneity.

For this purpose we assessed the effects of sex, age, study design (randomized controlled trials versus naturalistic studies), blinding, placebo use, class of antipsychotic drug (first- versus second-generation), fixed or flexible dosing schedule, sponsorship (whether the study was sponsored by the pharmaceutical company manufacturing the comparator drug or not), week of response assessment, first-episode patients only versus mixed or chronically ill populations, baseline severity, and duration of illness on the summary estimates of the primary outcome. We included these variables one by one as covariates in the bivariate model. The difference of each subgroup from the group without the specified feature was reflected by a delta estimate in percentage with a 95% confidence interval. A covariate was assumed to have a significant effect on the estimates of sensitivity and specificity and, thus, to explain some of the heterogeneity in the sample if the p value was <0.05.

In three post hoc sensitivity analyses we 1) examined amisulpride, haloperidol, olanzapine, and risperidone separately (the other drugs did not have enough data, i.e., more than three studies, to be entered in a diagnostic test meta-analysis), 2) excluded a single study of treatment-resistant patients (52), and 3) excluded the few studies without a quick titration of the dosage, defined as not reaching the target dosage at the third day. The third sensitivity analysis entailed mainly studies on first-episode patients and adolescents (53). For the statistical analyses, the program SAS 9.2, with NLMIXED, was used.

RESULTS

The results are presented in terms of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). As these terms could be confusing in a diagnostic test review for predicting nonresponse, explanations are provided in Figure 1. From a clinical point of view, specificity and PPV are more important than sensitivity and NPV since the diagnostic test should mainly 1) ensure that the antipsychotic is not changed unnecessarily when the patient still has a good chance of responding (confirmed by high values of specificity) and 2) predict nonresponse satisfactorily (confirmed by high values of PPV). For example, suppose that a diagnostic test for predicting nonresponse based on nonimprovement at 2 weeks has 86% specificity and 90% PPV, while sensitivity is 69% and NPV is 60%. That means that, from all responders at endpoint, 86% will be identified as such at 2 weeks on the basis of their early improvement (specificity) and that a patient showing nonimprovement at 2 weeks will have 90% probability of being a nonresponder at endpoint (PPV). On the other hand, from all nonresponders at endpoint, 69% will be identified as such at 2 weeks on the basis of their early nonimprovement (sensitivity), and a patient showing improvement at 2 weeks will have 60% probability of being a responder at endpoint (NPV). For simplicity, in the preceding example we had to assume that PPV and NPV referred to the mean prevalence of nonresponse of the studied population (65% in this example).

Search Results

Figure 2 presents the PRISMA flow diagram (based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (54) of the search. By the last search in February 2014, 34 studies were included in our analysis. Individual patient data were obtained for 32 of the 34 studies (94.1%).

Table S1 in the online data supplement presents characteristics of all individual studies. Most studies (N=29, 85.3%) were randomized (five open and 24 double-blind), whereas five (14.7%) were observational. Eight studies (23.5%) lasted 4 weeks, 23 (67.6%) lasted 6 weeks, and 3 (8.8%) lasted longer than 6 weeks. The participants' mean age was 34.9 (SD=6.6) years, and the mean duration of illness was 11.5 (SD=5.6)

FIGURE 3. Quality Assessment of Studies on Lack of Early Improvement As a Predictor of Nonresponse to Antipsychotics

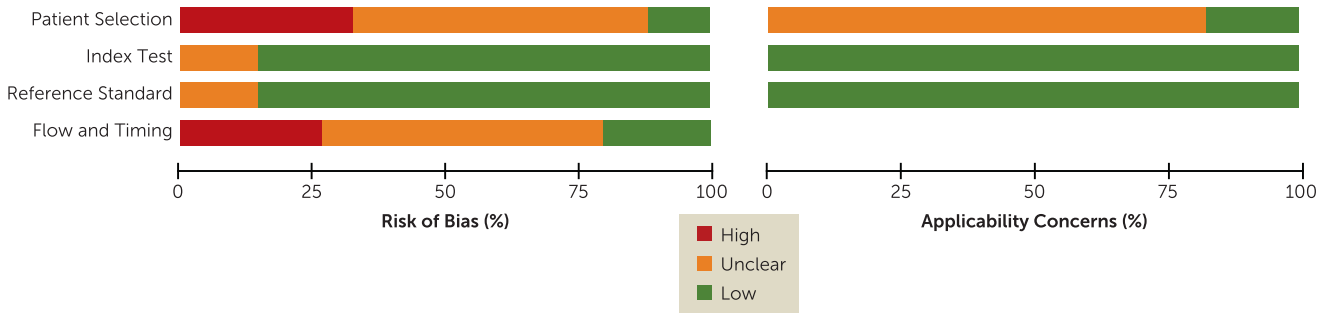
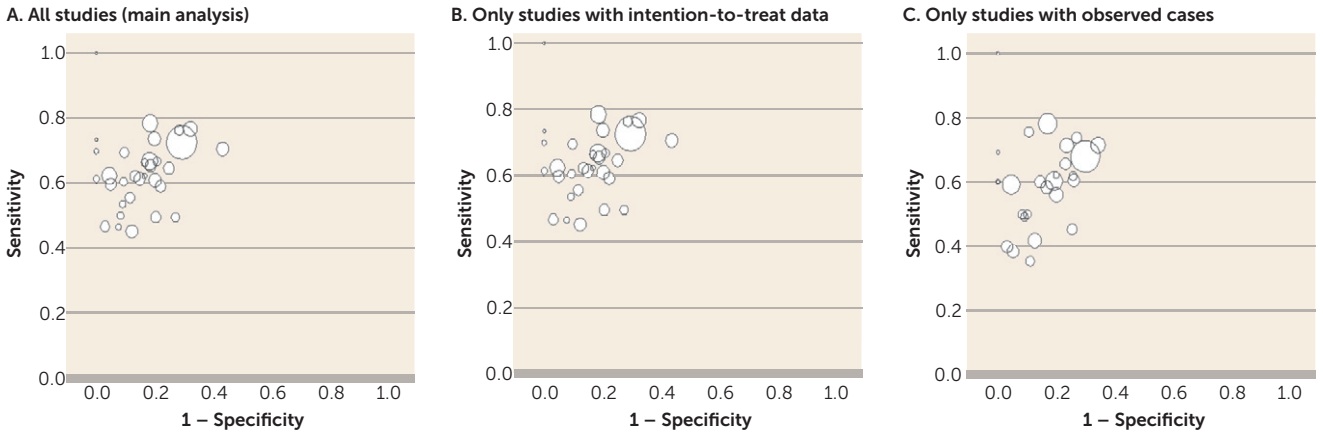


FIGURE 4. Estimates of Sensitivity and Specificity From Individual Studies on Lack of Early Improvement As a Predictor of Nonresponse to Antipsychotics[a]



[a] The analysis examined lack of improvement, defined as <20% symptom score reduction at 2 weeks (index test), as a predictor of nonresponse, defined as <50% symptom score reduction at endpoint (reference standard). For individual studies, sensitivity was plotted against 1 − specificity. The size of the bubbles indicates sample size.

years. Six studies (17.6%) included only first-episode patients, one study (2.9%) included patients with recent onset of schizophrenia, one study (2.9%) included treatment-resistant patients, and two studies (5.9%) included adolescents. Altogether, 9,460 patients were included in the meta-analysis. About half of the patients (52.6%, 4,976 of 9,457) did not show early improvement (i.e., they had <20% PANSS or BPRS score reductions at 2 weeks). From the patients who did not show early improvement, 88.4% (3,979 of 4,502) did not respond to treatment according to the criterion of 50% PANSS or BPRS reduction at endpoint, 83.7% (3,905 of 4,665) did not respond according to the symptomatic remission criteria of Andreasen et al. (49), and 58.8% (2,925 of 4,976) did not respond according to the criterion of 20% PANSS or BPRS reduction at endpoint.

Quality Assessment

Since the majority of included studies (29 of 34) were randomized controlled trials, it can be argued that there was a limitation in representativeness in the domain of patient selection. In terms of index tests and reference standards, the direct cooperation with the authors of almost all the original trials minimized the risk of bias and the concerns about applicability, because in contrast to what was originally reported, early improvement and later response could always be based on the same definitions. In the domain of flow and timing, the most common problem was a high drop-out rate (mean=29.2%) in many trials (see Figure 3).

Outcomes

1. Did a <20% reduction of total PANSS or BPRS score from baseline to week 2 predict nonresponse at endpoint (4–12 weeks)?

a. The primary outcome tested was whether a less than 20% PANSS or BPRS reduction at week 2 (pooled mean across all studies: 47.2%) predicted a less than 50% PANSS or BPRS reduction at endpoint (pooled mean across all studies: 63.7%). The diagnostic test indicated that it did so with a specificity of 86% and PPV of 90%. The sensitivity was 63% and the NPV was 53%, but as explained above the latter two parameters are less important for our question. The sensitivity analysis using data for observed cases did not substantially change the results; specificity was 86%, PPV 85%, sensitivity 59%, and NPV 60% (Figure 4).

TABLE 1. Meta-Analysis Summary Estimates for Lack of Early Improvement Cutoffs (0%–15%) As Predictors of Nonresponse (<20%) to Antipsychotics at Endpoint[a]

Cutoff Defining Lack of Early Improvement	Sensitivity	CI	Specificity	CI	PPV	CI	NPV	CI
≤0%	0.45	0.40–0.50	0.95	0.93–0.97	0.81	0.77–0.85	0.78	0.74–0.82
<5%	0.54	0.51–0.58	0.91	0.89–0.93	0.75	0.70–0.79	0.80	0.76–0.84
<10%	0.66	0.63–0.70	0.86	0.83–0.89	0.70	0.65–0.74	0.84	0.80–0.87
<15%	0.77	0.74–0.80	0.80	0.75–0.84	0.65	0.61–0.70	0.87	0.84–0.90

[a] Each diagnostic test meta-analysis examined lack of improvement at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard). Lack of early improvement and nonresponse refer to overall symptom score reduction on the PANSS or BPRS. PPV: positive predictive value; NPV: negative predictive value.

b. A secondary outcome was remission at endpoint. Remission was predicted with a specificity of 77%, a PPV of 88%, a sensitivity of 61%, and an NPV of 42%.

c. Another secondary outcome tested was whether a less than 20% PANSS or BPRS reduction at endpoint would be predicted. It was predicted with a specificity of 70%, a PPV of 55%, a sensitivity of 85%, and an NPV of 91%. As the prechosen index cutoff of 20% reduction at 2 weeks did not seem to work well enough for this secondary reference standard, we did a post hoc analysis of index cutoffs ranging from 0% to 15% PANSS or BPRS reduction at week 2 to predict 20% reduction at endpoint in order to obtain higher specificity and PPV. The results showed that a PANSS or BPRS reduction of less than 0% from baseline to week 2 predicted nonresponse adequately (specificity 95%, PPV 81%). The summary estimates for this secondary outcome are presented in Table 1.

2. Which factors influenced the accuracy of the diagnostic test and in what way?

a. We examined the time point at which nonresponse was assessed. When week 4 rather than week 6 or later was used as the endpoint, the specificity of the diagnostic test was higher (93% for week 4 versus 82% for week 6 or later, $p<0.0005$).

b. The effect of the baseline severity of illness (as measured by the PANSS total score) was another significant factor. Greater severity at baseline was associated with higher specificity of the diagnostic test ($p<0.0001$).

c. The duration of illness was another factor that affected the specificity of the test. A shorter duration of illness was associated with higher specificity ($p<0.04$). The initial investigation of significant moderators also included age. However, age and duration of illness were highly correlated. When age and duration of illness were included in one model, the effect of age was no longer significant; it was completely explained by the duration of illness ($p<0.02$).

d. Effects were nonsignificant for sex, study design (randomized controlled trials versus naturalistic studies), blinding, placebo use, class of antipsychotic drug (first- versus second-generation), fixed or flexible dosing schedule, sponsorship, and first-episode patients only versus mixed or chronically ill study groups. Results are presented in Table 2.

3. How well did the other cutoffs for reduction of the total PANSS or BPRS score from baseline to week 2 predict nonresponse at endpoint? The sensitivity, specificity, PPV, and NPV of each cutoff point of the index test, ranging from 0% to 50% PANSS or BPRS reduction, are given in Table 3. Higher cutoffs were associated with lower specificity and PPV, and higher sensitivity and NPV. Thus, the cutoff of 0% reduction from baseline to week 2 showed the highest specificity and PPV while the cutoff of 50% reduction showed the highest sensitivity and NPV (Figure 5). Figure 5 shows that the best trade-off between sensitivity and specificity lies around a cutoff of 20%−25% reduction, confirming our initial choice. As for our research question, high specificity and PPV are more important than sensitivity and NPV; thus, using 20% rather than 25% as a cutoff appears justifiable.

PPV and NPV depend on the prevalence of the condition (here nonresponse). We primarily used the actual prevalence of nonresponse in the included studies for the computation of predictive values. Figure 6 illustrates how these measures would fluctuate for different values of prevalence. As is typical for diagnostic tests, high prevalence of the condition (nonresponse at endpoint) leads to higher PPV and lower NPV (displayed as 1 − NPV in Figure 6). Figure S1 in the online data supplement shows the relation between predictive values and prevalence for all cutoffs.

4. Did the reduction in the PANSS or BPRS positive symptom score from baseline to week 2 predict nonresponse at endpoint more accurately? The test characteristics of each cutoff (ranging from 0% to 50%) for reduction on the PANSS or BPRS positive symptom subscale are given in Table 4. The use of the positive symptoms reduction yielded results similar to those for the total score reduction at week 2. A reduction of <20% in PANSS or BPRS positive symptoms score resulted in a sensitivity of 57%, a specificity of 87%, PPV of 87%, and NPV of 55%. As again specificity and PPV are most important for clinical purposes, the cutoff of <20% reduction in positive symptom score could also be considered to be a reasonable choice for the index test, but this analysis was conducted post hoc.

5. Post hoc sensitivity analyses. The results of the diagnostic test appeared to be equally applicable to the four individual antipsychotics that presented enough data to be examined separately (amisulpride, haloperidol, olanzapine, and risperidone; see Table 5). Removing the single study in treatment resistant patients (52) and excluding studies that did not follow a quick titration schedule did not markedly change the diagnostic test results (see Table 5).

DISCUSSION

We present a diagnostic meta-analysis with 34 studies and 9,460 participants that examined the question of whether nonimprovement at week 2 predicts later nonresponse to antipsychotics in patients with schizophrenia spectrum disorders. The major strength of this study is that we were able to obtain individual patient data for almost all trials. The analysis suggested that out of 100 patients showing nonimprovement at week 2 (<20% PANSS or BPRS score reduction), 90 will not show much improvement at endpoint (<50% PANSS or BPRS score reduction), 88 will not achieve symptomatic remission at endpoint, and 55 will not even minimally improve (<20% PANSS or BPRS score reduction).

A ≥50% PANSS/BPRS score reduction from baseline to endpoint is a clinically meaningful definition of response for patients with acute exacerbations of schizophrenia, because it roughly corresponds to "much improvement" as assessed with the CGI (44, 48). Contrary to common belief, symptomatic remission (49) has been shown to occur with a frequency similar to that for 50% PANSS or BPRS reduction (55) and, as a reference standard (here nonremission), yielded results similar to those for the <50% reduction in the diagnostic test meta-analysis. On the other hand, a ≥20% PANSS or BPRS reduction is a much looser definition of response, resulting in a higher number of responders at endpoint (the denominator of specificity) and a significant decrease in specificity and PPV. As a ≥20% PANSS/BPRS reduction reflects only "minimal improvement" (44, 48), it may not be a good indicator of response (compared with ≥50% and remission). However, a <20% reduction is an extremely stringent measure of nonresponse; most clinicians would change treatment for a patient not even minimally improved after 6 weeks. If one requires at least 80% specificity and PPV for that reference standard, the index cutoff of 0% PANSS or BPRS reduction at week 2 should be applied.

TABLE 2. Meta-Analysis Summary Estimates for Potential Moderators of Lack of Early Improvement (<20%) As a Predictor of Nonresponse (<50%) to Antipsychotics at Endpoint[a]

Potential Moderator	b[b]	SE	p	CI
Sex				
Sensitivity	−0.00	0.01	0.96	−0.02 to 0.02
Specificity	−0.02	0.01	0.18	−0.05 to 0.01
Study design (randomized controlled trial versus naturalistic study)				
Sensitivity	−0.18	0.23	0.42	−0.63 to 0.26
Specificity	−0.62	0.43	0.15	−1.47 to 0.22
Blinding				
Sensitivity	−0.15	0.17	0.37	−0.48 to 0.18
Specificity	−0.31	0.33	0.36	−0.95 to 0.34
Placebo use				
Sensitivity	−0.16	0.18	0.37	−0.51 to 0.19
Specificity	0.27	0.38	0.49	−0.48 to 1.02
Class of antipsychotic drug (first- versus second-generation)				
Sensitivity	0.03	0.15	0.84	−0.26 to 0.32
Specificity	−0.12	0.29	0.68	−0.70 to 0.45
Fixed versus flexible dosing				
Sensitivity	0.04	0.16	0.83	−0.28 to 0.35
Specificity	−0.45	0.30	0.14	−1.04 to 0.14
Sponsorship				
Sensitivity	0.28	0.20	0.17	−0.11 to 0.67
Specificity	−0.34	0.40	0.39	−1.12 to 0.44
Response assessment at week 4 versus week 6 or later				
Sensitivity	−0.03	0.18	0.87	−0.38 to 0.32
Specificity	1.07	0.30	<0.0005[c]	0.49 to 1.66
First-episode versus mixed or chronically ill patients				
Sensitivity	−0.33	0.20	0.10	−0.73 to 0.06
Specificity	−0.04	0.39	0.92	−0.80 to 0.72
Baseline severity				
Sensitivity	−0.01	0.01	0.12	−0.02 to 0.00
Specificity	0.05	0.01	<0.0001[c]	0.02 to 0.07
Age[d]				
Sensitivity	0.03	0.03	0.42	−0.04 to 0.09
Specificity	0.09	0.06	0.15	−0.03 to 0.20
Duration of illness[d]				
Sensitivity	0.01	0.03	0.71	−0.05 to 0.07
Specificity	−0.14	0.06	<0.02[c]	−0.25 to −0.03

[a] Each diagnostic test meta-analysis examined lack of early improvement at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard).
[b] Estimate of β covariate.
[c] Statistically significant ($p<0.05$).
[d] Age and duration of illness are included in the same model.

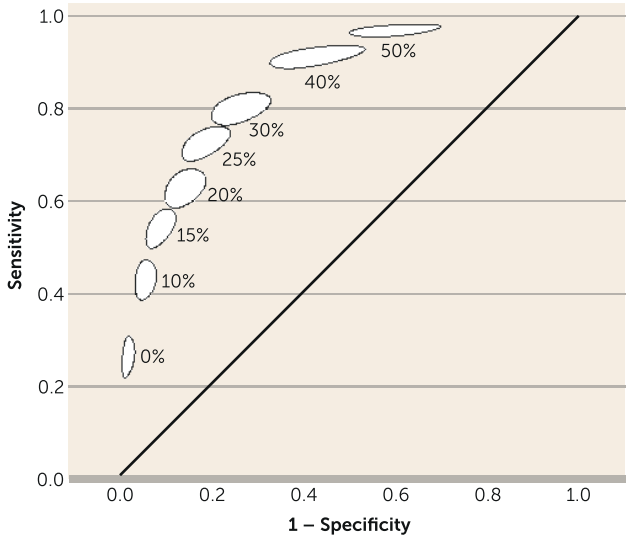
In research on the prediction of response to antipsychotics, many potential predictors have been identified, including early subjective response (56), severity of illness, homovanillic acid level (57, 58), structural changes shown by cranial imaging (59–61), and polymorphisms of brain receptor genes (62, 63). However, so far, none of these potential predictors has led to the development of a clinically useful decision-making tool. Early improvement in antipsychotic treatment is the strongest among those predictors (64–66), and it is now well replicated and could be implemented in clinical practice. Although many previous studies suggested an association between early improvement and later response (9, 11, 12, 15, 19, 20, 23, 26, 29, 67), a lack of consensus regarding the definitions of these benchmarks has prevented formulation of straightforward clinical recommendations. For example, if one study used 50% PANSS total score reduction to define ultimate response (65, 68) while another one used cross-sectional remission (69), it is difficult to summarize their findings. Moreover, the individual studies

TABLE 3. Meta-Analysis Summary Estimates for Lack of Early Improvement Cutoffs (0%–50%) As Predictors of Nonresponse (<50%) to Antipsychotics at Endpoint[a]

Cutoff Defining Lack of Early Improvement	Sensitivity	CI	Specificity	CI	PPV	CI	NPV	CI
≤0%	0.26	0.23–0.30	0.98	0.96–0.98	0.95	0.93–0.96	0.39	0.32–0.46
<10%	0.43	0.40–0.47	0.94	0.92–0.96	0.93	0.91–0.95	0.45	0.38–0.51
<15%	0.54	0.51–0.58	0.91	0.88–0.93	0.92	0.89–0.94	0.49	0.42–0.56
<20%	0.63	0.59–0.66	0.86	0.82–0.89	0.90	0.86–0.91	0.53	0.49–0.61
<25%	0.73	0.69–0.76	0.81	0.77–0.85	0.88	0.85–0.91	0.59	0.52–0.65
<30%	0.80	0.77–0.83	0.74	0.68–0.79	0.86	0.82–0.89	0.64	0.59–0.70
<40%	0.92	0.89–0.93	0.57	0.48–0.66	0.80	0.76–0.84	0.77	0.72–0.82
<50%	0.97	0.96–0.98	0.39	0.31–0.48	0.76	0.72–0.80	0.87	0.81–0.91

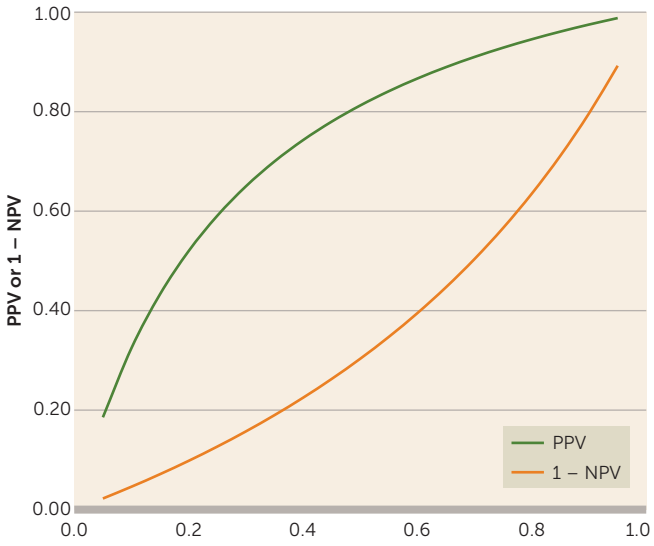
[a] Each diagnostic test meta-analysis examined lack of improvement at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard). Lack of early improvement and nonresponse refer to overall symptom score reduction on the PANSS or BPRS. PPV: positive predictive value; NPV: negative predictive value.

FIGURE 5. Meta-Analysis Summary Estimates of Sensitivity and Specificity for Lack of Early Improvement Cutoffs (0%–50%) As Predictors of Nonresponse (<50%) to Antipsychotics at Endpoint[a]



[a] Each diagnostic test meta-analysis examined lack of improvement at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard). Lack of early improvement and nonresponse refer to overall symptom score reduction on the PANSS or BPRS. Each ellipse represents the 95% confidence region.

FIGURE 6. Effect of Nonresponse Prevalence on Lack of Early Improvement As a Predictor of Nonresponse to Antipsychotics at Endpoint[a]



[a] The diagnostic test meta-analysis examined lack of improvement at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard). Lack of early improvement and nonresponse refer to overall symptom score reduction on the PANSS or BPRS. Lack of early improvement was defined as a reduction in score <20%; nonresponse was defined as a reduction in score <50%. As positive predictive value (PPV) and negative predictive value (NPV) depend on the prevalence of the condition (here nonresponse), we plotted the values of PPV (green curve) and 1 − NPV (orange curve) versus the prevalence of nonresponse. The plot shows that, as the prevalence of nonresponse increases, PPV increases whereas NPV (here shown as 1 − NPV) decreases.

usually attempted to derive the best cutoff by post hoc analyses. In the current review, improvement and response were defined a priori.

The specificity of the diagnostic test was shown to be influenced by three independent factors. First, the assessment of final nonresponse at week 4 was associated with higher specificity of the diagnostic test than was assessment of nonresponse at week 6 or later. The number of responders at endpoint (specificity's denominator) is expected to increase at later endpoints, and thus specificity decreases. Second, higher baseline illness severity was associated with higher specificity of the diagnostic test. For the mean baseline severity of the included patients (score of 97 points on PANSS items 1–7), the specificity was 86%; for 10 points lower baseline severity (87 points), it was 79%; and for 10 points higher (107 points), the specificity increased to 91%. Third, shorter illness duration was associated with higher specificity of the diagnostic test. For the mean illness duration of the included patients (11.5 years), the specificity was 87%; for a duration 5 years shorter, the specificity was 91%; and for a duration 5 years longer, it was 82%.

Our meta-analysis has several limitations, some of which are illustrated by the quality assessment with the QUADAS tool (Figure 2). Of the 34 included studies, 29 were

TABLE 4. Meta-Analysis Summary Estimates for Lack of Early Improvement in Positive Symptom Cutoffs (0%–50%) As Predictors of Nonresponse (<50%) to Antipsychotics at Endpoint[a]

Cutoff Defining Lack of Early Improvement in Positive Symptoms	Sensitivity	CI	Specificity	CI	PPV	CI	NPV	CI
≤0%	0.31	0.27–0.36	0.96	0.94–0.97	0.92	0.88–0.95	0.47	0.39–0.55
<5%	0.34	0.30–0.37	0.95	0.93–0.97	0.91	0.87–0.95	0.47	0.39–0.55
<10%	0.41	0.38–0.45	0.93	0.90–0.95	0.90	0.85–0.93	0.49	0.41–0.58
<15%	0.49	0.46–0.53	0.90	0.86–0.93	0.89	0.84–0.93	0.52	0.44–0.61
<20%	0.57	0.53–0.60	0.87	0.82–0.90	0.87	0.82–0.91	0.55	0.47–0.63
<25%	0.64	0.60–0.68	0.82	0.77–0.86	0.85	0.79–0.90	0.58	0.50–0.66
<30%	0.71	0.67–0.75	0.76	0.71–0.81	0.83	0.77–0.88	0.62	0.54–0.69
<40%	0.82	0.79–0.85	0.63	0.56–0.70	0.78	0.72–0.84	0.68	0.60–0.75
<50%	0.89	0.87–0.91	0.51	0.43–0.58	0.75	0.68–0.81	0.73	0.67–0.79

[a] Each diagnostic test meta-analysis examined lack of improvement in positive symptoms at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard). Lack of early improvement in positive symptoms and nonresponse refer to symptom score reduction (positive and overall symptoms, respectively) on the PANSS or BPRS. PPV: positive predictive value; NPV: negative predictive value.

TABLE 5. Summary Estimates From Sensitivity Analyses of Lack of Early Improvement (<20%) As a Predictor of Nonresponse (<50%) to Antipsychotics at Endpoint[a]

Analysis	Sensitivity	CI	Specificity	CI	PPV	CI	NPV	CI
Observed cases only	0.59	0.54–0.64	0.86	0.82–0.89	0.85	0.80–0.88	0.60	0.53–0.67
Without treatment-resistant patients	0.62	0.59–0.66	0.86	0.82–0.89	0.89	0.86–0.92	0.55	0.48–0.61
Quick titration only	0.62	0.58–0.66	0.86	0.80–0.91	0.89	0.85–0.93	0.54	0.44–0.63
Individual drugs								
Amisulpride	0.60	0.51–0.68	0.91	0.87–0.94	0.86	0.80–0.91	0.70	0.64–0.76
Haloperidol	0.62	0.55–0.68	0.89	0.81–0.94	0.91	0.86–0.95	0.54	0.41–0.67
Olanzapine	0.65	0.61–0.70	0.80	0.74–0.85	0.91	0.84–0.95	0.46	0.39–0.54
Risperidone	0.65	0.57–0.72	0.80	0.68–0.88	0.81	0.65–0.91	0.60	0.41–0.76

[a] Each diagnostic test meta-analysis examined lack of improvement at 2 weeks (index test) as a predictor of nonresponse at endpoint (reference standard). Lack of early improvement and nonresponse refer to overall symptoms score reduction on the PANSS or BPRS. PPV: positive predictive value; NPV: negative predictive value.

randomized controlled trials and may thus not accurately represent routine clinical practice (70, 71). However, whether the study was a randomized trial was not a significant moderator of the test performance. As for the high dropout rates usually seen in schizophrenia trials, the comparison of strict intention-to-treat and observed-case results in a sensitivity analysis did not show any significant difference, corroborating the validity of the results.

As all studies were pooled in the primary analysis, it is unclear whether the results apply to all antipsychotics. We had enough data on only four antipsychotics to allow for a comparison of the diagnostic test results among them, but these four drugs may represent a good selection because they cover drugs with quite different profiles. Amisulpride is a selective dopamine receptor antagonist that has no effects on histaminergic receptors and is not sedating. Haloperidol is a high-potency first-generation antipsychotic. Olanzapine and risperidone are frequently used second-generation antipsychotics that block serotonin 5-HT_{2a} receptors more than dopamine receptors, but risperidone produces more extrapyramidal symptoms and prolactin increase, while olanzapine has a higher risk of weight gain and has stronger effects on histaminergic receptors. No obvious difference among these antipsychotics was suggested, but additional analyses of other antipsychotics would be important.

Moreover, when a patient with schizophrenia is administered an antipsychotic medication, immediate anti-anxiety and anti-agitation effects, as well as side effects such as sedation, could be wrongly conceived as early improvement without necessarily an improvement in core symptoms of schizophrenia. In the same vein, the concomitant administration of benzodiazepines and/or adjunctive sleep medication, which were allowed in almost all included trials, could have biased the diagnostic test results, although this is similar to clinical practice, where such drugs are frequently coprescribed as well. We therefore examined whether the use of positive symptoms, instead of overall symptoms, as the index test would change the performance of the diagnostic test, but the results did not change markedly.

Furthermore, our data set contained mainly studies of chronically ill patients. Several studies have shown that response patterns in first-episode patients may differ from those of chronically ill patients, in that at least a subgroup can show later onset of response (17, 72, 73). Thus, although illness phase was not a significant moderator, this may have been due to an insufficient number of first-episode studies (N=6). Similarly, treatment-resistant patients were represented by only one study in our analysis, and its exclusion in a sensitivity analysis did not change the overall performance of the diagnostic test. Although there is some preliminary

evidence that the majority of improvement with antipsychotics occurs relatively early in the course of treatment for treatment-resistant patients as well (74), a number of studies suggest that longer-term trials are needed when investigating response in this particular subgroup (75–78). Therefore, the application of our results is more appropriate for patients who are neither in their first episode of schizophrenia nor exhibiting treatment resistance.

Finally, the translation and scalability of the findings of this meta-analysis to clinical care depend on the use of measurement-based approaches in usual care settings. Since the PANSS and BPRS are not routinely used by clinicians, the well-established correlation between the simple CGI improvement scale and the change in PANSS or BPRS total score (44, 46, 48) can be taken into account. These analyses have roughly showed that a 20% PANSS or BPRS reduction (our index test) corresponds to minimal improvement on the CGI and that a 50% score reduction (our primary reference standard) corresponds to much improvement. Indeed, a recent naturalistic study that used solely CGI improvement ratings of less than minimally improved at 4 weeks to predict ultimate nonresponse at 12 weeks, defined as less than much improved, confirmed the utility of this approach (67).

Despite the limitations, the current meta-analysis provides good evidence that nonimprovement at week 2 can be used for a clinically meaningful prediction of later nonresponse, saving patients from unnecessary long-term exposure to an antipsychotic that is unlikely to help them. Notably, some important treatment guidelines, such as those of PORT (31) and the World Federation of Societies of Biological Psychiatry (32), have already incorporated such statements. It is also crucial to emphasize that, before nonimprovement is established, patients should have received the antipsychotic at a sufficiently high dose. In this meta-analysis, dose titration schedule had no significant effect on the performance of the diagnostic test, but most studies followed a quick titration schedule (target doses were reached within 3 days). Therefore, in order to avoid premature changes of treatment, we caution that the results of this diagnostic test review should be applied only to patients who have received target doses (41)—we suggest even near the upper limits of these ranges—for at least 2 weeks. This is important, because in everyday clinical practice, doctors often titrate slowly because of tolerability issues, which can be an obstacle to rapid dosage increase. Plasma level measurements can also be useful, e.g., to rule out rapid metabolism due to cytochrome P450 polymorphisms, although plasma levels can vary substantially in individual patients and are not always directly correlated with efficacy (79).

What this meta-analysis has not explored and future studies need to address is which treatment strategies should be applied in case of nonresponse. Dosage increase is not well studied. Switching has been examined by only a few studies, some of which, clearly underpowered, were negative (68, 80), while the largest one was positive (18). Last, augmentation studies have usually focused on treatment-resistant patients at later stages, and they were mainly negative (81). Results of ongoing studies on switching strategies in patients without early improvement, such as SWITCH (82) and OPTIMISE (83), are awaited for replication of previous findings (18), but alternative and hopefully more effective strategies, pharmacologic and/or psychosocial in nature, that meet clinicians' and, above all, patients' needs are warranted.

AUTHOR AND ARTICLE INFORMATION

From the Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, the Netherlands; Dr. Kurt Fontheim's Hospital for Mental Health, Liebenburg, Germany; Chonbuk National University Medical School, Jeonju, Korea; the Department of Psychiatry, Marqués de Valdecilla University Hospital-IDIVAL, School of Medicine, University of Cantabria, Santander, Spain; CIBERSAM (Centro Investigación Biomédica en Red Salud Mental), Madrid; the Department and Institute of Psychiatry, University of São Paulo Medical School, São Paulo, Brazil; the Department of Psychiatry, Juntendo University Nerima Hospital, Tokyo; the Department of Psychiatry, Ludwig-Maximilians University, Munich; the Department of Psychiatry, University of Halle-Wittenberg, Halle, Germany; Hofstra North Shore Long Island Jewish School of Medicine, Hempstead, N.Y.; the Zucker Hillside Hospital, Glen Oaks, N.Y.; the Department of Psychiatry and Psychotherapy, Centre of Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Kaohsiung Municipal Kai-Syuan Psychiatric Hospital, Kaohsiung, Taiwan; the Department of Psychiatry, School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; University Hospital of Child and Adolescent Psychiatry, University of Bern, Bern, Switzerland; and the Department of Biomedical and NeuroMotor Sciences, University of Bologna, Bologna, Italy.

Address correspondence to Dr. Stefan Leucht (stefan.leucht@lrz.tum.de).

Dr. Samara and Dr. Claudia Leucht contributed equally to this work.

Funded by grant 01KG1118 from the German Federal Ministry of Education and Research.

The authors thank Samantha Roberts for conducting the literature search.

Dr. Stefan Leucht has received honoraria for consulting/advisory boards from Alkermes, Bristol-Myers Squibb, Eli Lilly, Janssen, Johnson & Johnson, Medavante, and Roche; he has received lecture honoraria from AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Essex Pharma, Janssen, Johnson & Johnson, Lundbeck Institute, Pfizer, and Sanofi-Aventis; and he has received medication from Eli Lilly for a trial for which he was the primary investigator; Dr. Claudia Leucht is his spouse, and so the same conflicts apply to her; she has no other conflicts. Dr. Anghelescu was a full-time employee at Janssen Pharmaceutica, Beerse, Belgium, from October 2009 to September 2012. Dr. Chung has received grant support from Janssen Korea and Korea Otsuka Pharmaceuticals as well as served as a consultant for AstraZeneca, Janssen Korea, Korea Otsuka Pharmaceuticals, and Sanofi-Aventis. Dr. Crespo-Facorro has received honoraria for consulting/advisory boards from Otsuka Pharmaceuticals and lecture honoraria from Janssen Johnson & Johnson, Lundbeck, Roche, and Otsuka Pharmaceuticals. Dr. Elkis has received research grant support from Janssen, Roche, and FAPESP (São Paulo Research Foundation) as well as travel support from Janssen and Roche; he has been on advisory boards of Roche and Janssen and has served on the speakers' boards of Cristalia and Novartis. Dr. Hatta has received lecture honoraria from Eli Lilly, Janssen, Dainippon-Sumitomo, Otsuka, Astellas, Tanabe-Mitsubishi, Takeda, and Meiji and has received honoraria for consulting from Eli Lilly and Janssen. Dr. Kane has received honoraria for lectures and/or consulting from Alkermes, Amgen, Bristol-Myers Squibb, Cephalon, Eisai, Boehringer Ingelheim, Eli Lilly, Forrest, Genentech, Intracellular Therapeutics, Janssen, Johnson & Johnson, Lundbeck, Merck, Novartis, Otsuka, Pfizer, Pierre Fabre, Proteus, Reviva, Roche, Sunovion, and Targacept; he is a shareholder of MedAvante.

Dr. Kayo has acted as a consultant for Abbott, Aché, Eurofarma, Genzyme, GlaxoSmithKline, Sankyo, Takeda, and Uniao Quimica; has received grant support from Roche; and has received travel support from Eli Lilly and GlaxoSmithKline. Dr. Lambert has been a consultant and/or advisor to or has received honoraria from AstraZeneca, Bristol-Myers Squibb, Lilly Deutschland, Janssen Cilag, Lundbeck, Otsuka Pharma, Roche Deutschland, and Sanofi-Aventis. Dr. Schimmelmann has been a consultant and/or advisor to Eli Lilly and Shire. Dr. Serretti is or has been consultant/speaker for Abbott, Abbvie, Angelini, AstraZeneca, Clinical Data, Boheringer, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Innovapharma, Italfarmaco, Janssen, Lundbeck, Naurex, Pfizer, Polipharma, Sanofi, and Servier. Dr. Correll has been a consultant and/or advisor to or has received honoraria from Actelion, Alexza, Bristol-Myers Squibb, Cephalon, Eli Lilly, Genentech, Gerson Lehrman Group, IntraCellular Therapies, Janssen/J&J, Lundbeck, Medavante, Medscape, Merck, Otsuka, Pfizer, ProPhase, Roche, Sunovion, Supernus, Takeda, and Teva; he has received grant support from the American Academy of Child and Adolescent Psychiatry, BMS, Janssen/J&J, the National Institute of Mental Health (NIMH), Novo Nordisk, Otsuka, and the Thrasher Foundation. The other authors report no financial relationships with commercial interests.

Received Sept. 24, 2014; revision received March 5, 2015; accepted March 12, 2015; published online June 5, 2015.

REFERENCES

1. Creese I, Burt DR, Snyder SH: Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. Science 1976; 192:481–483
2. Seeman P, Lee T, Chau-Wong M, et al: Antipsychotic drug doses and neuroleptic/dopamine receptors. Nature 1976; 261:717–719
3. Grace AA, Bunney BS: Induction of depolarization block in midbrain dopamine neurons by repeated administration of haloperidol: analysis using in vivo intracellular recording. J Pharmacol Exp Ther 1986; 238: 1092–1100
4. Grace AA: The depolarization block hypothesis of neuroleptic action: implications for the etiology and treatment of schizophrenia. J Neural Transm Suppl 1992; 36:91–131
5. Grace AA, Bunney BS, Moore H, et al: Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. Trends Neurosci 1997; 20:31–37
6. Kapur S, Remington G: Dopamine D(2) receptors and their role in atypical antipsychotic action: still necessary and may even be sufficient. Biol Psychiatry 2001; 50:873–883
7. Agid O, Kapur S, Arenovich T, et al: Delayed-onset hypothesis of antipsychotic action: a hypothesis tested and rejected. Arch Gen Psychiatry 2003; 60:1228–1235
8. Leucht S, Busch R, Hamann J, et al: Early-onset hypothesis of antipsychotic drug action: a hypothesis tested, confirmed and extended. Biol Psychiatry 2005; 57:1543–1549
9. Chen L, Kinon BJ, Ascher-Svanum H, et al: Predicting response to atypical antipsychotics based on early response in the treatment of schizophrenia (abstract). Biol Psychiatry 2007; 61(8 suppl):S81
10. Leucht S, Busch R, Kissling W, et al: Early prediction of antipsychotic nonresponse among patients with schizophrenia. J Clin Psychiatry 2007; 68:352–360
11. Lin C-H, Chou L-S, Lin C-H, et al: Early prediction of clinical response in schizophrenia patients receiving the atypical antipsychotic zotepine. J Clin Psychiatry 2007; 68:1522–1527
12. Kinon BJ, Chen L, Ascher-Svanum H, et al: Predicting response to atypical antipsychotics based on early response in the treatment of schizophrenia. Schizophr Res 2008; 102:230–240
13. Leucht S, Shamsi SA, Busch R, et al: Predicting antipsychotic drug response—replication and extension to six weeks in an international olanzapine study. Schizophr Res 2008; 101:312–319
14. Chen L, Ascher-Svanum H, Stauffer V, et al: Optimal thresholds of early response to atypical antipsychotics: application of signal detection methods. Schizophr Res 2009; 113:34–40
15. Jäger M, Schmauss M, Laux G, et al: Early improvement as a predictor of remission and response in schizophrenia: results from a naturalistic study. Eur Psychiatry 2009; 24:501–506
16. Lambert M, Schimmelmann BG, Naber D, et al: Early- and delayed antipsychotic response and prediction of outcome in 528 severely impaired patients with schizophrenia treated with amisulpride. Pharmacopsychiatry 2009; 42:277–283
17. Derks EM, Fleischhacker WW, Boter H, et al: Antipsychotic drug treatment in first-episode psychosis: should patients be switched to a different antipsychotic drug after 2, 4, or 6 weeks of nonresponse? J Clin Psychopharmacol 2010; 30:176–180
18. Kinon BJ, Chen L, Ascher-Svanum H, et al: Early response to antipsychotic drug therapy as a clinical marker of subsequent response in the treatment of schizophrenia. Neuropsychopharmacology 2010; 35:581–590
19. Pelayo-Terán JM, Perez-Iglesias R, Mata I, et al: Early response to antipsychotics as a marker of treatment response of psychosis. Schizophr Res 2010; 117:180
20. Schennach-Wolff R, Seemüller FH, Mayr A, et al: An early improvement threshold to predict response and remission in first-episode schizophrenia. Br J Psychiatry 2010; 196:460–466
21. Anghelescu I, Janssens L, Kent J, et al: Sustained treatment response in schizophrenia to JNJ-37822681 can be predicted within three days. Eur Neuropsychopharmacol 2011; 21:S490–S491
22. Giegling I, Porcelli S, Balzarro B, et al: Antipsychotic response in the first week predicts later efficacy. Neuropsychobiology 2012; 66: 100–105
23. Stauffer VL, Case M, Kinon BJ, et al: Early response to antipsychotic therapy as a clinical marker of subsequent response in the treatment of patients with first-episode psychosis. Psychiatry Res 2011; 187:42–48
24. Stentebjerg-Olesen M, Pagsberg AK, Fink-Jensen A, et al: Can early-onset antipsychotic effect predict later clinical effect of antipsychotic medication in children and adolescents? Schizophr Bull 2011; 37: 322–323
25. Leucht S, Zhao J: Early improvement as a predictor of treatment response and remission in patients with schizophrenia: a pooled, post-hoc analysis from the asenapine development program. J Psychopharmacol 2014; 28:387–394
26. Correll C, Werner P, Pikalov A, et al: Early improvement predicts endpoint response to lurasidone in schizophrenia: pooled analysis of five double-blind trials (abstract). Eur Neuropsychopharmacol 2013; 23(suppl 2):S485
27. Kayo M, Tassell I, Hiroce V, et al: Does lack of improvement in the first two weeks predict treatment resistance in recent-onset psychosis? Clinics (Sao Paulo) 2012; 67:1479–1482
28. American Psychiatric Association: Practice Guideline for the Treatment of Patients With Schizophrenia, 2nd ed. Am J Psychiatry 2004; 161(Feb suppl):1–56
29. Correll CU, Malhotra AK, Kaushik S, et al: Early prediction of antipsychotic response in schizophrenia. Am J Psychiatry 2003; 160: 2063–2065
30. Buchanan RW, Kreyenbuhl J, Kelly DL, et al: The 2009 schizophrenia PORT psychopharmacological treatment recommendations and summary statements. Schizophr Bull 2010; 36:71–93
31. Kreyenbuhl J, Buchanan RW, Dickerson FB, et al: The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2009. Schizophr Bull 2010; 36:94–103
32. Hasan A, Falkai P, Wobrock T, et al: World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Schizophrenia, part 1: update 2012 on the acute treatment of schizophrenia and the management of treatment resistance. World J Biol Psychiatry 2012; 13:318–378
33. Hasan A, Falkai P, Wobrock T, et al: World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Schizophrenia, part 2: update 2012 on the long-term treatment of schizophrenia and management of antipsychotic-induced side effects. World J Biol Psychiatry 2013; 14:2–44

34. Barnes TR, Schizophrenia Consensus Group of British Association for Psychopharmacology: Evidence-based guidelines for the pharmacological treatment of schizophrenia: recommendations from the British Association for Psychopharmacology. J Psychopharmacol 2011; 25:567–620
35. National Institute for Health and Care Excellence (NICE): Psychosis and Schizophrenia in Adults: Treatment and Management: NICE Clinical Guideline 178. London, National Institute for Health and Care Excellence, 2014 (https://www.nice.org.uk/guidance/cg178)
36. Reitsma JB, Glas AS, Rutjes AW, et al: Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005; 58:982–990
37. Macaskill P, Gatsonis C, Deeks J, et al: Analysing and presenting results, in Handbook for Systematic Reviews of Diagnostic Test Accuracy, version 1.0. Edited by Deeks J, Bossuyt P, Gatsonis C. Cochrane Collaboration, 2010 (http://srdta.cochrane.org/)
38. Kay SR, Fiszbein A, Opler LA: The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 1987; 13:261–276
39. Overall JE, Gorham DR: The Brief Psychiatric Rating Scale. Psychol Rep 1962; 10:799–812
40. Egger M, Zellweger-Zähner T, Schneider M, et al: Language bias in randomised controlled trials published in English and German. Lancet 1997; 350:326–329
41. Gardner DM, Murphy AL, O'Donnell H, et al: International consensus study of antipsychotic dosing. Am J Psychiatry 2010; 167: 686–693
42. Ascher-Svanum H, Nyhuis AW, Faries DE, et al: Clinical, functional, and economic ramifications of early nonresponse to antipsychotics in the naturalistic treatment of schizophrenia. Schizophr Bull 2008; 34:1163–1171
43. Guy W (ed): ECDEU Assessment Manual for Psychopharmacology: Publication ADM 76-338. Washington, DC, US Department of Health, Education, and Welfare, 1976, pp 218–222
44. Levine SZ, Rabinowitz J, Engel R, et al: Extrapolation between measures of symptom severity and change: an examination of the PANSS and CGI. Schizophr Res 2008; 98:318–322
45. Leucht S, Davis JM, Engel RR, et al: Defining 'response' in antipsychotic drug trials: recommendations for the use of scale-derived cutoffs. Neuropsychopharmacology 2007; 32:1903–1910
46. Leucht S, Rothe P, Davis JM, et al: Equipercentile linking of the BPRS and the PANSS. Eur Neuropsychopharmacol 2013; 23:956–959
47. Nicholson IR, Chapman JE, Neufeld RW: Variability in BPRS definitions of positive and negative symptoms. Schizophr Res 1995; 17: 177–185
48. Leucht S, Engel RR, Davis JM, et al. Equipercentile linking of the Brief Psychiatric Rating Scale and the Clinical Global Impression Scale in a catchment area. Eur Neuropsychopharmacol 2012; 22: 501–505
49. Andreasen NC, Carpenter WT Jr, Kane JM, et al: Remission in schizophrenia: proposed criteria and rationale for consensus. Am J Psychiatry 2005; 162:441–449
50. Whiting PF, Rutjes AW, Westwood ME, et al: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155:529–536
51. Leeflang MM, Deeks JJ, Rutjes AW, et al: Bivariate meta-analysis of predictive values of diagnostic tests can be an alternative to bivariate meta-analysis of sensitivity and specificity. J Clin Epidemiol 2012; 65: 1088–1097
52. Tollefson GD, Birkett MA, Kiesler GM, et al: Double-blind comparison of olanzapine versus clozapine in schizophrenic patients clinically eligible for treatment with clozapine. Biol Psychiatry 2001; 49:52–63
53. Buckley PF, Correll CU: Strategies for dosing and switching antipsychotics for optimal clinical management. J Clin Psychiatry 2008; 69(suppl 1):4–17
54. Moher D, Liberati A, Tetzlaff J, et al: Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. J Clin Epidemiol 2009; 62:1006–1012
55. Beitinger R, Lin J, Kissling W, et al: Comparative remission rates of schizophrenic patients using various remission criteria. Prog Neuropsychopharmacol Biol Psychiatry 2008; 32:1643–1651
56. Awad AG, Hogan TP: Subjective response to neuroleptics and the quality of life: implications for treatment outcome. Acta Psychiatr Scand Suppl 1994; 380:27–32
57. Chang WH, Hwu HG, Chen TY, et al: Plasma homovanillic acid and treatment response in a large group of schizophrenic patients. Schizophr Res 1993; 10:259–265
58. Akiyama K, Tsuchida K, Kanzaki A, et al: Plasma homovanillic acid levels and therapeutic outcome in schizophrenics: comparisons of neuroleptic-naive first-episode patients and patients with disease exacerbation due to neuroleptic discontinuance. Biol Psychiatry 1995; 38:639–648
59. Vyas NS, Patel NH, Nijran KS, et al: The use of PET imaging in studying cognition, genetics and pharmacotherapeutic interventions in schizophrenia. Expert Rev Neurother 2011; 11:37–51
60. Shergill SS, White TP, Joyce DW, et al: Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. JAMA Psychiatry 2014; 71:28–35
61. Szeszko PR, Narr KL, Phillips OR, et al: Magnetic resonance imaging predictors of treatment response in first-episode schizophrenia. Schizophr Bull 2012; 38:569–578
62. Lane HY, Hsu SK, Liu YC, et al: Dopamine D3 receptor Ser9Gly polymorphism and risperidone response. J Clin Psychopharmacol 2005; 25:6–11
63. Bondy B, Spellmann I: Pharmacogenetics of antipsychotics: useful for the clinician? Curr Opin Psychiatry 2007; 20:126–130
64. Agid O, Siu CO, Pappadopulos E, et al: Early prediction of clinical and functional outcome in schizophrenia. Eur Neuropsychopharmacol 2013; 23:842–851
65. Schennach-Wolff R, Jager M, Mayr A, et al: Predictors of response and remission in the acute treatment of first-episode schizophrenia patients–is it all about early response? Eur Neuropsychopharmacol 2011; 21:370–378
66. Emsley R, Rabinowitz J, Medori R, et al: Remission in early psychosis: rates, predictors, and clinical and functional outcome correlates. Schizophr Res 2007; 89:129–139
67. Stentebjerg-Olesen M, Jeppesen P, Pagsberg AK, et al: Early nonresponse determined by the Clinical Global Impressions scale predicts poorer outcomes in youth with schizophrenia spectrum disorders naturalistically treated with second-generation antipsychotics. J Child Adolesc Psychopharmacol 2013; 23:665–675
68. Hatta K, Otachi T, Sudo Y, et al: Difference in early prediction of antipsychotic non-response between risperidone and olanzapine in the treatment of acute-phase schizophrenia. Schizophr Res 2011; 128:127–135
69. Gade K, Köhler J, Klein P, et al: Predictors of symptomatic remission in first-episode psychosis outpatients treated with quetiapine: a naturalistic study. Int J Psychiatry Clin Pract 2013; 17:148–153
70. Riedel M, Strassnig M, Müller N, et al: How representative of everyday clinical populations are schizophrenia patients enrolled in clinical trials? Eur Arch Psychiatry Clin Neurosci 2005; 255:143–148
71. Hofer A, Hummer M, Huber R, et al: Selection bias in clinical trials with antipsychotics. J Clin Psychopharmacol 2000; 20:699–702
72. Emsley R, Rabinowitz J, Medori R: Time course for antipsychotic treatment response in first-episode schizophrenia. Am J Psychiatry 2006; 163:743–745
73. Gallego JA, Robinson DG, Sevy SM, et al: Time to treatment response in first-episode schizophrenia: should acute treatment trials last several months? J Clin Psychiatry 2011; 72:1691–1696
74. Suzuki T, Remington G, Arenovich T, et al: Time course of improvement with antipsychotic medication in treatment-resistant schizophrenia. Br J Psychiatry 2011; 199:275–280
75. Meltzer HY: Duration of a clozapine trial in neuroleptic-resistant schizophrenia (letter). Arch Gen Psychiatry 1989; 46:672
76. Meltzer HY, Bastani B, Kwon KY, et al: A prospective study of clozapine in treatment-resistant schizophrenic patients, I:

preliminary report. Psychopharmacology (Berl) 1989; 99(suppl): S68–S72
77. Zito JM, Volavka J, Craig TJ, et al: Pharmacoepidemiology of clozapine in 202 inpatients with schizophrenia. Ann Pharmacother 1993; 27:1262–1269
78. Lieberman JA, Safferman AZ, Pollack S, et al: Clinical effects of clozapine in chronic schizophrenia: response to treatment and predictors of outcome. Am J Psychiatry 1994; 151:1744–1752
79. Hiemke C, Baumann P, Bergemann N, et al: AGNP Consensus Guidelines for Therapeutic Drug Monitoring in Psychiatry: update 2011. Pharmacopsychiatry 2011; 44:195–235
80. Kinon BJ, Kane JM, Chakos M, et al: Possible predictors of neuroleptic-resistant schizophrenic relapse: influence of negative symptoms and acute extrapyramidal side effects. Psychopharmacol Bull 1993; 29: 365–369
81. Correll CU, Rummel-Kluge C, Corves C, et al: Antipsychotic combinations vs monotherapy in schizophrenia: a meta-analysis of randomized controlled trials. Schizophr Bull 2009; 35:443–457
82. ClinicalTrials.gov: The Switch Study—Efficacy of an Early Antipsychotic Switch in Case of Poor Initial Response to the Treatment of Schizophrenia. https://clinicaltrials.gov/show/NCT01029769
83. European Commission: Optimization of Treatment and Management of Schizophrenia in Europe (OPTIMISE). http://ec.europa.eu/research/health/medical-research/brain-research/projects/optimise_en.html