Patient-oriented randomisation: A new trial design applied in the Neuroleptic Strategy Study

Constanze Schulz1, Jürgen Timm1, Joachim Cordes2, Gerhard Gründer3, Bernd Mühlbauer1,4, Eckart Rüther5 and Martin Heinze6

Abstract

Background: The ‘gold standard’ for clinical studies is a randomised controlled trial usually comparing specific treatments. If the scientific study expands to strategy comparison with each strategy including various treatments, the research problems are increasingly complicated. The strategy debate in the psychiatric community is the starting point for the development of our new design. It is widely accepted that second-generation antipsychotics are the therapy of choice in the treatment of schizophrenia. However, their general superiority over first-generation antipsychotics could not be demonstrated in recent randomised controlled trials. Furthermore, we are becoming increasingly aware that the experimental conditions of randomised controlled trials, as in the European First Episode Schizophrenia Trial and Clinical Antipsychotic Trials of Intervention Effectiveness Phase 1 studies, may be inappropriate for psychiatric treatments. The high heterogeneity in the patient population produces discrepancies between daily clinical perception and randomised controlled trials results. The patient-oriented approach in the Cost Utility of the Latest Antipsychotic drugs in Schizophrenia Study reflects everyday clinical practice. The results, however, are highly dependent on the physicians’ preferences. The goal of the design described here is to take an intermediate path between randomised controlled trials and clinical studies such as Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study, combining the advantages of both study types.

Methods: The idea is to randomise two treatment pairs each consisting of one first-generation antipsychotic and one second-generation antipsychotic in a first step and subsequently, to involve the investigators in deciding for a pair most appropriate to the patients’ needs and then to randomise the allocation to one drug (first-generation antipsychotic or second-generation antipsychotic) of that chosen pair. This idea was first implemented in the clinical trial, the Neuroleptic Strategy Study, with a randomised design comparing efficacy and safety of two different strategies: either to use first-generation antipsychotics (haloperidol and flupentixol) or second-generation antipsychotics (olanzapine, aripiprazole and quetiapine) in patients suffering from schizophrenia.

Results: In the course of the Neuroleptic Strategy Study, feasibility of this design was demonstrated. All aspects of the new design were implemented: randomisation process, documentation of responses from investigators as well as patients and drug logistic experience. In implementing the design, furthermore, we could investigate its theoretical properties. The physicians’ preferences for specific drugs used for the respective patients were analysed.

Conclusion: The idea of patient-oriented randomisation can be generalised. In light of the heterogeneity and complexity of patient–drug interaction, this design should prove particularly useful.

Keywords

Neuroleptic Strategy Study, patient-oriented randomisation, strategy comparison, schizophrenia, first-generation antipsychotics, second-generation antipsychotics, ethics

1Competence Center for Clinical Trials Bremen, University of Bremen, Bremen, Germany
2Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany
3Department of Psychiatry, Psychotherapy and Psychosomatics, Medical Faculty, RWTH Aachen University, Aachen, Germany
4Department of Pharmacology, Klinikum Bremen Mitte, Bremen, Germany
5Department of Psychiatry and Psychotherapy, University of Göttingen, Göttingen, Germany
6Department for Psychiatry and Psychotherapy, Brandenburg Medical School, Immanuel Klinik Rüdersdorf, Berlin, Germany

Corresponding author:
Constanze Schulz, Competence Center for Clinical Trials Bremen, University of Bremen, Linzer Straße 4, 28359 Bremen, Germany.
Email: conschul@math.uni-bremen.de
Background

The 'gold standard' for clinical studies is a randomised controlled trial (RCT) usually comparing specific treatments, for example, a new drug, a comparator drug or placebo. This comparison is especially appropriate when efficacy and safety of new drugs are being investigated. The procedure becomes more complicated if the scientific problem expands to strategy comparison, in which each strategy includes various treatments. In the following article, we present the idea and the implementation of a new clinical trial design which can be used in problems involving strategy comparison. The starting point for the development of this design is a debate in the psychiatric community. Based on everyday practice, it is widely accepted that second-generation antipsychotics (SGAs) are considered to be the treatment of choice in the therapy of schizophrenic disorders. However, the general superiority of SGAs over the first-generation antipsychotics (FGAs) could not be clearly demonstrated in recent controlled clinical trials such as the European First Episode Schizophrenia Trial (EUFEST), Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) and Cost Utility of the Latest Antipsychotic drugs in Schizophrenia Study (CUtLASS). The methodological problems of those trials have been discussed extensively.

In the context of this article, we consider the advantages and disadvantages for a strategy decision (to use SGAs or FGAs). EUFEST is an open, randomised trial of haloperidol versus SGAs in first-episode schizophrenia. CATIE Phase 1 is a double-blind trial comparing perphenazine with four SGAs. One disadvantage of both studies is that only one representative was chosen in the group of FGAs to compare both strategies. The overall CATIE study design is more complex and allowed for patients who had discontinued one study antipsychotic drug to enter subsequent phases (2 and later 3) of the study and receive another antipsychotic drug. Thus, the CATIE study is a representative of the Sequential Multiple Assignment Randomised Trials, whose primary goal is to yield information for the development of adaptive intervention strategies. A further goal is to collect information on candidate tailoring variables. This differs slightly from the goal of comparing two one-step strategies (whether to use FGAs or to use SGAs). Thus, the CATIE study provided no clear answer to the question of strategy comparison. It does, however, 'provide more evidence to clinicians, patients, and policy makers that response to antipsychotic drug treatment is heterogeneous. The individual variations in response to medications are likely due to multiple factors that may be difficult to elucidate'. This difficulty underlines an additional condition important for our design: the treatments even in one strategy perform differently for different patients, but subgroups which might account for these differences are not identifiable. This implies that in a design with simple block randomisation, such as EUFEST and CATIE Phase 1, the results from the responders and non-responders may cancel each other out.

To measure the actual strategy effect, the physician should be involved in the decision concerning trial treatment within each strategy, taking risks and healing opportunities of different drugs for each patient into account. This means using a design close to everyday practice. Despite these points of criticism, however, we must not return to mere observational studies because, in fact, randomisation is the only way that effectively deals with the risk of bias by selection and confounding. The challenge is to combine the advantages of an RCT approach with the advantage of taking the physicians' expertise appropriately into account. The patient-oriented, straightforward RCT design for comparing two strategies in heterogeneous patient collectives was implemented in the CUtLASS study.

The CUtLASS study is an open-label trial, in which each psychiatrist chooses 1 of 13 FGAs and one of four SGAs before randomisation to one of these drugs. This patient-oriented approach for treatment selection and without blinding is very close to everyday practice. However, the fact that the doctor can freely choose from a large number of medications can lead to several problems partly visible in the CUtLASS study. For an ideal situation, every physician should be as thoroughly knowledgeable and experienced as possible with all 17 drugs to be able to decide objectively on the best treatment in each strategy. In practice, however, the physicians' preferences are often based on previous positive experience with a certain drug or on information gathered from courses, literature or advertising. Thus, the physician may miss the most effective drug of a strategy collection because he has never used it or has received misleading information. This selection bias will influence the study.

A related problem is the distribution of patients to the drugs tested, especially for the FGAs. Half of the patients in the FGA group received just one drug and only zero to three patients received any of the other 9 of 13 drugs. The question arises: 'Why aren't the physicians using the full range of drugs available?' Theoretically, the drugs within both strategies should reflect the real market distribution. This is obviously not the case. In the worst case scenario of the CUtLASS design, however, only one treatment in one strategy may be selected, resulting in a one-to-many drug comparison such as in EUFEST and CATIE, and not in a real comparison of strategies with patient-oriented treatment. Thus, the physicians' preferences may lead to poorly represented and thus underpowered comparisons, especially for new drugs or drugs with changed dosage or indications. With studies of the CUtLASS design, the study yields no evidence for drugs not used. Furthermore, blinding is an issue. In
the CUtLASS study, the physicians know if they had given a drug not belonging to their favoured strategy. This occurred in about 50% of cases.

In conclusion, the CUtLASS design already took a great step towards a more patient-oriented approach to testing, while EUFEST and CATIE continue to represent the ‘classical’ RCT approach, ensuring systematic comparisons independent of the physician’s preference. The goal of the new design described here is to take an intermediate path combining the advantages of both study types by randomising treatment pairs in a first step, involving the investigators in deciding for a pair most appropriate to the patients’ needs (described below) and then randomising the allocation to one drug (FGA or SGA) of that chosen pair. Through this procedure, the new design should reflect everyday practice better than EUFEST and CATIE studies, as well as yielding more systematic comparisons than CUtLASS.

### Methods

#### Study setting and design

The new design was implemented in the clinical trial ‘The Neuroleptic Strategy Study’ (NeSSy)\(^{14,15}\) comparing efficacy and safety of the strategies using either FGA or SGA in patients suffering from schizophrenia. NeSSy was a multi-centre, double-blind randomised study in which two FGAs and three SGAs were used. It was conducted from April 2010 to August 2013 at 20 clinical sites in Germany and was supported by the German Federal Ministry of Education and Research. More information is available at ClinicalTrials.gov, German Clinical Trials Register and EudraCT.

The design reflected the diversity of diagnoses and patient statuses by offering a limited choice out of two FGAs (haloperidol and flupentixol) and three SGAs (olanzapine, aripiprazole and quetiapine) as follows: six pairs of one FGA and one SGA were formed (see Table 1). The randomisation was performed in two steps. In the first step, the physician received a selection containing two pairs out of the total of six. These pairs were randomly selected from the set of all 15 possible two-pair combinations (see Table 2). The physician compared the benefits and risks of the drugs presented in the two-pair combination and chose the most suitable FGA–SGA pair for the patient. This decision involved either selection or deselection. Thereafter, he justified his choice in the case report form. In the second double-blinded step, a further randomisation allocated the patient either to the FGA or the SGA from the selected pair. The process is illustrated in Figure 1.

#### Objectives, outcomes and statistical analysis

In the NeSSy study, there are two primary endpoints: the Short-Form Health Survey (SF-36)\(^{16}\) for assessing the quality of life from the patient’s perspective and the Clinical Global Impression Score Improvement (CGI-I)\(^{17}\) for assessing the change in clinical status by the physician. The development of both criteria is measured by area under curve (AUC) on a logarithmic time scale. Main secondary outcome measures are subscales of SF-36, the Positive and Negative Syndrome Scale,\(^{18}\) the Subjective Well-Being under Neuroleptic Scale-short form\(^{19}\) and the Personal and Social Performance Scale.\(^{20}\)

The sample size calculation was based on a two-sided alpha level of 2.5% and a power of 80% for each of the primary criteria, on autocorrelation between visits, on literature data for SF-36\(^{21–23}\) and CGI-I\(^{4,22,24–27}\) resulting in AUC standard deviations of 20.38 (SF-36) and 0.99 (CGI-I), clinically relevant differences of 5 to 10 (SF-36) and 0.5 (CGI-I) defined prior to trial planning and an analysis of variance (ANOVA) contrast of 8 (SF-36) respectively 0.4 (CGI-I) to be detected if existent (20% below upper limit of clinical relevance). The special design of this trial implies equal sample sizes for

### Table 1. The six pairs formed out of the five basic drugs.

<table>
<thead>
<tr>
<th>Pairs</th>
<th>First-generation antipsychotics</th>
<th>Second-generation antipsychotics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Haloperidol</td>
<td>Olanzapine</td>
</tr>
<tr>
<td>2</td>
<td>Haloperidol</td>
<td>Aripiprazole</td>
</tr>
<tr>
<td>3</td>
<td>Haloperidol</td>
<td>Quetiapine</td>
</tr>
<tr>
<td>4</td>
<td>Flupentixol</td>
<td>Olanzapine</td>
</tr>
<tr>
<td>5</td>
<td>Flupentixol</td>
<td>Aripiprazole</td>
</tr>
<tr>
<td>6</td>
<td>Flupentixol</td>
<td>Quetiapine</td>
</tr>
</tbody>
</table>

Schulz et al. 253
Table 2. The 15 two-pair combinations and their frequency of use in the NeSSy study sorted by clear patient-oriented decisions (N = 149).

<table>
<thead>
<tr>
<th>Two-pair combination</th>
<th>Frequency of use</th>
<th>Random pair 1</th>
<th>Random pair 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear decision between haloperidol ↔ flupentixol</td>
<td></td>
<td>(haloperidol, olanzapine)</td>
<td>(haloperidol, olanzapine)</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>(haloperidol, aripiprazole)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>(haloperidol, quetiapine)</td>
<td>(flupentixol, aripiprazole)</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>(flupentixol, olanzapine)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>Clear decision between olanzapine ↔ aripiprazole</td>
<td></td>
<td>(haloperidol, olanzapine)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>(haloperidol, aripiprazole)</td>
<td>(flupentixol, aripiprazole)</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>(haloperidol, quetiapine)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>Clear decision between olanzapine ↔ quetiapine</td>
<td></td>
<td>(haloperidol, olanzapine)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>(haloperidol, aripiprazole)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>(flupentixol, aripiprazole)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>Clear decision between aripiprazole ↔ quetiapine</td>
<td></td>
<td>(haloperidol, olanzapine)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>(haloperidol, aripiprazole)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>(flupentixol, aripiprazole)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>No clear decision</td>
<td></td>
<td>(haloperidol, olanzapine)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>(haloperidol, olanzapine)</td>
<td>(flupentixol, aripiprazole)</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>(haloperidol, aripiprazole)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>(haloperidol, aripiprazole)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>13</td>
<td>8</td>
<td>(haloperidol, quetiapine)</td>
<td>(flupentixol, quetiapine)</td>
</tr>
<tr>
<td>14</td>
<td>10</td>
<td>(haloperidol, quetiapine)</td>
<td>(flupentixol, olanzapine)</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>(haloperidol, quetiapine)</td>
<td>(flupentixol, aripiprazole)</td>
</tr>
</tbody>
</table>

NeSSy: Neuroleptic Strategy Study.

Both strategies but unequal sample sizes for the five drugs. The physicians' decisions may lead to an imbalance between treatments within each strategy. In the worst case scenario, only 20% of the total FGA sample would receive haloperidol (see section 'Theoretical results'). This proportion depends on design parameters and has been adopted for the sample size calculation. These assumptions yielded sample sizes of 316 (SF-36) and 301 (CGI-I). A cautious dropout rate of 50% was deduced from the literature, doubling the number needed to 632. A balanced random list contained 630 cases (divisible by 30 because of the 15 two-pair combinations and two strategies). At the end of the regular study, an extension was requested based on a recalculated sample size. The previous assumptions for standard deviation, autocorrelation and imbalance were found to have been too cautious. Thus, the required sample size was smaller than 630 patients. A sample size of 150 patients could be conducted within the financed study duration, using the capacities of the participating centres (see above).

Statistical analysis was performed on the full analysis set consisting of all randomised patients with at least one use of study drugs. The primary efficacy analysis evaluated strategies by linear contrast in analysis of covariance (ANCOVA) with the baseline value as covariate and drugs as factors. The null hypothesis

\[ H_0 : \frac{1}{2} \mu_{\text{haloperidol}} + \frac{1}{2} \mu_{\text{flupentixol}} = \frac{1}{3} \mu_{\text{olanzapine}} + \frac{1}{3} \mu_{\text{aripiprazole}} + \frac{1}{3} \mu_{\text{quetiapine}} \]

of no difference between the strategy means was tested with a two-sided test for each primary endpoint. The multiple test problem (two primary criteria) was
addressed by the Bonferroni–Holm method. Missing values were imputed by last observation carried forward and linear interpolation.

Results

Practical realisation and drug logistics

NeSSy was the first study to implement a patient-oriented randomised design of this complexity. Treatment should start as soon as possible after randomisation. Both physician and patients should proceed blinded, using neutral capsules. To solve the problem of differences in the daytime/evening application (morning for aripiprazole and evening for the other drugs), a placebo was used. As a backup, morning and evening capsules differed in colour. Figure 2 illustrates the basic sequence of the randomisation process at one of the participating centres. After screening and written consent, the patient was admitted and assigned a random number (P1, Figure 2). Physician reported inclusion of the patient (P2) and received the two-pair combination information for the patient’s random number (P3, B2). The appropriate medication had already been centrally packed in two packages following the total random list (B1) and sent to the centre (B4). These packages were labelled with corresponding random numbers and respective encoded information (X and Y) about first or second pair choice (B3) (double-blinded FGA or SGA out of the selected pair). The physician then selected the most suitable pair from this two-pair combination for this patient (P4), informed the study nurse of the patient’s random number and his choice (P5) and handed over the first package of drugs (P6).

The specificity of the new study’s design necessitated that a greater number of study drugs be provided, as for each random number, two parcels with medication had to be delivered. Those managing the medication logistics tried to recycle the returned unused packages, but this procedure was of limited success, due to expiration dates and rates of patient recruitment. For the drug logistics (B4), a control system was implemented to monitor the medication.

Theoretical results

Independent of the randomisation process, the design has some further theoretical properties. For this purpose, we consider those two-pair combinations (random pair 1 versus random pair 2), in which either two identical FGAs or two identical SGAs were offered (see rows 1–9 in Table 2). For example, consider the two-pair combination 5 with pair 1 (flupentixol and olanzapine) and pair 2 (flupentixol and aripiprazole) or the two-pair combination 3 with pair 1 (haloperidol and quetiapine) and pair 2 (flupentixol and quetiapine). In the clear-decision cases, the physicians’ preference can be concluded from his choice of drug for the respective patient. In the examples, a decision for pair 1 in the first case indicated a preference for olanzapine, and in the second case, a preference for haloperidol. Whether the patient actually received the preferred medication was a matter of chance. After randomisation, a patient got either the first-generation or the second-generation component in the selection. The physician’s chance of giving the preferred drug was 50%.

Table 2 illustrates that in 9 of the 15 two-pair combinations (60%), the physician had a clear decision between two treatments of one strategy, independent of the randomisation between the two strategies. The different parts show the different decisions. In the first three lines, there is a clear decision in the group of FGAs between haloperidol and flupentixol (20%). In the group of SGAs, this number is smaller because of the higher number of drugs in this strategy. There are three kinds of treatment decision in this strategy: (olanzapine ↔ aripiprazole), (olanzapine ↔ quetiapine) and (aripiprazole ↔ quetiapine). For each of these pairs, there are two two-pair combinations (13.3%) indicating a clear medical decision (see Table 2).

The advantage of this design is not only the possibility of choice but also that each of the study drugs is certain to have been used with a minimum number of treated patients. If the physician were fully free to decide on the treatment, it might happen that unpopular treatments are omitted and thus not represented in the trial. As noted before, no statement could then be made about these drugs. The identical FGAs or identical SGAs in two-pair combinations ensure that in these cases, the physician has no choice within the respective strategy: in 50% of cases, the doubled drug would be the drug-to-get by randomisation. In our previous example, the patients received flupentixol in the first
case and quetiapine in the second case, with 50% probability. Each treatment was represented with a minimum allocation probability in the trial. For FGAs, the probability for identical treatments in a two-pair combination is 20% (three-fifteenth, see rows 4–9 of Table 2) and for SGAs 6.6% (one-fifteenth, see rows 1–3 of Table 2) due to the different treatment numbers in the strategies. The probability increases if the number of treatments in the other strategy increases or the number of treatments in its own strategy decreases. This aspect of the design was taken into account when planning the NeSSy study since haloperidol was unpopular, even if administered at a low dose. In the study, at least 20% of all randomised patients in the FGA strategy received haloperidol.

Randomisation and the patient-oriented decision

In the NeSSy study, 2374 patients were screened; 1863 were ineligible. Of the eligible patients, 149 could be randomised. The result of step 1 randomisation (two-pair combination assigned) is documented in Table 2. As a result of the second step of randomisation, 69 patients were assigned to FGAs (27 to haloperidol and 42 to flupentixol) and 80 patients to SGAs (25 to olanzapine, 23 to aripiprazole and 32 to quetiapine).

The number of two-pair combinations assigned with doubled haloperidol was 30 (12 + 8 + 10 in row 4, 6 and 8 of Table 2), so that the expected minimum number of patients to receive haloperidol (independent from the physician’s decision) was 15. Thus, more patients received haloperidol than was assumed in sample size calculation.

As shown in the section, ‘Theoretical results of the design’, the two-pair combinations with a doubled medication are especially interesting (see rows 1–9 of Table 2) when analysing the physician’s decision. In 93 out of 149 cases (62.4%), one can derive a clear patient-oriented decision from the physician between the preferred drugs of the physicians compared with the non-favoured alternative. The difference is insignificant, but illustrates that the physician’s decisions in favour of efficacy are a positive aspect of the study design.

<table>
<thead>
<tr>
<th>Clear patient-oriented decision from the physician between</th>
<th>N</th>
<th>Observed (%)</th>
<th>Expected (%)</th>
<th>Ratio</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haloperidol ↔ flupentixol</td>
<td>32</td>
<td>21.5</td>
<td>20</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>Olanzapine ↔ aripiprazole</td>
<td>21</td>
<td>14.1</td>
<td>13.3</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Olanzapine ↔ quetiapine</td>
<td>23</td>
<td>15.4</td>
<td>13.3</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Aripiprazole ↔ quetiapine</td>
<td>17</td>
<td>11.4</td>
<td>13.3</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>62.4</td>
<td>60.0</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Figure 3. Kaplan–Meier curves of patients receiving the physicians preferred drug and the remaining patients in the full analysis set.

Justification of the patient-oriented decision

In the NeSSy study, the physician was required to justify the decision for the selected pair in the case report form. The open-text justifications were analysed according to various factors. First, the origin of the decision was analysed. In 73.2% of cases, the physicians justified their decisions, reporting that in 34.8%
of the cases the patients were involved in the decision. In the remaining cases, a decision by chance (3.4%) or no justification was documented (23.5%). Second, the criteria for the choices were investigated, that is, whether the decision was based on the drug’s efficacy (25.5%) or whether a drug was deselected (48.3%) due to possible risk and side effects. Third, the reasons were recorded. The preferred or rejected drug was then documented, along with the physicians’ reasons for their choice. In some cases, the physicians stated only the preferred drug and gave no reason for the choice.

In the analysis of these documented decisions, it was found that the majority of physicians preferred SGAs in their choice. In the mixed two-pair combinations, the decision to select one pair or the other was based in more than half of the cases (53.6%) only on the SGAs, in 14.3% on the FGA and in 7.1% on both strategies. In the remaining cases (25%), the decision was not justified or was made at random.

**NeSSy study results**

The primary outcome was obtained by the Bonferroni-Holm test method. The mean AUC values for SF-36 were higher in the SGA than in the FGA group. The null hypothesis of no difference between the strategy means was rejected at the 2.5% level. The difference between both strategy means was also clinically significant. CGI-I values, which reflect judgment by an observer, did not differ between the two treatment strategies.

**Discussion**

**Practical experience**

Even though the study design used in NeSSy looks more complicated due to the integration of a clinical decision, the study was not so inconvenient for physicians and study nurses that they had difficulty following the procedures or were unwilling to collaborate. On the contrary, the personnel feedback shows that they felt themselves – their concerns for the patient and their expertise – to be taken more seriously. Patients felt that their health interests were at least partially recognised. Since the physicians knew that, after their clinical decision, only two of the five drugs were involved, some physicians felt that they were partially unblinded about the randomised treatment. Based on observation of side effects, physicians thought that they might be able to guess whether the substance used was an FGA or an SGA. In the end, however, the physicians reported that these concerns were not realised in most cases.

**Theoretical outlook**

This idea of patient-oriented randomisation can be generalised to different strategies containing different numbers of treatments. The first step generalises the concept to two strategies with arbitrary numbers of treatments in each strategy. The theoretical properties of a study can be determined analogous to the considerations of the NeSSy design. In addition to minimum allocation probability, we may also define a maximum allocation probability for each treatment in each design. While the probability of the block randomisation is always fixed, the physicians’ decisions ensure that the allocation probability for each treatment varies between these two limits (for more details, see Schulz and Timm). Whether the problem of heterogeneity can be overcome by this new design remains open. The issue of testing procedures between the strategies also requires further consideration.

The design provides randomised outcomes for 15 different two-pair combinations that represent a variety of two-drug contrasts. In principle, several hypotheses can be considered in this design. Patients may be clustered into different groups using different decision-making strategies. For example, patients of one group may be given a single drug compared with patients of another group receiving the drugs of the other strategy (see Table 2).

**Limitations**

Paradoxically, the main limitation of this design emerges from the issue it solves: it is the physician’s experience and knowledge that make strategy comparison in a heterogeneous patient group possible. Yet, the design is less effective if the physicians’ choices are not optimal. Furthermore, the study’s focus is on strategy comparison and not on single treatment comparison. If we want to check further hypotheses on the reasons for a physician’s choice, on efficacy or on safety of single drugs, then an increasing sample size might be necessary.

More limitations arise with an increasing number of drugs in each strategy. Because patients are excluded due to intolerance of a single drug, the minimal allocation probability for each drug also necessitates increasing the sample size which, in turn, increases the logistical effort involved.

**Conclusion**

This study design was developed to counter problems of ‘classical’ RCTs facing the high complexity of therapeutic situations. The heterogeneity of patients in psychiatric care makes strategy comparison especially difficult. Within NeSSy, the daily clinical perception of patients and the patient-oriented selection of treatment for schizophrenia were incorporated in an RCT-like design. The design achieves a meaningful comparison of the efficacy and tolerability of two strategies.
The psychiatric community tends to regard standardized studies as inadequate to provide a basis for generalized therapeutic conclusions. This dilemma could be addressed. The gap between methodological restrictions of evidence-based medicine and daily clinical impressions of physicians and psychiatric patients seems partially bridged.

An important result of the study is the demonstrated feasibility of this new biometric design, which has been evaluated very positively by the physicians at the trial centres. Physicians found the design suitable and practical in its benefits. The advantages of the design include a promising compromise between the conflicting demands for randomisation and for patient-oriented therapy choices. Additionally, the design provides the opportunity to learn more about therapeutic decisions in heterogeneous patient populations. Here, the differing individual medical histories, risk factors, and social and cultural perspectives are accounted for through analysing the documented reasons for therapeutic decisions based on these patient characteristics. Furthermore, patients participate more readily in the study design because they feel that their individual experience and needs are being taken into account. Therefore, this design may be used successfully to solve related problems in other fields with comparable problems. The NeSSy design responds to the need for a next step in evidence-based medicine, that of ‘integrating clinical expertise with the best available external clinical evidence from systematic research’.

**Acknowledgements**

The authors thank Astra Zeneca Deutschland, Bristol-Myer Squibb, Lundbeck GmbH and Lilly Deutschland GmbH for providing the study drug. ClinicalTrials.gov number NCT01164059; German Clinical Trials Register DRKS00000304; WHO ICTRUP UTN-Nr. U1111-1112-9727; EudraCT number 2009-01096647; Date of Ethics Committee Approval 19 November 2009.

**Declarations of conflicting interests**

Dr Gründer has served as a consultant for Cheplapharm (Greifswald, Germany), Eli Lilly (Indianapolis, Ind, USA), Lundbeck (Copenhagen, Denmark), Roche (Basel, Switzerland), Servier (Paris, France) and Takeda (Osaka, Japan). He has served on the speakers’ bureau of Eli Lilly, Gedeon Richter (Budapest, Ungarn), Janssen Cilag (Neuss, Germany), Lundbeck, Roche and Servier. He has received grant support from Boehringer Ingelheim (Ingelheim, Germany) and Roche. He is the co-founder of Pharma Image GmbH (Düsseldorf, Germany) and Brainfoods UG (Düsseldorf, Germany). Dr Cordes was a member of an advisory board of Roche, accepted travel or hospitality not related to a speaking engagement from Servier and he received support for symposia from Inomed, Localite, Magventure, Roche, Mag & More, NeuroConn, Synetika, FBI Mediotechnik, Spitzer Arzneimittel and Diamedic. Dr Rüther has served as a consultant for Eli Lilly, Lundbeck, Bayer, BMS, Otsuka. He has served on the speakers’ bureau of Eli Lilly, Lundbeck, Bayer, BMS, Otsuka, Servier. Mrs Schulz, Drs Heinze, Mühlbauer and Timm declare no conflicts of interest.

**Funding**

The study was financed by the German Federal Ministry of Education and Research (BMBF) under grant number 01KG0907. A part of the study medication used in the study was provided by the companies Astra Zeneca Deutschland, Bristol-Myer Squibb, Lundbeck GmbH and Lilly Deutschland GmbH.

**References**


