

Keywords: Schizophrenia; New drug; Clinical trial design.

Clinical trial designs of new medicinal products for treating schizophrenia: Discussion of EMA's Guideline and a Better Long Term Trial Design

Haibiao Jiang^a
Jingjing Zhang^{b,*}

^a Jiangsu Hansoh Pharmaceutical Co. Ltd.

^b Mental Health Center of Minhang District, Shanghai

CHINA

ABSTRACT – Background and Objectives: Schizophrenia is a severe chronic disease. End-point variables lack objectivity and the diagnostic criteria have evolved with time. In order to guide the development of new drugs, European Medicines Agency (EMA) issued a guideline on the clinical investigation of medicinal products for the treatment of schizophrenia.

Methods: Authors reviewed and discussed the efficacy trial part of the Guideline.

Results: The Guideline divides clinical efficacy trials into short-term trials and long-term trials. The short-term three-arm trial is recommended to replace the short-term two-arm active-controlled non-inferiority trial because the latter has sensitivity issues. The Guideline ultimately makes that three-arm trial a superiority trial. The Guideline discusses four types of long-term trial designs. The randomized withdrawal trial design has some disadvantages. Long-term two-arm active-controlled non-inferiority trial is not recommended due to the sensitivity issue. Extension of the short-term trial is only suitable for extension of the short-term two-arm active-controlled superiority trial. The Guideline suggests that a hybrid design of a randomized withdrawal trial incorporated into a long-term parallel trial might be optimal. However, such a design has some disadvantages and might be too complex to be carried out. Authors suggest instead a three-group long-term trial design, which could provide comparison between test drug and active comparator along with comparison between the test drug and placebo. This alternative could arguably be much easier to carry out compared with the hybrid design.

Conclusions: The three-group long-term design merits further discussion and evaluation.

Received: 3 September 2015

Revised: 27 November 2015

Accepted: 17 December 2015

Background

Schizophrenia is a severe chronic disease that is a heavy burden for patients, families, and society. There are many anti-schizophrenic drugs available on the market and several new agents are in development.

In order to guide the development of these new agents, the European Medicines Agency (EMA) issued the “Guideline on clinical investigation of medicinal products, including depot preparations in the treatment of schizophrenia”¹ in 2012 that went into effect at the beginning of 2013. This very important guideline covers clinical trial efficacy and provides a very useful guide for the clinical development of anti-schizophrenic agents.

From the perspective of clinical trial design, some important points need to be considered when evaluating the efficacy of a new anti-schizophrenic agent:

1. Schizophrenia is a chronic disease. As such, the long-term efficacy or effectiveness of the candidate drug may need to be evaluated in addition to the short-term efficacy.
2. There are no strictly objective tools for evaluating the efficacy of a candidate drug. The most widely used efficacy endpoints in schizophrenia studies are scores of Positive and Negative Syndrome Scale (PANSS) and other questionnaires. These endpoints might be subject to observational bias and require double-blind design to eliminate such bias.
3. The diagnostic criteria for schizophrenia have evolved over the years. An anti-schizophrenic drug approved years ago might have been tested in strict clinical trials before it was approved. However, the diagnostic criteria and clinical management in the trials might be different

from current clinical practice. These differences might challenge its role and value as a positive control in a study evaluating a new anti-schizophrenic drug.

The EMA’s guideline addresses these issues and attempts to provide solutions. The Guideline divides efficacy trials into short-term trials and long-term trials. The aim of the short-term trial is to determine whether or not the candidate drug has an anti-schizophrenic effect. The aim of a long-term trial is to determine whether or not this effect can be maintained for a relative long period.

In section 4.4 of the Guideline, several different designs of short-term and long-term clinical efficacy trials are discussed, most of which will be introduced and discussed herein.

Discussion of the three-arm short-term confirmatory trial

For the short-term confirmatory trial, the Guideline recommends two types of prospective, randomized, double-blind, and parallel trial designs. One design is a two-arm positive-controlled superiority study. The other design is a three-arm (e.g. placebo control, test product, and active control) trial (see Figure 1 for a schematic of a three-arm design). In both designs, the proposed treatment duration is 4 to 6 weeks and the primary endpoint is the response rate. The response rate suggested by the Guideline is *at least a 30% reduction on the total PANSS score compared to baseline*¹.

The Guideline does not recommend a two-arm positive-controlled non-inferiority (including equivalence) trial. This is because in many situations, such a trial has “sensitivity” issues. In other words, this type of trial cannot confirm whether the test agent is superior

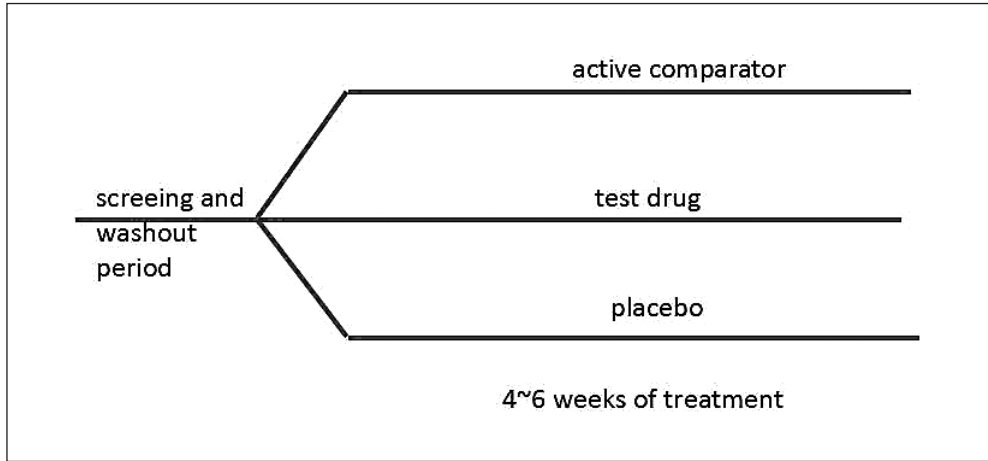


Figure 1. Three arm short term trial suggested by the guideline.

to placebo. This issue only exists in a two-arm non-inferiority (including equivalence) trial. In a two-arm positive-controlled superiority trial in which the test drug has been confirmed as superior to active comparator, most probably, the test drug would be superior to placebo even if the efficacy of the active comparator was uncertain.

A simplified model that further explains this phenomenon is provided in Figure 2. Suppose the efficacy of the placebo was zero, that of the active comparator was eight, and both the superiority and non-inferiority margins were five. In this situation, the active comparator would be effective. If the efficacy of the test drug was four, it is still non-inferior to the active comparator. However, it is not superior to the placebo because it did not exceed the superiority margin of the placebo.

There are two methods for solving this sensitivity issue. The first is adaptively adjusting the non-inferiority margin. In the simplified model mentioned above in which the efficacy of the placebo is zero, that of the active control is eight, and the superiority margin is five, we could set a non-inferiority margin

as three. Under these conditions, if non-inferiority to the active control is still confirmed, we would have confidence that the efficacy of the test drug would be no less than five and is superior to placebo. The example mentioned here is simplified for the sake of explanation. However, in real clinical trial settings, the adjustment of the non-inferiority margin is much more complex and should be carefully discussed with psychiatrists, clinical scientists, and bio-statisticians². Even so, this simplified model shows that this method requires reliable efficacy results for active comparator versus placebo before the non-inferiority margin can be adjusted. In order to reduce the bias of using historical data, a result from a recently completed high quality trial is preferable.

The second method is to introduce a placebo group into the trial. This transforms a two-arm active-controlled non-inferiority trial into a three-arm trial, as depicted in Figure 1. In addition to the comparison between test drug group and active control group, this design allows the test drug group to be directly compared with the placebo group in order to determine if the test drug is superior to the

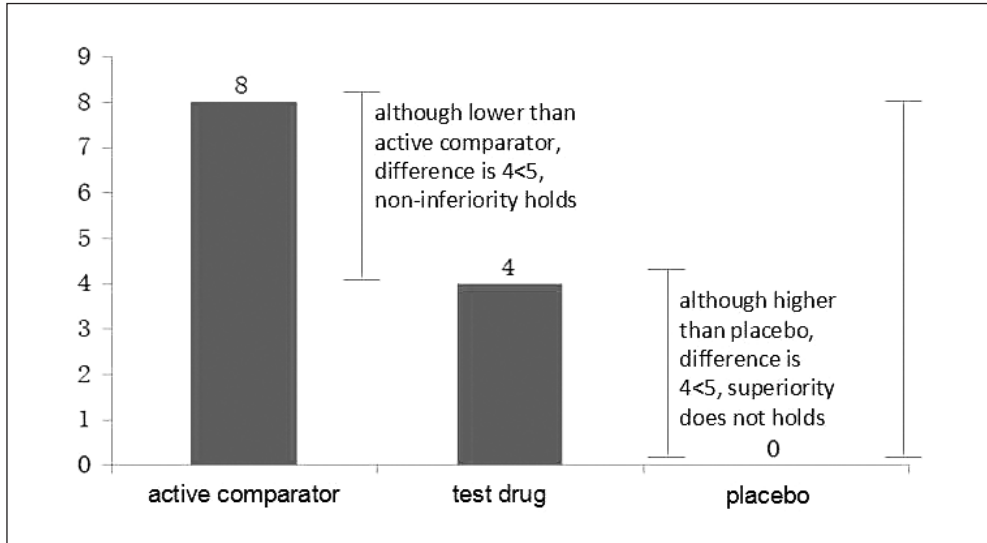


Figure 2. The sensitivity issue of a two arm active control non-inferiority trial.

placebo. Perhaps as a result of the evolution of the diagnostic criteria for schizophrenia and the fact that most marketed anti-schizophrenic drugs were launched years ago, the Guideline recommends this second method for a non-inferiority trial. The Guideline states:

“If the aim of the study is to demonstrate non-inferiority to an active comparator, then a three-arm study of placebo, test product and active comparator is recommended”¹.

However, some issues need to be noted for a three-arm study. First, there are several cautions regarding the placebo group. Placebo use might pose potential ethical problems. There might be risks of concern for disease progression because no active treatment would be provided to the patients in the placebo group. The Guideline also noticed these issues and stated:

“To avoid unnecessary risks for patients and others, placebo controlled studies should be performed in a highly controlled setting,

with stringent follow-up to apply predefined escape criteria, rescue medication and stopping rules”¹.

Assuming the treatment duration in a short-term trial is 4 to 6 weeks, it might be acceptable to use placebo under strictly controlled conditions (e.g., those suggested by the Guideline). In addition to these measures, another not mentioned in the Guideline might also be considered. This is the use of an imbalanced randomization allocation to reduce the number of subjects exposed to placebo. An example would be a test drug: active comparator: placebo ratio of 2:2:1.

Another issue which would be expected in a three-arm design is Type I error inflation due to multiple comparisons, or multiplicity. Typically, there will be three comparisons in a three-arm design: test drug versus active comparator, test drug versus placebo, and active comparator versus placebo. Several measures can be taken to control Type I error rate, according to the Guideline:

“The standard randomized placebo and active-controlled, parallel group trial design is intended to show superiority of the drug candidate to placebo and to quantify the efficacy of the drug candidate in comparison with a drug of known efficacy for the treatment of schizophrenia”¹.

This might suggest that the comparison of test drug versus placebo will be set as the primary endpoint and that other comparison, including the comparison between the test drug and the active comparator, would be secondary endpoints. Some trials used a similar design and primary endpoint in recent years^{3,4}. From the perspective of biostatistics, this means that this design is actually a superiority trial with the primary aim of confirming if the test drug is superior to placebo. The result of test drug versus active comparator will be exploratory. However, as discussed earlier, the original purpose of a three-arm design is to determine if the drug is non-inferior to an active comparator, and due to the intrinsic limitation of a two-arm active-controlled non-inferiority trial, a three-arm design is recommended by the Guideline:

“If the aim of the study is to demonstrate non-inferiority to an active comparator, then a three-arm study of placebo, test product and active comparator is recommended”¹.

Finally, the three-arm trial design becomes a superiority trial. It is not a non-inferiority trial.

Introduction and discussion on the long-term trial

Randomized withdrawal trial

The first kind of long-term trial design discussed by the Guideline is a randomized withdrawal trial (Figure 3). There are two stages in this trial design. In the first stage, all enrolled subjects will receive test drug treatment in an open manner. At the end of first stage, those who respond to treatment will enter the second stage and be randomized into two groups that receive either test drug or placebo treatment. In the second stage, those subjects whose disease relapses would withdraw from the study

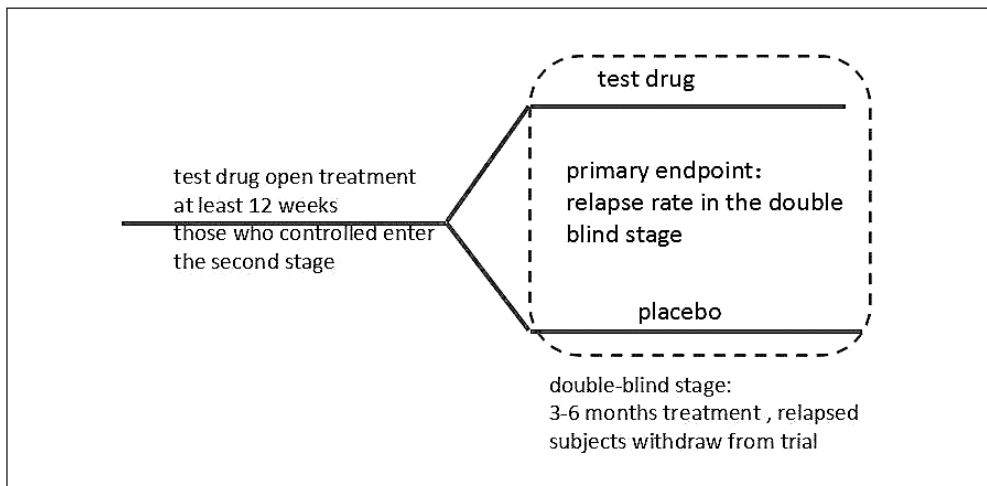


Figure 3. Randomized withdrawal trial design suggested by the guideline.

and receive effective treatment. Such a design minimizes the risks of long-term placebo treatment. However, such a design “cannot provide information on the relative efficacy of the test drug to other active treatments”¹. In addition, the randomized withdrawal design has the following limitations:

1. The evaluation of the efficacy of the test drug depends on the result of the second stage. However, among the initially enrolled subjects, only responders will enter the second stage. This means that this design is relatively inefficient and expensive.
2. Due to the “screening effect” or “enrichment effect” of the test drug treatment in the first stage⁴, those subjects who enter the second stage may have some specific features and may not fully represent the initially enrolled subjects. This may cause results in the second stage that are difficult to explain and may limit extrapolation to the initially enrolled subjects.
3. Open treatment in the first stage can influence the maintenance of blindness in the second stage. All of the subjects who enter the second stage have completed the open treatment in the first stage. If the test drug has a unique smell, taste, or side effects, then subjects, caregivers, and/or investigators may recognize that a subject is taking the test drug or placebo in the second stage. Thus, the blindness in the second stage might be compromised. Although this has rarely been discussed in the literature, this issue may warrant further investigation.

A stand-alone active comparator parallel trial

Due to the sensitivity issue discussed above (e.g. non-inferior to active comparator does

not necessarily mean superior to placebo), a non-inferiority active-controlled trial is not recommended by the Guideline.

Extension of the short-term trial

The Guideline does not support extension of an open label single arm trial because it can provide little evidence of long-term efficacy. For the extension of a short-term trial, the Guideline states:

“Extension of short term studies can however provide evidence of maintenance of effect if there is an active comparator control group and the double blind is maintained... The design of the study should be such that the comparison between test and active comparator treatments in the long term phase remains a comparison between truly random groups. For example selecting responders to short-term treatment with the test product and re-randomizing them to receive long term test or active comparator treatments would be unsatisfactory”¹.

These statements suggest that the extension trial is only for the two-arm positive-controlled superiority study. Long-term placebo treatment is probably ethically unacceptable in the absence of adoption of special management, such as withdrawal design. Therefore, for the extension of a three-arm short-term trial, the subjects in the placebo group have to stop placebo treatment and switch to the test drug or another effective treatment at the beginning of the extension stage. This indicates that the three-arm short-term trial cannot be extended, as the Guideline suggests.

A hybrid design of a randomized withdrawal trial incorporated into a long-term parallel trial

The Guideline claims that:

“Neither a randomised withdrawal trial nor a parallel group active comparator trial is ideal on its own for showing long term efficacy for treatment of schizophrenia. The former generally does not allow for a clear estimation of the magnitude of the treatment effect while the latter has issues with assay sensitivity and does not reliably demonstrate continuing benefit from treatment”¹.

The Guideline further suggests that:

“A hybrid design in which a randomized withdrawal period of 6 months is incorporated into a long term parallel group active comparator trial (after at least 12 months of treatment) could be an optimal approach as it could include the desired aspects from both of the basic designs”¹.

Although the hybrid design depicted in Figure 4 is complex, it can be regarded as a variation of the standard three-arm trial. However, for minimizing the ethical and clinical concerns of long-term use of placebo, the design makes some changes to the three-arm trial. In such a hybrid design, it is possible to:

1. Determine the long-term efficacy of the test drug via comparison between the test drug group and the placebo group in the second stage.
2. Determine the long-term efficacy of the active comparator via the comparison between the active comparator group and the placebo group in the second stage.
3. Compare the test drug and active comparator in the first stage.

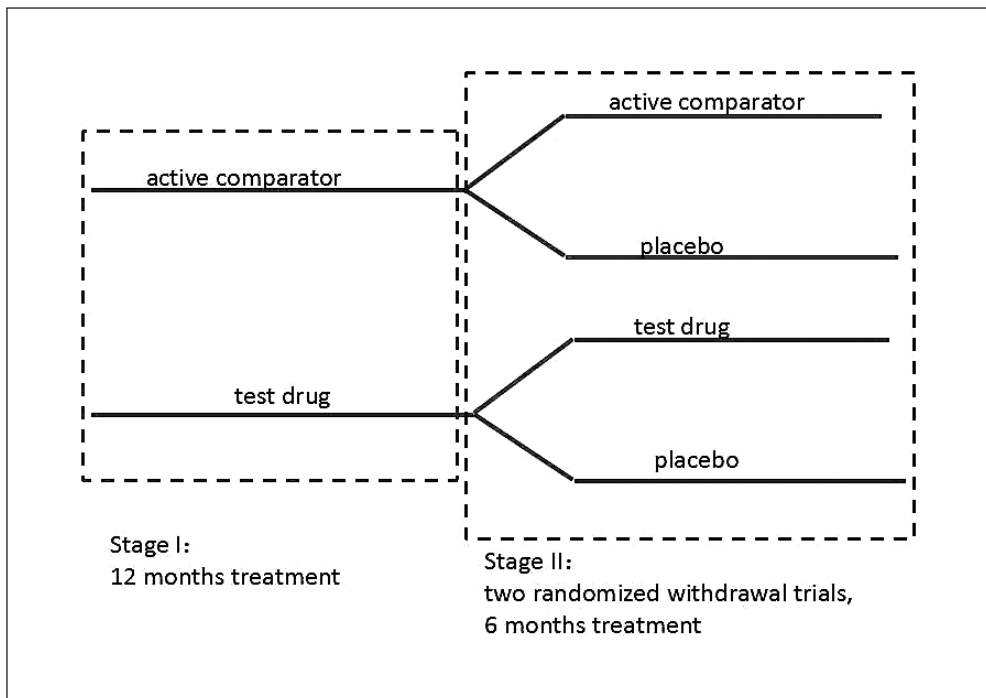


Figure 4. Hybrid design suggested by the guideline.

Similar to a three-arm design, this hybrid design may also have potential issues of multiple comparisons and Type I error inflation. Although this is not clearly stated in the Guideline, based on the suggestion raised by the Guideline on the short-term three-arm trial, we may infer that the primary endpoint of this hybrid design is the comparison between the test drug group and the placebo group in the second stage. Other comparisons will be secondary endpoints and the results will be exploratory.

Although the Guideline considered this hybrid design optimal, few trials have ever used it. This may be due to the fact that this Guideline was issued only 3 years ago. However, this design may also have limitations and disadvantages such that few sponsors want to do such a trial. The following are potential disadvantages:

1. This design has the same limitations and disadvantages as a randomized withdrawal trial. These are relatively low efficiency, expense, and potentially difficult to extrapolate the result of the second stage to subjects enrolled in the first stage due to the screening effect of drug intervention in the first stage. Also noteworthy is that the screening effect means that the three comparisons in this design (test drug versus placebo, active comparator versus placebo, and test drug versus active comparator) may be made in three populations that may not be the same.
 2. In the second stage of this design, the active comparator group will be compared with the placebo group. In other words, the sponsor of such a trial will have to verify the efficacy of the active comparator in the trial. This might not be acceptable to the sponsor, especially when the sponsor is not the license holder of the active comparator.
- Compared with the standard two-arm or three-arm designs, this design might be too complex to be implemented:
1. There will be two times of randomization in this design, one at the beginning of first stage and the other at the beginning of second stage. In other words, a patient would be first randomized to receive test drug or active comparator. If this patient responded to the allocated treatment, he/she would be randomized again to receive test drug or active comparator (depending upon on what kind of treatment the patient received in first stage) or placebo.
 2. Due to a lack of strictly objective endpoints, double-blind (sometimes even double-dummy) should be introduced, not only into the second stage, but also into the first stage. This suggests that a patient who responds to the first stage treatment might have to be unblinded before the randomization of the second stage. This is because a patient who received test drug in the first stage can only be randomized to receive test drug or placebo and cannot be randomized to receive active comparator (and vice-versa). Unblinding before the completion of study might raise concerns of potential bias and the integrity of data.
 3. It cannot be predicted which patient would respond to the treatment of first stage and would enter the second stage. This would make the randomization and blinding of the second stage difficult to implement. For each subject who completed the first stage of treatment and entered the second stage, either a new randomization code or a new record of blinding code should be generated. A new randomization code might be a very large (if not insurmountable) challenge in

the data management of such a trial, since it means one subject would have two randomization codes in one trial: one for the first stage and the other for the second stage. A new record of blinding code would be very difficult for copying and unblinding of randomization codes. For example, the pharmacovigilance group always needs to unblind the treatment when evaluating and reporting serious adverse events. If this group cannot obtain all of the blind codes from statisticians before the start of trial, but instead obtains this information on a case-by-case basis after a patient enters the second stage, it might be problematic.

In conclusion, this hybrid design might be too complex to carry out.

A new three-arm long-term design that might be better

The randomized withdrawal trial cannot compare the efficacy and safety of the test drug with that of an active comparator. In most ca-

ses, extension of the short-term trial is not possible because the active controlled superiority trial is not very common. Furthermore, the “optimal” hybrid design has its disadvantages and might not be very practical. For these reasons, authors suggest a new long-term design that can confirm the long-term efficacy of the test drug over placebo and also allow for comparison of the long-term efficacy between the test drug and an active comparator. This is a three-group design (Figure 5). In this design, each subject would be randomized to one of the three groups after enrollment:

1. Group 1 would receive test drug treatment for 12 months. Patients whose disease could not be controlled in the first 6 months would withdraw from the study. Those whose disease lost control in the second 6 months would also withdraw.
2. Group 2 would receive test drug treatment for 6 months. Those whose disease could be controlled by the test drug would receive a placebo of the test drug for another 6 months. Those whose disease

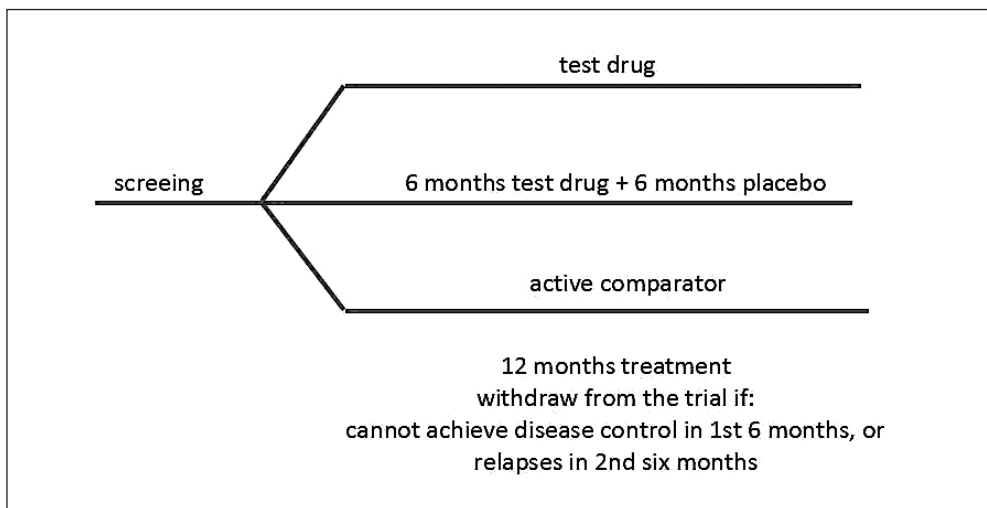


Figure 5. Three groups, design suggested by the author.

could not be controlled by test drug treatment in the first 6 months would withdraw from the study. Those whose disease lost control in the second 6 months (e.g. placebo treatment period) would also withdraw.

3. Group 3 would receive active comparator treatment for 12 months. This group would follow the same rules for withdrawal as Group 1.

In this design, the withdrawal rate (using the number of subjects whose disease is controlled at the end of the first 6 months of treatment as the denominator) in the second 6 months of Group 1 could be compared with that of Group 2. This result could be used to confirm the long-term efficacy of the test drug. For the comparison between test drug and placebo, the 6-month treatment period is consistent with the randomized withdrawal trial design mentioned above. The disease control rate at the end of 12 months treatment for Group 1 (using the number of subjects randomized to this group as the denominator) can be compared with that of Group 3. This could quantify the long-term efficacy of the test drug in comparison with a drug of known efficacy for the treatment of schizophrenia.

Compared with the hybrid design, this three-group design has the following advantages:

1. Each subject would be randomized only once and would have only one randomization code throughout the whole trial. This is important for trial management.
2. Although this design cannot confirm the efficacy of the active comparator over placebo, this design might be more acceptable to the sponsor of the trial.
3. There is no need to unblind subjects and implement new blindness and randomization for the second stage of the trial.

4. The treatment period in this three-group design is 12 months. This is shorter than the 18-month treatment in the hybrid design, but is longer than the treatment period in the randomized withdrawal trial design. A treatment period of 12 months is also common in trials evaluating the long-term efficacy of a test drug for a chronic disease.

Some scientists might have concerns that imbalance in sample size or differences in baseline data might exist when comparing the treatment effect of test drug and placebo in such a trial because subjects would not be re-randomized after 6 months of test drug treatment. However, all enrolled subjects would be randomized at the enrollment and treated in a double-blind manner in this trial. Subjects in Groups 1 and 2 would take the same test drug treatment for the same duration using the same withdrawal criteria. Therefore, there would most probably be few confounders and the imbalance on sample size would only be the result of chance and thus would not be a significant issue. The two times of randomization and blindness and two randomization codes for one subject in the hybrid long-term design make the trial too difficult to carry out. Therefore, it might be worthwhile to try this three-group design because it is much easier to carry out. It can also provide the most information (e.g. the efficacy of the test drug over placebo and the magnitude of the test drug) for the regulatory authority.

Although a three-group trial design could compare the test drug with active comparator and also compare test drug with placebo, one point should be noted. This design cannot completely eliminate the disadvantages of a randomization withdrawal trial design. One disadvantage is that the comparisons between the test drug and placebo and between the test drug and active comparator would probably not be done in the same populations. An-

other disadvantage is the efficiency of the comparison between Group 1 and Group 2. These two disadvantages probably could not be eliminated if confirmation of the long-term efficacy of the test drug over placebo was wanted for this disease. Due to ethical and safety considerations, subjects' disease should be controlled before they could receive long-term placebo treatment.

Discussion

In conclusion, due to the characteristics of schizophrenia, the lack of objective endpoints, the evolution of diagnostic criteria, and other considerations, it can be inferred that the Guideline does not support active control alone for both the short-term and long-term non-inferiority trials. Regardless of apparently not supporting the active control alone, it can be inferred that the Guideline supports comparison of the test drug to the active comparator. However, it would be preferred if the trial could re-test the efficacy of the active comparator as well, as observed in the three-arm short-term trial and the hybrid long-term trial. It may reasonable to assume that EMA, as the regulatory authority, would desire the following short-term and long-term results:

1. The efficacy and safety of the active comparator under current clinical practice.
2. The use of a placebo control to confirm the efficacy of the test drug.
3. The use of an appropriate active comparator to determine the relative efficacy and safety of test drug.

However, no trial design is perfect and each one has unique advantages and disadvantages. It is widely acknowledged that a

clinical trial has many constrains, including scientific considerations, ethical considerations, and management considerations. As such, it might not be wise to incorporate too many objectives into one trial.

If a two-arm active-controlled superiority trial is not suitable, a three-arm design might be a good alternative for a short-term trial. However, this is not actually a non-inferiority trial, but a superiority trial, and the comparison between the test drug group and the active comparator group is exploratory. For a long-term trial, if the short-term trial is a two-arm active control superiority trial, extension might be a good choice. However, if the sponsor declines to perform a two-arm active-controlled superiority trial, the long-term trial design might be problematic. A randomized withdrawal trial design has its own disadvantages and cannot provide a comparison on the efficacy and safety between the test drug and a known active anti-schizophrenic drug. A two-arm active-controlled non-inferiority trial has a sensitivity issue. The hybrid design of a randomized withdrawal trial incorporated into a long-term parallel trial, suggested by the Guideline as "optimal" for providing the information the regulatory authority desires to obtain, has the disadvantages of a randomized withdrawal trial design, and so might be too complex to carry out.

Compared with the randomized withdrawal trial design and the hybrid design of a randomized withdrawal trial incorporated into a long-term parallel trial, the three-group trial design suggested by the authors might be better. It not only provides long-term efficacy of the test drug versus the placebo, but also provides long-term efficacy of the test drug versus active comparator, and is easier to carry out. The authors suggest that the regulatory authority thoroughly discuss and evaluate this design.

Other Information

Sponsors: No.

This paper was supported by a Shanghai Municipal Health Bureau's Subject (Subject title: An investigation on reasons of outpatient depression patient's dropout and an evaluation on interventions targeted to control the dropout rate. Subject code: 20134147).

Conflict of interest

We did not receive any service from any third party for the submission of this manuscript. I (H.J) have no financial relationship with any entities from 24 months prior to the submission to present, except the company I am working for as an employee. However, this company has no relationship with the submission of this manuscript, and this manuscript does not stand for any views or positions of the company I am working for.

References

1. Guideline on clinical investigation of medicinal products, including depot preparations in the treatment of schizophrenia. EMA/CHMP/40072/2010 Rev. 1.
2. Chow SC. In Chow SC, Chang M (editors). Adaptive Design methods in clinical trials. Boca Raton, Florida: Chapman & Hall /CRC. 2007: 78-87.
3. Study of LY2140023 in Schizophrenia (NCT00149292). <https://www.clinicaltrials.gov/ct2/show/NCT00149292>
4. Efficacy and Safety of Asenapine With Placebo and Haloperidol (NCT00156104). <https://www.clinicaltrials.gov/ct2/show/NCT00156104>
5. Hewitt DJ, Ho TW, Galer B, Backonja M, Markovitz P, Gammaitoni A, Michelson D, Bolognese J, Alon A, Rosenberg E, Herman G, Wang H. Impact of responder definition on the enriched enrollment randomized withdrawal trial design for establishing proof of concept in neuropathic pain. *Pain*. 2011; 152(3): 514-521.

* Corresponding author:

Jingjing Zhang
No. 2500, Zhahang Road
Minhang District
Shanghai
China
Tel. 86-13651723955
E-mail: zjjzlc@hotmail.com