

On Statistical Principles for Clinical Trials in Pharmaceutical Development – A Review of China FDA Draft Guidance

Shein-Chung Chow¹, Zinan Chen¹ and Fuyu Song^{2*}

¹Duke University, School of Medicine, Durham, North Carolina, USA

²Peking University, Clinical Research Institute, Peking University Health Science Center, Beijing, China

Abstract

On August 21, 2015, the Center for Drug Evaluation (CDE) of China Food and Drug Administration (CFDA) circulated draft guidance on Statistical Principles for Clinical Trials in Pharmaceutical Development for public comments. The draft guidance is to assist the sponsor for providing accurate and reliable assessment of a test treatment under investigation in China. The draft guidance focuses on study designs, basic considerations for on-going trials, data management, and statistical principles for data analysis and reporting. In this article, we intend to comment on the draft guidance and provide constructive input and recommendations whenever possible.

Keywords: Adaptive design; Missing data imputation; Non-inferiority trial; Protocol amendment; Independent data monitoring committee (IDMC)

Introduction

In the past several decades, the quality, validity, and integrity of data collected from clinical trials conducted for pharmaceutical/clinical development of test treatments under investigation in China have been challenged by several regulatory agencies such as the United States Food and Drug Administration (FDA). To assist the sponsor for providing accurate and reliable assessment of a test treatment under investigation in China, the Center for Drug Evaluation (CDE) of China Food and Drug Administration (CFDA) circulated draft guidance on statistical principles for good statistical practice (GSP) in clinical trials for public comment on August 21, 2015 (CDE/CFDA, 2015). The draft guidance is intended to cover pharmaceutical/clinical development for chemical compounds, biological products, and traditional Chinese medicines (TCMs).

The draft guidance first provides some basic concepts regarding exploratory and confirmatory studies in clinical investigation and development program followed by the significant impact of bias/variation for controlling the overall type error rate at a pre-specified level of significance. For basic considerations of study design of an intended clinical trial, the draft guidance discusses several different types of clinical trials including multicenter trial, comparative trial and adaptive design methods including power calculation for sample size. During the conduct of the clinical trial, the draft guidance indicates the important of trial monitoring for assurance of data quality, protocol amendment and sample size adjustment, interim analysis, and role and responsibility of an independent data monitoring committee (DMC). Most importantly, similar to ICH E9, the draft guidance provides several statistical principles for data analysis and reporting in clinical trials, which include statistical analysis plan (SAP), analysis set, missing and outlying data, data transformation, statistical methods for data analysis, analysis for safety and tolerability, and statistical analysis report.

While we congratulate the effort that CDE/CFDA in developing the draft guidance, in this article, we would like to make an attempt to comment on the draft guidance and provide constructive input/recommendations whenever possible.

Overall Considerations for Clinical Trials

The draft guidance starts with the necessity for development of clinical program which include early phase of clinical development

such as proof-of-concept studies, dose finding studies, and late phase of clinical development such as phase III confirmatory studies. The draft guidance then focuses on study endpoints for measurement of therapeutic effect and distinguishes the difference between exploratory analysis and confirmatory analysis for substantial evidence. As indicated in the draft guidance, it is important to control the overall type I error rate at a pre-specified level of significance such as 5% and at the same time for achieving a desired power such as 80%.

The draft guidance indicates that there are different types of study endpoints including primary endpoint(s), secondary endpoints, composite endpoints, global assessment endpoint, surrogate endpoints, and qualitative endpoints. In clinical investigation, it is suggested that the intended study should be powered based on a single primary endpoint. In case there are multiple endpoints, a composite endpoint that incorporates the multiple endpoints may be considered. Surrogate endpoints should be used under that assumption that the surrogate endpoints are predictive of clinical outcomes (measured by the primary endpoint). It should be noted that sample size calculations for different types of endpoints (e.g., quantitative endpoint versus qualitative endpoint; primary endpoint versus secondary endpoints; surrogate endpoint versus regulatory clinical endpoint) may be very different [1]. Thus, in the interest of controlling the overall type I error rate at a pre-specified level of significance, it is suggested primary/secondary endpoints be clearly specified in the study power and power calculation should be performed based on the primary endpoint for clinical studies to be conducted as described in the clinical development program for an unbiased and reliable evaluation of the test treatment under investigation.

Basic Design Considerations

Types of clinical trial design

In the draft guidance, CDE/CFDA focuses on parallel-group

*Corresponding author: Fuyu Song, M.S., Peking University, Clinical Research Institute, Peking University Health Science Center, No. 38 Xueyuan Road, Haidian District, Beijing, 100191, China, Tel: 110 402 430 047; E-mail: fysong@hsc.pku.edu.cn

Received September 01, 2015; Accepted September 29, 2015; Published September 30, 2015

Citation: Chow SC, Chen Z, SongF (2015) On Statistical Principles for Clinical Trials in Pharmaceutical Development – A Review of China FDA Draft Guidance. Pharmaceut Reg Affairs 4: 152. doi:[10.4172/2167-7689.1000152](https://doi.org/10.4172/2167-7689.1000152)

Copyright: © 2015 Chow SC, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

design and crossover design (e.g., standard 2 x 2 crossover design and Williams' 6 x 3 crossover design for comparing three treatment), which are considered the most commonly employed study designs in pharmaceutical research and development. In addition, the draft guidance also indicates that factorial design may be useful for identifying possible interactions and studying drug products which contain multiple active components.

However, little was mentioned regarding confounding factors between demographics and patient characteristics. Thus, it is suggested that appropriate study designs for handling possible confounding factors be included in future revision of the draft guidance.

Multicenter studies

The draft guidance emphasizes the importance of center difference and treatment-by-center interaction which may have an impact on the evaluation of the test treatment under investigation. Thus, it is suggested that statistical test for treatment-by-center interaction be performed before the data from individual centers are combined for a combined analysis. In practice, however, a false positive effect of treatment-by-center interaction is most likely observed if there are too many heterogeneous small centers. In this case, the method of center grouping (either based on geographical location or random grouping) is usually recommended to form some bigger dummy centers and the treatment-by-center interaction is then tested based on these dummy centers for evaluation of the true treatment-by-center interaction.

As too many small centers could cause the issue of observing false treatment-by-center interaction, this raises the interesting question that how many centers should be considered in clinical trials. In practice, as rule of thumb for selection of the number of centers in a given clinical trial, it is suggested that the number of centers should not be greater than the number of subjects in each center for achieving the optimal statistical properties of the data collected from each center [2].

Types of comparative trials

For comparative trials, the draft guidance discusses superiority trial and non-inferiority and equivalence trials. As indicated by the draft guidance, one of the key issues in non-inferiority/equivalence trials is the selection of non-inferiority margin (or equivalence limit). Similar to FDA (2010a) [3], the draft guidance provides general principles for selection of non-inferiority margin and suggests considering either M_1 or M_2 approach. M_1 may be selected by considering the lower confidence limit of the effect of the active control agent, while M_2 may be selected depending upon the retention ratio (f) between the effect of test treatment and the effect of the active control agent. It is suggested that f should be selected based on clinical judgment.

Sample size

In clinical trials, sample size is often selected for achieving a desired power (say 80%) of correctly detecting a clinically meaningful difference at a pre-specified level of significance (say 5%) under a valid study design. The process is to first derive a statistical test under the null hypothesis. The derived statistical test is then evaluated under the alternative hypothesis for achieving the desired power. It is suggested the reference of Chow et al. [1], which includes formulas and/or procedures for sample size calculation for testing equality, superiority, non-inferiority, and equivalence under various designs with different data types (such as continuous, discrete, and time-to-event), be included in the revision of draft guidance for completeness.

Adaptive trial designs

For the potential use of adaptive design in clinical trial, the draft guidance by CDE/CFDA is similar to that of the US FDA (FDA, 2010b) [4]. However, in addition to the popular group sequential design, the draft guidance also discusses the potential use of response-adaptive design which is considered acceptable to the CDE/CFDA. Similar to the US FDA, CDE/FDA suggests that adaptive design should be used in early clinical development rather than later phase clinical development. Both CDE/CFDA and US FDA agree upon that (i) the overall type I error rate must be controlled for clinical trials utilizing adaptive trial designs, and (ii) sample size re-estimation should be conducted by an independent data monitoring committee (IDMC) in a blinded fashion.

The draft guidance did not mention the commonly considered adaptive trial designs such as adaptive dose finding design and two-stage seamless (e.g., phase I/II or phase II/III) adaptive design. Thus, it is suggested that design-specific regulatory guidance should be developed in order to assist the sponsors in their clinical trial utilizing complicated adaptive trial design.

Basic Considerations for Ongoing Trials

The draft guidance provides several statistical principles on protocol amendment and sample size adjustment, interim analysis, and independent data monitoring committee for on-going trials. Comments and recommendations on these statistical principles are briefly discussed below.

Protocol amendment

The draft guidance allows the issuance of protocol amendment by relaxing or modifying the inclusion/exclusion criteria if a notable portion of subjects fail to meet the eligible criteria or there is slow enrollment. It, however, should be noted that the original target patient population under study could have become a similar but different patient population if significant changes or modifications are made. This raises the concern regarding the validity of the statistical inference drawn based on data collected before and after protocol amendment [5].

In practice, there is a risk that major (or significant) modifications made to the trial procedures and/or statistical procedures could lead to a totally different trial, which cannot address the scientific/medical questions that the clinical trial is intended to answer. In clinical trials, most investigators consider protocol amendment is a God-sent gift which allows the flexibility to make any changes/modifications to the on-going clinical trials. It, however, should be recognized that protocol amendments have potential risks for introducing additional bias/variation to the on-going clinical trial. Thus, it is important to identify, control, and hopefully eliminate/minimize the sources of bias/variation. Thus, it is of interest to measure the impact of changes or modifications that made to the trial procedures and/or statistical methods after the protocol amendment. This has raised another concern regarding (i) the impact of changes made and (ii) the degree of changes that are allowed in a protocol amendment.

In current practice, standard statistical methods are applied to the data collected from the actual patient population regardless the frequency of changes (protocol amendments) that have been made during the conduct of the trial provided that the overall type I error is controlled at the pre-specified level of significance. This, however, has raised a serious regulatory/statistical concern that whether the resultant statistical inference (e.g., independent estimates, confidence intervals, and p-values) drawn on the originally planned target patient

population based on the clinical data from the actual patient population (as the result of the modifications made via protocol amendments) are accurate and reliable? After some modifications are made to the trial procedures and/or statistical methods, not only the target patient population may have become a similar but different patient population, but also the sample size may not achieve the desired power for detection of a clinically important effect size of the test treatment at the end of the study. In practice, we expect to lose power when the modifications have led to a shift in mean response and/or inflation of variability of the response of the primary study endpoint. As a result, the originally planned sample size may have to be adjusted. Thus, it is suggested that the relative efficiency at each protocol amendment be taken into consideration for derivation of an adjusted factor for sample size in order to achieve the desired power.

In addition, it is suggested that the impact of major changes (as described in protocol amendment) on statistical inference should be carefully assessed through clinical trial simulation with relevant sensitivity analysis. Also, it is suggested that the number of protocol amendments that are allowed for a given size of clinical trial should be specified to prevent the potential abuse of changing study protocol through the issuance of protocol amendments as frequent changes could lead to a similar but different study protocol which may not be able to address the scientific or medical questions the original protocol intended to answer.

Interim analysis and sample size re-estimation

For clinical trials utilizing a group sequential design with planned interim analyses, the guidance provides statistical principles regarding when and how the planned interim analyses would be performed. The draft guidance indicates that all interim analyses should be performed in a blinded fashion by an independent data safety monitoring committee (IDMC) to maintain the integrity of the trial. Appropriate alpha spending function should be used in order to fulfill with study objectives at interim analyses and to control the overall type I error rate at the pre-specified level of significance.

Although the draft guidance indicates that planned interim analyses allow the investigator to stop the trial early due to safety, futility and/or efficacy according to pre-specified stopping boundaries, it is suggested the stopping boundaries be viewed as statistical guide for consideration of stopping the trial which will give the IDMC the option to carefully perform risk/benefit assessment based on the entire clinical picture (performance) at interim.

Independent data monitoring committee (IDMC)

The draft guidance seems to suggest that interim analysis be performed by the independent data monitoring committee (IDMC) in an unblinded fashion. In practice, however, most IDMCs prefer performing interim analysis without unblinding the treatment codes whenever possible. Thus, it is suggested that “unblinded interim analysis” be changed to “blinded interim analysis whenever possible” for maintaining the integrity of the trial.

In addition, the draft guidance emphasizes that it is important to keep the IDMC independent of the clinical project (operation) team. In practice, since the IDMC members are selected and compensated by the sponsor and the IDMC charter is often developed by the sponsor, the independence of the IDMC has been challenged [6,7]. Regulatory guidance in this regard is needed.

Clinical Data Management

The draft guidance requires that clinical data management be in compliance clinical data management GCP published by CDE and State Food and Drug Administration (SFDA) to assure quality, accuracy, reliability, and completeness of the collected clinical data [8,9]. The draft guidance provides general principles for management of data collected from clinical trials.

In recent clinical trials, the use of electronic data capture (eDC) has become very popular. However, as eDC may be available in major cities such as Shanghai and Beijing in China, it may not be available and hence applicable in secondary cities. As a result, paperless eDC together with paper case report forms (CRFs) are necessarily employed for data capture. Thus, it is suggested that regulatory requirements for eDC with and without paper CRFs be provided in the draft guidance. In addition, a flowchart for clinical data management would be helpful to assist the sponsors in compliance with clinical data management GCP as required by CDE/CFDA.

Statistical Analysis and Report

Statistical analysis plan

The guidance suggests that statistical analysis plan (SAP) should be developed at the same time the study protocol is developed. The SAP should be modified to reflect changes made in the protocol amendments. SAP can only be changed prior to database lock or data unblinding. Although this requirement is similar to that of the United States Food and Drug Administration (FDA), it is not clear the SAP and the statistical section included in the study protocol is the same document.

In practice, a statistical section is developed for inclusion in the study protocol. After the initiation of the trial, a full SAP is then developed prior to database lock and/or data unblinding. If there are planned interim analyses, short versions of SAP will be developed for Data Monitoring Committee (DMC) analysis. Thus, it is suggested that the guidance clarify the concepts of statistical section for inclusion in the study protocol, DMC SAP, and full SAP for final data analysis to avoid possible confusion.

Analysis set

The draft guidance defines a full analysis set (FAS) as an analysis set that excludes subjects (i) who violate important inclusion/exclusion criteria, (ii) who got randomized but never receive treatment, and (iii) who do not have any post-randomization observations. The draft guidance seems to suggest that the full analysis set (FAS) should be the primary analysis set for evaluation of the test treatment under investigation. This actually contradicts to the intent-to-treat (ITT) analysis set suggested by the ICH E9 guideline which contains all randomized subjects regardless the compliance of the subjects [10]. The FAS is in fact a subset of the intent-to-treat (ITT) analysis set. The draft guidance indicates that the results based on FAS and PPS analysis sets should be compared for consistency. However, the comparison between the ITT and FAS is not mentioned. In practice, the analysis results based on ITT and FAS could be substantial. For example, hypothetically, suppose there are 10 subjects in the ITT analysis set and two subjects that meet one of the three criteria for FAS are excluded in the FAS analysis set. Suppose four subjects are considered responders. In this case, the ITT analysis estimates the response rate is 40%, while the FAS analysis gives an estimate of the response rate by 50%. The difference of 10% in response rate is not negligible.

If analysis based on FAS is considered the primary analysis, it is suggested that the following issues should be addressed before the test treatment can be evaluated accurately and reliably. First, what is the definition of important inclusion/exclusion criteria? Second, what if there is mix-up in randomization as this is commonly seen in randomized clinical trials? In other words, some subjects receive treatment A when they are randomized to treatment B and vice versa. Third, should these subjects be replaced for achieving the desired power? Finally, what if a significant difference between analysis results based on ITT and FAS is observed? For binary response, the selection of the denominator has an impact on the estimation of the response rate, which could alter the conclusions of the analysis results.

The draft guidance indicates that safety set (SS) should include all randomized subject who receive at least one treatment. In practice, it is possible that subjects receive partial treatment before dropout. Thus, it is suggested that the definition of SS be modified to include all randomized subjects who receive any amount of treatment.

Missing and outlying data

The draft guidance seems to suggest that under circumstances such as missing completely at random (MCAR) or missing at random (MAR), some statistical methods such as last observation carried forward (LOCF) can be used for missing data imputation. It should be noted that the validity of LOCF for missing data imputation is questionable as pointed out by many researchers [11-13].

One of major issues in missing data imputation is that the imputed missing data are usually estimates obtained from a statistical model/method based on the observed data. For a given estimate, there is variability associated with the estimate regardless which statistical model/method is used. As a result, the original overall type I error rate and desired statistical power cannot be preserved. This becomes more serious if the proportion of missing data increases. The guidance, however, did not discuss the potential impact on the accuracy and reliability of treatment assessment after missing data imputation. Thus, it is suggested that the guidance should focus on the prevention of missing data rather than providing methods for missing data imputation [14].

Regarding outlying data, the guidance suggests both statistical reasoning and clinical judgment should be employed for outlying detection. Statistical methods for outlying data detection and the handling of the identified outlying data should be pre-specified in the study protocol whenever possible. Clinical results with and without the identified outlying data should be compared for consistency. In case inconsistent results are observed, the treatment effect should be evaluated with caution. While these sound reasonable, no statistical methods for outlying data detection and handling of identified outlying data (which should be done in a blinded fashion as suggested by the guidance) were mentioned in the guidance.

In practice, it should be noted it is always very controversial that either it is an outlying data or the model is incorrect. Outlying data could be critical if the results could be altered with and without the outlying data. Thus, it is suggest that a non-parametric method be employed when there are potential outlying data.

Data transformation

The draft guidance suggests that data transformation such as log-transformation or square root transformation should be pre-specified based on prior knowledge of the data collected from similar studies. It, however, should be noted that data transformation for a valid analysis

depends upon the sampling distribution of the observed data. Thus, it is suggested that the draft guidance be modified by considering the commonly used method of Box-Cox transformation (which includes log-transformation and square root transformation) for an accurate and reliable assessment of the test treatment under investigation [15].

Statistical analysis

For statistical analysis, the draft guidance provides several statistical principles for a valid data analysis. These statistical principles include (i) descriptive statistics, (ii) statistical inference including estimation, confidence interval, and hypotheses testing, (iii) baseline comparability and analysis of covariance (ANCOVA), (iv) evaluation of interaction, (v) assessment of center effect, (vi) subgroup analysis, and (vii) adjustment for multiplicity. Among these principles, we would like to make comments on adjustment of multiplicity as follows.

As indicated in Chow (2011) [16], when conducting clinical trials involving multiple comparisons, the following questions are always raised:

- Why do we need to adjust for multiplicity?
- When do we need to adjust for multiplicity?
- How do we adjust for multiplicity?
- Is the family-wise error rate (FWER) well controlled?

To address the first question, it is suggested that the null/alternative hypotheses be clarified since the type I error rate and the corresponding power are evaluate under the null hypothesis and the alternative hypothesis, respectively.

Regarding the second question, it should be noted that adjustment for multiplicity is to ensure that the simultaneously observed differences are not by chance alone. For example, for evaluation of a test treatment under investigation, if regulatory approval is based on single endpoint, then no alpha adjustment is necessary. However, if regulatory approval is based on multiple endpoints, then α adjustment is must in order to make sure that the simultaneously observed differences are not by chance alone and they are reproducible? Conceptually, it is not correct that alpha needs to be adjusted if more than one statistical test (e.g., primary hypothesis and secondary hypothesis) is to be performed. Whether the alpha should be adjusted depends upon the null hypothesis (e.g., a single hypothesis with one primary endpoint or a composite hypothesis with multiple endpoints) to be tested. The interpretations of the test results for single null hypothesis and composite null hypothesis are different.

For questions (3) and (4), several useful methods for multiplicity adjustment are available in the literature [17-19]. These methods are either single-step methods (e.g., Bonferroni's method), step-down methods (e.g., Holm's method), or step-up methods (e.g., Hochberg's method). In the next section, some commonly employed methods for multiplicity adjustment are briefly described.

Westfall et al. [20] also pointed out that the controversial issues of multiplicity in clinical trials that are commonly encountered, which include (1) penalizing for doing more or good job (i.e., performing additional test), (2) adjusting α for all possible tests conducted in the trial, and (3) the family of hypotheses to be tested. Penalizing for doing good job is referred to adjustment for multiplicity for dose finding trials that include more dose groups. For adjusting α for all possible tests conducted in the trial, although the is controlled at the pre-specified level, it is over-killed because it is not the investigator's best interest to show that all of the observed differences simultaneously are not by

chance alone. In practice, it is very controversial for selecting appropriate family of hypotheses (e.g., primary endpoints and secondary endpoints for efficacy or safety or both) for multiplicity adjustment for clinical evaluation of the test treatment under investigation.

It should be added that the most worrisome impact of multiplicity on the inference for clinical trials is not only the control of FWER though that can be problematic, but also the power for correctly detecting a clinically meaningful treatment effect. One of the most controversial issues in multiplicity is having adequate control of FWER but may fail to achieve the desired power due to multiplicity.

As a result, it is suggested that the draft guidance should be revised to clarify the confusion of adjustment for multiplicity in clinical trials in order to address the above described questions.

Safety and tolerability assessment

The draft guidance suggests laboratory test values, sign and symptom, adverse events, and specific safety parameter such as QT_c prolongation for cardio-toxicity is assessed for safety. In clinical trials, the incidence rates of adverse events (AE) and severe (or serious) adverse events (SAE) (especially related to test treatment) are usually considered the primary endpoints for safety/tolerability assessment. The guidance suggests p-value in conjunction with confidence interval be used for key safety parameters. In practice, the assessment mean post-treatment change from baseline (either absolute change or relative change) may not accurately characterize the risk of the test treatment on individual subjects if the mean change falls within the therapeutic index and/or safety margin (e.g., normal ranges of the laboratory test values). Thus, it is suggested a shift table analysis that contains categorical shift (e.g., shift from normal to abnormal) be conducted for assessment of key safety parameters.

Statistical analysis report

The draft guidance points out that statistical analysis report is an important document which summarizes analysis results by including tables, figures and listings that generated using appropriate statistical software. It, however, is not clear whether a separate statistical report or an integrated clinical report is required for regulatory submission. Thus, it is suggested that the revision of the guidance be specific for avoiding possible duplicated effort in reporting.

Concluding Remarks

As indicated in CDE/CFDA (2015) [21], the draft guidance is intended for chemical compound, biological product, and traditional Chinese medicine (TCM). It, however, is not clear whether statistical principles described in the draft guidance can be directly applied to the traditional Chinese medicine (TCM) due to some fundamental differences between a Western medicine and a TCM [22]. For example, Western medicines (WMs) often contain a single active ingredient while TCMS may consist of up to 12-15 components. As a result, statistical methods for development of WM are necessarily modified before they can be applied to the development of TCMS.

The draft guidance does provide statistical principles for dealing with some key scientific factors such as the determination of non-inferiority margin, handling of missing and outlying data, the issue of multiplicity, and the importance of independent data monitoring committee (DMC). However, many practical issues in pharmaceutical/clinical development are still debatable. For example, many practical issues in multi-national (or multi-regional) trials that are commonly conducted in global pharmaceutical development. These issues

include, but are not limited to, the use of central laboratory versus local laboratory, potential differences in ethnic factors, sample size requirement at different countries (regions), and requirement for bridging studies.

References

1. Chow SC, Shao J, Wang H (2007) *Sample Size Calculation in Clinical Research* (2nd Edn) Chapman and Hall/CRC Press, Taylor and Francis, New York.
2. Shao J, Chow SC (1993) Two-stage sampling with pharmaceutical applications. *Statistics in Medicine* 12: 1999-2008.
3. FDA (2010a) *Guidance for Industry – Non-inferiority Clinical Trials*. The United States Food and Drug Administration, Rockville, Maryland, USA.
4. FDA (2010b) *Guidance for Industry – Adaptive Design Clinical Trials for Drugs and Biologics*. The United States Food and Drug Administration, Rockville, Maryland, USA.
5. Chow SC, Song FY (2015) Statistical considerations for traditional Chinese medicine clinical trials. *Journal of Case Studies* 4: 134-142.
6. Chow SC, Corey R, Lin M (2012) On independence of data monitoring committee in adaptive clinical trial. *J Biopharm Stat* 22: 853-867.
7. Hess C, Roe MT, Gibson CM, Temple R, Pencina M (2014) Independent data monitoring committees: Preparing a path for the future. *Am Heart J* 168: 135-141.
8. SFDA (2003) *Good Clinical Practice (GCP)*, China Food and Drug Administration, Beijing China.
9. CDE (2012) *Guidance on Data Management for Clinical Trial*. Center for Drug Evaluation, Beijing China.
10. ICH E9 (2008) *Statistical Principles for Clinical Trials*. International Conference on Harmonization, Geneva, Switzerland.
11. Molnar FJ, Hutton B, Fergusson D (2008) Does analysis using "last observation carried forward" introduce bias in dementia research? *CMAJ* 179: 751-753.
12. Cheng B, Chow SC (2012) Validity of LOCF. In *Encyclopedia of Biopharmaceutical Statistics* (3rd edn) CRC Press, New York, 1427-1432.
13. Bredemeier M (2012) Last observation carried forward approach threatens the validity of intent-to-treat analysis in fibromyalgia trials: Comment on the article by Arnold et al. *Arthritis and Rheumatism* 64: 2809-2810.
14. Roderick J Little, Ralph D'Agostino, Michael L Cohen, Kay Dickersin, Scott S Emerson (2010) *The Prevention and Treatment of Missing Data in Clinical Trials*. *N Engl J Med* 367: 1355-1360.
15. Box GEP, Cox DR (1964) An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 26: 211-252.
16. Chow SC (2011) *Controversial Statistical Issues in Clinical Trials*. Chapman and Hall/CRC Press, Taylor and Francis, New York, New York.
17. Hsu JC (1996) *Multiple Comparisons – Theory and Methods*. Chapman and Hall, London, UK.
18. Westfall PH, Tobias RD, Rom D, Wolfinger RD, Hochberg Y (1999) *Multiple Comparisons and Multiple Tests Using the SAS* (2nd edn) SAS Institute, Gary, North Carolina.
19. Chow SC, Liu JP (2013) *Design and Analysis of Clinical Trials – Revised and Expanded* (3rd edn) John Wiley and Sons, New York, New York.
20. Westfall P, Bretz F (2010) Multiplicity in clinical trials. In *Encyclopedia of Biopharmaceutical Statistics* (3rd edn) Taylor and Francis, New York.
21. CDE/CFDA (2015) *Guidance on Statistical Principles for Clinical Trials in Pharmaceutical Development*. Center for Drug Evaluation, China Food and Drug Administration, Beijing, China.
22. Chow SC (2015) *Quantitative Methods for Traditional Chinese Medicine Development*. Chapman and Hall/CRC Press, Taylor and Francis, New York.

Citation: Chow SC, Chen Z, SongF (2015) On Statistical Principles for Clinical Trials in Pharmaceutical Development – A Review of China FDA Draft Guidance. *Pharmaceut Reg Affairs* 4: 152. doi:[10.4172/2167-7689.1000152](https://doi.org/10.4172/2167-7689.1000152)