The Technical Guideline for the Application of Real-world Data in Clinical Evaluation of Medical Devices

(Draft for Comments)

Contents

The Technical Guideline for the Application of Real-world Data in Clinical Evaluation of Medical Devices

（Draft for Comments）

In order to solve the problems affecting innovation, quality and efficiency of medical devices, accelerate the modernization of the medical device management system and improve governance capacity, NMPA launched “Scientific Method of Medicine Supervision of China”, which is based on the reality of China and focused on the reform and innovation of the review and approval system. This guideline is based on the findings of the first project of the action plan-" Research on How to Apply Real-world Data in Clinical Evaluation of Medical Device Scientifically".

This guideline aims at standardizing and guiding the application of real-world data in clinical evaluation of medical devices, providing technical guidance for the application of real-world data in registrations and evaluation of medical devices.

This guideline is a technical guidance documents for applicants and reviewers, so administrative process like registration approval is not involved. This guideline is not compulsory. Any other methods that can meet the requirements of laws and regulations can be adopted too. The guideline is based on the existing scientific power and cognitive level and it will be improved and revised as the development of real-world data research and changes of relevant laws, policies and standards.

1. Overview

1.1 Real-world Data & Evidences

The real-world data in this guideline refers to data from real medical cases and its resources are not traditional clinical trials. These data can reflect the health status of patients and the medical service process in the actual diagnosis and treatment.

Real-world data research involves real-world data, many different disciplines, multi-disciplinary methods and techniques in Epidemiology, Biostatistics and Evidence-Based Medicine. The research process includes collecting real-world data in clinical conventions and using rational epidemiological design and statistical analysis under pre-established research assumptions. It can be future-oriented or conclusion-oriented and it completes traditional clinical trials. The vast amount of clinical data generated by clinical practices which is supported by information-technology laid a solid foundation for real-world data research.

Real-world evidences refer to clinical evidences related to product use and potential risks/benefits, which are based on the analysis of real-world data. However, due to the different sources and types of real-world data, data quality and information they covered

vary greatly, not all real-world data are applicable to clinical evaluation of medical devices. In the context of related requirements, real-world evidence can show the risk and benefit features of medical devices in the whole service life, which may help in decision-making in supervision.

1.2 Advantages & Disadvantages of Real-world Data

With fewer restrictions on patients' conditions, real-world data studies have a wider coverage of the population than traditional clinical trials. Real-world data studies are carried out in a realistic health care environment and its research conclusions are easy to use in speculation. Real-world research emphasizes the comprehensive use of data from different resources, for instance, hospital electronic medical records, enrollment data, regional health care data, medical insurance data, etc. This means that we can get clinical conclusion data on a long-term basis. Real-world studies can also be used in observing rare serious adverse events, answering questions about rare diseases, and evaluating differences in clinical outcomes among different people.

There are two disadvantages of real-world research. Firstly, there are many data sources in the real-world. In collecting and storing data, measurement/classification errors or data missing are common occurrences. The data is not structured and the data quality needs to be evaluated. Connection between different data sources is lame. Secondly, studies based on real-world data often have many biases and confusions, and the results is difficult to use in stipulation.

2. Common Resources and Classification of Real-world Data

Common sources of real-world data include registry databases, hospital electronic medical records, regional health care data, medical insurance data, health records, routine public surveillance data, self-reported patient data (including home environments), and data from other health tests (e.g. mobile devices). Real-world data suitable for medical devices also include data from whole service life of a medical device (manufacturing, marketing, transportation, storage, installation, use, maintenance, retirement and disposal), e.g. acceptance report, maintenance record, user feedback, operation environment, calibration records, performance log, original image, etc.).

In terms of clinical evaluation of certain devices, the above data sources can be divided into two categories according to the relationship between time and research launch time. The first type is the existing data resources, it means the data resources already existed when the research carried out. Based on the differences in the data generating process, this type of data resource can be divided into two cases as bellow:

a) First type of data came from the process of providing health care services and payment, such as electronic medical record, medical insurance records, health records, etc.

b) Second type of data refers to database which is established in an orderly manner and it is based on certain research purpose, data standard and data collection method. For instance, registration data, database based on the effectiveness clinical trials.

The second type is for the purpose of clinical evaluation on certain devices, it usually has clear data standards and data collection models. In short, these data are established in an orderly manner. Representative data includes devices registration data and effectiveness clinical trials data.

3. Evaluation of Real-world Data

The quality of real-world data directly affects the reliability of evidence from real-world research. So, evaluation of real-world data is the foundation of carrying out real-world research. In addition, corresponding reliability assurance system and evaluation measures need to be established and implemented. To evaluate the data, researchers are required to consider reliability of data source and control data quality throughout the whole research process.

Firstly, in using the data from existing data sources for clinical evaluation, it is necessary for researchers to evaluate whether the existing data includes the all targeted population, key variables and follow-up duration. Researchers should also consider the accuracy and integrity of medical device identification information, IFU, etc. Secondly, in specific research design, researchers should filter and extract existing data. In this way, we can get a real-world database.

During the data collection, firstly, we need to consider the rationality and feasibility of the research design. There are many inclusion and exclusion criteria in real-world data, we need to ensure quality of follow-up investigation and its time. Secondly, it is necessary to ensure the authenticity, accuracy and traceability of data, establish systematic follow-up standards, then, train and supervise researchers. It is important to identify the possible research biases and confounding factors before taking action. At the same time, we should measure and record relevant confounding factors in the preparation process. To control and minimize the effects of confounding factors, we can use hierarchical analysis, multi factor analysis and tendentiousness score in data analysis.

For standardize evaluation of data management, we can consider factors like management processes, personnel, information systems, and data standardization. In term of data standardization, we should establish a standardized document format and data structure, in addition, we need a standardized variable dictionary.

In evaluation, we should consider the relevance and reliability of the data. In terms of relevance between data and the research, we should evaluate sufficiency of individual variable and consider whether it can adequately answer the clinical questions base on the research purpose; in terms of adequacy, we should put our focus on accuracy of data

collection which includes collection scope, variables, data dictionary, and collection methods (such as data extraction tables). In this way, we can minimize errors, and ensure the authenticity and integrity of data.

4. Design Principles, Common Types & Statistical Analysis Techniques of Real-world Data Research

4.1 Design Principle of Real-world Data Research

In the clinical evaluation of medical devices, we should use the real-world data in a practical manner and design the research based on the specific research purpose. The overall research plan includes clarifying research questions, confirming data sources, determining how data is generated, filtering data and forming research teams. The overall research design includes determining the design type, clarifying the research objects and research variables, identifying the source of confounding factors, conducting reasonable control, and formulating a statistical analysis plan. In planning, designing and implementing real-world research, we should also pay full attention to ethical factors and data security.

We should pay attention to biases in using real-world data, no matter what type it is. This may limit the reasoning and interpretations of research results. To minimize potential biases, researchers need to identify them during the planning, design, implementation, and analysis stages of the research, and formulate corresponding measures in advance to form a research plan and analysis plan carefully before acquiring and analyzing data. Applicants can choose different research designs according to their needs. They can also choose different research designs at the same time if it is necessary.

4.2 Common Types of Real-world Data Research Design

The types of real-world data research can be mainly divided into 3 types: Effectiveness test, observational research and other design types.

4.2.1 Effectiveness Test

pRCT (Pragmatic Randomized Controlled Trials) is typical in interventional research and it is also an important part of effectiveness test. pRCT means using random or contrasting methods in real or less-real medical environment, and compare results of different intervention methods. Being different from traditional contrast tests, pRCT are usually conducted in real clinical environment--the research object may have comorbidities, the intervention processes is highly similar to real clinical case, however, the processes are influenced by skills or the experiences of intervenors. Therefore, we should consider the research planning and designing comprehensively.

Because of different choices of groups, pRCT usually includes the following test types:

a) Individual pRCT: It means that individual is intervened and monitored as one random grouping unit. In clinical evaluation, the unit means the patient.

b) Group pRCT: It means that the group (such as hospitals, clinics, districts and schools) is considered as one intervention unit. In analyzing the real-world data, we should consider both group effect and evaluate group members case by case.
c) Hierarchical pRCT: It is a special type of pRCT, group members accept intervention randomly in batches at different time.

Selection of outcome indicators of pRCT is based on research targets which includes safety, effectiveness, treatment adherence, healthy economics and so on. Generally, the clinical indicators which have important clinical meaning for patients (or the user of this research result) will be chose as outcome indicators. Intermediate indicators like biology indicators and imaging science indicators are often excluded. If the intervention cannot be done randomly, we recommend you to chose outcome indicators such as stroke and tumor size which are not influenced easily by intervention. In common cases, the sample calculation should be done according to specific research design.

Generally, we choose routine treatments, standard treatments and effective treatments as contrasting group, placebo effect treatments are not applicable. The evaluation of long-term outcome is the focus of pRCT, so we should conduct multi-time outcome measurement. Accordingly, the follow-up time should be longer and the follow-up frequency should be lower than routine pRCT.

The focus of pRCT is mainly intervention effects in real clinical case, however, the research environment and conditions must be combined with disease features and clinical reality for the final judgement. The research objects should be typical in patient groups.

4.2.2 Observational Research

In observational research which is based on real-world data, the real-world data quality varies because of different data resources. The measurement of outcome and exposure may different from research definition. Generally, doctors allocate different treatments to patients, so it is not arbitrary. The conclusion is, it is an important to recognize and control confounding biases in design and analyzing stage in observational research. Otherwise, biases may limit reasoning. In using real-world data in observational research, to clarify relationship between exposure and outcomes, we suggest researchers to consider carefully about key factors and procedures and then, design research plan and statistical analysis plan.

There are 3 steps for designing the observation research in real-world, ①confirming the research purpose ②resolving the problems of the devices occurred in clinical evaluation ③clarifying the research hypotheses. While establishing research hypotheses, the key factors of research should be emphasized, it means that if P (Population), I (Intervention), C (Control), O (Outcome), T (Timing) can be generated based on the data from real-world, including①if the data of study population meeting the requirements will be picked up from the data of real-world. ② if the unified intervention plan or standard intervention plan will

be made, ③if the comparable contrast can be set, ④ if the outcome indicators and measure results which are necessary for research will be included.

Observation research including array research (prospective, retrospective, deque), Case-control study and derivative design (Nested case-control study), as well as self-contrast case series and other design types. The applicator can choose property study design as well as the other designs at the same time, according to research purpose and the features of data from real-world. The applicator should have a comprehensive identification for possible biases during the process (selection bias, measurement bias, etc.) , establish effective measures to control measures.

4.2.3 Other Design Types

Using real-world data as external control for single unit trial is a design type, from which we can get clinical evidences. Among them, the historical control is usually not comparable because of the differences in clinical practice, the changes in follow-up time, and the lack of consistency in the standards of diagnosis and outcome measurement. It is an effective way to improve these limitations by selecting the contemporaneous control rather than the historical control and collecting the detailed and accurate information of relevant variables.

4.3 Statistical Analysis Techniques

In real-world data research, researchers need to apply reasonable statistical methods based on the research purpose, data type and research design type. See Appendix 2 for common statistical analysis methods.

The common research designs in real-world research mainly include experimental research and observational research. Among them, the pragmatic randomized controlled trial (pRCT) is typical, which combines the advantages of randomization and real-world data, and it can control the influence of confounding factors and biases well. There is no essential difference between the statistical analysis methods in this kind of research and the traditional clinical trials. The statistical analysis plan includes data set definition, analysis principle and strategy, missing data processing, analysis index and analysis method, sub-group or hierarchical analysis, sensitivity analysis, supplementary analysis and result report, etc. The basic principle of statistical analysis is the principle of intention to treat (ITT) analysis. Commonly used analysis methods include parameter test, nonparametric test, stratified analysis, regression analysis and other methods.

Observational research based on real-world data is more likely to produce confounding factors and biases. The key of data analysis is to use statistical analysis technology to minimize the biases caused by confounding factors. In the observational research of real-world data, researchers are also recommended to make a statistical analysis plan before the statistical analysis, so as to reduce the probability of false positive results. Commonly-used methods are traditional hierarchical analysis and multivariate analysis methods, in addition, tendency scoring and other methods is also suitable in this situation.

5. Common Situations in which Real-world Data Can be Considered for Clinical Evaluation of Medical Devices

Real-world evidence based on real-world data can support the whole-service-life clinical evaluation of medical devices, including pre-market clinical evaluation and post-market clinical evaluation. Common situations in which real-world evidence can be used in clinical evaluation of medical devices are as follows:

5.1 Provide Clinical Evidences in Congeneric Clinical Evaluation Process

The clinical evaluation process of congeneric medical devices is based on congeneric clinical data of medical devices. The required clinical data include clinical data from congeneric medical devices or products under registration process.

Real-world data is an important resource of congeneric medical device evaluation, which helps to confirm the safety and effectiveness of the product in the real medical cases. It also helps us to identify the potential risks of the product (such as rare serious adverse events), to understand results of congeneric medical device in different patient groups, even identify types of patients who get the best results. We can also understand average curative level of the congeneric medical device, it provides useful information like pre-market risk evaluation and post-marketing surveillance for products under registration process. Real-world data from legal resources can be used in confirming the differences between congeneric medical device and products under registration process. And it do no harm to safety and effectiveness evaluation of product under registration process.

5.2 Support Medical Device Registration as Supplementary Evidences

Most medical devices can’t enter markets of different countries at the same time, because of different policies and manufactures marketing strategy. Applicants should consider product design, application areas, existing clinical data and different requirements for clinical evidences in regions comprehensively. Based on real-world data collected in regions where the product already entered the market, applicants get series of real-world evidences which can support the registration in China. In this way, applicants can avoid additional clinical trials.

5.3 Real-world Data from Chartered Medical Devices Can be Used as Supplementary Evidences in Registration

According to the national unified deployment, according to relative regulations and clinical technical requirements, after strict data collection, systematic processing, scientific statistical analysis and multi-dimensional evaluation, real-world data from chartered medical devices in institutions of some regions which are urgently needed for clinical use and have not been approved in China can support registration as a supplementary evidence. For clinical trials conducted overseas, if any evidence shows that Chinese patients may have different results, we should use this data in connection research.

5.4 External Reference Materials in Single Unit Test

In the design of single group clinical trial, cases comparable to the experimental group and their clinical data can be extracted from a quality-controlled real-world database as an external control。The external control usually comes from the registration database with a good quality management system， It can be accept the evaluation of the sponsor and the supervisor, etc. To confirm the relevance and reliability of their data. It is suggested to use external control in the same period, such as using historical data to control, introduce multiple biases due to time differences, and reduce the evidence intensity of clinical trials.

5.5 Provide Clinical Data for Setting Target Value of Single Unit

The target value is the minimum standard that must be reached for the effectiveness / safety evaluation index of a certain type of medical device recognized in the professional field, including objective performance standards and performance targets. It is based on previous clinical data and is used for test devices. Comparison and evaluation of main evaluation indicators. Real-world data can be used as a data source for constructing or updating target values.

5.6 Provide Basis for Revise of Application Area and Indications

After the medical device is put on the market, the real-world data obtained can be used to support the modification of application scope and indications under the premise of legal use based on the relevant laws and regulations of the regulatory region. Possible situations include finding additional efficacy in the process of legal use, or using outside the scope of approval permitted by laws and regulations of some overseas regulatory areas.

5.7 Provide Basis for Revise of Clinical Statement in IFU

Real-world evidence of medical devices which are already in the market can be used to support the modification of clinical claims in the labeling. For example, for medical devices that measure and calculate patient physiological parameters and functional indicators, some physiological parameters and functional indicators are mainly concerned with the accuracy of poor measurements and calculations during pre-market evaluation, and their clinical value has not been fully explored. Real-world data can be used to construct physiological parameters and functional indicators, or to make casual inferences between clinical treatment decisions and clinical outcomes based on them, thereby modifying the clinical claims of the product in the description.

5.8 Support Post-marketing Research of Registered Products with Conditions

Medical devices that are used to treat rare diseases and life-threatening diseases which do not yet have effective cure, and devices which are needed in responding public health

emergency can use the real-world data in post-marketing research. It can support the completion of the matters stated in the registration certificate.

5.9 Long-term Safety & Effectiveness Evaluation for Medical Devices with High-risk Implants and Others

Medical devices with high-risk implants and others, especially the high-risk implants that have appeared on the market for the first time, during pre-market clinical evaluation, are difficult to confirm their long-term curative effect and risks of the product as well as hardly identify rare serious adverse events. Post-market studies of this type of product can be conducted using real-world data to determine the long-term safety and effectiveness of the product and complete a life-cycle clinical evaluation of the product.

5.10 Life-cycle clinical evaluation of medical devices used to treat rare diseases

-speed up their market launch, and meet patient needs

Real-world data can quickly bring medical devices that treat rare diseases to market. If a pre-market clinical trial is planned, the real-world data can be used as an external control for a single unit of trials or used to setting target values. After conditional approval, real-world data can be used to confirm the effectiveness of the product, identify product risks, and conduct product risks/benefits evaluation.

5.11 Post-market Surveillance (PMS)

Post-market surveillance of products, including surveilling of adverse events and re-evaluation of product safety and effectiveness, is an important part of the whole service life clinical evaluation of medical devices. Real-world data should play an important role in post-market surveillance. For example, by collecting and extracting risk signals, conducting attribution analysis of adverse events, detecting and controlling the risks of listed medical devices in a timely manner, analyzing and promoting the design and improvement of listed products by manufacturers, and promotes the development of new products.

Appendix 1: Common Statistical Analysis Methods

1. Statistical Methods for Pragmatic Randomized Controlled Trials(pRCT)

The statistical analysis ideas of Pragmatic Randomized Controlled Trials(pRCT) are similar to the statistical analysis ideas of traditional randomized controlled trials, including the need to implement statistical analysis plans, such as considering adjustment of covariates, control center effects and group effects to build statistical models, subgroup analysis, sensitivity analysis, and so on. The difference is that pRCT is carried out in the

actual clinical medical environment, so the standardization of patients receiving interventions is reduced, compliance may also be lower than in the traditional trial environment, and loss of follow-up may increase. Intentional analysis is a commonly used analysis method, but attention should be paid to the outcome of patients who are lost to follow-up. The treatment of patients who are lost to follow-up and the reasons need to be clarified in advance. Compared with traditional randomized controlled trials, the results of pRCT are more likely to tend to invalid hypotheses. Therefore, pRCT design should be used with caution when designing non-inferiority trials.

In addition, since pRCT may vary according to individual differences or the choice of clinical professionals after randomization, the interventions that patients receive may change, resulting in new confounding (often referred to as random post-confounding), which requires adjustment of covariates. The corresponding statistical models are often used for adjustment, such as multiple linear regression, Logistic regression, Cox regression, Poisson regression, etc. pRCT may come from multiple centers, so the central effect needs to be controlled. When the main outcome variable is a continuity indicator, a covariance analysis method can be used; when the main outcome variable is a categorical indicator, the Cochran-Mantel-Haenszel method considering the central effect can be used; When there are other covariates to consider, a random effects model can be used.

There is no need to consider group effects for individual pRCT. However, cluster-randomized control trials(cRCT) and stair-wedge randomized control trials (swRCT) have different analysis contents and methods due to the phased introduction of interventions and random allocation units for groups and interventions. cRCT analysis can use mixed effect models, multilevel / hierarchical modeling techniques. At the same time, considering the influence of group, individual level and inter-group characteristics, Bayesian hierarchical modeling can also be used to obtain reasonable interval estimates of intervention effects. The swRCT analysis mostly uses the group random effect model to adjust the effect of time effect for intervention effect analysis. In the pRCT statistical analysis, it is recommended to pay attention to sensitivity analysis to evaluate the robustness of statistical inference.

2. Statistical analysis methods commonly used in observational study

In observational studies of real-world data, the key to data analysis is the use of statistical analysis techniques to maximize control of bias caused by confounding factors. The common analysis methods are as follows:

2.1 Stratified analysis

Stratified analysis refers to dividing data into multiple layers according to possible confounding factors. Stratified analysis refers to dividing data into multiple layers according to possible confounding factors. The Mantel-Haenszel method is a commonly used stratified analysis method to evaluate the effect of confounding factors on the results. This

analysis can determine whether external factors are confounding or effect modification, or which effect is dominant, and determine the size and direction of confounding or the size of effect modification. However, stratified analysis can only control a small number of confounding factors. If the number of confounding factors is too large, it may lead to over-stratification and make the sample size in the stratum small. For continuous variables, only the hierarchical stratification method can be used, which often causes unreasonable grouping.

2. 2 Multivariate regression models

multivariate regression model is the most common statistical analysis method for controlling confounding factors. It is often used in observational studies. Logistic regression, linear regression, Poisson regression, and Cox proportional hazard regression are selected according to the characteristics of the outcome variables. Considering whether to choose a multilevel model based on the existence of a hierarchical structure of the data, for those data with repeated measurements, a GLMM（Generalized Linear Mixed Model）and a GEE (generalized estimating equation) can be applied. However, when applying these models, the assumptions and applicability of the models still need to be considered.

2.3 Propensity Score Analysis Method

Propensity score analysis is the most common analysis method used in causal inference in observational studies. It is an adjustment approach which aiming to solve multi-confounding factors, especially suitable for studies with common exposure factors and rare outcomes, or with multiple outcomes variables. Common application methods of propensity score include PSM (propensity score matching), PSS (propensity score stratification), IPTW (inverse probability of treatment weighting) and the method of incorporating propensity score as the only covariate into the statistical model for adjustment analysis. Significantly, if investigators applied the propensity score approach in the evaluation of treatment outcomes, First, in the research plan or analysis plan, they should first specify the variables used to establish the propensity score model in the research plan or analysis plan, and the criteria for judging the goodness of fit of the model and the prediction effect. More importantly, when establishing a propensity score model for baseline indicators, the "blindness" of the outcome indicators should be maintained, not until the propensity score model is completely established and determined, the outcome indicators could be introduced in order to directly evaluate the results, aiming to avoid readjusting propensity score models based on comparative results of efficacy outcomes thus to achieve "ideal" or "expected" results.

When applying propensity score for analysis, it is necessary to report the results before and after using propensity scores simultaneously, and also be sure to consider the impact might be caused by analysis on final results after processing the propensity score method on the final results, for example, the estimated accuracy could be reduced after applying PSM (due to a decrease in sample size); or when PSW (propensity score weighting) is applied, individual samples with extremely large weights may have a greater impact on the

analysis results. What needs to be emphasized is that the propensity score method can only deal with observable confounding factors instead of controlling the potential impact of confounding factors that uncollected in the study. It is recommended that the evaluation results shall be reasonably interpreted and discussed in the study and carried out the possible quantitative analysis.

2.4 Instrumental Variable Analysis Method

The limitation of using the above-mentioned methods including stratified analysis, multivariate regression model, and propensity score to control confounding is that it can only control the influence of measured confounding factors, but unable to adjust for unmeasured ones. The causal effect analysis method using instrumental variables does not involve the specific adjustment of confounding factors/covariates. It can control the impact of unknown confounding factors, hence estimate the causal effects of intervention/exposure factors and outcomes. If a variable is related to the level of the intervention factor (exposure), and the effect on the outcome variable can only be achieved by affecting the intervention/exposure factors, and it is not related to the confounding factors of exposure and outcomes, therefore the variable can be referred as an instrumentation variable of its exposure factors. Once the instrumental variables are defined, even if there are unknown unmeasured confounding factors, the causal effects of exposure on outcomes can be estimated by separately estimating the effects of instrumental variables on both exposure and outcomes. The major difficulty in estimating causal effects using instrumental variables is to find appropriate instrumental variables that meet the above assumptions. Where possible, it is recommended to select and use multiple instrumental variables with descriptive reasons. Through the sensitivity analysis, the robustness of the results will be verified.