PharmaSUG China 2024 - Paper SA-045

Implementing Statistical Methods in Companion Diagnostic: A Case Study in NSCLC

Danxia Hu, Dizal Pharma;
Mandy Gao, Dizal Pharma;
Jianyong Tong, Dizal Pharma

ABSTRACT

Companion diagnostics (CDx) are increasingly important in oncology studies with the development of precision medicine, as they provide essential information for the safe and effective use of corresponding drugs or biological products. If the diagnostic test is inaccurate, the treatment decisions based on that test may not be optimal. In real-world scenarios, clinical trials often use the clinical trial assay (CTA) tests other than the candidate CDx. In such cases, performing a bridging study between the candidate CDx and CTA can expedite the marketing of the CDx, ultimately benefiting patients. However, this process presents challenges, particularly in statistical programming, data manipulation, and analysis, due to differing regulations, data standards, and statistical methods required for both drug and CDx development. The purpose of this paper is to introduce the implementation of statistical methods in CDx using SAS from the perspective of a single study. It covers the processing CDx analysis data, statistical modeling, the preparation of an NDA submission package, and challenges related to SAS techniques.

INTRODUCTION

Companion diagnostic (CDx), as defined by the FDA, are medical devices or tests that provide information essential for the safe and effective use of a corresponding drug or biological product. CDx tests are crucial in precision medicine, particularly in oncology, as they help identify patients who will benefit most from specific treatments based on their diagnostic tests. This targeted approach not only improves treatment outcomes but also reduces adverse effects. Furthermore, the use of CDx can also result in cost savings for healthcare systems by ensuring that resources are allocated to patients who are most likely to benefit from them [1].

Conducting a bridging study between CDx and CTA, along with pursuing Pre-Market Approval (PMA), is an effective strategy to accelerate the market introduction of these diagnostics. This method ensures that both the diagnostic test and the drug are available at the same time, providing timely treatment options for patients. For example, Roche's Ventana PD-L1 (SP142) Assay for urothelial carcinoma patients, developed as a CDx for the drug Tecentriq® (atezolizumab), illustrates the success of this approach. This collaboration shows how bridging speeds up the availability of effective treatments and diagnostics, ensuring patients receive the appropriate therapy based on their specific biomarkers [2].

Developing and validating CDx involves careful data management, thorough statistical analysis, and strict regulatory compliance. Statistical programmers play a crucial role in this process. They manage large amounts of data, ensure its accuracy, and perform complex analyses to validate biomarkers and diagnostic tests. Their work is critical for preparing regulatory submissions that meet the stringent requirements of agencies like the FDA and EMA.

Given the complexity and importance of CDx development, statistical programmers face several key challenges, including the integration of CDx analysis data, sophisticated statistical analysis, and regulatory compliance. This paper discusses these challenges and explores the implementation of statistical methods in CDx using SAS, focusing on a case study in NSCLC that is being developed using a bridging study. By examining these topics, we aim to highlight the critical role of statistical programming in developing and obtaining regulatory approval for companion diagnostics, thereby supporting the advancement of personalized medicine.

PROCESSING CDX ANALYSIS DATA

The processing of CDx analysis data involves multiple steps to integrate data from various sources, ensuring they are clean, standardized, and ready for analysis. This comprehensive process involves handling clinical trial data (including clinical characteristics, clinical trial assays, and clinical outcomes) as well as CDx testing data. The goal is to create an analysis-ready dataset that can support robust statistical analysis and meet the regulatory requirements for traceability and accuracy.

CLINICAL TRIALS DATA

Clinical characteristics data refer to patient baseline information collected prior to entry the clinical trial. This includes demographic information (age, sex, race, pathology diagnosis), medical history, and other relevant clinical parameters. CTA data include testing results and assay associated data. Clinical outcome data capture the effectiveness of the therapeutic intervention. In oncology studies, this includes endpoints such as overall survival (OS), progression-free survival (PFS), and response rates which depends on the clinical trial design for the therapeutics.

Table 1.1 lists example data required for each category.

Clinical Characteristics	Clinical Trial Assays	Clinical Outcomes
Subject name or identifier	Mutation Status	Study Arm/Treatment Dose
Internal id for the site	Mutation Specification Nucleotide change	Confirmed Response per IRC
Age	Mutation Specification Amino Acid Change	Duration of Response (Month)
Sex	Mutation Analysis Method	Censor (censoring status for duration of response).
Race	Test/Kit Name	Full Analysis Set Population Flag
Country of Enrollment	…	…
Smoking Status at Informed Consent	…	…
Stage/AJCC Stage	…	…
Cancer Type at Initial Diagnosis	…	…

Table 1.1

Clinical trials data are standardized into formats CDISC SDTM (Study Data Tabulation Model) or ADaM (Analysis Data Model). For this case, key analysis datasets include ADSL (Subject-Level Analysis Dataset) and other ADaM datasets including study efficacy endpoints. Each of these datasets plays a crucial role in the analysis and reporting.

CDX DATA

CDx data refer to the testing results of CDx tests and associated baseline information. Table 1.2 lists the example data required.

Extraction and Quantification	Pathology	Sequencing
Subject name or identifier	Biopsy Type	Sample Repeat
Site Number	Tumor Content	Final DNA Passing Run?
Sample ID	Percent Necrosis	CDx Test Results
Sex	Mutation Analysis Method	Enrolling CTA
Race	Slide Surface Area	Results Specification 1
Stage/AJCC Stage	…	Results Specification 2
Final DNA Concentration	…	…
Final RNA Concentration	…	…

Table 1.2

DATA INTEGRATION

Integrating clinical trial data with CDx data is essential for comprehensive analysis and reporting. This integration enables a holistic view of patient data, correlating clinical outcomes with diagnostic test results, and enhancing the understanding of treatment efficacy and safety.

To facilitate this integration, the **ADAM OTHER** class can be used because the dataset does not fit the typical structures (ADSL, BDS, OCCDS) and has unique requirements. The integration process involves several key steps:

1. **Data Harmonization:** This step involves standardizing data formats and terminologies used in clinical trial and CDx data. Clinical trial data often follows CDISC (Clinical Data Interchange Standards Consortium) standards, while CDx data might have different formats and terminologies. Harmonizing these ensures that data from both sources can be seamlessly combined. For example, ensuring that date formats, units of measurement, and medical terminologies are consistent across datasets.
2. **Mapping Patient Identifiers:** This involves mapping patient identifiers such as subject ID, sample ID, or any other unique identifiers. This step is crucial to ensure that data from clinical trials and diagnostic tests are accurately linked to the same patient.
3. **Merge Datasets:** This step is to merge the clinical trial datasets (e.g., ADSL - Subject Level Analysis Dataset and efficacy ADaM datasets including study efficacy endpoints) with CDx data based on common identifiers.
4. **Data Validation:** Conduct rigorous validation to ensure data accuracy and consistency post-integration. Verify that all relevant data points are correctly mapped and there are no discrepancies.

Table 1.3 and 1.4 below illustrates an example of an **ADCDX** dataset, showing some key variables:

Subject Identifier	Sex	Age	Race	AJCC Stage	Treatment	Confirmed Response per IRC	Duration of Response	Censor	CTA Testing Results
A9999103	F	55	ASIAN	Stage IIIA					Negative
A9999103	F	55	ASIAN	Stage IIIA					Negative
A9999105	F	67	ASIAN	Stage IIIA					Negative
A9999106	M	63	ASIAN	Stage IIIA					Negative
A9999107	F	47	ASIAN	Stage IIIA					Negative
A9999153	F	63	WHITE	Stage IV	Treat 1	PR	210	1	Positive
A9999153	F	63	WHITE	Stage IV	Treat 1	PR	210	1	Positive
A9999153	F	63	WHITE	Stage IV	Treat 1	PR	210	1	Positive
A9999156	F	78	WHITE	Stage IV	Treat 2	PR	195	0	Positive
A9999157	F	88	WHITE	Stage IV	Treat 2	CR	416	1	Positive

Table 1.3

Sample Repeat Number	CDx Testing Results	Tumor Content	Biopsy Type	Slide Surface Area	Percent Necrosis
0	#N/A	10	Resection	110	2
1	Negative	10	Resection	110	2
0	Negative	30	Resection	180	5
0	Negative	70	Resection	290	10
0	Positive	80	Resection	100	0

0	#N/A	90	FNA	87	2
1	#N/A	90	FNA	87	2
2	Positive	90	FNA	87	2
0	Positive	5	FNA	78	0
0	Positive	5	Resection	90	0

Table 1.4

The ADCDX analysis dataset facilitates detailed subgroup analyses, allowing researchers to explore the relationships between clinical outcomes and diagnostic test results. In bridging studies, it is used to demonstrate the similarity between CDx and CTA.

STATISTICAL MODELING

In bridging studies, missing data can significantly impact the demonstration of similarity between CDx and CTA, including the accuracy of both tests and the clinical response categorized by CDx. Conducting sensitivity analysis helps in understanding the robustness of the results against different assumptions and methods for handling missing data. This is particularly important for ensuring that treatment decisions based on CDx data are reliable and valid [3]. This section focuses on sensitivity analysis methods tailored for the CDx development, using SAS for implementation.

COMBINED USE OF MI AND BOOTSTRAP

Multiple Imputation (MI) is a commonly used statistical method for handling missing data. Instead of filling in a single value for each missing data point, MI creates multiple complete datasets by imputing values based on the distribution of observed data. These datasets are then analyzed separately, and the results are combined to produce estimates that reflect the uncertainty associated with the missing data. MI is particularly useful in scenarios where the missing data mechanism is assumed to be at random [4].

Bootstrap methods are another popular technique used to assess the variability and confidence intervals of estimates. The bootstrap approach involves repeatedly resampling the observed data with replacement to create many "bootstrap samples." Each sample is analyzed to obtain an estimate, and the distribution of these estimates is used to infer confidence intervals and other measures of uncertainty. Bootstrap methods are valuable because they do not rely on strong parametric assumptions and can be applied to complex data structures [5].

While MI and bootstrap methods are both widely used independently, their combined application is relatively uncommon, especially in the context of CDx. Combining MI with bootstrap techniques leverages the strengths of both methods: MI addresses the uncertainty due to missing data, and bootstrap provides robust confidence intervals for the estimates [6].

CASE STUDY ANALYSIS

In the NSCLC case study, we utilized the combined MI and bootstrap to handle missing data and estimate confidence intervals. Take computing the Positive Percent Agreement (PPA) as an example.

1. Positive Percent Agreement (PPA) Definition

To calculate the PPA, a 2x2 contingency table, as shown in Table 2.1, is constructed to display the frequencies of the CDx and CTA results.

CDx Testing Result	**CTA Testing Result**	
	Positive	Negative
Positive	n_{11}	n_{12}
Negative	n_{21}	n_{22}

Table 2.1

Based on the notation of Table 2.1, the PPA is computed as the proportion of biomarker-positive specimens as identified by the CTA, which are also biomarker-positive as identified by the CDx. The PPA is calculated using the formula:

$$PPA = \frac{n_{11}}{n_{11} + n_{21}} \times 100\%.$$

In real-world data, testing results can sometimes be missing. Before computing the PPA, we need to handle any missing results. Table 2.2 provides an example of an input dataset where two subjects (A9999103 and A9999104) have missing test results.

Subject Identifier	Sex	CTA Testing Results	CDx Testing Results	Tumor Content	Biopsy Type	Slide Surface Area	AJCC Stage
A9999103	**F**	**Negative**		**34**	**Resection**	**130**	**Stage IIIA**
A9999104	**M**	**Negative**		**10**	**Resection**	**110**	**Stage IIIA**
A9999105	F	Negative	Negative	15	Resection	180	Stage IIIA
A9999106	M	Negative	Negative	15	Resection	290	Stage IIIA
A9999107	F	Negative	Positive	15	Resection	100	Stage IIIA
A9999108	F	Negative	Negative	34	Resection	150	Stage IV
A9999109	M	Negative	Negative	10	Resection	99	Stage IIIA
A9999110	F	Negative	Negative	15	Resection	87	Stage IIB
A9999111	M	Negative	Negative	15	Resection	65	Stage IIIB
A9999112	F	Negative	Negative	15	Resection	43	Stage IB

Table 2.2

The following steps outline the process of calculating confidence intervals for PPA.

2. Multiple Imputation

We used the PROC MI procedure in SAS to create multiple imputed datasets. Each dataset was imputed based on the observed data's distribution, considering variables such as patient demographics, clinical characteristics, and diagnostic test results.

Example code:

```
proc mi data=dsn seed=1305417 nimpute=50 out=mi_out;
       class EX20ODXN EX20CTAN SEXN TNMSTAGN BIOPTYPE;
       fcs nbiter=20
              logistic(EX20ODXN = EX20CTAN SEXN TNMSTAGN BIOPTYPE SSA TUMRCONT
              /details likelihood=augment link=logit;
       fcs discrim(TNMSTAGN=EX20CTAN SEXN EX20ODXN BIOPTYPE SSA TUMRCONT
              /details classeffects=include;
       fcs discrim(BIOPTYPE=EX20CTAN SEXN EX20ODXN TNMSTAGN SSA TUMRCONT
              /details classeffects=include;
       var EX20ODXN TNMSTAGN EX20CTAN SEXN BIOPTYPE SSA TUMRCONT;
run;
```

Table 2.3 is an example of the imputed dataset (**M times**), here is 50:

Imputation	Subject Identifier	Sex	CTA Testing Results	CDx Testing Results	Tumor Content	Biopsy Type	Slide Surface Area	AJCC Stage
1	**A9999103**	**F**	**Negative**	**Positive**	**34**	**Resection**	**130**	**Stage IIIA**
1	**A9999104**	**M**	**Negative**	**Positive**	**10**	**Resection**	**110**	**Stage IIIA**
1	A9999105	F	Negative	Negative	15	Resection	180	Stage IIIA

1	A9999106	M	Negative	Negative	15	Resection	290	Stage IIIA
1	A9999107	F	Negative	Positive	15	Resection	100	Stage IIIA
1	A9999108	F	Negative	Negative	34	Resection	150	Stage IV
1	A9999109	M	Negative	Negative	10	Resection	99	Stage IIIA
1	A9999110	F	Negative	Negative	15	Resection	87	Stage IIB
1	A9999111	M	Negative	Negative	15	Resection	65	Stage IIIB
1	A9999112	F	Negative	Negative	15	Resection	43	Stage IB
…	…	…	…	…	…	…	…	…
50	**A9999103**	**F**	**Negative**	**Negative**	**34**	**Resection**	**130**	**Stage IIIA**
50	**A9999104**	**M**	**Negative**	**Positive**	**10**	**Resection**	**110**	**Stage IIIA**
…	…	…	…	…	…	…	…	…

Table 2.3

3. Bootstrap Resampling

For each imputed dataset, we applied bootstrap resampling using the PROC SURVEYSELECT procedure in SAS. We generated numerous bootstrap samples for each imputed dataset to assess the variability of the estimates.

Example code:

```
proc surveyselect data=mi_out NOPRINT seed=1 out=br_out
      method=urs
      samprate=1
      outhits
      reps=1000;
      strata _imputation_;
run;
```

Table 2.4 is an example of the output dataset (**M×B times**), here is 50x1000:

Replicate	Imputation	Subject Identifier	Sex	CTA Testing Results	CDx Testing Results	Tumor Content	Biopsy Type	Slide Surface Area	AJCC Stage
1	1	A9999107	F	Negative	Positive	15	Resection	100	Stage IIIA
1	1	A9999108	F	Negative	Negative	34	Resection	150	Stage IV
1	1	A9999109	M	Negative	Negative	10	Resection	99	Stage IIIA
1	1	A9999110	F	Negative	Negative	15	Resection	87	Stage IIB
1	1	A9999111	M	Negative	Negative	15	Resection	65	Stage IIIB
1	1	A9999111	M	Negative	Negative	15	Resection	65	Stage IIIB
1	1	A9999111	M	Negative	Negative	15	Resection	65	Stage IIIB
1	1	A9999112	F	Negative	Negative	15	Resection	43	Stage IB
1	1	A9999112	F	Negative	Negative	15	Resection	43	Stage IB
1	1	A9999112	F	Negative	Negative	15	Resection	43	Stage IB
…	…	…	…	…	…	…	…	…	…
1000	50	A9999103	F	Negative	Positive	34	Resection	130	Stage IIIA
1000	50	A9999106	M	Negative	Negative	15	Resection	290	Stage IIIA
1000	50	A9999107	F	Negative	Positive	15	Resection	100	Stage IIIA

1000	50	A9999107	F	Negative	Positive	15	Resection	100	Stage IIIA
1000	50	A9999108	F	Negative	Negative	34	Resection	150	Stage IV
1000	50	A9999109	M	Negative	Negative	10	Resection	99	Stage IIIA
1000	50	A9999109	M	Negative	Negative	10	Resection	99	Stage IIIA
1000	50	A9999110	F	Negative	Negative	15	Resection	87	Stage IIB
1000	50	A9999112	F	Negative	Negative	15	Resection	43	Stage IB
1000	50	A9999112	F	Negative	Negative	15	Resection	43	Stage IB

Table 2.4

4. Analysis and Combination

1) Create a Distribution of PPA: After computing the PPA for all bootstrap samples, you will have a distribution of PPA values.

2) Determine the Confidence Intervals: Sort the PPA values and extract the 2.5th and 97.5th percentiles to form the 95% CI. These percentiles represent the lower and upper bounds of the 95% confidence interval, respectively.

   Example code:

```
proc univariate data=ppa noprint;
      var ppa;
      output out=ci
            pctlpts =2.5  97.5
            pctlpre=CI95_
            pctlname=Lower Upper;
run;
```

CHALLENGES RELATED TO SAS TECHNIQUES

Implementing statistical methods in CDx using SAS involves several challenges. Two significant challenges are efficiently processing large datasets and absence of a comprehensive SAS procedure for integrating MI and bootstrap techniques. Following are the strategies corresponding to the challenges.

- **Efficient Data Processing**

Handling large volumes of data can significantly increase processing time. Here are several strategies to optimize data processing in SAS.

1) Write efficient SAS code by minimizing the number of data reads and writes. Use WHERE clauses to subset data early in the process, rather than IF statements, to reduce the amount of data being processed.

2) Avoid Using ODS OUTPUT: The ODS OUTPUT statement can be resource-intensive and slow. Instead, use procedures that directly create datasets. For example, many statistical procedures have options to create output datasets without using ODS OUTPUT.

```
proc freq data=br_out noprint;
      by SampleID imputation;
      tables CDX*CTA/ out=sum outexpect sparse missprint;
run;
```

3) Compression: Use the COMPRESS option to reduce the size of datasets, which can improve computer performance and decrease processing time.

- **Absence of a comprehensive SAS procedure**

Currently, SAS does not provide a single procedure that fully integrates Multiple Imputation (MI) and bootstrap methods. Combining these techniques requires a multi-step, custom implementation. Developing custom macros or scripts can streamline this process, ensuring accuracy and efficiency in

the combined use of MI and bootstrap methods. Advanced SAS programming skills and thorough validation are essential for successful implementation.

PREPARATION OF AN NDA SUBMISSION PACKAGE

Preparing an NDA submission package for a CDx requires meticulous and precise documentation. This process is typically undertaken in collaboration with in vitro diagnostic (IVD) company. The submission package must include two distinct parts: one for the PMA registration, and the another for NDA. This section outlines the necessary components from a programming perspective to ensure the NDA submission package is thorough and compliant.

CONTENTS FOR PMA REQUIRED

- **Excel listing of clinical trial data information.**
  1) Clinical Characteristics Data: Patient baseline information, including demographic information (age, sex, race, pathology diagnosis), medical history, and other relevant clinical parameters.
  2) Clinical Trial Assay Data: Testing Results and assay associated data.
  3) Clinical Outcome Data: Data capturing the effectiveness of the therapeutic intervention, including endpoints such as overall survival (OS), progression-free survival (PFS), and response rates etc.
- **Integrated Analysis Dataset**

  Prepare an integrated analysis dataset that merges clinical trial data with CDx data. This dataset must be ready for statistical analysis and reporting.

  However, there are no specific data standard requirements for CDx submissions. The existing framework used for drug data submission can be adapted for CDx submissions, ensuring consistency and standardization in data presentation and documentation.
- **Key Tables for Reporting**
  1) Concordance between CTA and CDx
  2) Patients demographics, clinical disease characteristics and sample characteristics
  3) Drug efficacy for primary efficacy population and stratified analysis based CDx testing results.
  4) Sensitivity analysis
     a. **Best Case**: Assume all missing CDx results to be concordant with CTA result.
     b. **Worst Case**: Assume all missing CDx results to be discordant with CTA results.
     c. **Likely Case**: Assuming missing CDx results are missing at random. A multiple imputation model including CTA result, clinical outcome, demographic variables (age, gender, race), baseline, and sample characteristics (e.g., tumor content) that are imbalanced across CDx evaluable and unevaluable.

CONTESNTS FOR CSR REQUIRED

This section can follow the general drug submission standards (e.g., CDISC), incorporating the CDx as part of the clinical study report (CSR) for NDA. This integration ensures that the CDx is meticulously documented and presented in alignment with the overall drug submission, providing a comprehensive package for regulatory review. The general contents required for the CSR include:

- **SDTM Datasets:** Follow the standards for creating and validating SDTM datasets, ensuring accuracy and readiness for statistical evaluation.
- **Metadata Documentation**: Provide detailed metadata for each dataset, including variable descriptions, derivations, and controlled terminologies. Use define.xml to ensure comprehensive

documentation. Clearly outline how the CDx data were used to support the clinical trial results and their interpretation.

- **Quality Control**: Implement rigorous validation and quality control procedures to ensure data accuracy and consistency. This includes cross-checking data entries, verifying calculations, and confirming that all data points are correctly mapped and documented.

CONCLUSION

This paper highlights the innovative combination of Multiple Imputation (MI) and Bootstrap techniques in the context of CDx for NSCLC. The strategy of performing bridging study not only accelerates marketing authority but also enhances the precision of treatment decisions, which is pivotal in the era of personalized medicine.

Key Contributions:

- **Innovative Methodology**: The combined use of MI and Bootstrap methods addresses the challenges of missing data in CDx studies, providing robust confidence intervals and leveraging the strengths of both techniques.
- **Case Study Implementation**: The NSCLC case study illustrates the practical application of these methods using SAS, detailing the steps of data processing, integration, and statistical analysis.
- **Regulatory Compliance**: The discussed methods ensure that data processing and analysis meet the stringent requirements of the FDA, facilitating the approval and market entry of CDx.

Future research could explore the application of these combined methods in other therapeutic areas and types of diagnostics. Additionally, developing more streamlined and automated processes for integrating MI and Bootstrap techniques could further enhance their practical utility and efficiency.

In conclusion, the innovative combination of MI and Bootstrap methods for handling missing data in CDx studies represents a significant advancement in statistical programming and personalized medicine. By addressing the challenges of data integration, regulatory compliance, and robust statistical analysis, this approach supports the development and approval of effective CDx, ultimately benefiting patients and healthcare systems.

REFERENCES

[1]. U.S. Food and Drug Administration. "In Vitro Companion Diagnostic Devices: Guidance for Industry and Food and Drug Administration Staff." FDA, 2014. https://www.fda.gov/media/81309/download.

[2]. Roche. "VENTANA PD-L1 (SP142) Assay." Accessed June 25, 2024. https://diagnostics.roche.com/global/en/products/params/ventana-pd-l1-sp142-assay.html.

[3]. Little, Roderick J. A., and Donald B. Rubin. Statistical Analysis with Missing Data. 2nd ed. Hoboken, NJ: Wiley, 2002.

[4]. Carpenter, James R., and Michael G. Kenward. Multiple Imputation and its Application. Chichester: John Wiley & Sons, 2013.

[5]. Efron, Bradley, and Robert Tibshirani. An Introduction to the Bootstrap. New York: Chapman & Hall, 1993.

[6]. Schomaker, Michael, and Christian Heumann. "Bootstrap inference when using multiple imputation." Statistical Methods and Applications (2018).

ACKNOWLEDGMENTS

I would like to thank my team members for their invaluable contributions. Special thanks to our programmers for managing complex datasets and implementing statistical methods, and to our statistician for their rigorous analysis. I am also grateful to our Scientist, TGx, for their insightful guidance and innovative ideas. Their dedication and support were crucial to the success of our work. Finally, I appreciate everyone who contributed to this work, both directly and indirectly.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Danxia Hu
Danxia.Hu@dizalpharma.com

Mandy Gao
Mandy.Gao@dizalpharma.com

Jianyong Tong
Jianyong.Tong@dizalpharma.com