Paper DH07

Mind the Gap: Imputing Missing Dosing Times in QD & BID Regimens Using ADPPK Guidelines

Aschwin van der Giesen, Certara, Oss, The Netherlands
Gerly van der Vleuten, Certara, Oss, The Netherlands

ABSTRACT

Missing dosing times in pharmacokinetic datasets present a frequent challenge in clinical data programming, potentially affecting dataset consistency and interpretability. The CDISC® ADPPK guidelines provide a standardized framework for imputing these missing values while preserving data integrity and traceability. This paper explores practical imputation strategies for once-daily (QD) and twice-daily (BID) dosing regimens, including interval-based methods, protocol-defined administration window alignment, and PK sampling time–based approaches. A case study illustrates the application of these techniques within a reproducible programming workflow, demonstrating improved dataset completeness and regulatory compliance. By adhering to ADPPK standards and documenting imputation decisions explicitly, programmers can enhance dataset consistency, minimize ambiguity, and support scientifically robust pharmacometric analysis.

INTRODUCTION

Accurate dosing time information is critical for preparing high-quality pharmacokinetic (PK) analysis datasets. However, in real-world clinical trials, missing or incomplete dosing times are a common challenge—particularly in Phase I–III studies where data are collected across diverse sites, sometimes with varying levels of precision and protocol adherence.

When dosing times are missing, assumptions must be made to allow dataset generation to proceed. These assumptions, if not standardized and well-documented, can compromise data traceability, reduce confidence in the derived variables, and raise regulatory concerns during submission. In particular, inconsistent imputations can distort exposure calculations and complicate alignment between dosing and sampling records.

To address these challenges, the Clinical Data Interchange Standards Consortium (CDISC) introduced the ADaM dataset for population PK analyses—ADPPK. The ADPPK guidelines provide a standardized structure for capturing analysis-ready PK data, including strategies for handling missing dosing information while preserving transparency and scientific rigor.

This paper explores practical approaches for imputing missing dosing times in once-daily (QD) and twice-daily (BID) dosing regimens in alignment with ADPPK principles. It outlines commonly used imputation techniques, discusses their implications for dataset preparation, and provides a programming-focused workflow to ensure transparency, reproducibility, and regulatory compliance. Through a case study, we demonstrate how systematic and well-documented imputations improve data quality and traceability in PK datasets.

OVERVIEW OF ADPPK GUIDELINES

The Analysis Data Model for Population Pharmacokinetics (ADPPK) is a specialized ADaM extension developed by CDISC to support the needs of pharmacometricians and clinical programmers engaged in PopPK modeling. It aims to standardize how dosing and PK sampling data are structured, enabling more consistent, traceable, and analysis-ready datasets for nonlinear mixed-effects modeling tools like NONMEM.

One of the primary motivations behind ADPPK is to address inconsistencies in how dosing records—particularly missing or imputed values—are handled across trials and sponsors. Missing dosing times are a recurring issue, and how these gaps are managed can significantly affect exposure predictions and subsequent modeling.

ADPPK outlines several key principles relevant to handling missing dosing times:

Preservation of Traceability: Imputed values must be clearly distinguishable from collected data. ADPPK recommends using dedicated variables such as EXMDEF (Explanation for Modified Dosing Time) and EXMFL (Imputation Flag) to indicate any modifications or imputations.

Documentation of Logic: Imputation decisions must be reproducible and fully documented. This includes providing a rationale for the imputation method used (e.g., assuming fixed intervals, using protocol-defined dosing windows, or

leveraging sampling times).

Model Compatibility: The resulting dataset should be formatted in a way that facilitates direct use in modeling tools, often requiring precise alignment between dose events (EVID=1) and observed concentrations (EVID=0), with accurate timing in relative and nominal units.

Avoiding Hidden Assumptions: Any assumptions made during imputation—such as defaulting to a fixed interval without evidence—must be explicit and justifiable. ADPPK encourages minimizing the use of "magic numbers" (e.g., always assuming 08:00) unless protocol or context clearly supports it.

In addition, ADPPK supports the use of derived variables to assist in modeling (e.g., DOSEGRP, TAD, NDose) while maintaining clarity around their source. The guidelines also emphasize sponsor flexibility within the standard framework—allowing implementation choices that suit specific study designs, provided traceability and justification are maintained.

As missing dosing times can have a disproportionate impact in QD and BID regimens—due to the extended or more frequent dosing cycles—this paper focuses on practical ways to apply the ADPPK principles to those specific contexts.

The most recent implementation guidance for ADPPK is available from the CDISC standards library (CDISC ADaM for Population PK [ADPPK], Version 1.0, 2023).

CHALLENGES IN QD/BID REGIMENS

Although protocols define target dosing windows (e.g., morning and evening doses), actual administration times often vary due to site logistics, patient compliance, or operational factors. This variability complicates the choice of imputation methods and increases the risk of inconsistency if assumptions are too rigid.

Incorrect or inconsistent imputation of dosing times can lead to misalignment between dosing and PK sampling data, potentially affecting derived timing variables (e.g., time after dose, nominal time). These inconsistencies can reduce the interpretability and reusability of the dataset and complicate downstream analysis steps.

Therefore, imputation strategies must balance between protocol-defined nominal dosing times and real-world data trends observed in recorded dosing or sampling times, ensuring both consistency and transparency. Adopting methods that respect ADPPK principles is essential to maintain data quality and traceability.

IMPUTATION APPROACHES

When facing missing dosing times in QD and BID regimens, several practical imputation strategies can be employed to maintain dataset integrity and support robust PopPK analyses. The choice of approach depends on data availability, protocol details, and the extent of missingness. The following methods align with ADPPK guidelines and promote transparency and traceability.

INTERVAL-BASED IMPUTATION

Interval-based imputation uses the expected dosing interval to estimate missing times. For QD regimens, a fixed 24-hour interval is assumed, while for BID regimens, a 12-hour interval is applied. This method typically imputes missing dosing times relative to the last known dose time or the first dose in the series.

Advantages: Simple, protocol-aligned, and easy to implement programmatically.

Limitations: Assumes strict adherence to dosing schedule; may not capture real-world timing variability.

PROTOCOL-DEFINED ADMINISTRATION WINDOW ALIGNMENT

If the clinical protocol defines specific dosing windows (e.g., 08:00–10:00 for morning dose, 20:00–22:00 for evening dose), missing dosing times can be imputed using the midpoint or start of the respective window. This approach leverages protocol information to ground imputations in expected administration times.

Advantages: Reflects study design; minimizes guesswork; supports traceability.

Limitations: Assumes patient dosing falls within defined windows, which may not hold in practice.

PK SAMPLING TIME-BASED IMPUTATION

When dosing times are missing but pharmacokinetic (PK) sample collection data are available, dosing times can sometimes be inferred or constrained using a combination of planned and actual sampling information. This approach leverages study design metadata—particularly PCTPTREF (Planned Collection Time Point Reference)—together

with the observed sampling timestamps (PCDTC), to estimate the most plausible dosing time consistent with the planned post-dose interval.

In this method, the expected elapsed time between a dose and its corresponding PK sample (as defined by PCTPTREF or PCTPT) is used to back-calculate the dosing time when it is missing. For example, if the planned timepoint indicates “2 hours postdose” and the actual sample was collected at 10:15, the missing dose time can be imputed as approximately 08:15. Similarly, when multiple samples are available around the same dosing event, median or mode-based approaches can be used to refine the imputation and minimize the influence of timing deviations.

Advantages:
Incorporates observed PK data, improving biological and operational plausibility.

Aligns with the ADPPK principle of traceability by referencing planned timing metadata (PCTPTREF, PCTPT).

Can yield more accurate dosing alignment than interval-based or nominal window methods, particularly in early-phase or intensive sampling studies.

Limitations:
Requires consistent mapping between sampling and dosing occasions (e.g., using OCCASION, PCTPTREF, or nominal visit information).

Assumes adherence to the protocol-defined PK sampling schedule; deviations can propagate into the imputed dosing time.

Validation is essential to ensure that imputed dosing times remain biologically reasonable and consistent across subjects.

Implementation Tip
When applying this method, programmers should document both the planned post-dose interval used for imputation and the source of evidence (e.g., “imputed using PCTPTREF = 2H_POSTMORNING_DOSE and actual sample time 10:15”). These details should be captured in EXMDEF, while EXMFL = 'Y' flags the imputed records.
Limitations: Dependent on strict protocol adherence; requires careful validation.

HANDLING EDGE CASES AND MULTIPLE MISSING DOSES

In situations where multiple consecutive doses are missing or dosing intervals are irregular, combining the above methods with sensitivity analyses or flagging the affected data for exclusion may be necessary. Explicit flags and explanatory variables should document these scenarios in the dataset.

Best Practice: Regardless of the approach, imputation must be explicitly flagged in the dataset using ADPPK variables (e.g., EXMFL) and accompanied by clear documentation explaining the method, assumptions, and rationale. This ensures reproducibility and regulatory compliance.

IMPLEMENTATION STRATEGY

Translating imputation approaches into a reproducible and traceable programming workflow is crucial for ensuring data integrity and regulatory compliance. This section outlines a stepwise strategy to programmatically impute missing dosing times in QD and BID regimens aligned with ADPPK standards.

DATA PREPARATION AND IDENTIFICATION OF MISSING DOSING TIMES

Flagging Missing Values: Start by scanning the dosing dataset for missing or incomplete dosing time fields. Create binary flags (e.g., MISSING_DOSE_TIME) to identify records requiring imputation.

Baseline and Reference Times: Identify baseline doses or known anchor points (e.g., first dose administration time) to serve as references for interval-based imputations.

SELECTION OF IMPUTATION METHOD

Use decision logic based on protocol specifics, data completeness, and dosing regimen:

If dosing windows are clearly defined and missing dose falls within a known window, apply protocol-window imputation.

If prior dose time exists, use interval-based imputation relative to that dose.

If neither is available but PK sampling times exist, calculate backward from sampling time considering protocol-defined post-dose sampling intervals.

Include fallback rules for complex or multiple missing doses, such as setting missing times to NA and flagging for sensitivity analysis.

In some cases, a combination of protocol-window and interval-based imputations may be applied within the same subject. For example, when the first missing dose occurs within a known dosing window (e.g., evening dose), the midpoint of that window can be used. However, if subsequent consecutive doses are also missing, interval-based imputation may be applied relative to the previous imputed time to preserve dosing continuity within the regimen. This ensures temporal consistency across doses while maintaining transparency that later imputations are derived, not observed.

APPLYING IMPUTATION AND SETTING FLAGS

Perform imputations by updating the dosing time variable, ensuring all modifications are traceable.

Set ADPPK-specific flags (e.g., EXMFL = 'Y') and explanatory text variables (EXMDEF) detailing the imputation method used.

DOCUMENTATION AND TRACEABILITY

Maintain a detailed log of imputation decisions, preferably automated and output as a separate metadata dataset or report.

Include programming comments and annotated code to aid reviewers and auditors in understanding the imputation logic.

QUALITY CHECKS AND VALIDATION

Cross-validate imputed times by comparing imputed dose intervals with protocol specifications.

Visualize dosing timelines alongside PK sampling times to detect anomalies.

Conduct sensitivity analyses to evaluate the impact of imputations on key PK parameters.

CASE STUDY: IMPUTING MISSING DOSING TIMES IN A BID REGIMEN

To illustrate the practical application of the imputation approaches and programming workflow, we present a simulated case study involving a twice-daily (BID) dosing regimen with missing dosing times.

DATASET DESCRIPTION

The dataset includes dosing records and PK sampling times for 50 subjects enrolled in a Phase II clinical trial evaluating a new oral drug. The dosing regimen is BID, with protocol-defined dosing windows of 08:00–10:00 for the morning dose and 20:00–22:00 for the evening dose.

Approximately 10% of dosing time entries are missing due to site data collection errors. PK samples are collected at multiple timepoints post-dose. The focus of this case study is the programming and traceability process used to impute these missing dosing times, rather than subsequent modeling activities.

Table 1 below illustrates a simplified example of dosing records before and after imputation. Missing times were flagged and imputed according to the protocol-defined morning and evening dosing windows (08:00–10:00 and 20:00–22:00). All imputations were marked with ADPPK-standard flags and explanations to ensure transparency and traceability.

IMPUTATION WORKFLOW

Identification: Missing dosing times were flagged (MISSING_DOSE_TIME = Y).

Method Selection: For doses with missing times but known dosing occasion (morning/evening), the midpoint of the protocol-defined dosing window was used for imputation (e.g., 09:00 for morning dose).

Fallback: For doses with missing occasion labels, interval-based imputation was applied, calculating dosing times 12 hours apart relative to the previous known dose.

Flagging: Imputed dosing times were marked with EXMFL = 'Y' and explanatory notes added to EXMDEF.

Following imputation, the dataset was standardized according to ADPPK conventions. Table 2 illustrates the final structure of the exposure dataset, where all dosing times are represented in ISO 8601 format (EXSTDTC). Imputed doses are clearly flagged using EXMFL and documented through EXMDEF, ensuring full transparency of assumptions and enabling reproducibility. Derived timing variables such as TAD were recalculated to maintain alignment between dose and concentration records

TRACEABILITY AND COMPLIANCE
All imputation steps were documented in programming logs, and flags were reviewed during dataset validation. The workflow facilitated audit readiness by clearly demonstrating the rationale and execution of missing data handling. By using ADPPK-compliant variables and automated metadata reporting, the programming process ensured traceability and reproducibility across datasets and studies.

REGULATORY CONSIDERATIONS

Handling missing dosing times in pharmacokinetic datasets is a critical aspect of data quality and transparency expected by regulatory agencies such as the FDA and EMA. Submissions lacking clear documentation or standardized imputation methods risk delays or requests for additional analyses.

Key regulatory expectations include:

- **Traceability**: All imputations must be clearly flagged and documented within the analysis datasets. ADPPK variables such as EXMFL and EXMDEF provide a standardized mechanism for this purpose.

- **Justification of Methods**: The rationale for chosen imputation approaches should be detailed in submission documents, explaining why specific assumptions and methods were applied.

- **Reproducibility**: Imputation workflows should be reproducible and auditable, with programming code and logs maintained as part of submission packages.

- **Sensitivity Analyses**: Agencies often recommend conducting sensitivity analyses to evaluate the impact of imputation assumptions on key PK parameters and conclusions.

- **Alignment with Standards**: Using CDISC ADPPK guidelines reinforces adherence to industry standards, which is viewed favorably during regulatory review.

By proactively integrating imputation strategies aligned with ADPPK and maintaining thorough documentation, sponsors can improve the robustness of their PopPK analyses and facilitate smoother regulatory interactions.

CONCLUSION AND RECOMMENDATIONS

Missing dosing times in QD and BID pharmacokinetic datasets pose a significant challenge to the integrity and interpretability of analysis-ready datasets. The CDISC ADPPK guidelines offer a structured framework to impute these missing values while ensuring transparency, traceability, and regulatory compliance.

This paper presented practical imputation approaches—including interval-based, protocol-defined window alignment, and PK sampling time–based methods—tailored for common dosing regimens. Implementing these strategies within a standardized, well-documented programming workflow facilitates reproducible and auditable data processing.

The case study demonstrated how careful imputation improves dataset completeness and traceability, supporting robust downstream pharmacometric analysis. Furthermore, adhering to ADPPK principles helps meet regulatory expectations by explicitly flagging imputations and justifying assumptions.

To optimize PK dataset quality, we recommend:
- Early planning of imputation strategies during study design.
- Consistent application of ADPPK standards across studies.
- Comprehensive documentation of imputation logic and assumptions.
- Performing sensitivity analyses to understand imputation impact.
- Fostering collaboration between programmers, pharmacometricians, and data managers.

By minding the gaps in dosing times and applying standardized imputations, programming teams can ensure scientifically sound, traceable, and regulatory-compliant PK datasets that form a solid foundation for subsequent analyses.

CONTACT INFORMATION

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

Author Name: Aschwin van der Giesen
Company: Certara
Address: Kloosterstraat 9, 5349 AB Oss
Email: aschwin.vandergiesen@certara.com
Website: www.certara.com

Author Name: Gerly van der Vleuten
Company: Certara
Address: Kloosterstraat 9, 5349 AB Oss
Email: gerly.vandervleuten@certara.com
Website: www.certara.com

USUBJID	VISIT	DOSE_ID	DOSE_DATE	DOSE_TIME	MISSING_DOSE_TIME	IMPUTED_DOSE_TIME	EXMFL	EXMDEF	OCCASION
1001	VISIT 2	1	2024-05-01	08:15	N	08:15	N	Actual collected dosing time	Morning
1001	VISIT 2	2	2024-05-01	—	Y	21:00	Y	Imputed using midpoint of evening window (20:00–22:00)	Evening
1001	VISIT 3	1	2024-05-02	—	Y	09:00	Y	Imputed using 12-hour interval from previous dose	Morning
1002	VISIT 2	1	2024-05-01	09:05	N	09:05	N	Actual collected dosing time	Morning
1002	VISIT 2	2	2024-05-01	21:30	N	21:30	N	Actual collected dosing time	Evening
1003	VISIT 1	1	2024-05-01	—	Y	09:00	Y	Imputed using midpoint of morning window (08:00–10:00)	Morning
1003	VISIT 1	2	2024-05-01	—	Y	21:00	Y	Imputed using midpoint of evening window (20:00–22:00)	Evening

Table 1. Example of raw dosing data before and after imputation.
Abbreviations: EXMFL = Imputation Flag (Y = imputed, N = collected); EXMDEF = Explanation for Modified Dosing Time; MISSING_DOSE_TIME = programming flag identifying missing dosing times prior to imputation.
All dosing windows were defined per protocol (morning: 08:00–10:00; evening: 20:00–22:00).
Imputed dosing times were derived using either protocol-defined window midpoints or 12-hour intervals relative to the previous known dose.

USUBJID	EXSEQ	EXTRT	EXDOSE	EXSTDTC	EXMFL	EXMDEF	OCCASION	TAD (h)	COMMENTS
1001	1	Study Drug A	100 mg	2024-05-01T08:15	N	—	Morning	0.0	Actual recorded dosing time
1001	2	Study Drug A	100 mg	2024-05-01T21:00	Y	Imputed using midpoint of evening window (20:00–22:00)	Evening	11.8	Imputed from evening dosing window
1001	3	Study Drug A	100 mg	2024-05-02T09:00	Y	Imputed using 12-hour interval from previous dose	Morning	23.9	Interval-based imputation
1002	1	Study Drug A	100 mg	2024-05-01T09:05	N	—	Morning	0.0	Actual recorded dosing time
1002	2	Study Drug A	100 mg	2024-05-01T21:30	N	—	Evening	12.4	Actual recorded dosing time
1003	1	Study Drug A	100 mg	2024-05-01T09:00	Y	Imputed using midpoint of morning window (08:00–10:00)	Morning	0.0	Protocol window-based imputation
1003	2	Study Drug A	100 mg	2024-05-01T21:00	Y	Imputed using midpoint of evening window (20:00–22:00)	Evening	12.0	Protocol window-based imputation

Table 2. Example of final ADPPK-compliant exposure dataset after imputation.
Abbreviations: EXSTDTC = Start Date/Time of Dose (ISO 8601 format); EXMFL = Imputation Flag; EXMDEF = Explanation for Modified Dosing Time; OCCASION = Dose event within BID schedule; TAD = Time After Dose (hours).
All imputed dosing times are explicitly flagged (EXMFL = 'Y') and justified in EXMDEF, in alignment with CDISC ADPPK recommendations for transparency and traceability.