PharmaSUG 2026 - Paper PK-387

A systematic approach for imputing missing dose information in population pharmacokinetic analysis datasets

Prema Sukumar, Erin Dombrowsky, Renuka Hegde, and Neelima Thanneer, Bristol Myers Squibb

ABSTRACT

Accurate recording of dosing events is essential for population pharmacokinetic (popPK) analysis. Missing dose dates or times frequently occur due to limitations in case report forms or incomplete data reconciliation, potentially compromising data integrity and analysis quality. Standardized imputation rules have been developed for oral, intravenous (IV), and subcutaneous (SC) doses to ensure complete dosing histories for popPK datasets.

Dose records lacking both start and end dates are excluded. Oral, IV, and SC studies use PK collection and adjacent dosing times for time imputation. IV and SC studies also use protocol-defined durations for imputation when dates or times are missing. ADDL, II are also used in Oral studies for missing dose records.

All imputed records are flagged, and the algorithm is documented in both dataset specifications and pharmacometric reports. This systematic approach enables pharmacometric programmers to consistently handle data deficiencies, improving analysis quality and reproducibility.

This paper will illustrate comprehensive examples for IV, SC, and oral dosing scenarios, serving as a reference for adapting standard imputation rules in popPK dataset preparation. Adoption of these standards increases efficiency and aligns with Clinical Data Interchange Standards Consortium Analysis Data Model popPK Implementation Guide, supporting industry-wide best practices.

INTRODUCTION

Population pharmacokinetic (popPK) analyses play a critical role in understanding drug behavior within the body, optimizing dosing strategies, and supporting regulatory decisions. Reliable popPK modeling requires a complete and accurate record of dosing events. However, clinical datasets often contain missing dose date and time information due to factors such as patient non-compliance, protocol changes, or limitations in data collection methods (Figure 1). When standardized imputation procedures are lacking, these gaps can introduce inconsistencies, decrease efficiency, and impact analysis quality, leading to incorrect pharmacokinetic parameter estimates. This paper outlines a systematic and transparent method for imputing missing dose information in popPK datasets to maintain data integrity and support reproducible pharmacometric analyses (Thanneer 2014).

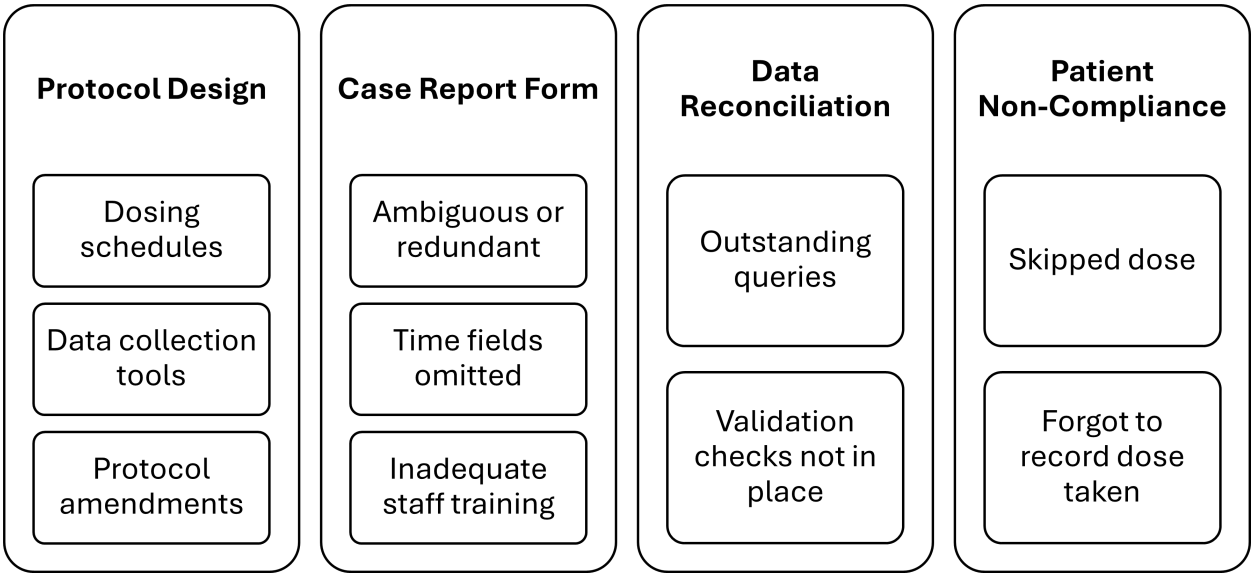

Figure 1 . Possible Causes for Missing Dose Dates and Times

PopPK datasets are derived from the Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) standards. These datasets include detailed exposure records, such as start and end dates and times for intravenous (IV) and subcutaneous (SC) dosing, and interval-based records for oral dosing. Despite rigorous data management processes, missing dose information remains a common challenge.

The FDA Population Pharmacokinetics Guidance for Industry (FDA 2022) emphasizes the importance of addressing missing data in popPK analyses to ensure efficient regulatory review. The guidance recommends that procedures for handling missing data, including dosing information, are defined and documented and all assumptions made during imputation are clearly stated.

DOSE IMPUTATION METHODS

Standardized dose imputation rules (Figure 2) are applied in popPK datasets when complete dosing details are missing from the exposure dataset (EX). Dosing intervals are typically recorded in the EX, while dose dates and times relative to PK samples may come from EX or other sources. Typically, when dose records are added from these datasets, flags are not assigned at the time of insertion because these represent actual dosing events as recorded in the source data. Instead, flags might be added afterward if it becomes necessary to impute missing dose times.

Insert Dosing Records

Include dose data directly available in SDTM/ADaM:

- Start date and time of dosing period in EX/ADEX
- End date and time of dosing period (if not the same as start date)
- Dose recorded on prior or same day of PK sample collection

Add the following if a dosing record doesn't already exist on the same date:

- Start and end dates of dosing gap or interruption (AMT = 0)
- Day prior to and after dose interruption
- Treatment start and end dates from ADSL
- On the same date of a PK sample
- Concomitant medication start and end dates

If a dosing record is missing both start and end date, it should not be included.

⇩

Dose Time Imputation

1. Infusion stop time is available, but infusion start time is missing.	→	Subtract the protocol-defined infusion duration from the infusion stop time.
2. Trough or day-one pre-dose PK sample on the same date as dosing record with missing time.	→	Use the time of the trough or day-one pre-dose PK sample.
3. End of infusion or post-dose PK sample on the same date as dosing record with missing time.	→	Use the end of infusion or post-dose PK sample time minus the protocol-defined duration after previous dose.
4. No PK sample on the same date as dosing record with missing time.	→	Use the dose time from the previous (or next if it is the first dose). Apply recursively if the dose time is missing for multiple dosing occasions.

⇩

Implicit Dose Expansion

For oral dosing, impute non-recorded doses with the number of additional doses (ADDL) and inter-dose interval (II) variables.

Figure 2 . Workflow for Including Dosing Records in PopPK Datasets

INSERT DOSING RECORDS FROM EXPOSURE

The following outlines the steps for adding dosing records from EX (or ADEX):

1. If both start and end date/time are available and not the same, then expand each dosing period into two records.
2. If start and end date/time are the same, only keep one record with that date.
3. If there are gaps in dosing period, insert those dates and assign AMT = 0.

For example, in the EX dataset below, continuous dosing from 2019-07-01 until 2019-08-09 is shown. An interruption from 2019-08-10 to 2019-08-18 is not shown in data. A single dose is recorded on 2019-08-26 and then continuous dosing from 2019-08-27 until 2019-09-23.

USUBJID	EXTRT	EXSTDTC	EXENDTC	EXSTDY	EXENDY	EXDOSE
TA123001-01-001	ABC-123	2019-07-01T12:09	2019-07-29	1	29	5
TA123001-01-001	ABC-123	2019-07-30T11:20	2019-08-09	30	40	2.5
TA123001-01-001	ABC-123	2019-08-19T10:00	2019-08-25	50	56	2.5
TA123001-01-001	ABC-123	2019-08-26T10:15	2019-08-26	57	57	2.5
TA123001-01-001	ABC-123	2019-08-27T12:05	2019-09-23	58	85	2.5

In the popPK dataset, a record is added for each unique start and end date (and time, if available) and two records are added for interruption from 2019-08-10 to 2019-08-18 (highlighted).

USUBJID	DATE	TIME	AMT
TA123001-01-001	7/1/2019	12:09	5
TA123001-01-001	7/29/2019		5
TA123001-01-001	7/30/2019	11:20	2.5
TA123001-01-001	8/9/2019		2.5
TA123001-01-001	8/10/2019		0
TA123001-01-001	8/18/2019		0
TA123001-01-001	8/19/2019	10:00	2.5
TA123001-01-001	8/25/2019		2.5
TA123001-01-001	8/26/2019	10:15	2.5
TA123001-01-001	8/27/2019	12:05	2.5
TA123001-01-001	9/23/2019		2.5

INSERT DOSING RECORDS FROM OTHER SOURCES

When dose records are added from sources other than EX, dosing information is used only if the dose date/time falls within the dosing interval in EX. A new record is not added if one already exists for that date for QD, or if two already exist for BID. For BID, an AM record is not added if one already exists for that day, nor is a PM record added if one already exists.

1. If the dates for the first and/or last dose records in EX differ from the treatment start and/or end dates in subject level datasets (like DM/ADSL), add a dose record reflecting the treatment start and/or end date.
2. When dose interruptions are available and fall within the treatment period, add records for the start and end dates of the interruption with AMT set to 0. Insert additional records on the day before and the day after the interruption using the dose amount from the dosing period.

   For example, the following interruption is recorded in the EC dataset (from ECSTDTC to ECENDTC):

USUBJID	ECSTDTC	ECENDTC	AMT
TA123001-01-001	2019-10-02	2019-10-06	0

   This single record is integrated into the popPK dataset as four records (highlighted below):

USUBJID	DATE	TIME	AMT
TA123001-01-001	2019-09-30	11:29	6
TA123001-01-001	2019-10-01		6
TA123001-01-001	2019-10-02		0
TA123001-01-001	2019-10-06		0
TA123001-01-001	2019-10-07		6
TA123001-01-001	2019-10-08	5:46	6
TA123001-01-001	2019-10-09	9:45	6
TA123001-01-001	2019-10-15	5:38	6
TA123001-01-001	2019-10-16	10:06	6
TA123001-01-001	2019-10-24	5:49	6

3. If the PK dataset contains the prior day or visit day dose date and time, add a dose record for each. When a dose record is missing on the date of a PK trough or pre-dose PK sample, add a dose record with the same date and time as the PK trough or pre-dose PK sample.

DOSE TIME IMPUTATION

Standard rules for imputing dose time apply to all popPK dataset specifications, covering IV, SC, and oral dosing (including QD and BID regimens). Records where dose time has been determined according to these guidelines should be flagged to ensure transparency. The following section outlines the dose time imputations across all administration types.

For a non-missed dose, if dose date is available but time is missing, impute the dosing time using the following (apply rules in the order presented):

1. For infusion dose only: If infusion stop time is available but infusion start time is missing, use the protocol-defined duration to determine the start of infusion, and vice versa if stop time is not available.
2. If both infusion stop and start times -- or dose time for oral dose -- are missing on any day other than day 1:

a) If a trough sample was collected on the same day, use the trough sample time as the dosing time to impute dose start time.

b) If there is no trough but a post-dose sample was collected on the same day, use the earliest post-dose sample time minus nominal duration after previous dose to impute dose start time.

3. If both infusion stop and start times -- or dose time for oral dose -- are missing on day 1, use pre-dose sample time or end of infusion sample time on the same day minus nominal infusion duration or post-dose sample time minus nominal duration after previous dose to impute dose start time.
4. If there are no concentrations associated with the missing dose time, use the dose time from the previous (or next, if it is the first dose) occasion as the current dosing time. Apply this rule recursively if the dose time is missing for multiple dosing occasions.

DOSE TIME IMPUTATION IMPLICIT DOSE EXPANSION

NONMEM, a widely used software for popPK modeling allows the variables ADDL and II in the popPK dataset to enable a concise representation of dosing regimens, particularly when not every dose is recorded. ADDL indicates how many additional doses follow the initial recorded dose, while II defines the time interval between these additional doses.

An example of ADDL and II follow for QD dosing:

USUBJID	DATE	TIME	AMT	ADDL	II
TA123001-01-001	2019-09-30	11:29	6	7	24
TA123001-01-001	2019-10-08	5:46	6	0	0
TA123001-01-001	2019-10-09	9:45	6	5	24
TA123001-01-001	2019-10-15	5:38	6	0	0
TA123001-01-001	2019-10-16	10:06	6	7	24
TA123001-01-001	2019-10-24	5:49	6	0	0

MAINTAINING TRANSPARENCY

Dosing records with imputed dose times are flagged and documented in the dataset specification (Chen 2024) and the pharmacometric report (Dombrowsky 2016) to maintain transparency during dataset assembly.

CONCLUSION

In this paper, we developed and implemented standardized dose imputation rules to enhance efficiency in the preparation of popPK datasets. Our approach provides pharmacometric programmers with a systematic method to address deficiencies in source data, ensuring accurate characterization of pharmacokinetic parameters in popPK analyses. By establishing these rules, we support consistent and reliable dataset preparation, which ultimately facilitates robust pharmacometric modeling.

Similar standards have been incorporated in the CDISC ADaM popPK Implementation Guide (CDISC 2023), aiming to promote these practices as industry-wide best standards.

REFERENCES

CDISC. 2023. "Basic Data Structure for ADaM popPK Implementation Guide v1.0." Retrieved from https://www.cdisc.org/standards/foundational/adam/basic-data-structure-adam-poppk-implementation-guide-v1-0

Thanneer, N., A. Roy, P. Sukumar, J. Bandaru, and E. Carleen. 2014. "Best Practices for Preparation of Pharmacometric Analysis Data Sets." Poster presented at the 5th American Conference on Pharmacometrics, Las Vegas, NV, October 12–15.

Dombrowsky, E., P. Sukumar, J. Bandaru, A. Roy, and N. Thanneer. 2016. "Best Practices for Preparation of Submission Quality Data Sets for Pharmacometric Analysis." Poster presented at the 7th American Conference on Pharmacometrics, Bellevue, WA, October 23–26.

Chen, L., E. Dombrowsky, B. Boyle, C. Tang, and N. Thanneer. 2024. "PmWebSpec: An Application to Create and Manage CDISC-Compliant Pharmacometric Analysis Dataset Specifications." *AAPS Journal* 26:39. https://doi.org/10.1208/s12248-024-00910-0

U.S. Food and Drug Administration. 2022. “Population Pharmacokinetics: Guidance for Industry.” Silver Spring, MD: U.S. Department of Health and Human Services. https://www.fda.gov/media/128793/download

CONTACT INFORMATION

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

Prema Sukumar
Associate Director, Data Science, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Princeton, NJ, USA
Prema.Sukumar@bms.com

Erin Dombrowsky
Director, Data Science, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Tampa, FL, USA
Erin.Dombrowsky@bms.com

Renuka Hegde
Associate Director, Data Science, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Princeton, NJ, USA
Renuka.Hegde@bms.com