PharmaSUG China 2026 - Paper DS154-2026

A Standardized Programming Framework for Daily Dose Dataset Construction from Multi-Source Glucocorticoid Data to Support Cumulative Exposure Analysis

Li, Xianjun and Lu, Chengye, Novartis

ABSTRACT

Glucocorticoids (GCs), such as prednisone, are widely used in clinical trials and are often central to evaluating steroid-sparing treatment effects. However, deriving reliable daily and cumulative GC exposure is technically complex because dose information may originate from multiple sources, including study-defined taper schedules, concomitant systemic GC medications, and rescue or escape treatments. This paper presents a standardized ADaM-based programming framework for constructing a daily GC dose dataset by aligning multi-source GC exposure on a subject-day timeline, converting all eligible GC doses to prednisone equivalents, and retaining separate dose components before aggregation. The resulting dataset supports consistent derivation of daily and cumulative exposure endpoints, estimand-aligned intercurrent event handling, and downstream analyses such as cumulative dose through Week 52 or adjusted annual cumulative dose. The framework improves traceability, reproducibility, reviewability, and cross-study consistency in GC exposure analysis for regulatory submission deliverables.

INTRODUCTION

Glucocorticoids remain widely used in clinical development, but prolonged or high cumulative exposure is associated with clinically relevant toxicities. In recent years, increasing interest in steroid-sparing treatment effects has created new analysis needs for summarizing GC exposure across multiple sources, including study-defined tapering, non-protocol GC use, and rescue or escape treatment. To support meaningful comparison of treatment effects, these different sources need to be organized on a common scale, such as prednisone-equivalent dose, while also considering intercurrent events. This paper proposes ADCUMD as an innovative ADaM-style daily dose structure that can help organize these requirements at the subject-day level and provide a transparent, reproducible basis for cumulative GC exposure analysis.

The ADCUMD dataset described in this paper extends this concept also to the daily dose level. Instead of deriving cumulative dose only at an endpoint visit, the framework creates one record per participant per study day, allowing daily prednisone-equivalent exposure, component-level dose contributions, and endpoint-ready cumulative variables to be traced within the same structure.

PROGRAMMING CHALLENGES IN GC DOSE INTEGRATION

Constructing a reliable daily GC exposure dataset requires more than a direct summation of collected dose records. In practice, GC information is captured across heterogeneous source domains, uses different time scales, and may include overlapping treatments or inconsistent dose units. These programming challenges were a major motivation for developing a unified ADCUMD structure that can harmonize multi-source exposure information before cumulative endpoints are derived.

- **Heterogeneous data sources:** GC exposure may be recorded across exposure and medication sources, such as study treatment records and concomitant medication data, with different structures, variable conventions, and levels of detail.
- **Non-aligned time scales:** Study-defined taper schedules are typically interval-based, while concomitant medication records may start and stop irregularly, requiring transformation to a common subject-day timeline.
- **Overlapping treatments:** Multiple GC drugs or dose components may be used concurrently, so the framework needs to retain component-level information before deriving the total daily dose.
- **Dose inconsistency:** Different corticosteroids, routes, dose units, and frequencies require standardized conversion and distribution rules before they can be combined meaningfully.

Addressing these challenges at the programming level helps ensure that cumulative exposure is derived from a consistent and traceable daily-dose foundation. The ADCUMD framework therefore serves as an organizing layer between source data and downstream analysis datasets, supporting both accurate exposure derivation and efficient review.

Figure 1 summarizes the data flow from source data to ADCUMD and downstream cumulative GC exposure outputs.

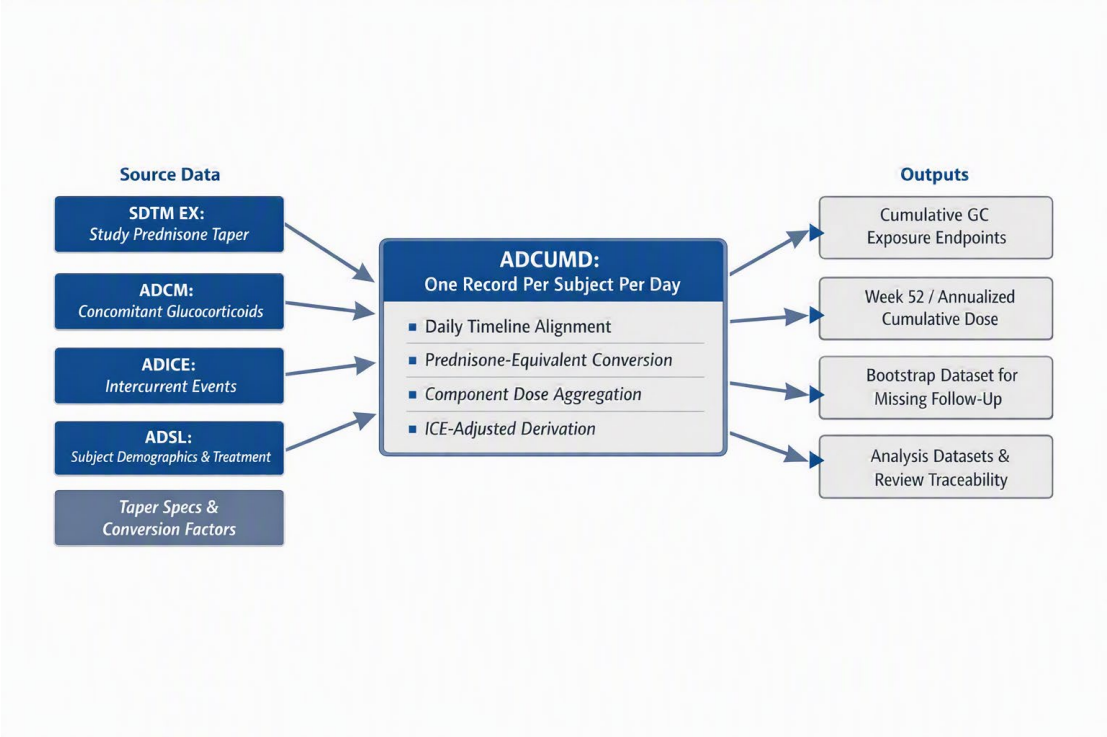

Figure 1. Data Flow for Constructing ADCUMD and Downstream Cumulative GC Exposure Outputs

ADCUMD DATASET DESIGN

The ADCUMD framework is guided by four design principles. First, exposure is organized at daily granularity so that planned taper schedules, concomitant medication records, rescue or escape dates,

and discontinuation timing can be aligned consistently. Second, dose components are retained separately before aggregation, allowing users to distinguish planned taper, escape GC, rescue GC, and other systemic GC exposure. Third, eligible systemic GC exposure is converted to prednisone-equivalent dose using pre-specified conversion factors. Fourth, intercurrent event handling is incorporated into derivation logic so that analysis-ready values can be created in alignment with the target estimand.

Core Dataset Structure

ADCUMD is constructed as one record per subject per day, typically covering the treatment or primary analysis period such as Day 1 through end of treatment period, depending on the study-level analysis window. This daily structure is useful for programmers because it provides a clear scaffold for merging study-defined taper dose, concomitant GC medication exposure, rescue or escape treatment, and intercurrent event information. It also supports reviewer-guide style traceability by keeping daily dose components, cumulative dose variables, and ICE-adjusted values within the same analysis dataset.

Table 1 lists key ADCUMD variables that help programmers and reviewers understand the daily-dose structure. **Variable**	**Description**
USUBJID	Subject identifier.
ADT / ADY	Analysis date and analysis day.
GCDOSEP	Daily planned study prednisone dose (protocol).
GCDOSET	Daily actual study prednisone dose (protocol).
GCDOSEE	Daily escape GC dose.
GCDOSER	Daily rescue GC dose.
GCDOSEO	Daily other systemic GC dose, typically derived from concomitant medication records.
GCDOSE	Total daily GC dose, calculated as the sum of component daily doses.
GCCUMP	Cumulative planned study prednisone dose (protocol).
GCCUMT	Cumulative actual study prednisone dose (protocol).
GCCUMC	Cumulative concomitant medication GC dose.
GCCUMTOT	Cumulative GC dose.
GCDOSEIE	ICE-adjusted daily GC dose used for estimand-aligned analysis.
GCCUMIE	ICE-adjusted cumulative GC dose used for endpoint derivation.

Table 1. Key ADCUMD Variables and Descriptions

Source Data Inputs

- Study actual prednisone taper information from SDTM exposure sources.
- Concomitant systemic GC medications from concomitant medication analysis data.
- Intercurrent event information from the intercurrent event analysis dataset.
- Subject-level treatment, population, and timing variables from ADSL.
- Study-defined taper schedules and dose conversion tables from study-level derivation specifications, original from protocol and SAP.

DERIVATION METHODOLOGY

Daily Dose Derivation

For each subject and each analysis day, GC exposure is derived from multiple dose components. GCDOSET represents the actual study-defined prednisone taper dose, GCDOSEE represents escape treatment dose, GCDOSER represents rescue treatment dose, and GCDOSEO represents other systemic GC medication dose. The total daily GC dose is calculated as:

GCDOSE = GCDOSET + GCDOSEE + GCDOSER + GCDOSEO

All dose components are expressed in prednisone equivalents so that different corticosteroid types and sources can be combined on a comparable scale.

Concomitant GC Processing

Concomitant systemic GC medications are processed separately before being merged into the daily ADCUMD structure. This step identifies eligible GC records, aligns each record to the analysis window, expands the exposure period to daily records, and derives prednisone-equivalent daily dose. The daily values are then summed across active records to derive the concomitant GC component for each subject day.

GC Selection Criteria

GC medications are selected from concomitant medication analysis data using drug classification and route restrictions. In a typical implementation for example, systemic corticosteroids are included based on ATC class H02, while topical corticosteroids such as H02C and mineralocorticoids such as H02AA are excluded. Oral, intravenous, and intramuscular routes are included, and records with unknown or unsupported dosing frequency are excluded or flagged for review according to study-level rules. Notice these are based on specific study definition, your study might have different selection criteria.

Exposure Window Alignment

For each selected GC record, the analysis start date is defined as day 1 of the treatment period. The analysis end date is based on the recorded medication end date when available; otherwise, it is imputed using a study-level cut-off such as end of study, Week 52, or another predefined analysis boundary. Each selected medication record is then expanded to daily-level exposure records within the aligned analysis window.

Dose Standardization, Conversion, and Frequency Adjustment

Eligible GC doses are standardized to a common prednisone-equivalent scale before daily aggregation. Prednisone and prednisolone typically use a conversion factor of 1, while other corticosteroids are converted using predefined study-level factors, such as methylprednisolone 1.25, hydrocortisone 0.25, dexamethasone 20/3, and deflazacort 5/6. After conversion, each dose is adjusted by dosing frequency to derive the applicable daily prednisone-equivalent dose.

$$\text{dose\_pred} = \text{dose} \times \text{conversion\_factor}$$
$$\text{daily\_dose} = \text{dose\_pred} \times \text{frequency\_factor}$$

Applying a common prednisone-equivalent scale allows taper, escape, rescue, and other concomitant GC components to be combined meaningfully. Frequency adjustment is then applied to translate the recorded administration pattern into a daily amount. For example, once-daily dosing uses a frequency factor of 1, twice-daily dosing uses a factor of 2, every-other-day dosing may use a factor of 0.5, and weekly dosing may be distributed across 7 days using a factor of 1/7, depending on the study-level convention. Frequencies that are missing, ambiguous, or not supported by predefined rules should be reviewed or flagged before aggregation rather than automatically assumed. When multiple eligible GC records are active on the same analysis day, their frequency-adjusted daily prednisone-equivalent doses are summed to derive the daily concomitant GC component.

Cumulative Dose Derivation

Cumulative GC dose is derived as the running sum of daily prednisone-equivalent dose over time. Component-level cumulative variables are retained to support traceability, for example GCCUMT for cumulative taper dose and GCCUMC for cumulative concomitant medication GC dose. The total cumulative dose can then be derived as GCCUMTOT = GCCUMT + GCCUMC, or as the cumulative sum of total daily dose, depending on the study-level derivation convention. The cumulative dose may represent total GC exposure through Week 52, through Day 364 or Day 365, or through an earlier study discontinuation date when applicable.

INTERCURRENT EVENT HANDLING

Intercurrent event handling is central to the ADCUMD framework. The same observed GC dose may be interpreted differently depending on the estimand. Escape treatment is generally handled using a treatment policy strategy, because the GC dose selected by the investigator is considered clinically informative and reflective of disease management. Rescue treatment and certain prohibited medication use are generally handled using a composite strategy, because post-event GC exposure may no longer represent the same clinical question. Treatment discontinuation may be handled using treatment policy when follow-up GC data are collected, with additional imputation when follow-up is incomplete.

Composite Strategy Imputation

For participants with rescue treatment or prohibited medication handled using the composite strategy, post-event GC dose may be imputed using a sufficiently unfavorable value. One possible approach is to derive each participant’s average daily GC dose during a clinically relevant reference period and then apply a high percentile of those average daily doses by treatment group or analysis stratum. The post-ICE imputed dose is calculated by multiplying this daily value by the number of remaining days to the end of the treatment period or the relevant analysis cut-off. The resulting analysis-ready cumulative dose equals observed cumulative dose before the composite ICE plus the imputed post-ICE dose.

The post-ICE imputed daily dose should be pre-specified in the SAP and scientifically justified based on the treatment context, patient population characteristics, and disease indication.

Study Discontinuation and Missing Follow-Up

If a participant discontinues study treatment but continues follow-up and GC dose data are collected through the relevant analysis period, observed values can be used under the treatment policy strategy. If follow-up is incomplete, the missing portion can be imputed using bootstrap samples or another SAP-defined method based on participants with clinically comparable observed GC exposure. Each bootstrap sample produces a completed cumulative dose value, and the downstream analysis can combine results across samples using Rubin's rules where required. Different approaches to impute missing follow-up not due to ICE can be defined in the SAP according to the study context.

IMPLEMENTATION EXAMPLE

The following pseudo-code illustrates the major programming steps. The exact implementation may vary by study, data standards, and internal macros, but the logic provides a reusable blueprint for ADCUMD derivation:

```
/* 1. Create daily calendar */
create one record per subject per study day from Day 1 to min(Week 52, EOS);

/* 2. Merge planned taper */
map baseline prednisone dose and taper regimen to daily planned dose;

/* 3. Select systemic glucocorticoids */
select eligible CM records by ATC class and route;
derive medication start and end dates within the analysis period;

/* 4. Convert and distribute daily dose */
prednisone_equiv_dose = original_dose * conversion_factor;
daily_cm_dose = prednisone_equiv_dose adjusted by frequency;

/* 5. Aggregate daily components */
total_daily_gc = taper_daily + escape_daily + rescue_daily + other_daily;
cumulative_gc = cumulative sum of total_daily_gc by subject and study day;

/* 6. Apply ICE handling */
derive earliest relevant ICE date and strategy;
if composite strategy then impute post-ICE daily dose using unfavorable value;
derive total_daily_after_ice and cumulative_after_ice;

/* 7. Create endpoint-ready variables */
derive cumulative dose at Week 52 and adjusted annual cumulative GC dose where
applicable;
```

Short R Code Example

The simplified R code below illustrates the core programming pattern used to construct a daily-dose analysis dataset. In a production implementation, study-specific functions, controlled terminology checks, date imputation rules, medication frequency parsing, and traceability variables would be added according to study-level derivation specifications and submission documentation expectations. The analysis period shown in the example should be replaced by the pre-specified analysis window for the target study.

```
# Load required packages
library(dplyr)
library(tidyr)

# 1. Create daily calendar from first dose to Week 52 or end of study
daily_calendar <- adsl %>%
  mutate(end_ady = pmin(365, eosdy, na.rm = TRUE)) %>%
  select(usubjid, trtsdt, eosdy, end_ady) %>%
  rowwise() %>%
  mutate(ady = list(seq(1, end_ady))) %>%
  unnest(ady) %>%
  ungroup() %>%
  mutate(adt = trtsdt + ady - 1)

# 2. Convert eligible systemic GC records to prednisone-equivalent dose
cm_gc_daily <- adcm %>%
  inner_join(gc_conversion, by = "gcdrug") %>%
  filter(route %in% c("ORAL", "INTRAVENOUS", "INTRAMUSCULAR")) %>%
  mutate(
    pred_eq_dose = cmdose * conv_factor,
    daily_cm_gc = pred_eq_dose * freq_factor
  ) %>%
  rowwise() %>%
  mutate(adt = list(seq(cmstdt, cmendt, by = "day"))) %>%
  unnest(adt) %>%
  ungroup() %>%
  left_join(adsl %>% select(usubjid, trtsdt), by = "usubjid") %>%
  mutate(ady = as.integer(adt - trtsdt + 1))

# 3. Aggregate concomitant GC dose by subject and analysis day
cm_gc_sum <- cm_gc_daily %>%
  group_by(usubjid, ady, adt) %>%
  summarise(other_gc_dose = sum(daily_cm_gc, na.rm = TRUE), .groups = "drop")

# 4. Merge daily dose components and derive total daily dose
adcumd <- daily_calendar %>%
  left_join(taper_daily, by = c("usubjid", "ady", "adt")) %>%
  left_join(cm_gc_sum, by = c("usubjid", "ady", "adt")) %>%
  mutate(
    across(c(taper_dose, escape_gc_dose, rescue_gc_dose, other_gc_dose), ~replace_na(.x,
0)),
    daily_gc_dose = taper_dose + escape_gc_dose + rescue_gc_dose + other_gc_dose
  ) %>%
  arrange(usubjid, ady) %>%
  group_by(usubjid) %>%
  mutate(cum_gc_dose = cumsum(daily_gc_dose)) %>%
  ungroup()
```

Bootstrap Dataset Construction for Missing Follow-Up After Study Discontinuation

For participants without ICE but with study discontinuation and incomplete GC follow-up, a bootstrap-based dataset can be generated to support multiple-imputation style analysis. The implementation uses ADCUMD as the daily-dose source and ADICE to identify relevant intercurrent events. Participants requiring imputation are identified from post-discontinuation daily records with missing or zero GC dose, while the donor pool is constructed from participants with clinically comparable observed GC exposure. Within each treatment stratum, donor average daily doses are sampled with replacement and assigned to discontinued participants for each bootstrap replicate. The imputed cumulative GC dose is then combined

with the observed cumulative dose before discontinuation to create analysis-ready records for each replicate.

```
# 1. Prepare daily ADCUMD source within the Week 52 analysis window
dt0 <- adcumd %>%
  filter(ady <= 365) %>%
  mutate(gcdose = rowSums(across(c(gcdoset, gcdosec)), na.rm = TRUE))

# 2. Identify discontinued subjects requiring post-discontinuation imputation
eos_imp_subj <- dt0 %>%
  group_by(usubjid, trt01pn) %>%
  summarise(
    eos_ady = max(ady[eosfl == "Y"], na.rm = TRUE),
    obs_cum_before_eos = max(gccumie[ady <= eos_ady], na.rm = TRUE),
    missing_dur = 365 - eos_ady,
    .groups = "drop"
  ) %>%
  filter(missing_dur > 0)

# 3. Build donor pool from subjects with observed escape-period GC exposure
escape_donor <- dt0 %>%
  filter(ice01rea == "Escape period") %>%
  group_by(usubjid, trt01pn) %>%
  summarise(
    dur = n(),
    donor_daily = sum(gcdose, na.rm = TRUE) / dur,
    .groups = "drop"
  )

# 4. Bootstrap donor daily doses within treatment stratum
set.seed(123456)
n_rep <- 1000

boot_donor <- expand_grid(
  impute = seq_len(n_rep),
  eos_imp_subj
) %>%
  group_by(impute, trt01pn) %>%
  mutate(
    sampled_daily = sample(
      escape_donor$donor_daily[escape_donor$trt01pn == first(trt01pn)],
      size = n(),
      replace = TRUE
    )
  ) %>%
  ungroup()

# 5. Derive imputed post-discontinuation exposure and completed cumulative dose
eos_imp <- boot_donor %>%
  mutate(
    gccumbt = sampled_daily * missing_dur,
    gccumies = obs_cum_before_eos + gccumbt,
    impufl = "Y",
    imptype = "BOOTSTRAP"
  )

# 6. Repeat observed Week 52 records for subjects not requiring imputation
observed_w52 <- dt0 %>%
  filter(ady == 365) %>%
  anti_join(eos_imp_subj, by = c("usubjid", "trt01pn")) %>%
  select(usubjid, trt01pn, gccumie, gccump) %>%
  crossing(impute = seq_len(n_rep)) %>%
```

```
  mutate(
    gccumies = gccumie,
    impufl = "N",
    imptype = "OBSERVED"
  )

# 7. Stack imputed and non-imputed records for analysis
adcumybt <- bind_rows(eos_imp, observed_w52) %>%
  mutate(gccumiea = gccumies - gccump) %>%
  arrange(impute, usubjid, trt01pn)
```

TRACEABILITY AND IMPLEMENTATION CONSIDERATIONS

The daily-dose framework improves traceability by retaining source identifiers, component-level dose variables, and intercurrent event information within ADCUMD. This allows reviewers to distinguish planned taper, escape, rescue, and other GC exposure, while linking derivation decisions directly to study-level analysis rules.

Key implementation considerations include:

- Build derivations at daily granularity, even when endpoints are cumulative or annualized.
- Keep dose components separate before aggregation to simplify review and troubleshooting.
- Define ICE handling, steroid conversion, and frequency rules consistently before programming.
- Use ADCUMD as the common source for downstream exposure analyses to reduce inconsistent cumulative dose derivations.

DISCUSSION

Cumulative GC exposure analysis illustrates how programming, clinical interpretation, and estimand thinking intersect in modern clinical trials. A simple summation of collected GC doses may not be sufficient when post-rescue exposure, prohibited medication use, treatment discontinuation, or incomplete follow-up changes the interpretation of the endpoint. By organizing these decisions within ADCUMD, the proposed framework provides a transparent bridge from study-level derivation requirements to analysis deliverables.

The framework is intentionally configurable. Differences in endpoint definitions, taper durations, source selection rules, stratification factors, and analysis models can be handled through study-level parameters while preserving the same underlying daily-dose structure. This approach can reduce duplication and improve consistency across studies without requiring every study to follow an identical derivation path.

CONCLUSION

A standardized daily cumulative GC dose dataset provides a practical and reviewable foundation for GC exposure analyses in clinical trials with steroid-sparing objectives. By integrating planned taper schedules,

concomitant medication conversion, intercurrent event handling, and endpoint derivations, ADCUMD supports transparent and reproducible cumulative exposure analysis. The framework can be adapted across studies where cumulative steroid exposure is an important measure of treatment benefit, safety risk, or both.

ACKNOWLEDGMENTS

The authors thank the clinical, statistical, and programming colleagues who contributed to the development and review of the cumulative GC dose derivation strategy and the supporting programming framework. The authors especially acknowledge Fen Wu and Yijie Huang for their foundational preparation work, which supported the development of this framework. The authors also acknowledge Novartis for providing the coding environment and Copilot AI assistant that supported the preparation of this work.