Paper DS05

From ATC to ACG: Implementing the Concomitant Medication Grouping (ACG) Dataset in Clinical Trial Analysis

Michelle Starkell, Hoffman-La Roche, Mississauga, ON, Canada

ABSTRACT

The Concomitant Medication Grouping Definitions Dataset (ACG) provides a standardized approach to categorizing medications into therapeutic groups, enabling consistent evaluation of drug interactions, safety signals, and treatment patterns. Clinical trials often rely on ATC classification codes, but interpreting these hierarchies can be challenging. Leveraging Standard Drug Groupings (SDGs) and defining Custom Drug Groupings (CDGs) helps ensure that ACG implementation aligns with analytical needs. A project-specific ACG supports reporting across activities, dataset version control, and the interpretability of protocol-defined events. This paper presents an approach to constructing an ACG dataset, illustrating its integration into downstream ADaM structures to support intercurrent event determination. In particular, the ability to identify prohibited or confounding concomitant medications is essential for ensuring alignment with trial estimands, as these directly impact the interpretation of treatment effects. The methodology outlines how to apply SDGs and CDGs to maximize consistency and clarity in clinical trial analytics.

INTRODUCTION

CONCOMITANT MEDICATION GROUPING AND THE ROLE OF ACG

The Concomitant Medication Grouping Definitions Dataset (ACG) is a metadata dataset used to categorize concomitant medications into clinically meaningful drug groups at predefined classification levels. ACG supports the derivation of grouped summaries of concomitant medication use and standardizes the application of WHO Drug Dictionary classifications, primarily via Anatomical Therapeutic Chemical (ATC) codes, to enable consistent downstream analyses.

Within the CDISC framework, concomitant medication information is captured in the SDTM CM (Concomitant Medications) domain, which records medications taken by a study participant before or during the trial, excluding the investigational product. The CM domain includes attributes such as medication name, dosage, route of administration, start and end dates, and indication. For analysis purposes, the Analysis Dataset for Concomitant Medications (ADCM) is derived from CM and enriched with additional flags and variables to facilitate grouped summaries, safety analyses, and protocol-defined evaluations of medication use. The ACG dataset provides the foundational definitions required to consistently derive these grouped variables within ADCM and other downstream ADaM datasets.

ATC CLASSIFICATION

The WHO Drug Dictionary classifies medications using the Anatomical Therapeutic Chemical (ATC) system, which organizes drugs across four hierarchical levels. The first level (ATC1) represents 14 anatomical main groups (e.g., Nervous System). The second level (ATC2) reflects broad therapeutic or pharmacological use (e.g., Antihypertensives). The third level (ATC3) groups drugs by chemical or pharmacological class (e.g., Beta-blockers), while the fourth level (ATC4) further refines the classification based on mechanism of action or specific therapeutic subgroup (e.g., Selective beta-blockers).

This hierarchical structure enables consistent categorization of medications across studies and supports analyses focused on drug classes rather than individual products, which is critical for identifying potential safety signals, drug interactions, and treatment patterns. Importantly, a single drug may be assigned to multiple ATC paths, reflecting differences in indication, mechanism of action, route of administration, or formulation, as well as combination

products. As a result, ATC coding allows a single medication to traverse multiple classification pathways simultaneously.

However, ATC assignments are not always available at all four levels. Many WHO Drug Dictionary preferred terms are associated with mixed or incomplete ATC classifications (e.g., ATC1 to ATC3 only). At the dataset level, this means that a single CM record may correspond to multiple analysis records in ADCM, each representing a distinct ATC path. This multiplicity introduces both flexibility and complexity when defining analysis-ready drug groupings.

Table 1. Overview of the ATC Classification Hierarchy

ATC Level	Classification	Example
ATC1	Anatomical level	Nervous system
ATC2	Pharmacological or therapeutic properties	Anesthetics
ATC3	Chemical	Opioids
ATC4	Pharmacological or therapeutic subgroups	Natural opium alkaloids

PRACTICAL CHALLENGES IN DRUG CLASSIFICATION

The implications of multi-path ATC classification are illustrated using the example of methylprednisolone. While methylprednisolone appears as a single record in the CM domain, it can map to multiple ATC1–ATC4 classifications based on formulation and therapeutic context. Consequently, the corresponding ADCM dataset may contain multiple records for the same medication, each representing a unique ATC classification path.

This behavior has direct implications for the definition of drug groupings. Depending on study objectives, groupings may be defined at varying levels of specificity: for example, at the preferred drug name level to capture specific chemotherapies, at the ATC2 level for broader therapeutic categories (e.g., antihistamines), or at the ATC4 level for more targeted subgroups (e.g., COVID-19 vaccines). *Table 2* demonstrates the multiple ATC classification paths associated with methylprednisolone and highlights the need for a structured approach to managing these mappings.

Table 2. Multiple ATC Classification Paths for Methylprednisolone

CMDECOD	ATC1	ATC2	ATC3	ATC4
METHYLPREDNISOLONE	DERMATOLOGICALS	CORTICOSTEROIDS, DERMATOLOGICAL PREPARATIONS	CORTICOSTEROIDS, PLAIN	CORTICOSTEROIDS, WEAK (GROUP I)
METHYLPREDNISOLONE	DERMATOLOGICALS	CORTICOSTEROIDS, DERMATOLOGICAL PREPARATIONS	CORTICOSTEROIDS, PLAIN	CORTICOSTEROIDS, POTENT (GROUP III)
METHYLPREDNISOLONE	SYSTEMIC HORMONAL PREPARATIONS, EXCL. SEX HORMONES AND INSULINS	CORTICOSTEROIDS FOR SYSTEMIC USE	CORTICOSTEROIDS FOR SYSTEM USE, PLAIN	GLUCOCORTICOIDS

METHYLPREDNISOLONE	SENSORY ORGANS	OPHTHALMOLOGICALS	ANTIINFLAMMATORY AGENTS	CORTICOSTEROIDS, PLAIN

STANDARD AND CUSTOM DRUG GROUPINGS

WHO Drug Dictionary provides predefined drug baskets known as Standard Drug Groupings (SDGs), which represent cross-industry consensus groupings and can be readily implemented using standard programming approaches. In contrast, Custom Drug Groupings (CDGs) are organization- or study-specific definitions created to address analytical needs not sufficiently covered by SDGs.

The ACG dataset serves as the mechanism through which both SDGs and CDGs are defined, versioned, and applied consistently across analysis datasets. While SDGs are often sufficient for common therapeutic categories, CDGs are necessary when study-specific considerations require more tailored definitions.

STUDY-SPECIFIC ACG CONSTRUCTION

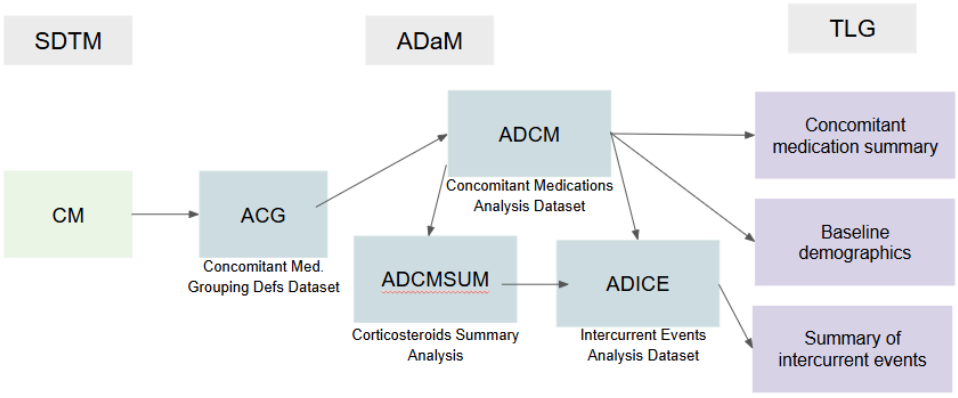
In the case study presented in this paper, certain medications are prohibited from initiation after study enrollment due to their potential to confounding treatment effects or interference with the interpretation of estimands. Existing SDGs do not fully capture the range of medications relevant to patients; for example, newly approved agents or drugs used for specialized indications. As a result, reliance on SDGs alone is insufficient to identify prohibited medications with the required level of completeness and precision.

Defining study-specific CDGs within ACG therefore represents a critical first step in creating an analysis-ready framework for evaluating concomitant medication use. A project-specific ACG enables transparent documentation of grouping logic, supports consistent application across datasets, and facilitates alignment between protocol-defined restrictions and downstream analyses.

This paper presents a structured approach to constructing an ACG dataset and illustrates its integration into downstream ADaM datasets. Using methylprednisolone as a running example, the paper traces the data flow from SDTM through ACG (drug grouping definitions), ADCM (application of drug baskets), ADCMSUM (calculation of prednisone-equivalent dose), and ADICE (identification of intercurrent events). Programming considerations and common challenges are discussed using additional examples to highlight potential pitfalls and best practices for implementation.

DATA FLOW

Figure 1. illustrates the flow of information from the SDTM CM domain through ACG and into the ADCM, ADCMSUM, ADICE datasets, and resulting TLGs.



The ACG dataset serves as the linkage between raw concomitant medication records in SDTM and analysis-ready drug group flags in ADaM. At a high level, drug group definitions are first specified in ACG using either Standard Drug

Groupings (SDGs) or Custom Drug Groupings (CDGs). These definitions are then applied to subject-level concomitant medication records in CM to derive grouped indicators within ADCM, enabling consistent downstream analyses and reporting. ADCM is used as an input to ADCMSUM (Corticosteroids Summary Analysis Dataset) and ADICE (Intercurrent Events Analysis Dataset) to determine prednisone-equivalent dosing records and intercurrent events triggered by the start of protocol-prohibited medication, respectively. These ADaM datasets are used to inform final TLG analysis such as summary of concomitant medications, baseline demographics of patients taking certain medication groups, and summary of intercurrent events.

ACG IMPLEMENTATION EXAMPLE

Table 3 presents example ACG records for two drugs: methylprednisolone and diphenhydramine hydrochloride. Each record in ACG represents a single drug grouping definition, identified by a unique PREFIX, and linked to a drug or drug class via TERMNAME and TERMID. A single drug may appear multiple times in ACG if it belongs to more than one drug grouping.

In this example, methylprednisolone is associated with two SDGs (Corticosteroids and Immunosuppressant drugs), while diphenhydramine hydrochloride is associated with a single SDG. The ability for a drug to be part of one, multiple, or no drug groupings is relevant when ACG is integrated with CM records to create drug grouping variables in ADCM.

Table 3. Example ACG Records for Selected Drugs and Standard Drug Groupings

PREFIX	TERMNAME	TERMID	GRPTYPE	GRPNAME	VERSION
SDG01	METHYLPREDNISOLONE	00049601001	SDG	Corticosteroids	WHODRUG GLOBAL B3 MARCH 1, 2025
SDG02	METHYLPREDNISOLONE	00049601001	SDG	Immunosuppressant drugs	WHODRUG GLOBAL B3 MARCH 1, 2025
SDG02	DIPHENHYDRAMINE HYDROCHLORIDE	00000402001	SDG	Immunosuppressant drugs	WHODRUG GLOBAL B3 MARCH 1, 2025

CONCOMITANT MEDICATION RECORDS IN SDTM

Table 4 shows the corresponding records to *Table 3* in the SDTM CM domain. Each subject has a single CM record per reported medication, identified by the preferred drug name (CMDECOD) and its associated drug name code (CMPNCD).

Table 4. Example SDTM CM Records for Selected Drugs

USUBJID	CMDECOD	CMPNCD
XXX-1001	METHYLPREDNISOLONE	00049601001
XXX-1002	DIPHENHYDRAMINE HYDROCHLORIDE	00000402001

IMPLEMENTATION OF ACG IN ADCM

Using ACG as an input, drug group indicators are derived in the ADCM dataset. *Table 5* illustrates how SDG flags are populated for each subject–drug record. Drugs may be flagged in multiple groupings, a single grouping, or none, depending on their presence in ACG.

Table 5. Example ADCM Records with Derived Drug Group Flags

USUBJID	CMDECOD	CMPNCD	SDG01NAM	SDG02NAM	CQ01NAM
XXX-1001	METHYLPREDNISOLONE	00049601001	Corticosteroids	Immunosuppressant drugs	
XXX-1002	DIPHENHYDRAMINE HYDROCHLORIDE	00000402001		Immunosuppressant drugs	

The drug name code (CMPNCD) is the preferred key for merging CM with ACG and ADCM, as drug names may exceed character limits and be split across variables in SDTM datasets. Using CMPNCD ensures accurate and robust classification of medications into drug groupings.

A single drug may fall into multiple baskets, as demonstrated by methylprednisolone, while other drugs may be present in only one basket or none. Conversely, when defining drug groupings based on ATC codes rather than individual drugs (e.g., COVID-19 vaccines), a single ACG record may correspond to multiple drugs. In such cases, the associated ATC code can be stored in TERMID, and any ADCM record matching that ATC code is included in the grouping. This approach avoids reliance on drug name matching and reduces the risk of inadvertent exclusions due to truncation or formatting inconsistencies.

PROGRAMMING APPROACHES FOR SDGs and CDGs

SDGs can be programmatically defined in ACG using the open-source admiral package. For each SDG, a query object is created and later combined into the ACG dataset. An example query for the “Corticosteroids” SDG is shown below:

```
sdg01 <- query(
  prefix = "SDG01",
  add_scope_num = TRUE,
  id = auto,
  definition = basket_select(
    name = "Corticosteroids",
    scope = "BROAD",
    type = "sdg"
  )
)
```

Key arguments include:

- Prefix: populates the PREFIX variable in ACG
- name: defines the GRPNAME:
- id: populates GRPID
    - when set to auto: derived directly from the SDG query identifier within the WHO Drug Dictionary
- basket_select(): specified to select a query from the SMQ database

CDGs may be defined using alternative approaches, such as custom queries or external specifications (e.g., spreadsheets listing included drugs or ATC codes). When defining CDGs, careful consideration of ATC hierarchy is required, as groupings may be constructed at varying levels of specificity. Higher ATC levels represent broader drug categories, while lower levels define more narrowly scoped groupings. CDG definitions should be developed in collaboration with clinical and statistical stakeholders to ensure they accurately reflect protocol intent and analytical requirements.

Once defined, the `admiral::create_query_data()` function is used to combine individual query objects into a single ACG dataset, which serves as a reusable input for downstream ADaM derivations.

DERIVING DRUG GROUP VARIABLES IN ADCM

Drug group indicators in ADCM are derived using the `admiral::derive_vars_query()` function. For each unique PREFIX value in ACG, corresponding CDISC variables (e.g., SDGzzNAM, SDGzzCD, SDGzzSC) are created in ADCM. A record-level match is performed by comparing the TERMCHAR or TERMNUM values in ACG to the specified source variables in ADCM.

While implementation details may vary by study, this approach provides a scalable and standardized mechanism for applying drug group definitions to concomitant medication data. The resulting ADCM dataset can then be used directly in downstream ADaMs and tables, listings, and graphs (TLGs), or as input for intermediate datasets such as ADCMSUM. Using the same running example of *methylprednisolone*, the next section demonstrates how these derived drug groupings support intercurrent event identification and estimand alignment.

SUPPORTING INTERCURRENT EVENT IDENTIFICATION USING ACG

Intercurrent event (ICE) definitions frequently depend on nuanced medication-based criteria that extend beyond the simple presence or absence of a drug. Protocols may define ICEs based on cumulative exposure, single-administration thresholds, route of administration, duration of use, or dose expressed in standardized units such as prednisone-equivalent dosing. Implementing these definitions directly from SDTM CM using free-text drug names or ad hoc programming logic can be complex, error-prone, and difficult to reproduce across studies.

ACG enables structured and transparent implementation of these definitions by providing standardized drug group indicators in ADCM. In practice, multiple prohibited medication “baskets” may be required. For example, one basket may include drugs that belong to both an SDG and a CDG and additionally meet protocol-defined conditions such as systemic route and a minimum prednisone-equivalent dose sustained over a specified duration. Other baskets may include drugs captured exclusively through a CDG that fall outside standard drug groupings. ACG supports these scenarios by allowing a single drug to appear in multiple groupings and by enabling downstream logic to reference grouping flags rather than hard-coded drug lists.

For drug classes requiring dose standardization, such as systemic corticosteroids, ACG facilitates accurate prednisone-equivalent dose calculations by identifying medications requiring conversion. These derived dose metrics can then be summarized in intermediate datasets such as ADCMSUM and evaluated against protocol-specified ICE thresholds.

Table 6 illustrates an example ADCMSUM record derived for methylprednisolone, where prednisone-equivalent dosing is aggregated at the subject level and assessed against a protocol-defined criterion (e.g., dose exceeding a specified threshold for a defined number of consecutive days). This summarized information can be used directly to flag the occurrence of an intercurrent event and to support estimand-aligned analyses.

Table 6. Example ADCMSUM Records Supporting ICE Identification

USUBJID	CMDECOD	ADOSE	PARAM
XXX-1001	METHYLPREDNISOLONE	5.2	
XXX-1001			Dose above X amount for at X consecutive days

This case study demonstrates how drug groupings defined in ACG function as a core analytical input in downstream ADaM derivations. Rather than embedding drug-specific logic directly within analysis programs, ACG enables intercurrent event identification to be driven by standardized SDG and CDG indicators derived in ADCM.

Because intercurrent event definitions may differ by drug class, ACG supports tailored analytical logic while preserving a consistent implementation framework. Drugs flagged through SDG and CDG variables (e.g.,

SDG01NAM, SDG02NAM) can be selectively processed in downstream datasets, allowing event criteria to be applied only to relevant medications. This approach supports complex rules such as cumulative exposure, single-administration thresholds, duration of use, and standardized dose calculations (e.g., prednisone-equivalent dosing), while maintaining alignment with protocol-defined specifications.

CONCLUSION

The Concomitant Medication Grouping Definitions (ACG) dataset provides a standardized framework for classifying concomitant medications, supporting consistent derivation of ADCM variables and downstream analyses, including intercurrent event determination and protocol-aligned dosing. By centralizing SDGs and CDGs, ACG addresses challenges from complex ATC hierarchies, multi-path drug classifications, and study-specific medication restrictions. By centralizing drug classification, it simplifies downstream derivations, improves interpretability, and reduces implementation burden, enabling more efficient and robust evaluation of concomitant medication effects across studies.

The structured approach presented in this paper has several key advantages:

1. **Consistency and Reproducibility**: By externalizing drug classification into ACG, complex conditional logic is modular, transparent, and consistently applied across analyses. This reduces reliance on free-text drug names, minimizes ad hoc programming, and lowers the risk of inconsistencies that could arise from manual curation.

2. **Protocol Alignment and Analytical Flexibility**: ACG allows intercurrent event definitions to reflect protocol-specified rules while supporting study-specific nuances. Custom Drug Groupings (CDGs) can be adapted to incorporate newly approved medications or drugs with unique mechanisms of action, ensuring analyses remain aligned with study objectives.

3. **Reusability Across Trials**: Beyond the presented nephrology case study, a well-constructed ACG provides a reusable framework that can be adapted for other indications or trial designs. Existing SDGs and CDGs can be modified or extended, promoting standardization across programs and facilitating traceability of drug classification logic from protocol to analysis.

IMPACT

ACG improves analytical consistency, reproducibility, and efficiency across trials, enabling scalable, estimand-aligned analyses and reducing reliance on ad hoc programming. Its reusable, study-agnostic structure makes it a practical tool for robust and transparent concomitant medication evaluation in clinical research.

REFERENCES

Lindquist M, Bate A, Edwards IR. The WHO Drug Dictionary and standardised drug groupings in pharmacovigilance. Drug Safety. 2020;43(9):817–828. doi:10.1007/s43441-020-00130-6.

OpenAI. (2026). ChatGPT (GPT-4o, January version) Large language model. https://chatgpt.com/

Pharmaverse. derive_vars_query — Derive Query Variables. Pharmaverse; 2025. Available from: https://pharmaverse.github.io/admiral/cran-release/reference/derive_vars_query.html. Accessed January 31, 2026.

Uppsala Monitoring Centre. WHODrug Standardised Drug Groupings (SDGs). UMC; 2025. Available from: https://who-umc.org/whodrug/whodrug-standardised-drug-groupings-sdgs/. Accessed January 7, 2026.

WHO Collaborating Centre for Drug Statistics Methodology. Anatomical Therapeutic Chemical (ATC) Classification System. WHO; 2025. Available from: https://www.who.int/tools/atc-ddd-toolkit/atc-classification. Accessed January 7, 2026.

CONTACT INFORMATION

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

Michelle Starkell
Hoffman-La Roche
7070 Mississauga Rd, Mississauga, ON L5N 5M8
michelle.starkell@roche.com