Journal of the Society for Clinical Data Management

Gasparyan SB, Major N, Bäckberg C, Ravikiran S, Wani P, Karpefors M. Basic Data Structure for Hierarchical Composite Endpoints: An Application to Kidney Disease Trials. *Journal of the Society for Clinical Data Management*. 2024; 4(1): 2, pp. 1–10. DOI: https://doi.org/10.47912/jscdm.265

ORIGINAL RESEARCH

# Basic Data Structure for Hierarchical Composite Endpoints: An Application to Kidney Disease Trials

Samvel B. Gasparyan*, Nicole Major†, Christoffer Bäckberg*, Srivathsa Ravikiran†, Parag Wani* and Martin Karpefors*

**Introduction:** Hierarchical composite endpoints (HCE), including the recently introduced kidney HCE, are complex endpoints that are usually analyzed by win statistics and are visualized using novel maraca plots. As a result of its novelty and the complexity of the analyses of HCE using win statistics, the construction of analysis datasets that conform to the fundamental principles put forward by the Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model (ADaM) is not straightforward.
**Objectives:** We show that in the case of a fixed follow-up it is possible to construct an analysis dataset that conforms to Basic Data Structure principles and is analysis-ready for conducting multiple analyses, including win statistics generation and visualization of HCE using maraca plots.
**Methods:** We use theoretical justification for the fixed follow-up designs to show that the pair-wise comparisons of participants for the win statistics analyses can be reduced to a participant-level ranking, and use the fundamental principles put forward by CDISC and Tidy principles of the data science community to derive an ADaM-compliant dataset.
**Results:** In the setting of fixed follow-up designs, we construct an ADaM-compliant dataset for conducting win statistics analyses and visualization using maraca plots, with the required metadata traceability.
**Conclusions:** Based on the growing importance of HCEs in clinical trials, and the difficulty in creating ADaM-compliant datasets for these analyses, we provide principles to create such datasets, to prompt the clinical community and CDISC to work towards standardization of analysis datasets for hierarchical composite endpoints.

**Keywords:** hierarchical composite endpoints; win statistics; maraca plots; CDISC ADaM; Tidy data; basic data structures

## Introduction

Hierarchical composite endpoints (HCEs) are complex endpoints[1-4] that are analyzed using win statistics and visualized using maraca plots.[6] An HCE has a hierarchical structure and uses the most clinically severe event of a participant in studies with a fixed follow-up design. This results in an ordinal endpoint, similar to the severity scale endpoints. As a result of its hierarchical nature, an HCE can combine outcomes of different types into a composite, for example, clinical events of death and hospitalization with numerical laboratory variables or symptom summary scores.[7,8] In addition, the clinical events may contribute to the composite with the time of the corresponding event, as an additional layer of severity. This means that participants having an event of the same severity are compared using the timing of the event, with a later event signifying a better outcome. Overall, the ordering is done so that a higher order means a better outcome. A characteristic of ordinal endpoints is that the concepts of better or worse are defined but not the quantitative magnitude of how much better or worse (unlike a continuous endpoint). HCEs are implemented in different therapeutic areas: COVID-19,[9,10] heart failure,[8,11] and chronic kidney disease (CKD),[7] to name a few.

Due to its novelty and the complexity of the analyses involving HCE, the construction of analysis datasets conforming to the fundamental principles put forward by the Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model (ADaM)[12] is not straightforward nor is it apparent whether it is possible. These fundamental principles were suggested with the purpose of providing standardization of the datasets across various stakeholders included in the conduct, analysis, and reporting of clinical trials in order to achieve transparency in analyses, as well as in communication and review.[13]

* Late-Stage Development, Cardiovascular, Renal, and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, SE

† Late-Stage Development, Cardiovascular, Renal, and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, US

Corresponding author: Samvel B. Gasparyan (samvel.gasparyan@astrazeneca.com)

ADaM is one of the implementations of these fundamental principles; other implementations of similar principles are known to the data science community as Tidy data principles.[14]

Win statistics[5] (win ratio,[15] win odds[16,17] or win ratio with ties,[18,19] net benefit[20]) are statistical methods for analyzing HCE and are based on the principle of comparing each participant in the active group with each participant in the control group using multiple outcomes and differing follow-ups for these outcomes. Construction of an ADaM compliant analysis dataset is therefore a challenge facing every clinical trialist involved in the analysis and reporting in a regulatory setting where such data structures are a requirement.

Using theoretical justification in the case of a fixed follow-up, we show that it is possible to construct an analysis dataset, ADHCE, that conforms to ADaM principles using the Basic Data Structure (BDS) that is analysis-ready for conducting win statistics analyses. In other words, this dataset can be used for performing the analyses without having to manipulate data first. The created BDS for the HCE analysis will therefore allow the separation of the analysis data creation from the analysis result generation (as is the intention of ADaM datasets), even for such complex analyses as win statistics calculations.

Traceability between analysis data values and their specific predecessor records is provided in the form of data point traceability. Traceability facilitates transparency of analysis conduct and allows for its replication. Detailed traceability is particularly important for the HCE derivation as it involves multiple outcomes derived through complex data manipulations from different datasets. Construction of a single ADHCE dataset that follows the BDS and is analysis-ready is important for clear communication of results and software development for analysis and reporting.

## Background

### *Basic data structures for common analysis methods*

An ADaM dataset is a particular type of analysis dataset that follows the ADaM fundamental principles defined in the ADaM[12] and is compliant to ADaM defined structures or follows as closely as possible to the ADaMIG variable naming and other conventions.[13] Currently, ADaM has three structures: Subject Level Analysis Dataset (ADSL), Basic Data Structure (BDS), and Occurrence Data Structure (OCCDS). An ADaM dataset contains both source and derived data; it is therefore important to clearly document the variable derivations and how to use them for obtaining the analysis results. ADSL is a required, participant-level dataset that contains participants' baseline and demographic characteristics, population flags that indicate the participant's inclusion in different analysis populations, planned and actual treatment variables for each period, and important dates. The BDS datasets contain endpoints and data that vary over time during the course of a study and are organized as one or more records per subject per analysis parameter per analysis timepoint. It is often optimal to have more than one BDS analysis dataset, but not necessarily one dataset per analysis. The BDS datasets are the main data structures used for complex statistical analyses but are not designed to support analysis of incidence of adverse events or other occurrence data. Analysis of such data is supported in the OCCDS. For commonly used analysis methods (eg, analysis of variance or covariance, logistic regression and so on) the BDS implementation is straightforward. A more complex analysis method for time-to-event analyses has its own standardized BDS, ADTTE, that is well developed[21] and widely used. Although the BDS supports most statistical analyses, it does not support all statistical analyses. For example, it does not support simultaneous analysis of multiple dependent (response/outcome) variables or a correlation analysis across a range of response variables.

In the ADaM design, at a minimum, the analysis datasets should contain the datasets needed for the recreation of specific statistical methods. There is no requirement that every analysis has its own dataset, but rather, a single dataset can support multiple analyses to achieve the optimal number of analysis datasets. Each analysis dataset should contain all the analysis-enabling variables required for performing the statistical analysis it is designed to support (it can even contain supportive variables not needed for the analysis but that are of interest for traceability purposes). This can lead to redundancy, that is, the same data appearing in multiple datasets, but this is necessary for having analysis-ready datasets. Analysis-ready does not mean that the results can be generated in a single statistical procedure, but rather that each of the summary statistics included in the results can be derived with minimal programming effort using standard statistical procedures with the dataset as input.

We briefly describe the fundamental principles governing the structure of BDS in connection to Tidy data principles and discuss the structure of ADLB (analysis datasets for laboratory values) that is used for the ANCOVA-type analyses and ADTTE for time-to-event analyses, as these two datasets, alongside the participant-level ADSL, are the source datasets for ADHCE. Then, following the BDS principles, we construct the ADHCE dataset, which is analysis-ready for multiple analyses (with its metadata traceability describing the source datasets and variables) and provide the minimal steps required to perform these analyses using ADHCE.

The methodology provided here is applicable only for fixed follow-up settings. For settings without fixed follow-up, we explore the challenges associated with the derivation of an analysis dataset that conforms to the BDS principles.

## Methods

### *The kidney hierarchical composite endpoint: the definition and the algorithm for construction*

Consider the case of two treatment groups, with active and control treatments, and assume that all participants have the same follow-up and there are no dropouts, meaning all participants were followed for all events of interest until the end of the fixed follow-up. The kidney HCE[4,7] has the following construction: during a fixed follow-up, participants are followed for one of the six dichotomous events in the provided hierarchy described in **Table 1**.

**Table 1:** The outcomes in the kidney HCE.

| Rank | Outcome | Subcategorization | Favorability | Source dataset |
|---|---|---|---|---|
| 1. | Death | Timing (later is better) | Worst | ADTTE |
| 2. | Dialysis | Timing (later is better) | | ADTTE |
| 3. | Sustained eGFR <15 | Timing (later is better) | | ADTTE |
| 4. | Sustained >=57% decline in eGFR | Timing (later is better) | | ADTTE |
| 5. | Sustained >=50% decline in eGFR | Timing (later is better) | | ADTTE |
| 6. | Sustained >=40% decline in eGFR | Timing (later is better) | | ADTTE |
| 7. | Individual rate of change of GFR | Actual values (higher is better) | Best | ADLB |

eGFR = estimated glomerular filtration rate.

If a participant experiences death, they are ranked in the category one and the timing of the death is used to determine the ranking within that category, with an earlier death being a worse outcome (a lower rank is assigned). Otherwise, if the participant is alive at the end of the follow-up, then the next event in the hierarchy is considered for ranking this participant and so on. If the participant did not experience any of the six events, then they fall into category seven in which the individual rate of change of glomerular filtration rate (GFR) is used to further rank the participants, with a lower rate of kidney decline being a better outcome (ranked higher).

The time-to-event (TTE) analysis dataset, ADTTE, is an ADaM BDS dataset that includes additional TTE variables designed for survival analyses. The distinguishing feature of survival data is that at the end of the observation period the event of interest may not have occurred for all subjects. The single ADTTE dataset can support multiple survival analyses, for example, Cox proportional hazards regression, Log-rank test and so on. For a given analysis parameter value (PARAM or the short name of the analysis parameter value PARAMCD), ADTTE has one record per subject and the two variables used in all models of survival analyses: the analysis value, AVAL, which shows the timepoint until when the participant was observed for the event of interest and the censoring variable, CNSR, which indicates whether or not the event of interest occurred. The variable ADTTE.AVAL therefore shows either the timing of the occurrence of the event (if CNSR=0) or the length of the fixed follow-up duration for participants without an event (CNSR=1). ADTTE should also include the subject identifier (SUBJID) and the treatment variable showing planned treatment allocation (TRTP) in a randomized, controlled trial. The fixed-follow up duration is stored in Primary Analysis Day (PADY), which is inherited from ADSL, since this variable is a common analysis date for all participants and is needed across multiple datasets. The ADTTE dataset contains the six dichotomous events of interest (**Table 1**), each having a unique PARAM value.

The BDS for laboratory data, ADLB, has one row per subject per visit per analysis parameter value and contains GFR measurements under a specific analysis parameter, PARAM, and the variables AVISIT, which indicates the timepoint of measurements (categorical variable with visit names); analysis day ADY for the number of days relative to an anchor date (in this case, the date of randomization); the analysis values AVAL, which contain the GFR measurements at each visit; and the BASE variable for the baseline GFR values for each subject. In addition, the individual rate of change of GFR over time can be derived (see the supplementary material)[7] in ADLB.AVAL corresponding to a new analysis parameter value (PARAM = "Rate of change of GFR").

### *An HCE analysis results metadata – win statistics and maraca plot*

An HCE can be analyzed using the methods for ordinal endpoints, for example, rank ANCOVA,[22] ordinal logistic regression[23] or win statistics.[5] We consider the win odds[17] but the same principles can be applied to other win statistics. Based on the hierarchy defined above, based on which each participant in the active group is compared with each participant in the control group using each participant's clinically most severe outcome. Hence, first we select the clinically most severe outcomes of the participants from the given fixed follow-up duration, then compare participants based on those outcomes. If the participant in the active group has a less severe outcome than the participant in the control group, then this is a "win" for the participant in the active group. Forming all possible comparisons of participants in the active group with participants in the control group, we derive the total number of wins, losses, and ties of the active group. The win odds of the active group against control is formed as the total number of wins (plus half of all ties) divided by the total number of losses (plus the second half of the ties). Win odds greater (less) than 1.0 is indicative of the treatment effect in the active (control) group, while win odds of 1.0 is indicative of no difference between groups.

To visualize HCEs, maraca plots (so named after their visual similarity to the musical instrument) were introduced.[6] On the maraca plot for a kidney HCE, the x-axis is divided into the seven HCE component categories in severity order from left to right. The six TTE components are visualized with adjoined cumulative Kaplan-Meier plots. For the continuous component, the x-axis corresponds to the annualized rate of change of GFR and a beneficial effect on the continuous component is characterized by a shift to the right. The associated vertical dashed lines show the median values for the annualized

rates of changes of GFR among participants without dichotomous outcomes in the two treatment groups. Each participant contributes to the HCE with one event, and the width of each category (dichotomous or continuous outcomes) corresponds to the percentage of that category in the composite. An illustration of analysis results with win odds is provided in **Table 2**, with the corresponding maraca plot in **Figure 1**.

## Results

### *ADHCE as an analysis-ready BDS*

The win odds compares every participant in the active group with every participant in the control group (a cartesian product) and hence requires these pair-wise comparisons in a dataset so that the summary of wins/losses/ties is calculated. But a dataset with that structure will not be an ADaM compliant analysis dataset and, in fact, will have a very *messy* structure according to Tidy principles, since each row will not be an observation, but a combination of observations from two treatment groups. Like BDS principles, the data science community uses Tidy principles,[14] according to which each variable should form a column, each observation should form a row, and each type of observational unit should form a dataset. Any violation of these principles results in *messy* datasets, for example, if column headers are values, not variable names or if variables are stored in both rows and columns. The Tidy principles are like the BDS principles, but they also describe in detail how these principles can be violated. The use of pair-wise comparisons in a dataset would therefore result in two columns representing the treatment groups and hence having the column names as analysis values (because the treatment group is used as an analysis value), violating another Tidy principle.

**Table 2:** Win statistics analysis example.

| Endpoint | Timepoint | Group | Participants with event n (%) | Comparison of treatment groups | | |
|---|---|---|---|---|---|---|
| | | | | Estimate | 95% CI | p-value |
| Kidney hierarchical composite endpoint | 3 years | Active N = 750 | 118 (15.7) | 1.33 | (1.18, 1.50) | <0.001 |
| | | Control N = 750 | 172 (22.9) | | | |

n (%) shows the number and percentage of participants with a dichotomous event. The percentage is calculated using the number of participants in each treatment group as a denominator.

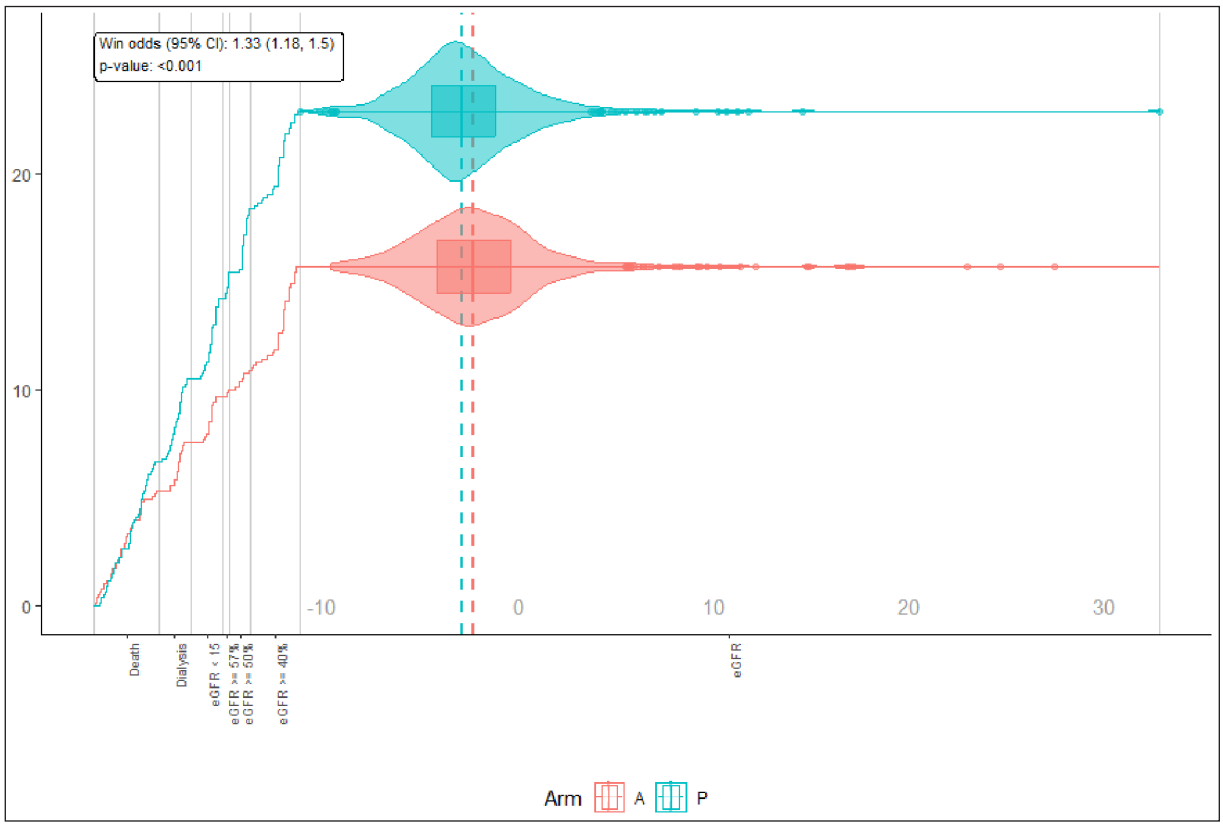



**Figure 1:** A maraca plot for HCEs.

Another possible structure for the analysis dataset would be to keep only the number of wins/losses/ties for each participant as a counting response variable. But this would mean having multiple response variables, which is also non-compliant with ADaM principles. Keeping only the wins for each participant plus half the number of ties allows a compliant dataset to be created, but limits analysis to only win odds analysis. For a win ratio analysis, a different definition of the analysis value would be needed to keep only the number of wins without ties. Importantly, different types of analyses, eg maraca visualization or ordinal regression, cannot be performed using these analyses' values.

We derive an ADaM compliant dataset (see **Figure 2**), ADHCE, with a single analysis variable that is analysis-ready for multiple analyses. The theoretical justification for this is that the number of wins of a participant can be derived using the rank of the participant in the overall dataset (both treatment groups combined) and the rank of that participant in their own treatment group.[17,18] Therefore, the participant-level ranking from the worst outcome to the most favorable can help to create an analysis value for the win statistics calculation. This methodology is applicable only in the cases of fixed follow-up durations since in case of differing follow-ups between participants comparison issues may arise, known as transitivity issues,[4,24] which would lead to comparisons not being on the participant level (impossibility to rank participants using their outcomes).

To derive AVAL in ADHCE (**Table 4**), first identify participants with any of dichotomous outcomes by selecting the PARAM value in ADTTE corresponding to this event (for example, selecting ADTTE.PARAM= "All-cause death" and ADTTE.CNSR=0). Then select the most severe event of a participant and the corresponding timing of the event from ADTTE.AVAL. If ADTTE.PADY shows the length of the fixed follow-up, then the algorithm for AVAL for each participant is shown in **Box 1**

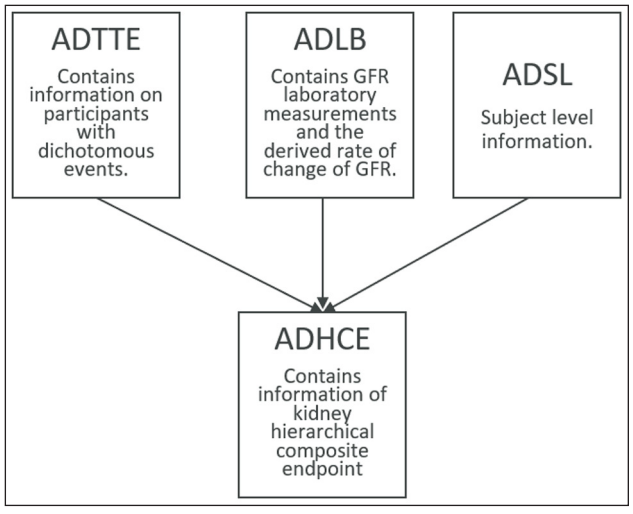



**Figure 2:** Schematic representation of relationship of ADHCE source data.

**Box 1:** Derivation of ADHCE.AVAL for dichotomous outcomes.

- if ADTTE.PARAM="All-cause death" and ADTTE.CSNR=0 then ADHCE.AVAL = 1*ADTTE.PADY + ADTTE.AVAL,
- else if ADTTE.PARAM="Dialysis" and ADTTE.CSNR=0 then ADHCE.AVAL = 2*ADTTE.PADY + ADTTE.AVAL and so on.

For participants without any dichotomous outcomes, we use the individual rate of change of GFR from ADLB, which can be negative. Regardless of their rate of change, a participant without any outcomes should have a higher AVAL than any other participant in all other categories, as shown in **Box 2**.

**Box 2:** Derivation of ADHCE.AVAL for the continuous outcome.

- ADHCE.AVAL = 7*ADTTE.PADY + ADLB.AVAL(PARAM = "Rate of change of GFR") – m + 1,

  where m is the minimum of all values ADLB.AVAL(PARAM = "Rate of change of GFR") for participants who did not have any of the dichotomous events.

The categorization of AVAL, AVALCAT1, contains the type of the event (presented in **Table 1**), while AVALCA1N is the numeric order of this categorization. As part of the traceability, we provide an illustration of the ADHCE dataset (**Table 3**), the metadata of analysis variables (including analysis parameter values) included in ADHCE (**Table 4**). For full traceability between the results, the analysis datasets and the source datasets the analysis results metadata is presented in **Table 5** (for results in **Table 2** and **Figure 1**).

### *Analysis and visualization using ADHCE*

The dataset ADHCE (**Table 3**) is analysis-ready for win odds analysis and visualization using maraca plots. Win odds in the SAS® software[25] (using the procedures *freq* or *npar1way*) is provided in the Appendix of Gasparyan et al.[17] For example, using *proc freq* the win odds can be calculated as follows (caution should be made to select the control group as the reference). See **Box 3**.

**Table 3:** Illustration of analysis dataset ADHCE.

| SUBJID | TRTP | AVAL | AVALCAT1 | AVALCA1N | PADY | PARAM | PARAMCD |
|---|---|---|---|---|---|---|---|
| 001 | A | 21 | Death | 0 | 1080 | Kidney Hierarchical composite endpoint | KHCE |

**Table 4:** Illustration of ADHCE Analysis Variable Metadata, Including Analysis Parameter Value.

| Dataset Name | Parameter Identifier | Variable Name | Variable Label | Variable Type | Display Format | Codelist/Controlled Terms | Source/Derivation |
|---|---|---|---|---|---|---|---|
| *file name of the analysis dataset* | *PARAMCD or *ALL* or *DEFAULT** | *name* | *description* | *type* | *display information* | *valid values or codes and decodes* | *where the variable came from in the source data or how the variable was derived* |
| ADHCE | *ALL* | SUBJID | Subject Identifier for the Study | Char | $11 | | ADSL.SUBJID |
| ADHCE | *ALL* | TRTP | Planned Treatment | Char | $2 | A, P | ADSL.TRT01P |
| ADHCE | *ALL* | AVAL | Analysis Value | Num | 3.2 | | First, identify participants with any of the 1–6 dichotomous events by selecting the PARAM value in ADTTE corresponding to these events. Then select the most severe event of a participant and the corresponding timing of the event.<br>If ADTTE.PARAM=" All-cause death" and ADTTE.CSNR=0 then ADHCE.AVAL = 1*ADTTE.PADY + ADTTE.AVAL<br>Else if ADTTE.PARAM="Dialysis" and ADTTE.CSNR=0 then ADHCE.AVAL = 2*ADTTE.PADY + ADTTE.AVAL and so on. Here we are using the numeric rank of each type of an event, 1 for death, 2 for dialysis and so on, following the order of the outcomes in Table 1. If the participant did not experience any of the outcomes in 1–6 then the participant falls into category 7. For this participant select the record from ADLB with PARAM = "Rate of change of GFR" and derive AVAL as ADHCE.AVAL = 7*ADTTE.PADY + ADLB.AVAL − m+1, where m is the minimum of all values ADLB.AVAL(PARAM = "Rate of change of GFR") for participants who did not have any of the dichotomous events. |
| ADHCE | *ALL* | AVALCAT1 | Analysis Value Category 1 | Char | $11 | "Death", "Dialysis", "eGFR < 15", "eGFR >= 57%", "eGFR >= 50%", "eGFR >= 40%", "eGFR" | If the result comes from ADTTE, then set to ADTTE.PARAM<br>Else if ADLB.PARAM = "Rate of change of GFR" then AVALCAT1 = "eGFR" |
| ADHCE | *ALL* | AVALCA1N | Analysis Value Category 1 (N) | Num | 3.0 | | if AVALCAT1 = "Death" then AVALCA1N = PADY<br>Else if AVALCAT1 = "Dialysis" then AVALCA1N = 2*PADY<br>Else if AVALCAT1 = "eGFR < 15"then AVALCA1N = 3*PADY<br>Else if AVALCAT1 = "eGFR >= 57%" then AVALCA1N = 4*PADY<br>Else if AVALCAT1 = "eGFR >= 50%" then AVALCA1N = 5*PADY<br>Else if AVALCAT1 = "eGFR >= 40%" then AVALCA1N = 6*PADY<br>Else if AVALCAT1 = "eGFR" then AVALCA1N= 7*PADY |
| ADHCE | *ALL* | PADY | Primary Analysis Day | Num | 3.0 | | ADSL.PADY |

**Table 5:** Analysis Results Metadata.

| Metadata Field | *Definition of field* | Metadata |
|---|---|---|
| **DISPLAY IDENTIFIER** | *Unique identifier for the specific analysis display* | Table 14.1.1 |
| **DISPLAY NAME** | *Title of display* | Primary Endpoint Analysis: Kidney hierarchical composite endpoint by Day 1080 – win statistics |
| **RESULT IDENTIFIER** | *Identifies the specific analysis result within a display* | Comparison of treatment group |
| **PARAM** | *Analysis parameter* | Kidney Hierarchical composite endpoint |
| **PARAMCD** | *Analysis parameter code* | KHCE |
| **ANALYSIS VARIABLE** | *Analysis variable being analyzed* | AVAL |
| **REASON** | *Rationale for performing this analysis* | Primary efficacy analysis as pre-specified in protocol |
| **DATASET** | *Dataset(s) used in the analysis.* | ADHCE |
| **SELECTION CRITERIA** | *Specific and sufficient selection criteria for analysis subset and/or numerator* | FASFL='Y' and PARAMCD= "KHCE" |
| **DOCUMENTATION** | *Textual description of the analysis performed* | The kidney hierarchical composite endpoint by Day 1080 is analyzed using win odds |
| **PROGRAMMING STATEMENTS** | *The analysis syntax used to perform the analysis* | PROC FREQ DATA = ADHCE;<br>TABLES TRTP * AVAL / MEASURES;<br>ODS OUTPUT MEASURES = MEASURES0;<br>RUN;<br>DATA MEASURES;<br>SET MEASURES0;<br>WP = (VALUE + 1) / 2 ;<br>ASE = ASE / 2 ;<br>ALPHA = 0.05 ;<br>C = PROBIT (1 – ALPHA / 2);<br>WO = WP/(1-WP);<br>LCL0 = WP – C * ASE;<br>UCL0 = WP + C * ASE;<br>LCL = LCL0/(1- LCL0);<br>UCL = UCL0/(1- UCL0);<br>Z = ABS (WP – 0.5) / ASE;<br>P = 2 * (1 – PROBNORM (Z));<br>KEEP WO LCL UCL P;<br>RUN; |

**Box 3:** SAS implementation of win odds.

```
proc freq data = ADHCE;
     tables TRTP * AVAL / measures;
     ods output Measures = Measures0;
run;
data measures;
     set measures0;
     WP = (value + 1) / 2 ;
     ASE = ASE / 2 ;
     alpha = 0.05 ;
     C = PROBIT (1 – alpha / 2);
     WO = WP/(1-WP);
     LCL0 = WP – C * ASE;
     UCL0 = WP + C * ASE;
     LCL = LCL0/(1- LCL0);
     UCL = UCL0/(1- UCL0);
     Z = abs (WP – 0.5) / ASE;
     P = 2 * (1 – PROBNORM (Z));
     keep WO LCL UCL P;
run;
```

Similarly in the R software, the package *hce*[26] can be used to derive the win odds. This confirms that the analysis dataset ADHCE is analysis-ready for win odds analysis since it is possible to perform the calculations without first having to manipulate the data (**Table 5**). The package *maraca*[27] in R can be utilized for producing **Figure 1** from the dataset ADHCE with minimal programming. The maraca package recognizes the ADHCE data structure as of class "adhce", meaning that it expects all the variables mentioned in the dataset's derivation above and hence can effortlessly produce the plot as shown in **Box 4**.

**Box 4:** R implementation of maraca plots.

```
library(ggplot2)
library(maraca)

class(ADHCE) #adhce
plot(ADHCE)
```

The maraca plots are *ggplot2*[28] objects and hence allow for customization. The maraca plots have the functionality of also producing an associated analysis dataset that can be used for validating this output.[29]

## Discussion

The most important question in creating BDS datasets is the decision of when to keep the required analysis value as a new variable (column) in the dataset or as a new record (row). A similar rule exists in creating Tidy datasets, which states that the column headers should not be values, but variable names.[14] In the ADaM implementation, the analysis values are stored in a column called AVAL, and the rules for adding new variables that contain analysis values are stricter. The main rule is to keep all analysis values in AVAL and to group them by the analysis parameter (PARAM) values. There are some permitted deviations though. For example, the BASE variable contains the values of AVAL corresponding to the baseline (initial timepoint). While AVALCATy (eg, AVALCAT1, AVALCAT2, and so on) and AVALCAyN are parameter variant categorizations of analysis values to categorical and numerical categories, respectively. Additional variables for analysis can be created, only if they follow the fundamental rule of adding new columns to a BDS, according to which a parameter-invariant (calculated the same way for all parameters for which the variable is populated in a dataset) function of AVAL and BASE can be derived into a new variable if it does not involve a transformation of BASE. For example, the variable CHG (change from baseline), which is derived as CHG = AVAL − BASE, is parameter-invariant and does not include a transformation of BASE, so CHG can be a new column in the analysis dataset. But a transformation of analysis values that does not meet this condition should be added as a new parameter, and AVAL should contain the transformed values. Therefore, the fundamental principle of BDS is that only one analysis variable per participant can be derived as a column in the dataset (in any other case not covered by the permitted deviations and by the fundamental rule of adding new columns described above), while multiple analysis values need to be retained in the same variable under different analysis parameter values.

An ADaM dataset is a particular type of analysis dataset that follows the ADaM fundamental principles defined in the ADaM and is compliant to ADaM defined structures or follows as closely as possible to the ADaMIG variable naming and other conventions.[13] ADTTE (Time-to-event analysis dataset)[21] is a special case exception. It does not strictly follow the fundamental principle of basic data structure as it essentially has two analysis values: length of the follow-up (AVAL variable) and a censoring variable showing whether an event happened during that follow-up (CNSR variable). This flexibility allows two dependent variables that can be used in statistical modelling. CDISC standardization of this dataset makes this a widely used and ADaM compliant dataset. Time-to-event analyses are common in clinical trials (including as a primary analysis), hence standardization of this dataset was important and is helpful for implementation.

To follow the BDS fundamental principles for the hierarchical composite endpoints in the absence of fixed follow-up is difficult since the participants are compared using their shared follow-up approach.[15] This leads to transitivity issues[16,24] and consequently participants cannot be compared on a common clinical scale, hence the impossibility to derive one analysis value per participant. All relevant events of the participant along with the maximum length of follow-up for each participant therefore need to be retained as analysis values. Different analysis values from these multiple values would contribute to analysis that depend upon which participants are compared. Therefore, this may potentially lead to multiple analysis values per participant, hence to the creation of a non-compliant analysis dataset. This would mean that an analysis dataset for win statistics analyses with variable follow-ups will either follow the BDS principles but not be analysis-ready (multiple data transformations should be done on this dataset before win statistics can be calculated) or the dataset will be analysis-ready but will not be ADaM compliant.

The presence of a fixed follow-up is of course a restriction, but it solves different statistical issues (for example, the analysis results can be interpreted on a participant level which may be more clinically meaningful) and, as described in this paper, solves issues of having multiple analysis variables as columns, hence creating the possibility to derive a dataset that conforms to the fundamental principles of ADaM and is analysis-ready for multiple analyses.

## Conclusion

We have provided the principles of constructing an analysis dataset for the hierarchical composite endpoints in a fixed follow-up setting. As an example, we have used the novel kidney HCE, but the same principles can be applied for HCEs in different therapeutic areas as well. We demonstrated that the constructed analysis dataset conforms to the fundamental principles of BDS, and so it is an ADaM compliant dataset. It is analysis-ready for multiple analyses, including generating win statistics and visualization using maraca plots. The purpose of this paper is to highlight the principles and to provide an example for content illustration with only key variables included. The constructed ADHCE dataset should not be considered as a standardization of the structure and appearance of the dataset. In line with the general note in CDISC guidance documents, eventual implementation of the dataset may follow the same principles but have a different display and contents.

Here we want to highlight the growing importance of hierarchical composite endpoints in clinical trials, including their use as a primary endpoint, and we urge the clinical community and CDISC to work together to derive a standardized analysis dataset for hierarchical composite endpoints and for win statistics analyses in general, similar to the ADTTE dataset. We hope that this paper serves as the first modest step in this direction.

## Acknowledgements

We would like to thank *Damian Kruszewski* for valuable discussions on this topic. We thank *Finn Landell* for their guidance and the overall support of this project.

## Disclaimer

The contents of this paper are the work of the authors and do not necessarily represent the opinions, recommendations, or practices of AstraZeneca. Any brand and product names are trademarks of their respective companies.

## Competing Interests

The authors have no competing interests to declare.

## References

1. **Packer M.** Proposal for a new clinical end point to evaluate the efficacy of drugs and devices in the treatment of chronic heart failure. *Journal of cardiac failure*. 2001; 7(2): 176–182. DOI: https://doi.org/10.1054/jcaf.2001.25652
2. **Packer M.** Development and evolution of a hierarchical clinical composite end point for the evaluation of drugs and devices for acute and chronic heart failure: a 20-year perspective. *Circulation*. 2016; 134(21): 1664–1678. DOI: https://doi.org/10.1161/CIRCULATIONAHA.116.023538
3. **Gasparyan SB,** et al. Hierarchical Composite Endpoints in COVID-19: The DARE-19 Trial, in Case Studies in Innovative Clinical Trials. *Chapman and Hall/CRC*. 2023; 95–148. DOI: https://doi.org/10.1201/9781003288640-7
4. **Little DJ,** et al. Validity and utility of a hierarchical composite endpoint for clinical trials of kidney disease progression: A review. *Journal of the American Society of Nephrology*. 2023; 34(12): 1928–1935. DOI: https://doi.org/10.1681/ASN.0000000000000244
5. **Dong G,** et al. Win statistics (win ratio, win odds, and net benefit) can complement one another to show the strength of the treatment effect on time-to-event outcomes. *Pharmaceutical Statistics*; 2022. DOI: https://doi.org/10.1002/pst.2251
6. **Karpefors M, Lindholm D, Gasparyan SB.** The maraca plot: A novel visualization of hierarchical composite endpoints. *Clinical Trials*. 2022; 20(1): 84–88. DOI: https://doi.org/10.1177/17407745221134949
7. **Heerspink HJ,** et al. Development and Validation of a New Hierarchical Composite End Point for Clinical Trials of Kidney Disease Progression. *Journal of the American Society of Nephrology*. 2023; 34(12): 2025–2038. DOI: https://doi.org/10.1681/ASN.0000000000000243
8. **Kondo T,** et al. Use of Win Statistics to Analyze Outcomes in the DAPA-HF and DELIVER Trials. *NEJM Evidence*. 2023; 2(11): EVIDoa2300042. DOI: https://doi.org/10.1056/EVIDoa2300042
9. **Kosiborod M,** et al. Effects of dapagliflozin on prevention of major clinical events and recovery in patients with respiratory failure because of COVID-19: Design and rationale for the DARE-19 study. *Diabetes, Obesity and Metabolism*. 2021; 23(4): 886–896. DOI: https://doi.org/10.1111/dom.14296
10. **Kosiborod MN,** et al. Dapagliflozin in patients with cardiometabolic risk factors hospitalised with COVID-19 (DARE-19): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Diabetes Endocrinology*. 2021; 9(9): 586–594. DOI: https://doi.org/10.1016/S2213-8587(21)00180-7
11. **Pocock SJ,** et al. The win ratio method in heart failure trials: lessons learnt from EMPULSE. *European journal of heart failure*; 2023. DOI: https://doi.org/10.1002/ejhf.2853
12. CDISC, *Analysis Data Model (ADaM)*; 2009.
13. CDISC, *Analysis Data Model Implementation Guide*; 2021.
14. **Wickham H.** Tidy data. *Journal of Statistical Software*. 2014; 59(10): 1–23. DOI: https://doi.org/10.18637/jss.v059.i10
15. **Pocock SJ,** et al. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *European Heart Journal*. 2012; 33(2): 176–182. DOI: https://doi.org/10.1093/eurheartj/ehr352
16. **Brunner E, Vandemeulebroecke M, Mütze T.** Win odds: An adaptation of the win ratio to include ties. *Statistics in Medicine*. 2021; 40(14): 3367–3384. DOI: https://doi.org/10.1002/sim.8967
17. **Gasparyan SB,** et al. Power and sample size calculation for the win odds test: application to an ordinal endpoint in COVID-19 trials. *Journal of Biopharmaceutical Statistics*. 2021; 31(6): 765–787. DOI: https://doi.org/10.1080/10543406.2021.1968893
18. **Gasparyan SB,** et al. Adjusted win ratio with stratification: calculation methods and interpretation. *Statistical Methods in Medical Research*. 2021; 30(2): 580–611. DOI: https://doi.org/10.1177/0962280220942558
19. **Dong G,** et al. The win ratio: on interpretation and handling of ties. *Statistics in Biopharmaceutical Research*; 2019. DOI: https://doi.org/10.1080/19466315.2019.1575279
20. **Buyse M.** Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. *Statistics in Medicine*. 2010; 29(30): 3245–3257. DOI: https://doi.org/10.1002/sim.3923
21. **CDISC.** *The ADaM Basic Data Structure for Time-to-Event Analyses*; 2012.
22. **Stokes ME, Davis CS, Koch GG.** *Categorical data analysis using SAS*. Third ed. 2012: SAS institute.
23. **Harrell FE.** *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. Vol. 608. 2001: Springer. DOI: https://doi.org/10.1007/978-1-4757-3462-1
24. **Gasparyan SB,** et al. Design and Analysis of Studies Based on Hierarchical Composite Endpoints: Insights

from the DARE-19 Trial. *Ther Innov Regul Sci*. 2022; 56(5): 785–794. DOI: https://doi.org/10.1007/s43441-022-00420-1

25. **SAS Institute Inc.** *The SAS System. Version 9.4*. 2013, SAS Institute Inc., http://www.sas.com/: Cary, NC.
26. **Gasparyan, SB.** *hce: Design and Analysis of Hierarchical Composite Endpoints.* R package version >=0.5.0. 2022. https://CRAN.R-project.org/package=hce.
27. **Karpefors M, Gasparyan SB, Huhn M.** maraca: The Maraca Plot: Visualization of Hierarchical Composite Endpoints in Clinical Trials. R package version >=0.5.0. 2023. https://CRAN.R-project.org/package=maraca.
28. **Wickham H.** *ggplot2: Elegant Graphics for Data Analysis*. Use R! 2016: Springer New York, NY. DOI: https://doi.org/10.1007/978-0-387-98141-3
29. **Major N,** et al. *Validating novel maraca plots–R and SAS love story.* https://www.pharmasug.org/proceedings/2023/SA/PharmaSUG-2023-SA-068.pdf, in *PharmaSUG 2023*. 2023: San Francisco.

**How to cite this article:** Gasparyan SB, Major N, Bäckberg C, Ravikiran S, Wani P, Karpefors M. Basic Data Structure for Hierarchical Composite Endpoints: An Application to Kidney Disease Trials. *Journal of the Society for Clinical Data Management*. 2024; 4(1): 2, pp. 1–10. DOI: https://doi.org/10.47912/jscdm.265

**Submitted:** 09 August 2023 **Accepted:** 10 January 2024 **Published:** 05 January 2024

**Copyright:** © 2024 SCDM publishes JSCDM content in an open access manner under a Attribution-Non-Commercial-ShareAlike (CC BY-NC-SA) license. This license lets others remix, adapt, and build upon the work non-commercially, as long as they credit SCDM and the author and license their new creations under the identical terms. See https://creativecommons.org/licenses/by-nc-sa/4.0/.

*Journal of the Society for Clinical Data Management* is a peer-reviewed open access journal published by Society for Clinical Data Management.

OPEN ACCESS