PharmaSUG 2025 - Paper SS-070

Identification of Domains Containing Screen Failure Participants in SDTMs and ADaMs for Reviewer's Guides

Sumanjali Mangalarapu, Merck & Co., Inc., Rahway, NJ, USA.

ABSTRACT

This paper addresses the challenge of efficiently identifying screen failure participants within submission datasets, as outlined in the Clinical Submission Data Reviewer's Guide (CSDRG) and the Analysis Data Reviewer's Guide (ADRG). Statisticians and programmers often face time-consuming manual reviews of numerous SDTM and ADaM domains to pinpoint screen failure participants. The popular software, Pinnacle 21 community version, which performs compliance checks, cannot trace screen failure in all SDTM/ADaM datasets. This paper presents a programmatic solution that automates the detection of screen failure data, significantly reducing the workload for data analysts while ensuring the reliability of results.

This approach will benefit statistical programming departments in both clinical research organizations and the pharmaceutical industry. This macro demonstrates how automation enhances data management in clinical trials, improving the overall quality of submissions and ensuring compliance with regulatory requirements.

INTRODUCTION

The preparation for regulatory submissions entails multiple components and is characterized by stringent timelines, especially in the context of Real-Time Oncology Review (RTOR) submissions. Notably, the CSDRG includes a direct question in the Subject Data Description section, requiring the author to specify which submission datasets include screen failure participants. Similar information is needed for the ADRG.

This paper provides a solution that significantly supports the programmer in preparing e-submission reviewer guide components. It operates under the premise that the input datasets—specifically the SDTMs and ADaMs—are finalized and stable, with no further updates expected. The macro intentionally excludes trial design domains such as Trial Visits (TV), Trial Summary (TS), Trial Exclusion (TE), Trial Inclusion (TI), and Trial Arms (TA), as these domains do not contain subject-level data according to the SDTM Implementation Guides (SDTMIGs).

The scrn0check.sas macro is applicable to both SDTM and ADaM datasets, enhancing the efficiency of the process to achieve optimized results. The key information required to identify screen failures can be derived from the SDTM DM domain; consequently, the macro skips the DM SDTM domain when processing the SDTM datasets.

To fully leverage this methodology, readers should have intermediate proficiency in SAS, SAS/SQL, and experience in Windows, Linux, or Unix environments, SDTM and ADaM standards published by CDISC.

MACRO SCRN0CHECK.SAS

MACRO OVERVIEW

The macro logic outlined below in Figure 1 begins with identifying participants' screen failure information in the DM dataset by examining the participant's enrollment status through the DM.ARM variable. For example, in SDTM v3.3, the macro verifies whether the ARM is NULL and ARMNRS values are "SCREEN FAILURE" "NOT ASSIGNED". For earlier versions, such as SDTM v3.2, the code examines the ARM variable to see if values are NULL or "Screen Failure".

Once the screen failure participants are identified, their IDs are output to a temporary dataset. The macro then merges the list of screen failure participant IDs with all other SDTM domains (and ADaM datasets, if specified in the macro call) while retaining only the participant ID throughout the macro logic. The macro

generates output files contain a list of domains with information on participant who failed screening, saving them as .txt files for SDTM domains and, if applicable, ADaM domains.

The figure below outlines the step-by-step processing logic of the macro.

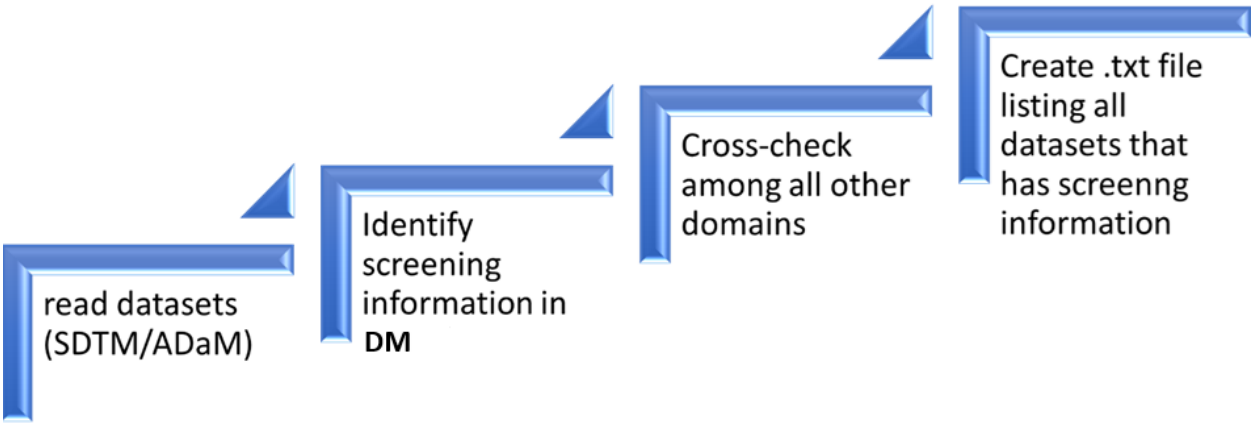


Figure 1. Flow chart of the process behind scrn0check.sas.

MACRO PARAMETERS

Macro parameters are described in Table 1 and provide detailed information to users. Among the six parameters, only “SDTM” and “SDTMV” parameters are required. By default, the macro performs checks on the SDTM domains. For ADaM checks, entries for both the “SDTM” and “ADaM” parameters are required, and macro parameter CHECK should be assigned the value “adam”.

Parameter	Description	Default	Required	Valid Values
SDTM	Location of SDTM datasets.		Y	Valid location
SDTMV	SDTMIG version number to identify the screen failure subjects		Y	Example 1: <upcase(arm) in (“SCRNFAIL” “NOT ASSIGNED” “”)> for SDTM v3.2
ADAM	Location of ADaM datasets.		N	Valid location.
CHECK	Checks SDTM or ADaM datasets library.	Performs check on SDTM datasets.	N	sdtm or adam if “adam” is entered, then “sdtm” location must parse in sdtm parameter.
PARTID	Participant ID	USUBJID	N	USUBJID, SUBJID

OUTLOC	Location to save output file as .txt format.	Defaults to SDTM or ADaM dataset location depending on type of check. if no output location is specified, defaults to SDTM library location.	N	Valid location.

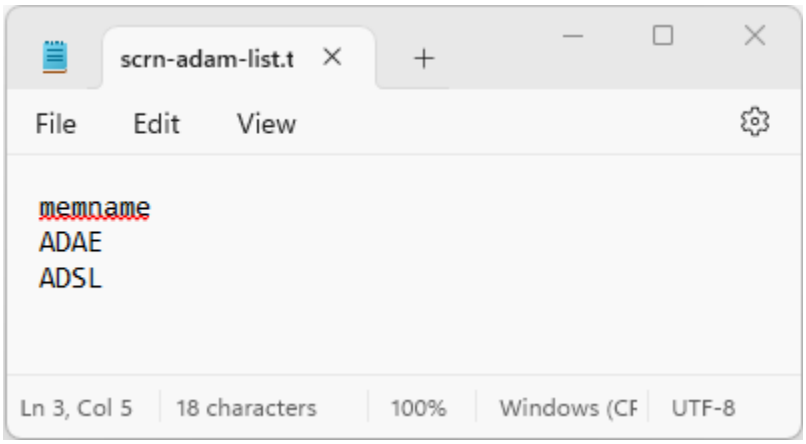
Table 1. List of macro parameters.

MACRO OUTPUT

The output file generated as .txt file is shown in Display 1 with list of ADaM datasets containing the screen failure information.

Display 1. TXT output generated by Macro.

The sample contents of output .txt file is below. The content of .txt will vary depending on the checks performed. The list here displays the list of ADaM datasets that contain screen failure participants.



Display 2. Contents of .TXT output generated by Macro.

MACRO CODE SNIPPETS

The following code snippets consider a scenario where the SDTMIG version 3.4 is the input for SDTM source data, and the checks are performed on ADaM datasets. This scenario is used to explain macro processing behind the scrn0check.sas.

Step 1: Identify screen failure information.

This code snippet identifies the screen failure participants using the SDTM DM domain.

```
proc sort data=sdtm.dm out=ids(keep=&partid.);
  by &partid.;
       %if %str(%sysevalf(&sdtmv)) =3.2 %then %do;
       where missing(arm) or upcase(arm) in ("SCREEN FAILURE" "NOT
      ASSIGNED");
       %end;
```

```
        %else %if %str(%sysevalf(&sdtmv)) >=3.3 %then %do;
        where missing(arm) and upcase(armnrs) in ("SCREEN FAILURE" "NOT
       ASSIGNED");
        %end;
run;
```

Program 1. Identification of screen failure information.

Step 2: Merge DM with SDTM and ADaM datasets.

The next code snippet shows the merge of screen failure information with all other SDTM tabulation domains and ADaM analysis domains.

Additionally, this example includes code if the user assigns the ***check*** parameter as “adam”. In this case, the macro will continue to check for screen failure information in the ADaM analysis datasets.

```
%do i=1 %to &max. ;
      %if &check.=sdtm %then %do;
            proc sort data=sdtm.&&ds&i. out=&&ds&i. (keep=&partid.)
            nodupkey;
             by &partid.;
            run;
      %end;
      %else %if &check.=adam %then %do;
            proc sort data=adam.&&ds&i. out=&&ds&i. (keep=&partid.)
            nodupkey;
             by &partid.;
            run;
      %end;

      data &&ds&i.(keep=&partid.);
            merge ids(in=a) &&ds&i. (in=b);
            by &partid.;
            if a and b;
      run;

%end;
```

Program 2. Merge DM with other SDTM and ADaM datasets.

Step 3: Create .txt files.

After step 2, all datasets, excluding the trial design domains and the SDTM DM domain, are consolidated and stored in the temporary _obs_ dataset. The information in this dataset is finally exported to .txt file.

```
proc export data=_obs_
    outfile="&outloc./scrn-&check.-list.txt"
    dbms=tab
    replace;
run;
```

Program 3. Create output file in .txt format.

Step 4: Update ADRG.

Analysis domains listed in external file `scrn-adam-list.txt` will be used to address the reviewer guide question “Are data for screen failures, including data for run-in screening (for example, SDTM values of ARMCD=”SCRNFAIL”, or “NOTASSGN”) included in ADaM datasets?” as noted in Display 3.

5. Analysis Dataset Descriptions

5.1 Overview

Are data for screen failures, including data for run-in screening (for example, SDTM values of ARMCD='SCRNFAIL', or 'NOTASSGN') included in ADaM datasets?

Yes, screen failure subjects are included in the ADSL and ADAE datasets.

Display 3: Overview of section 5.1 in ADRG.

CONCLUSION

The structured approach presented in this paper ensures that accurate and efficient extraction of key information, whether from SDTM or ADaM domains, thereby enhancing the quality of clinical trial submissions.

This paper serves as a guide for statistical programmers aiming to leverage automation in e-submission preparation, aligning with the increasing demand for efficiency in the late-stage clinical trial process.

REFERENCES

1. Winsor H. B. and Widel M. 2017. "Good versus better SDTM: Including Screen Failure Data in the Study SDTM?" *Proceedings of the PharmaSUG*, Baltimore, MD. Available at https://pharmasug.org/proceedings/2017/SS/PharmaSUG-2017-SS04.pdf.
2. SDTMIG versions published by the CDISC website SDTMIG | CDISC
3. SDTM validation rules published by the Pinnacle SDTM Validation Rules | Pinnacle 21
4. ADaM validation rules published by the Pinnacle ADaM Validation Rules | Pinnacle 21

ACKNOWLEDGEMENTS

I would like to thank Merck & Co., Inc., Rahway, NJ, USA. for this opportunity to present at the PharmaSUG conference. Sincere thanks to Amy Gillespie, Mary Varughese, and Lisa Pyle for reviewing the paper and providing valuable feedback.

CONTACT INFORMATION

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

Author Name: Sumanjali Mangalarapu
Company: Merck & Co., Inc., Rahway, NJ, USA.
Address: 351 N Sumneytown Pike, North Wales, PA 19454, USA.
Email: Sumanjali.mangalarapu@merck.com