

Paper SM03

Group, Collapse, Repeat: A Dynamic SAS Macro for Stratification Pooling in Oncology Efficacy Analysis



INVENTING
FOR LIFE

PHUSE2026

Venkatesh Burla
Lulu Wang

Author Details

- Venkatesh Burla
 - Associate Principal Scientist, Stat programming at Merck & Co., Inc., Rahway, NJ, USA
 - 13 Years of experience in SAS programming
 - TA: Oncology
- Lulu Wang
 - Senior Scientist, Stat programming at Merck & Co., Inc., Rahway, NJ, USA
 - 10+ Years of experience in SAS programming
 - TA: Oncology

Agenda

1

Introduction

2

Pooling Method

3

ADSTRATA (Pooled Stratification Analysis) dataset

4

Analysis Results

5

Conclusion

6

Q and A

Introduction

It is often necessary to conduct stratified analysis in time-to-event and response-based efficacy analysis such as:

- progression-free survival (PFS)
- overall survival (OS)
- objective response rate (ORR)

ADSTRATA



The ADSTRATA dataset contains pooled stratification variables used for efficacy analysis across endpoints such as progression-free survival (PFS), Overall Survival (OS), and Objective Response Rate (ORR).



Stratification factors:

- Sarcomatoid – sarcomatoid status (csam),
- IMDC risk group (IRTIMDC)
- Region (CREGION)



This dataset is created using a single macro (%ADSTRATA) that calls two utility macros:

- %get_group_sheet
- %my_pooling.

%GET_GROUP_SHEET



This macro extracts the stratification factors and corresponding event counts by combining information from the ADSL and ADTTE datasets.



For time-to-event endpoints such as PFSIRC, and OS, the macro filters for non-censored subjects (cnsr=0) to ensure that strata are evaluated based on the observed events only.

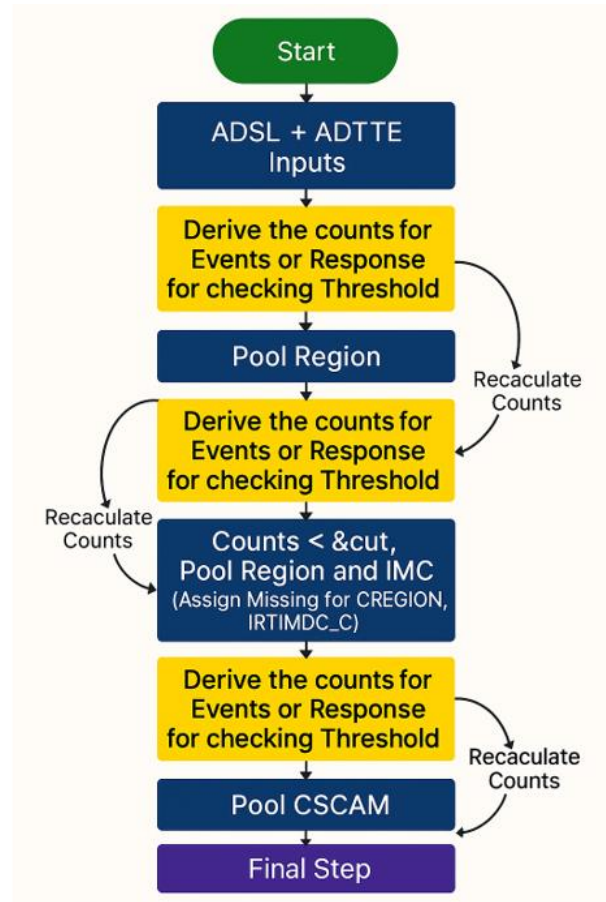


For response-based endpoints such as ORR, the macro summarizes counts directly from ADSL.

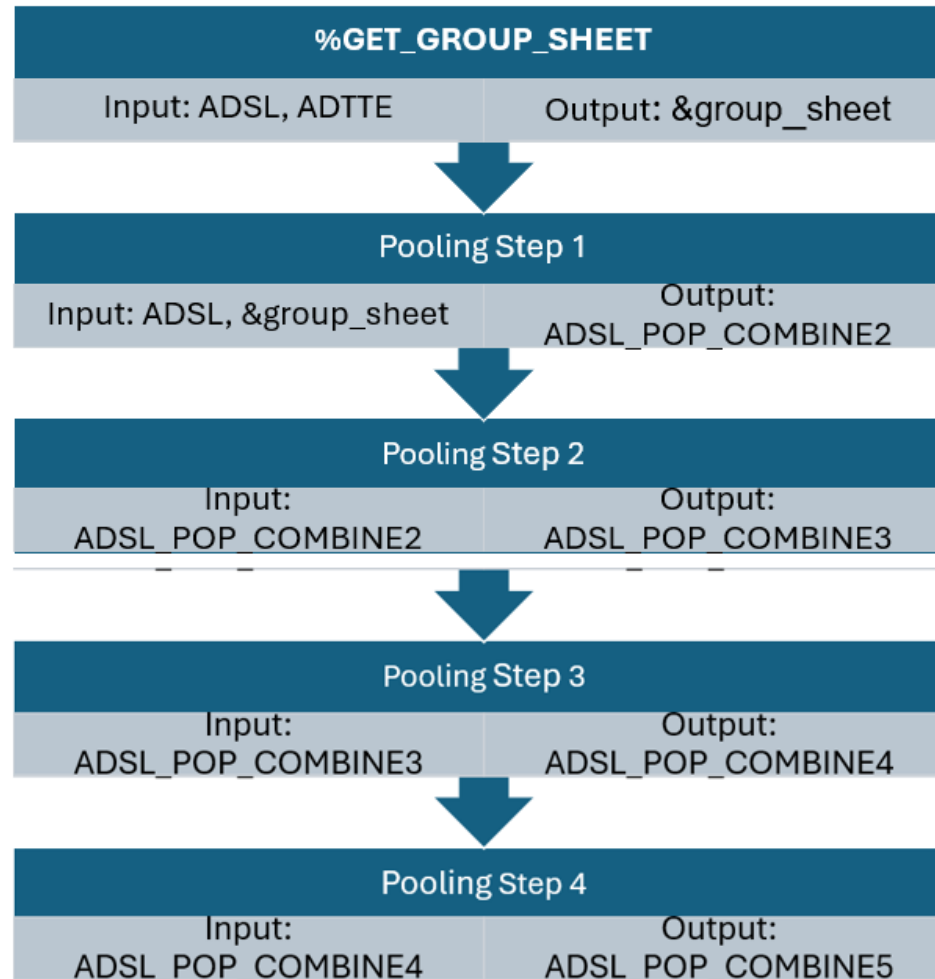


The output dataset, a “&group_sheet” provides a stratification-level summary of events across combinations of CSCAM, IRTIMDC, and CREGION.

%MY_POOLING:

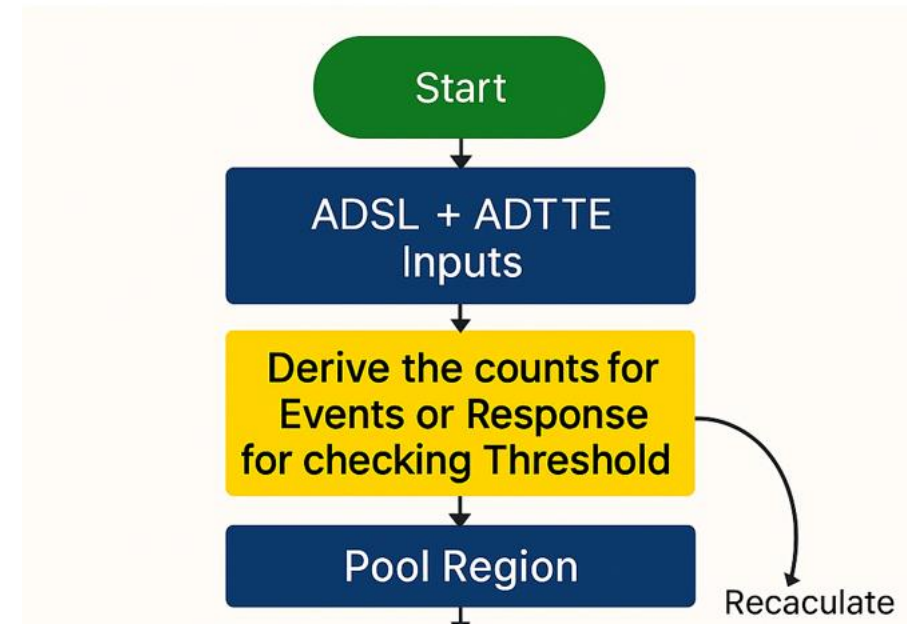


%GET_GROUP_SHEET



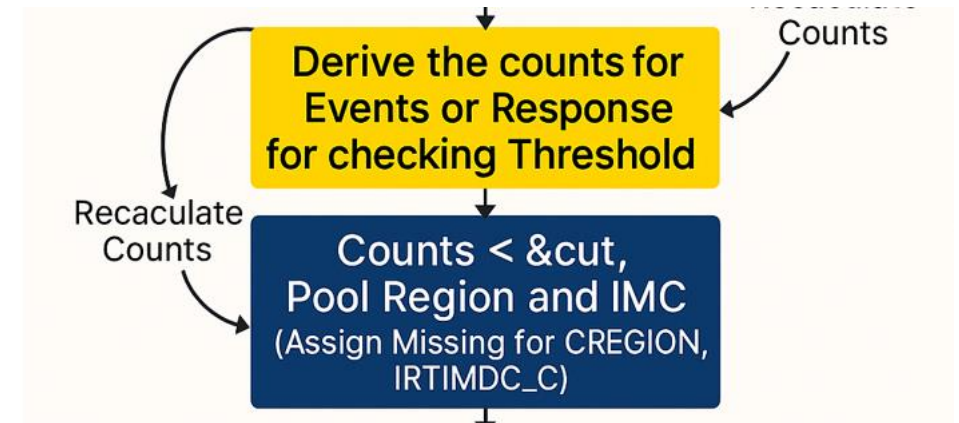
Pooling Step 1

- Creating ADSL_POP_COMBINE
- Flag the events within the Threshold
- Calculate the number of flagged strata within each CSCAM, IRTMDC combination.
- Assign the corresponding Region (CREGION) to missing if any low-count strata are detected.
- The result of step 1 is ADSL_POP_COMBINE2.



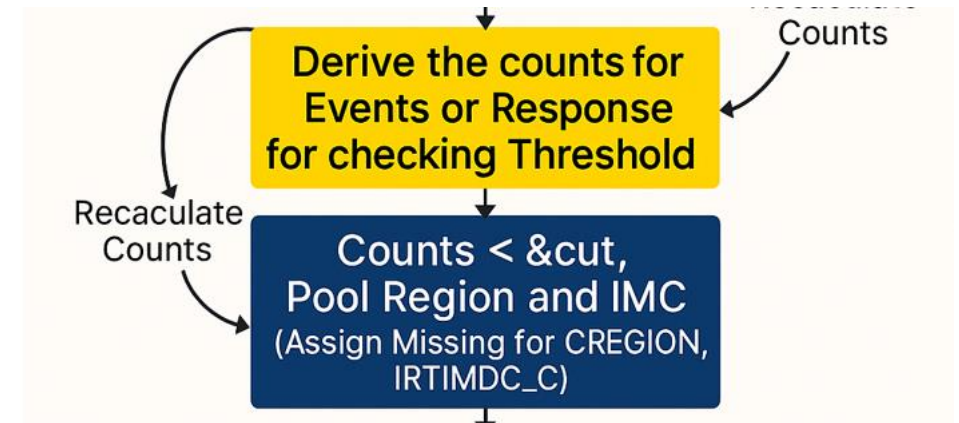
Pooling Step 2

- Call %GET_GROUP_SHEET macro with ADSL_POP_COMBINE2 as in input (from step 1)
- Flag any strata where event counts remain below the specified threshold (&cut) and create a combined IMDC category (Favorable/Intermediate) to merge those two risk groups when needed.
- Evaluate the number of flagged strata within each CSCAM and IMDC combination.
- If low-count strata persist, assign the missing value for Region (CREGION), and combine the values of IRTIMDC (Fav/Inter).
- The result of step 2 is ADSL_POP_COMBINE3.



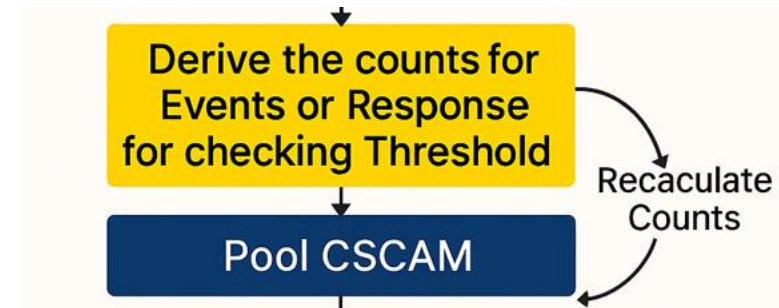
Pooling Step 3

- Run %GET_GROUP_SHEET based on ADSL_POP_COMBINE3 (result of step 2) to get GROUP_SHEET3 to reassess event counts.
- Completely drop the IRTIMDC if there are still small strata.
- Reassess the event and flag strata that still below the threshold (&cut.). If present, then both CREGION and IRTIMDC variables are reset to missing.
- The result of step 3 is ADSL_POP_COMBINE4.



Pooling Step 4

- Run %GET_GROUP_SHEET based on ADSL_POP_COMBINE4 (result of step 3) to get GROUP_SHEET4 to reassess event counts.
- Completely drop the CSCAM if there are still small strata.
- Reassess the event at the CSCAM-IMDC-CREGION group and flag strata that still below the threshold (&cut.). If present, set REGION, IRTIMDC, CSCAM to missing.



Before Pooling:

PFS/OS:

CSAM	IRTIMDC	CREGION	EVENT_COUNT
No	Favorable	North America/Western Europe	12
No	Favorable	Rest of the World	3
No	Intermediate	North America/Western Europe	18
No	Intermediate	Rest of the World	10
No	Poor	North America/Western Europe	6
No	Poor	Rest of the World	2
Yes	Favorable	North America/Western Europe	8
Yes	Favorable	Rest of the World	4
Yes	Intermediate	North America/Western Europe	20
Yes	Intermediate	Rest of the World	15
Yes	Poor	North America/Western Europe	7
Yes	Poor	Rest of the World	5

After Pooling:

CSAM	IRTIMDC	CREGION	EVENT_COUNT
No	Favorable		15
No	Intermediate	North America/Western Europe	18
No	Intermediate	Rest of the World	10
No	Poor		8
Yes	Favorable		12
Yes	Intermediate	North America/Western Europe	20
Yes	Intermediate	Rest of the World	15
Yes	Poor		12

ORR:

CSAM	IRTIMDC	CREGION	EVENT_COUNT
No	Favorable	North America/Western Europe	9
No	Favorable	Rest of the World	7
No	Intermediate	North America/Western Europe	25
No	Intermediate	Rest of the World	18
No	Poor	North America/Western Europe	6
No	Poor	Rest of the World	5
Yes	Favorable	North America/Western Europe	8
Yes	Favorable	Rest of the World	6
Yes	Intermediate	North America/Western Europe	20
Yes	Intermediate	Rest of the World	14
Yes	Poor	North America/Western Europe	7
Yes	Poor	Rest of the World	4

CSAM	IRTIMDC	CREGION	EVENT_COUNT
No	Favorable		16
No	Intermediate	North America/Western Europe	25
No	Intermediate	Rest of the World	18
No	Poor		11
Yes	Favorable		14
Yes	Intermediate	North America/Western Europe	20
Yes	Intermediate	Rest of the World	14
Yes	Poor		11

Pooling Final Step

A single, descriptive character variable

“Sarcomatoid: No, IMDC: Favorable/Intermediate, Region: Rest of the World”

USUBJID	CSCAM	IRTIMDC	CREGION	STRATP1	STRATP2	STRATP3
1001	No	Favorable	North America/Western Europe	Sarcomatoid: No, IMDC: Favorable, Region: Combined	Sarcomatoid: No, IMDC: Favorable	Sarcomatoid: No
1002	Yes	Intermediate	Rest of the World	Sarcomatoid: Yes, IMDC: Intermediate, Region: Combined	Sarcomatoid: Yes, IMDC: Intermediate	Sarcomatoid: Yes
1003	No	Poor	North America/Western Europe	Sarcomatoid: No, IMDC: Poor, Region: Combined	Sarcomatoid: No, IMDC: Poor	Sarcomatoid: No
1004	Yes	Favorable	North America/Western Europe	Sarcomatoid: Yes, IMDC: Favorable/Intermediate	Sarcomatoid: Yes, IMDC: Favorable/Intermediate	Sarcomatoid: Yes
1005	No	Poor	Rest of the World	Sarcomatoid: No, IMDC: Poor, Region: Combined	Sarcomatoid: No, IMDC: Poor	Sarcomatoid: No

Additional Functions

Support multiple efficacy endpoints with different treatment comparisons

Dynamically apply thresholds

Export the pooled events into Excel workbook

Conclusion



The development of the ADSTRATA dataset and its supporting pooling macro provides a standardized and automated approach to handling stratification factors in efficacy analysis.

This approach eliminates repetitive manual programming, reduces potential for error, and enhances consistency across studies.

Overall, the ADSTRATA macro supports efficient implementation of stratified analysis in Oncology trials and can be easily adapted for use in other therapeutic areas.

Acknowledgements

Amy Gillespie, Sai Govind Chenna, Himanshu Patel, Dileep Penmetsa, and Li Liu

Contact Information

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

Venkatesh Burla
Merck & Co., Inc., Rahway, NJ, USA
215-631-5295
Email: venkatesh.burla@merck.com

Lulu Wang
Merck & Co., Inc., Rahway, NJ, USA
267-305-6745
Email: lulu.wang2@merck.com

THANK YOU



MERCK



QUESTIONS