

From SAS to R for TLGs: A Case Study on Implementing an 'R-First, SAS-QC' Workflow for Time-to-Event Analysis

Adler Moisés Fallas Rodríguez
Intego Clinical
San José, Costa Rica

The problem we are actually solving

QC effort should not be dominated by presentation differences

Formatting $\neq$ statistics

QC must be scalable



Rendered Table ✗



Statistical Values ✓



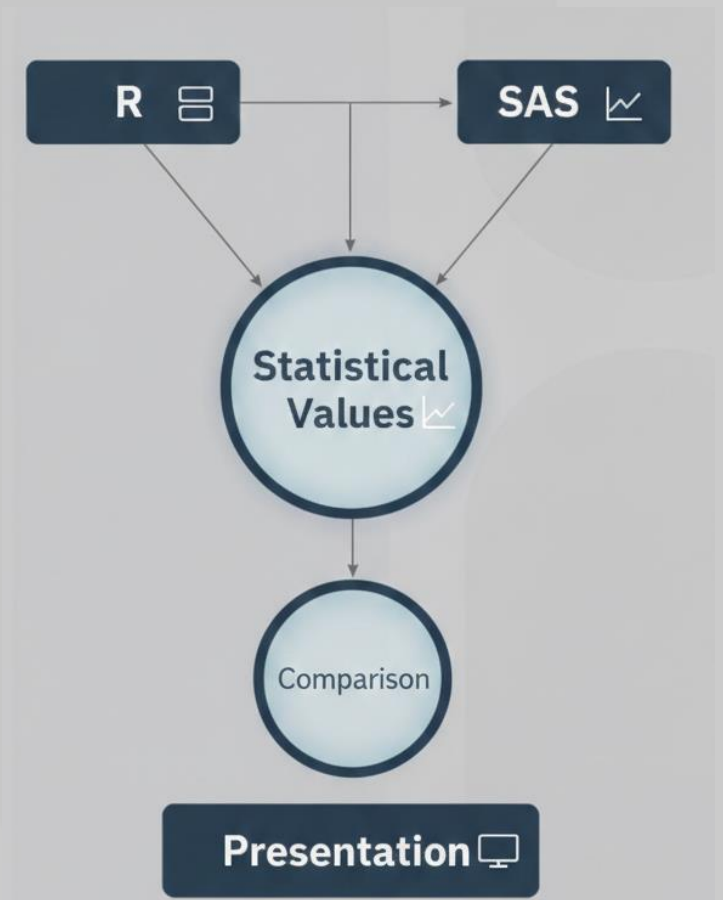
Statistical Values ✓

Design principle

Separate statistical content from presentation

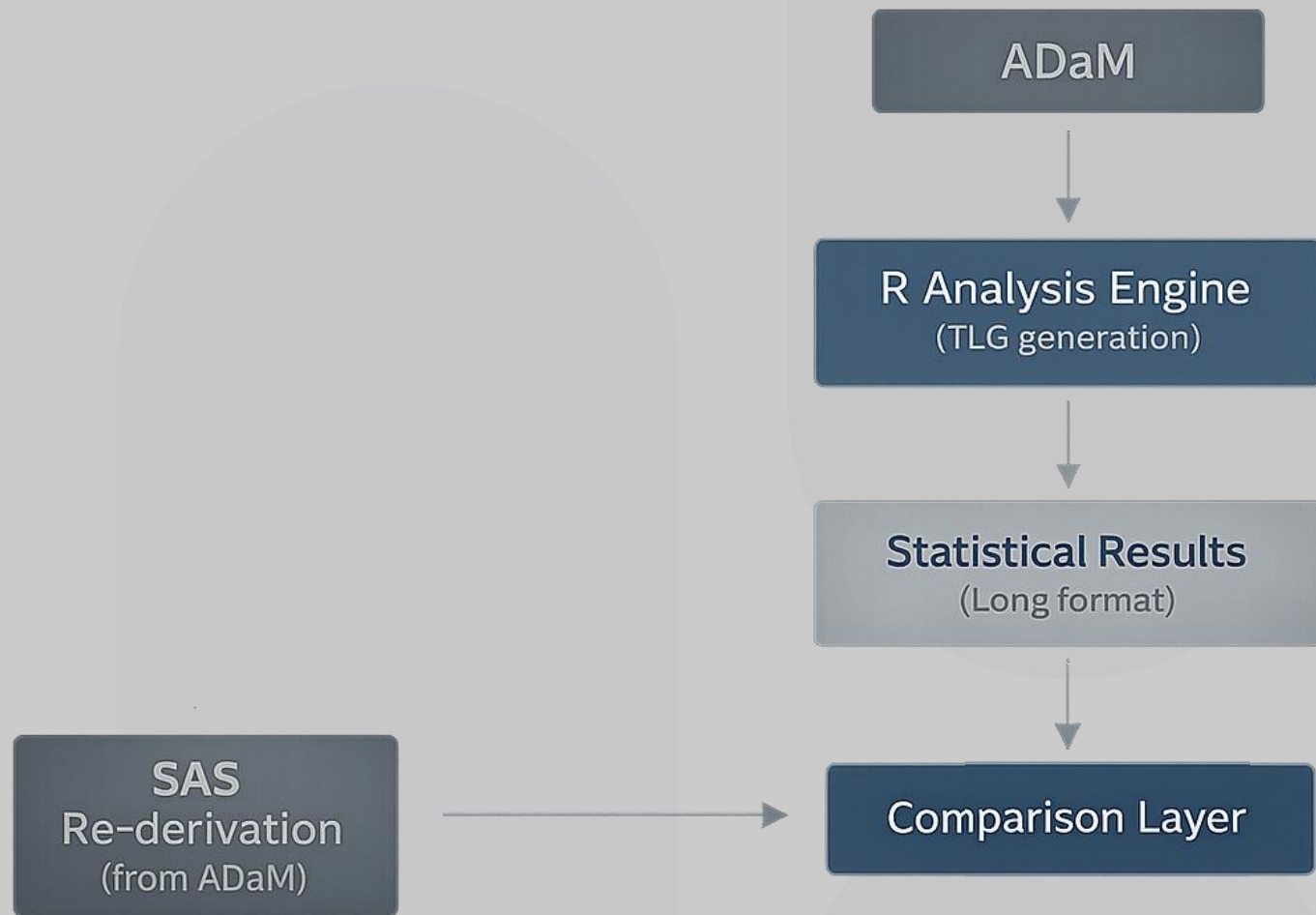
Compare statistical outputs, not rendered tables

Presentation becomes a downstream concern



High-level workflow overview

R-first production with independent SAS QC





The bridging dataset

A presentation-agnostic comparison contract

Machine-readable statistical outputs

Long-format (one statistic per row)

Stable identifiers, layout-independent

Not a rendered table

TRT	ROW_NUM	LABEL_NAME	STAT	VAL
Treatment A	1	Patients with event (%)	count	42
Treatment A	1	Patients with event (%)	percent	35.0
Control	1	Patients with event (%)	count	51
Control	1	Patients with event (%)	percent	43.2
Treatment A	2	Median Time to Event (Months)	count	14.8
Control	2	Median Time to Event (Months)	count	12.3

Generating the bridging dataset in R

**Statistical results extracted directly from
the table object**

Reshaped into a unified long format

One row = one statistic

Persisted as a platform-neutral Parquet file

TRT	ROW_NUM	LABEL_NAME	STAT	VAL
Treatment A	1	Patients with event (%)	count	42
Treatment A	1	Patients with event (%)	percent	35.0
Control	1	Patients with event (%)	count	51
Control	1	Patients with event (%)	percent	43.2
Treatment A	2	Median Time to Event (Months)	count	14.8
Control	2	Median Time to Event (Months)	count	12.3

```
r

qc_df <- as_result_df(final_tlg) |>
  reshape_to_long()

arrow::write_parquet(qc_df, "tte_qc.parquet")
```

Using the bridging dataset in SAS

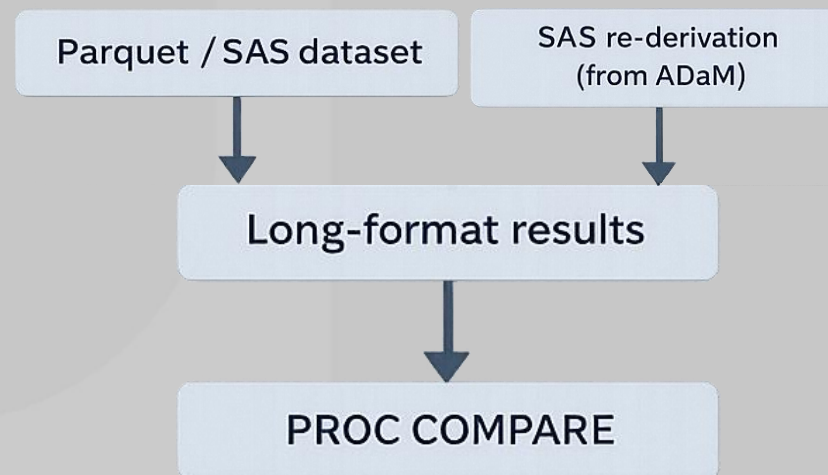
Independent re-derivation and comparison

**Bridging dataset read as a standard
SAS dataset**

**SAS independently re-derives
statistics from ADaM**

**Results assembled into the same
long-format structure**

**Comparison performed on numeric
values only**



Automated comparison and tolerances

Focusing QC on statistical agreement

Comparison performed on numeric values only

Matching on stable identifiers (TRT, ROW, STAT, ...)

Predefined tolerances for floating-point differences

No dependence on table layout or formatting

```
sas

proc compare
  base=t_ef_tte
  compare=qc_t_ef_tte
  %str(warn)ings
  listall
  method=absolute
  criterion=0.001; * tolerate minor SAS-R numeric differences (rounding to .01);

  id descending trt row_num stat label_name;
run;
```

Outcome: Meaningful discrepancies only

What this enables

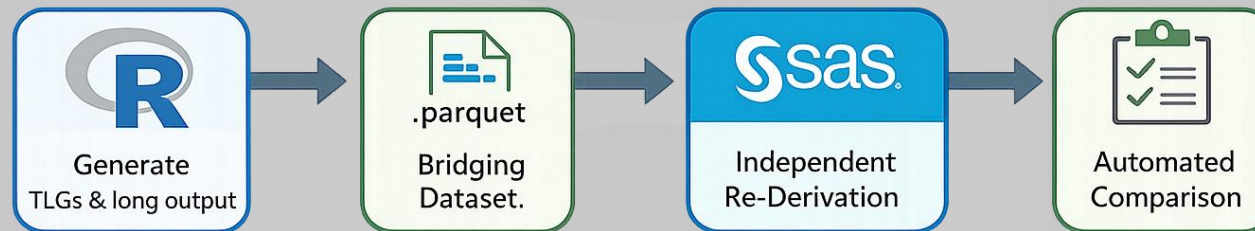
From manual QC to scalable verification

Scalable QC across multiple TLGs

Reduced manual table review

Reusable comparison logic

Clear separation of responsibilities



R-first, SAS-QC workflow

Case study

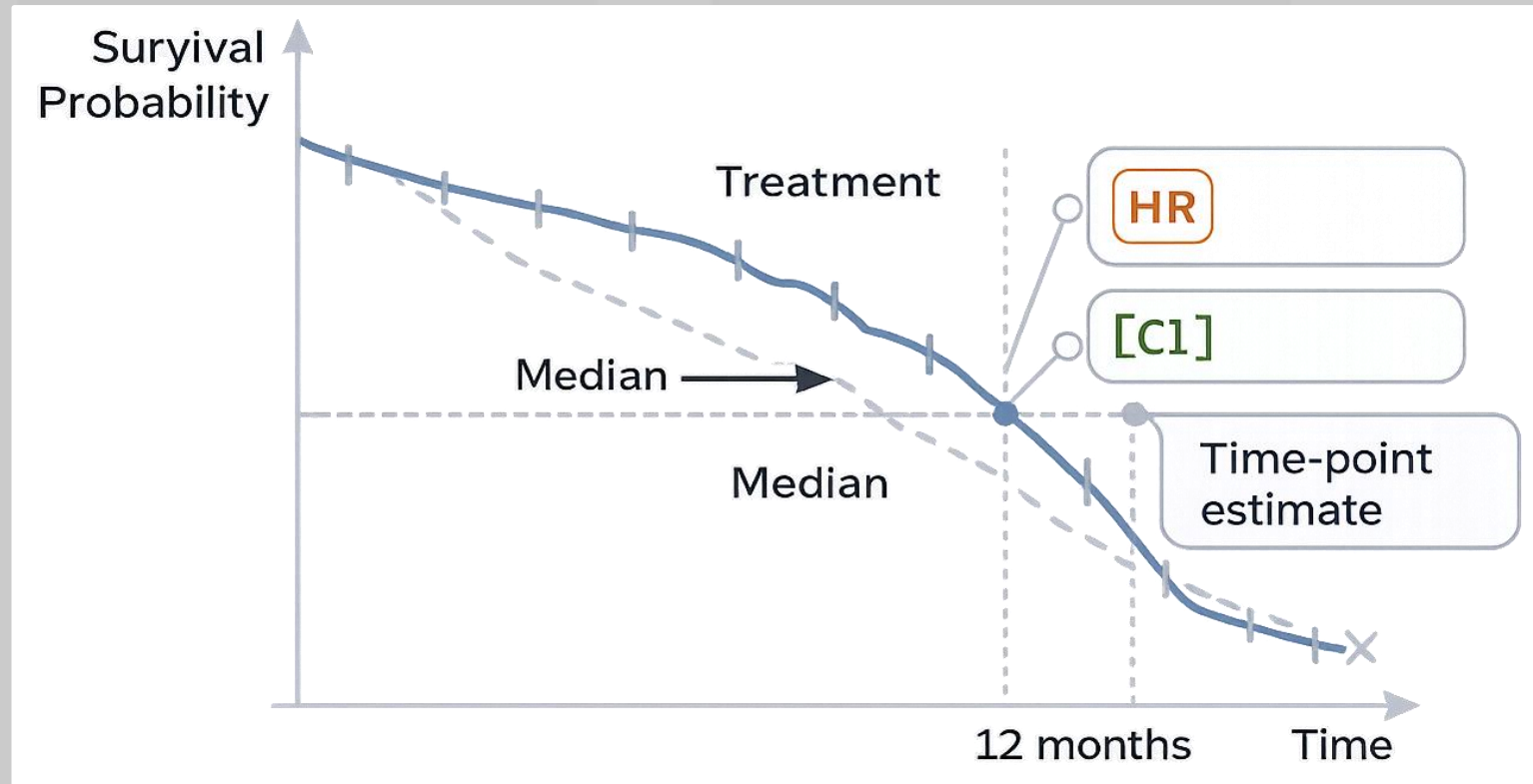
Time-to-event analysis (TTE)

A complex statistical reporting scenario

Multiple statistics per output row

Derived values sensitive to implementation choices

High QC burden under traditional workflows



Why this case stresses QC

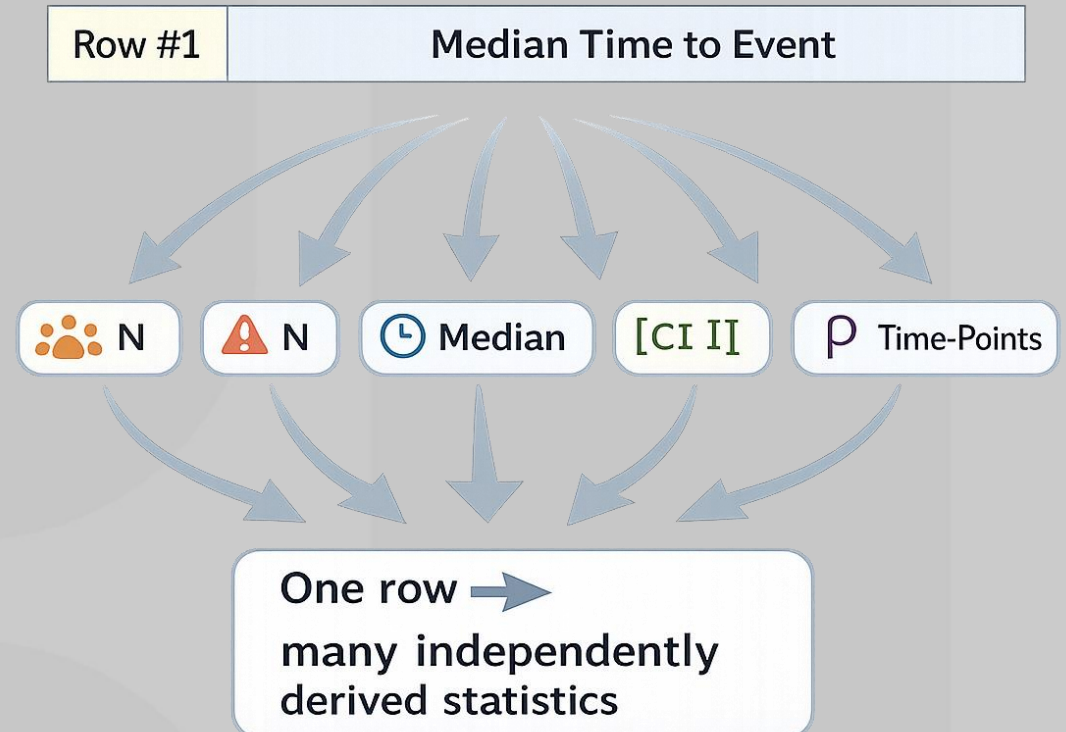
Stress-testing the comparison strategy

One-to-many relationship between rows and statistics

Numeric sensitivity to rounding and censoring rules

Presentation-heavy traditional QC

Why this case stresses QC

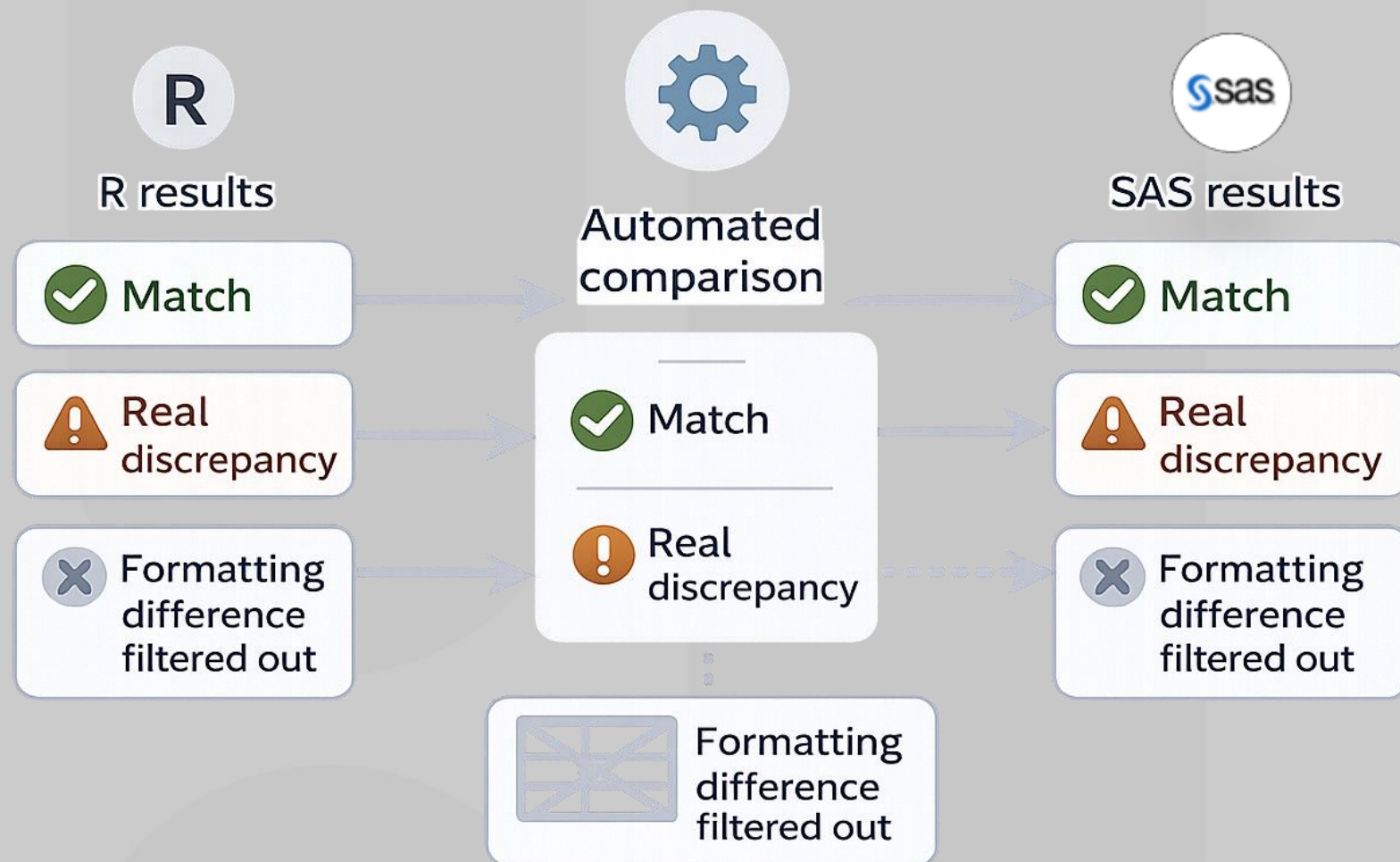


QC findings

Automated discrepancies identified

Differences traced to real derivation choices

No false positives due to formatting



Lessons from the case

**Stable comparison contract
was critical**

**Tolerances required explicit
justification**

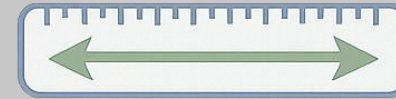
**Independence between
systems was preserved**

**These lessons generalize
beyond TTE—they apply to any
complex, statistics-heavy TLG**

Lessons from the case



Comparison contract



Tolerances



Independence

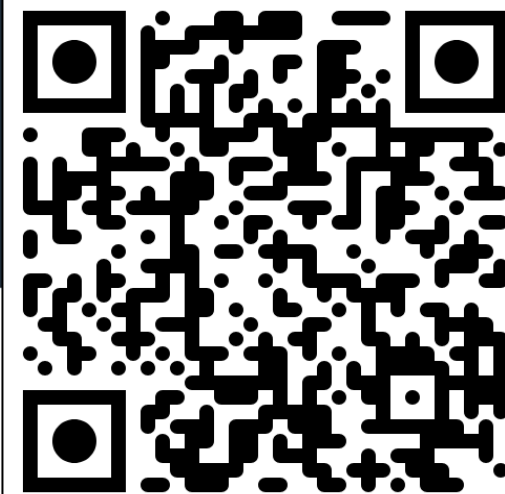
Questions?

Thank you!



Adler Moisés Fallas Rodríguez ✓

Clinical Statistical Programmer | Professor of Medicine |
Microbiologist and Computer Scientist



Let's continue the conversation: