Paper SA01

From Data to Dossier: Lessons from Cross-Company Regulatory Submissions

Gayathri Mahadevan, Daiichi Sankyo, Inc, Durham, USA
Nikita Joseph, AstraZeneca, Toronto, Canada

ABSTRACT

Cross-company pooling and submission presents significant challenges requiring agile strategy adaptation, unique problem-solving skills, and predictive insight. Drawing on personal experiences collaborating in cross-company projects this paper will share practical lessons to drive a successful regulatory submission for products co-developed by pharma companies. We will examine complexities around such pooling strategies, and how shifting company-driven priorities impact programming deliverables. We will explore data challenges caused by differing company standards through examples from adverse events, exposure, pharmacokinetics, and immunogenicity data. Particularly, challenges around data integration and alignment, standard adherence, and maintaining analysis consistency will be discussed. We will then shift our focus to programming review of submission documents like pooling strategy, SAP, CSR, and SmPC. Building on these experiences, we will demonstrate how early alignment, proactive risk monitoring, and systematic review processes will improve efficiency and quality. Attendees will learn strategies to drive successful submissions in the collaborative drug development space.

1. INTRODUCTION

Integrated safety analyses play a central role in regulatory submissions by combining safety information across multiple clinical studies into a unified narrative. When these studies span **two independent co-sponsors**, the integration effort becomes considerably more complex. Each organization brings its own standards, data structures, operational workflows, and regulatory expectations. While the therapeutic context may be aligned, the underlying assumptions governing data collection, derivations, and analysis often differ substantially.

In this submission, the Integrated Safety Summary (ISS) combined one pivotal study with three co-sponsor studies. These data sources varied in CRF design, SDTM/ADaM implementation, timing of data availability, and analytical philosophies regarding TEAE windows, exposure adjustments, PK sampling, ADA interpretation, and adjudication processes. As a result, the pooled analysis required both technical harmonization and cross-functional collaboration to ensure that the final outputs presented a consistent and scientifically defensible safety profile.

The objective of this paper is to provide practical insights into how teams can navigate such collaborative environments. We describe the planning foundations needed to establish a stable pooling strategy, the project management considerations that shaped timelines and responsibilities, the technical challenges encountered across key data domains, and the lessons and quality practices that supported efficient delivery. The aim is not only to document differences, but to offer a framework for managing future co-sponsored submissions where integrating diverse data sources is essential.

2. COMPLEXITIES OF POOLING CROSS-COMPANY STUDIES

2.1 POOLING STRATEGY

Establishing a clear and stable pooling strategy is essential in any ISS, but it becomes particularly critical when the contributing studies originate from two co-sponsors with different priorities and regulatory timelines. In this environment, the strategy must account not only for what studies will be pooled, but also *when* each dataset becomes available, how the definition of the safety population may evolve, and how regulatory interactions influence the scope of analysis.

In our submission, the pooling plan was shaped by several factors. Each sponsor was managing its own submission activities, leading to different expectations for when interim or final data would be ready. Differences in labeling negotiations, safety concerns under review, and evolving guidance from regulatory agencies also influenced which analyses were necessary for the pooled safety narrative. Programming teams therefore needed to monitor ongoing cross-company discussions closely to anticipate changes that could alter shells, SAP language, and derivations.

A planning window of **12–16 weeks before clinical data lock** proved necessary to establish a stable foundation for pooling. This period allowed for updating specifications, harmonizing shells across sponsors, reconciling derivations, and conducting an initial dry run. Because pooling strategy changes can cascade into multiple areas—population flags, analysis windows, AESI definitions, subgroup breakdowns—programming leads played an active role in providing feedback on feasibility, risks, and operational implications.

A recurring theme was the need for **early visibility and proactive communication**. As priorities shifted on either side, programming leads were often the first to identify how these changes would affect timelines or increase development effort. Establishing open channels across regulatory, statistics, clinical, and safety teams helped ensure that impacts were understood and incorporated into planning cycles.

2.2 PROJECT MANAGEMENT

Managing an ISS across two co-sponsors requires close coordination, clear responsibilities, and continuous communication. Unlike single-sponsor submission, ownership of data, derivations, and timelines was distributed across multiple internal teams and external vendors. Early mapping of this ecosystem was essential to prevent delays and misrouted questions.

A central project management activity was establishing **who owns what**. The pivotal study involved internal programmers, statistical leads, and vendor partners, while the co-sponsor operated separate programming teams for ISS/ISI, PK, and ADA domains. Clarifying these roles early helped avoid ambiguity during technical issue resolution, particularly for data domains requiring repeated clarification (e.g., PK sampling, ADA assay interpretation, ILD adjudication logic).

Kick-off meetings between sponsors set the foundation for collaboration. These sessions aligned expectations on data transfer timing, documentation requirements, deliverable sequencing, and review cycles. They also established escalation paths, which were critical in a dual-company model where decisions often required approval or input from multiple functions.

Throughout the project, **cross-functional transparency** proved vital. Updates in labeling, RMP strategy, statistical considerations, or clinical interpretation often trigger modifications to shells, SAP statements, or derivations. Regular engagement between programming, statistics, clinical, safety, and regulatory teams ensured decisions were made with shared context and were implemented consistently.

In short, effective project management acted as the coordination framework that allowed the technical integration work to proceed smoothly. Clear roles, structured communication, and proactive risk management ensured that both sponsors could progress toward a unified pooled submission despite differing internal processes and priorities.

2.3 TECHNICAL COMPLEXITIES AND ISSUE RESOLUTION

Integrating safety data across two co-sponsors revealed substantial differences in how studies were designed, conducted, and analyzed. Many challenges originated from variations in raw data collection, while others stemmed from differences in analysis, company standards, and process workflows. Harmonizing these elements requires not only technical adjustments but also deep engagement with statisticians, safety scientists, regulatory strategists, and data management teams to ensure alignment with clinical interpretation and submission expectations.

2.3.1 RAW DATA COLLECTION DIFFERENCE

ADVERSE EVENTS

The largest discrepancy between sponsors was in AE collection. The co-sponsor's studies logged every shift in AE severity and seriousness as a separate record, with continuous dates even when no clinical break occurred. For example, Grade 1 nausea escalating to Grade 3 and later improving to Grade 2 would appear as three AE rows. In contrast, the pivotal study captured severity/seriousness changes within a single AE record and created new rows only if the AE ended and restarted at least a day later, resulting in a more collapsed structure.

Directly pooling these structures would inflate AE counts in the co-sponsor data and distort worst-grade assessments, while the pivotal structure risked missing clinically meaningful severity transitions. For the ISS, we reconstructed AE episodes in the co-sponsor studies by concatenating sequential rows representing the same clinical event based on term, timing, and severity pattern. Only after this reconstruction could worst-grade, duration, and AESI flags be applied consistently.

This work also revealed differences in relatedness rules: the pivotal study defaulted missing AEREL to "RELATED," while the co-sponsor derived relatedness based on timing to treatment start. These were harmonized to avoid inflating ADR counts in one dataset, with the ISS adopting a unified rule and documenting exceptions for transparency.

EXPOSURE

Exposure data differed not just in format but also in conceptual representation. The co-sponsor captured infusion interruption explicitly via a dedicated CRF question asking whether infusion was interrupted. Meanwhile, the pivotal study encoded interruptions indirectly as drug interruptions within the EX domain. In addition, the two organizations applied different code lists for dose delay, dose reduction, and drug withdrawal. Some categories existed only in one sponsor's CRF design, while in the other they were combined or mapped differently.

The differing exposure modification categorizations had clear analytical consequences: the co-sponsor's structure distinguished dose delay from drug interruption, while the pivotal study combined them into one code. This meant subjects with dose delay in the pivotal study could not be directly compared to those labeled "dose delay" in the co-sponsor's studies. Harmonization required a defined mapping strategy in which categories unique to one sponsor were marked as "not applicable" in the other, and footnotes in exposure modification tables were necessary to prevent misinterpretation of numerical differences resulting from these structural discrepancies.

BASELINE BRAIN METASTASIS

Baseline brain metastasis was an important subgroup criterion for solid tumor studies involved in the pooling. The co-sponsor study used BICR baseline evaluation to derive the subgroups. Whereas the pivotal used investigator data logged into non RECIST disease characteristics form to derive the subgroup.

There were several outputs at subgroup level which relied on the subgroup. As this was a gap at data collection level and could be corrected post CDL, derivations were kept separate. Process alignment across the two companies was noted for future studies, especially for critical subgroups.

PHARMACOKINETICS (PK): Sampling Schedules

Pharmacokinetic harmonization was needed because the two sponsors employed fundamentally different sampling schedules. The co-sponsor collected a rich series of timepoints after infusion - including 30 minutes, 1 hour, 3 hours, 5 hours, 7 hours, 24 hours, and beyond. In contrast, the pivotal study collected only pre-dose and one post-dose sample (nominally one hour after infusion).

Moreover, the ISS's analytical framework required analyzing only the pre-dose and post-dose (end of infusion) record for pooled PK outputs, which was straightforward for the pivotal study but required careful selection from multiple potential timepoints in the co-sponsor's datasets.

The selection of EOI samples for pooled outputs required clinical consultation to determine whether the earliest valid post-dose sample or the sample corresponding to the 1-hour nominal window should be used. Ultimately, the team agreed that for comparability, the earliest valid post-dose sample would be used, with clear documentation in the ADRG regarding differences in data richness.

ADA (IMMUNOGENICITY): Assay differences and Titer interpretations

Immunogenicity harmonization required reconciling the co-sponsors use of two assay formats, one for frozen liquid DP with an Lower limits of quantification (LLOQ) of 90, and one for lyophilized DP with an LLOQ of 30 with no qualitative result available in the ADaM (ADIS) to interpret the titer. This led to the co-sponsor pooled ADaM having analysis values such as "<30", "<90" with no qualitative result to interpret the result.

Developing a harmonized pooled analysis required obtaining qualitative results from the individual co-sponsor studies to derive the titer values needed for pooled analysis.

Snippet of the Programming specification for reference: *"Merge AVALC of PARAMCD = <qualitative parameter> to respective AVALC of PARAMCD = <quantitative parameter> on AVISIT, if AVALC = "POSITIVE" and AVAL contains inexact values like "> x", "≥ x", "< x" or "≤ x", x is taken for AVAL.*

2.3.2 COMPANY LEVEL ANALYSIS DIFFERENCE

While collection differences typically require reconstruction, company-level analysis differences require conceptual alignment. These distinctions arise when two organizations apply different philosophies to what constitutes a derivation, analysis window, or population definition. Harmonizing these definitions required careful statistical deliberation and repeated cross-company collaboration.

ADVERSE EVENTS: Treatment-emergent and Infusion Related Reactions

TEAE definitions illustrate these discrepancies. The co-sponsor defined TEAEs as events occurring or worsening within 35 days after the last dose, while the pivotal study truncated the window at the earlier of 35 days or initiation of new anticancer therapy. This affected the number of TEAEs attributed to study treatment, particularly when subjects quickly moved to next line therapy. Rather than imposing one definition on both datasets - which could misrepresent either sponsor's regulatory position - the ISS retained a harmonized TEAE definition in outputs: "A TEAE is defined as an adverse event with a start or worsening date on or after the start of study treatment until xx days since the last dose (co-sponsor pool), or earliest of xx days since last dose and start of 1st subsequent therapy (pivotal)."

Infusion-related reaction (IRR) summaries also posed challenges. The co-sponsor concatenated consecutive IRRs across infusions for number of events per subject but did not do so for subject-wise counts, which was inconsistent. For the ISS, we applied concatenation uniformly at the pool level, leading to differences between ISS and co-sponsor counts. IRR timing definitions also differed: within a day for the co-sponsor versus within 24 hours for the pivotal study.

ILD ADJUDICATION

Interstitial lung disease (ILD) adjudication illustrated not only the technical complexities of pooling but also the operational nuances of vendor managed processes. The co-sponsor maintained a sophisticated adjudication pipeline that differentiated cases based on whether they could be merged back into the EDC, whether the adjudicated preferred term aligned with investigator reported PTs, and whether ILD-related death was confirmed by adjudicators. This information was spread across multiple SDTM datasets; AE supplemental qualifiers (SUPPAE) for cases that were mergeable back to EDC, and FA for cases not mergeable back to EDC. The pivotal study, however, stored only adjudicated ILD cases that could be linked directly to EDC records; no additional adjudication metadata was preserved.

In practical terms, this meant that the ISS could not rely solely on AE or adjudication datasets from one sponsor to infer pooling rules applicable to both. To align on this, the co-sponsors approach of creating, a new dataset, ADXI, was extended to the pivotal study. This required merging ILD-related data from AE, SUPPAE, SUPPDD, FA, and RAW adjudication vendor files while reproducing the co-sponsor's adjudication logic. The "existing in EDC" flag was constructed for the pivotal study by identifying cases traceable back to EDC sources. This enabled consistent presentation of ILD listings and summary tables, preventing discrepancies that would otherwise be questioned by regulatory reviewers.

LABORATORY: Harmonization and Renal Impairment

Both sponsors used Cockcroft–Gault to calculate creatinine clearance, but thresholds for "severe renal impairment" differed: the pivotal study defined <=30 mL/min as severe, while the co-sponsor used 15 - 30 mL/min. Though numerically small, this caused noticeable discrepancies in baseline renal impairment and renal shift tables. Resolving it required more than applying one rule universally; a harmonized definition had to reflect clinical reasoning, regulatory precedent, and the pivotal study's CSR standards. After cross-functional review, a unified categorization was adopted for the pooled analysis, with observed differences between pooled and individual study outputs documented.

PHARMACOKINETICS (PK): Visit Windowing

The pivotal study did not derive windowing flags for timing validation. This difference had implications for identifying pre- and post-dose windows in the pooled analysis. To address these inconsistencies, ISS-level windowing rules were developed for the pivotal study, while the co-sponsor's existing windowing flags were retained. Some pivotal study samples occurred exactly at infusion end, complicating classification, while missing exposure timestamps further limited precise determination of dosing windows.

ADA: Analysis Derivation

The pivotal study derived ADA-related categories, such as treatment-emergent ADA, treatment-boosted ADA, and neutralizing antibody status - as CRIT flags within a BDS structured ADaM dataset (ADMI), while the co-sponsor maintained subject-level ADA categories in a custom ADSL dataset (ADISL). For the pooled ISS, a unified ADISL dataset was created following the co-sponsor's approach, aligning ADA categories across both companies by mapping pivotal CRIT flags to the corresponding subject-level flags. This also involved supplementing missing ADA categories from the pivotal datasets by deriving additional flags, ensuring all regulatory expected ADA categories were included.

REGION CLASSIFICATION

Countries were classified differently between the pivotal study and co-sponsor pool. The co-sponsor summarized regions as “Japan”, “USA”, “Western Europe” and “Rest of World”, whereas the pivotal study categorized regions into “US, Canada, Europe” and “RoW”. To address this, we aligned the co-sponsor regions to match the pivotal study at the ISS pool level and categorized them based on their individual countries.

3. LESSONS LEARNED

Start with early programming involvement.
Programming leads should participate in regulatory strategy and submission planning discussions from the outset to anticipate changes in pooling strategy, analysis scope, and derivation needs.

Promote transparency across functions and sponsors.
Sharing the broader submission context helps teams understand interdependencies, enabling more efficient planning and reducing delays caused by misaligned expectations.

Strengthen cross-functional collaboration.
Conduct regular risk-monitoring meetings with the GST; complex or unresolved issues should be tracked in a centralized Risk-Decision Log to ensure visibility and timely escalation.

Ensure the right representation in document reviews.
Submission documents, including SAPs, shells, labeling sections, RMP components, and medical writing deliverables must be reviewed by all relevant stakeholders to ensure alignment in terminology, definitions, and analytical expectations.

Lead technically through informed SAP review.
Programming leads should check for alignment with the pivotal study, identify analysis differences between co-sponsors, highlight missing or unclear derivations, and ensure footnotes accurately describe necessary nuances.

Foster constructive partnerships with co-sponsor teams.
Building strong working relationships helps facilitate clarification of company-specific processes, discuss interpretation differences, and resolve issues that span multiple functional areas.

Communicate risks clearly and proactively.
Programming leads should promptly flag risks such as delayed data transfers, unclear specifications, or evolving strategy expectations, as these can significantly impact timelines.

Plan ahead for strategy finalization.
Pooling strategy should be confirmed 12 to 16 weeks prior to CDL to allow sufficient time for up-versioning specifications, conducting dry runs, and incorporating cross-functional feedback.

Clarify project management and team structure early.
Mapping the outsourcing and vendor model for each sponsor enables efficient routing of questions and avoids bottlenecks caused by uncertainty around ownership.

Establish structured data transfer processes.
Cross-company data sharing should follow a formal governance model, including leadership approval and well-documented transfer pathways to support future traceability.

4. QA/QC RECOMMENDATIONS

Document all cross-company derivation differences in the pooling strategy.
This ensures technical transparency and provides reviewers with clear justification for any deviations between studies.

Use an independent third-level QC review.
A reviewer outside the core programming team can identify gaps or misalignments at the SAP or shell level before programming begins.

Incorporate study-specific columns in pooled outputs.
This supports rapid comparison and reconciliation of subject counts, derivations, and safety populations across studies, especially when pooling concurrent pivotal trials.

Validate company-specific derivations directly with SMEs.
Any complex or proprietary derivations (e.g., ADR, ADA categories, ILD adjudication) should be confirmed and documented to ensure consistency.

Cross-check safety signals programmatically across sponsors.
Ensuring that both sponsors detect and incorporate the same safety signals avoids discrepancies in ADR identification or interpretation.

Run programmatic checks for completeness and missingness.
These checks help verify that all required subsets (e.g., populations, subgroups, PK samples) were transferred correctly and included in pooled analyses.

Compare pooled datasets across versions.
Version-to-version comparisons help identify unexpected changes or shifts caused by updates to pooling strategy, source data corrections, or specification revisions.

Identify and investigate outliers systematically.
Automated flags for implausible values (e.g., AGE <18, RDI >200%) ensure early detection of data quality issues or derivation errors.

Ensure pooled population counts align with individual studies.
Deviations should be documented, and justification provided when following a partner's process leads to intentional differences.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please feel free to contact us below,

Gayathri Mahadeva, Daiichi Sankhyo, USA (gayathri.mahadevan@daiichisankyo.com)

Nikita Joseph, AstraZeneca, Canada (nikita.joseph@astrazeneca.com)

DISCLAIMER

Contributions by the authors and contributors are solely their own and are not intended to express the views of their organizations.