

Presenter Bio



Gayathri is a multilingual statistical programming expert with 15 years of industry experience. She is a recognized subject-matter expert in global regulatory submissions (FDA, PMDA, EMA) and integrated summaries, with a proven track record of delivering high-quality, compliant, and efficient study packages in high-stakes regulatory environments. She has successfully led multiple regulatory submissions, often in cross-company collaborative settings, with a focus on standards alignment and harmonization. Gayathri has been an active contributor to the PHUSE community—serving as a working group member since 2022 and joining the PHUSE Connect committee in 2023. Beyond her professional work, Gayathri is an avid hiker who enjoys exploring U.S. National Parks and heritage sites.



Nikita Joseph is a Senior Statistical Programmer at AstraZeneca, bringing more than eight years of experience in the pharmaceutical industry. She is currently based out of Toronto, Canada. Her career spans both CROs and sponsor companies, including AstraZeneca and Eli Lilly. Nikita specializes in statistical programming for clinical trials, supporting regulatory submissions and data analysis across multiple therapeutic areas. She has previously attended PHUSE SDEs and values the opportunity for knowledge sharing and professional networking they provide. This year marks a special milestone as it is the first time she will attend PHUSE US Connect as a presenter, and she looks forward to contributing to the vibrant programming community. Outside of work, Nikita is passionate about baking—an interest that becomes especially lively during the festive Christmas season. She is also an avid hiker and enthusiastic reader, currently re-reading Harper Lee’s “Go Set a Watchman.”



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

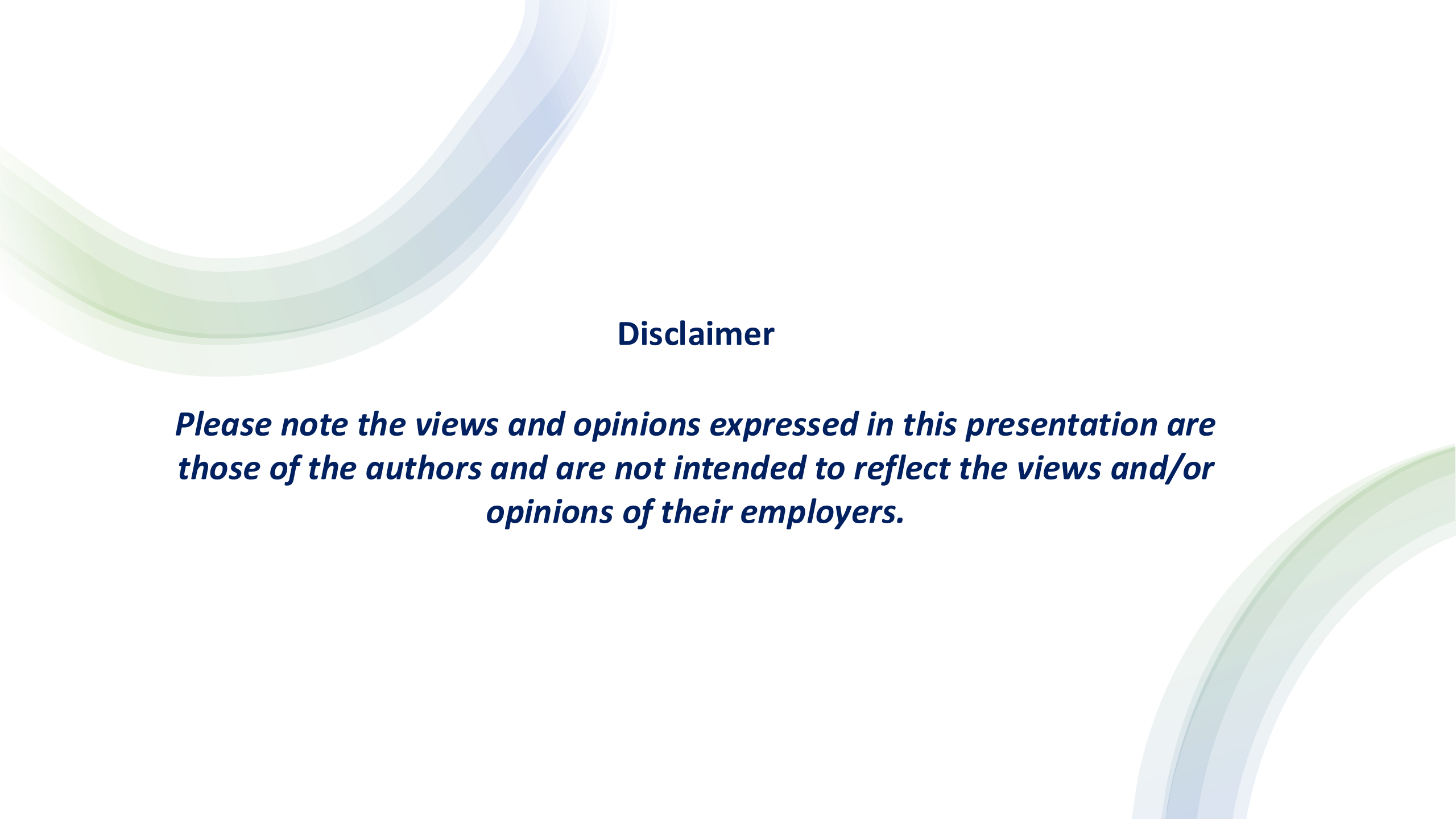
Presented by,

Gayathri Mahadevan, AD Statistical Programming, Daiichi Sankyo, Inc

Nikita Joseph, Senior Statistical Programmer, AstraZeneca

PHUSE US Connect 2026

22nd March – 26th March 2026



Disclaimer

Please note the views and opinions expressed in this presentation are those of the authors and are not intended to reflect the views and/or opinions of their employers.

AGENDA

- Overview: ISS
- Complexities of safety pooling in cross-company studies
 - Pooling strategy
 - Project management
 - Technical complexities and issue resolution
 - Raw data collection difference
 - Company level analysis difference
- Lessons learned
- QA/ QC Recommendations
- Questions



ISS OVERVIEW

- Safety pooling is required for NDA for FDA, expected for MAA for EMA and NDA for PMDA.
- Safety data pooling refers to combining safety information from different (but generally similar) studies to provide a larger sample size for data analysis.
- Safety data pooling supports the labelling process for ADRs by helping identify ADRs. Safety pooling may also help in addressing other aspects such as warnings, precautions, SmPC report and RMP report.
- Pooling helps achieve larger sample size to detect safety signals (common/rare events, long term safety, sub-population effects) and compare differences in rates.

The background features two large, curved, overlapping bands. One band is a light blue color, and the other is a light green color. They are positioned in the top-left and bottom-right corners, creating a sense of movement and depth.

COMPLEXITIES: OVERVIEW



Strategy

Parallel submissions and conflicting priorities

Ever-evolving submission landscape

Pool and study data availability

Pooling strategy dependency on company strategy and priorities



Submission Documents

SAP finalization delay due to pooling strategy delay and cross company alignment delay

TLF Shell delay due to scope finalization delay across two leadership/ regulatory teams

Longer review process and timeline due to larger stake holder management

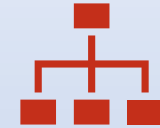


Technical Complexities

Different company level raw data collection and handling make standardization tricky

Company specific analysis, criterion and algorithms, often creating mismatch and gaps at output level

Different company level processes and SOPs



Project Management

Multiple vendor and outsourcing models across companies means more process and SOPs to navigate

Need to navigate between more than one team and point of contact for issue resolution and escalation

Data sharing dependency for trials conducted by partner, for pooling

TECHNICAL COMPLEXITIES



CASE SCENARIOS...



Example 1: “Company X collected grade changes/ severity in AE as a new log line with AE dates flowing in a continuous manner whereas Company Y collected grade changes in a collapsed manner.”



Category: Raw data collection difference.



Severity: High. Affects TEAE derivation, subject level counts, outcome within worst grade, duration, number of event count per subject, event rate within cycle



Resolution: Non collapsed AEs coming from company X were collapsed as per company SOPs. Any difference in handling of worst outcomes or worst grade due to collapsing against company Y was detailed in footnotes.

CASE SCENARIOS...

- Example 2: *“Company X captured additional code list value of dose delay for action taken with drug in AE form. Company Y uses drug interrupted and dose delay interchangeably”*
- **Category:** Raw data collection difference.
- **Severity:** Medium. Specific output summarizing dose delays caused by AEs proposed by company X team were impacted. No 1:1 categorization between the companies for action taken with a drug.
- **Resolution:** Outputs not programmable were removed from ISS formal outputs. Company X categories were combined at higher level to make mapping 1:1.

CASE SCENARIOS...



Example 3: “Company X used different vendor for ADA sampling and analysis which used frozen liquid and lyophilized sample. Company Y only used frozen liquid sample for the assay. This resulted in different LLOQ and derivation for ADA positive subsets”



Category: Raw data collection difference.



Severity: High. There were several outputs which were based on quantitative results being $\geq$ or $<$ LLOQ values. There were also qualitative characterization described in pooling SAP which was based on numerical results, which could not be implemented due to difference in numerical LLOQ.



Resolution: Difference in LLOQ were noted in SAP and additional text were added to accommodate differing LLOQ values for qualitative groups.

CASE SCENARIOS...

- Example 8: *“Company X had different SOPs and procedures around PK sample collection and timepoints vs Company Y. Therefore similar timepoints were not available across studies to pool and summarize for ISI”*
- **Category:** Raw data collection difference.
- **Severity:** High. There were several outputs which were based on pooled PK analysis.
- **Resolution:** To align the results between the studies from both companies, analysis timepoints had to be created based on SAP assumptions. For example, multiple latest of 30minutes and 1 hour post dose timepoints were taken from Company Y studies and re labelled as post infusion, as standard infusion time is 60min.

CASE SCENARIOS...



Example 5: “Company X implemented a detailed adj. process for an AESI (keeping all transfer cases/ identifying events mergeable and not to EDC / sensitivity analysis etc.) Company Y only analyzed those cases which were adj. & which was flagged in EDC.



Category: Company level analysis difference.



Severity: High. The AESI in question was a critical safety data point and adj. result and its sensitivity analysis were also critical, so was the analysis of adjudicated result outside of EDC verdict. Due to the simple process within Company Y, gaps existed in analysis



Resolution: To align on this, Company X approach of creating, a new dataset—ADXl—was extended to the pivotal study. This required merging AESI related data from AE, SUPPAE, SUPPDD, FA, and **RAW adjudication vendor files** at pooling level.

CASE SCENARIOS...

- Example 6: *“Baseline brain metastasis was an important subgroup criteria for solid tumor studies involved in the pooling. Company X used BICR baseline evaluation to derive the subgroups. Company Y used investigator data logged into non RECIST disease characteristics form to derive the subgroup ”*
- **Category:** Raw data collection difference.
- **Severity:** High. There were several outputs at subgroup level which relied on the subgroup.
- **Resolution:** As this was a gap at data collection level and could not be corrected post CDL, derivations were kept separate. Process alignment across the two companies were noted for future studies especially for critical subgroups.

CASE SCENARIOS...



Example 7: “*Company X had company specific procedure to categorize countries into region. Company Y had different procedure and different categorization. REGION was a critical subgroup.*”



Category: Company level analysis difference.



Severity: High. There were several outputs which were based on REGION subgroup.



Resolution: To align the results with pivotal and historical regional representation, Company X derivation were amended to match Company Y. No further footnotes or SAP text were added.

CASE SCENARIOS...

- Example 4: *“Company X applied windowing logic to derive acceptance flag at timepoint level for PK samples and company Y did not”*
- **Category:** Company level analysis difference.
- **Severity:** High. There were several outputs which were based on pooled PK analysis.
- **Resolution:** To align the results between the studies from both companies, team tried to apply windowing logic to company Y samples. However, since analysis timepoint were re-derived, the windowing logic had to be applied at original timepoints and only those satisfying the window were then selected for deriving analysis timepoint. This also meant acceptance flags did not match pivotal and PK summary table numbers also mismatched with pivotal. This was explained in DEFINE.

LESSONS LEARNED



LESSONS LEARNED

- **Begin right:** Programming lead should get involved in regulatory and strategy discussions from the outset.
- **Foster Collaboration:** Recognize that strategy updates can impact outputs, tables of contents, and key variable derivations. Build strong relationships with all team members.
- **Communicate Risks:** Proactively flag and discuss potential risks as soon as they arise.
- **Plan Ahead:** Finalize pooling strategies 12–16 weeks before the Clinical Data Lock (CDL).
- **Project Management:**
 - Map out team structures and outsourcing models on both sides to direct queries efficiently.
 - Hold kick-off meetings with both internal and external teams for alignment.
 - Set and meet clear milestones.
- **Data Transfer:**
 - Formalize cross-company data sharing.
 - Automate global package transfers post-agency submission, with leadership approval.

LESSONS LEARNED

- **Promote Transparency:** Share the broader submission context, driving efficiency and engagement.
- **Cross functional collaboration:** Hold regular risk meetings with GST; log unresolved complex issues into Risk decision Log.
- **CST Alignment:** Use submission Strategy meetings to get cross-team milestone alignment (PS, Clinical, Regulatory, Stats & Prog, MW, Labelling, RMP, PM, QCP).
- **Right representation:** Ensure comprehensive stakeholder review of submission documents and shells (labeling, RMP, PS, QCP, MW, eSUB).
- **Lead Technically:** The programming lead should prioritize SAP review focusing on:
 - Alignment with pivotal, identifying analysis differences between partners.
 - Identifying any potential gaps in SAP against previous company specific pooling documents.
 - Following up at the GST level for open questions, especially those around company specific process and handling.
 - Reviewing any gaps around the necessary footnotes to highlight key analysis differences between the two companies.



QA/ QC RECOMMENDATIONS

QA/ QC RECOMMENDATIONS

Pooling strategy document should clearly document derivation differences between companies.

Use a third-level QC by someone outside the core team to catch gaps at SAP/ Shell level.

Add study-specific columns to outputs for comparison and consistency, especially when pooling concurrent pivotal across partners.

Complex company specific derivations and their interpretation should be checked and agreed by partner SME via formal emails.

Programmatically compare critical safety signals for a compound between partners. This ensures all signals coming from one company is used for partner pooling and vice versa.

Programmatic checks should be included for flagging large or completely missing data (at the population level, subgroup level, and stratification level) across all studies, to ensure data subset was transferred.

Programmatic checks against the previous version of pooled data should also be included to identify any critical changes or unexplained large shift in results.

Programmatic checks should also be done to check for outlier data (e.g.: aage < 18, RDI > 200%).

Pooled dataset should replicate individual study population numbers where possible and should be detailed if following partner process explains a deviation.

Contact us...

- Gayathri Mahadevan

Ass. Director, Statistical Programming

Gayathri.Mahadevan@daiichisankyo.com

- Nikita Joseph

Senior Statistical Programmer

nikita.joseph@astrazeneca.com

THANK YOU