Paper AD06

Harnessing the Power of Safety Database

Author - Ravichandra Hugar, AstraZeneca India Private Limited, Bangalore, India

Co-Authors - Aparna Sarju, Vinodkumar Ramagiri, Pravallika Kesari, AstraZeneca India Private Limited, Bangalore, India

ABSTRACT

In the realm of Data Analytics and Statistical Programming, a significant challenge for analytics team involves establishing a real-time, project-level safety database. It is critical to review safety of the drug and there is not established process to do it across project(s). This critical resource would serve to bolster safety signal management and facilitate the ongoing evaluation of a product's benefit-risk profile. Imagine a meticulously structured database designed to house analysis datasets, incorporating intricate derivations, advanced methodologies, sophisticated data handling techniques, and robust imputation algorithms, all alongside the precise analysis-level definitions inherent in a Statistical Analysis Plan (SAP). Furthermore, this envisioned database would seamlessly integrate pooled design requirements, encompassing various study indications, dosing regimens, developmental phases, platform studies across subjects.

1. INTRODUCTION TO SAFETY DATABASE

Across a compound’s development, particularly when targeting multiple indications or spanning diverse therapeutic areas, and even within larger programs with fewer indications planning continuous, integrated safety reviews and aggregate analyses can be complex.

Maintaining consistent definitions and analytical methods across studies is critical. Unblinded safety reviews can be harmonized through a compound level Data Monitoring Committee, while post study consistency can be guided by a well-designed Safety Analysis Plan. That said, prioritizing ongoing safety review during clinical development is critical to fully characterize the compound’s evolving benefit–risk profile.

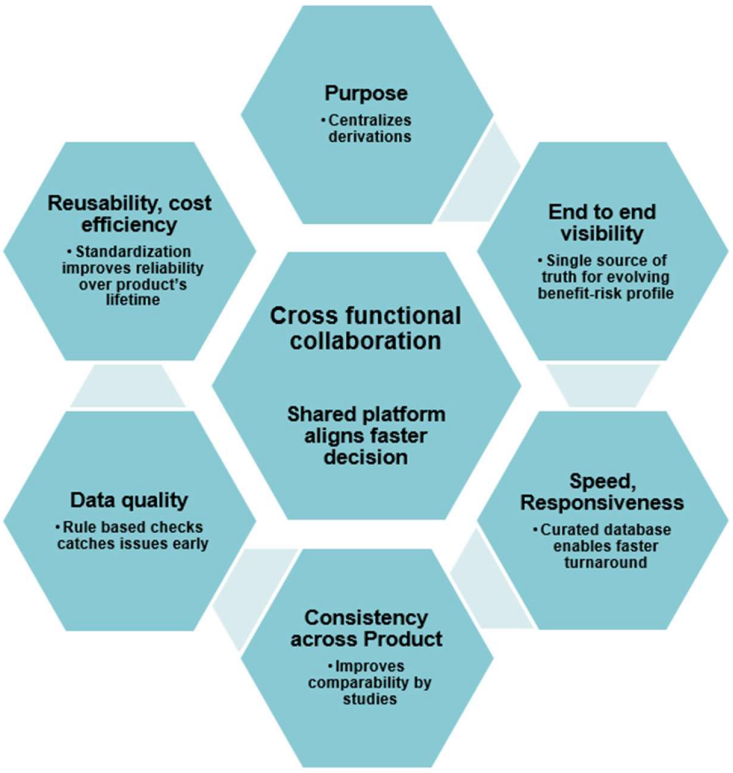
Such a comprehensive safety database, readily accessible in real-time and updated periodically, would empower stakeholders to continuously monitor and visualize emerging safety data, assessing the effectiveness of Toxicity Management Guidelines (TMGs) to support routine safety surveillance. It would also efficiently address the potential analysis needs for required periodical and annual safety reports, including Investigation Brochures, while constantly accommodating exploratory analyses and requests from regulatory bodies. The design and construction of such a well-structured database necessitate the collaboration of diverse teams and a strategic process where the expertise and actionable insights from key stakeholders are paramount in guiding decisions to ensure the successful achievement of the end goal.

This paper aims to explain the process and strategies involved in building a robust safety database. It will also provide technical insights into developing real-time data flow from source systems to the Analysis Data Model (ADaM) datasets and other possible sources, adopting innovative methodologies, and utilizing data to analyse safety trends and patterns for more effective decision-making. Additionally, this paper will illustrate how this application/database can deliver significant improvements in operational efficiency over the long-term including working a template to use other sources of data like SDTM etc.

2. What is Safety Database?

Safety database is a pooled clinical safety data from all studies that use the investigate product. Includes Demographics, Adverse Events, Labs, Medical History, Concomitant Medications, Exposure summary etc used for timely Safety surveillance at project level structured, governed data repository that consolidates safety-related information about a medicinal product across sources and time, so teams can monitor, analyse, and report on its benefit risk profile. It’s the backbone for reliable, compliant safety analyses and reporting, enabling timely insight into risks and the evolving benefit risk profile.

3. Purpose of Safety Database



Purpose: Centralizes adverse events, Con Meds, exposures, labs, and case reports to support signal detection, aggregate reporting (e.g., DSUR/PSUR/PBRER/IB/Ad hoc request), inspections, and decision making.

End-to-end visibility: Traditional methods silo safety data by study and system, forcing manual reconciliation. A safety database centralizes clinical trials, labs, and concomitants, giving a single source of truth for an evolving benefit–risk profile.

Speed and responsiveness: Manual extracts and ad hoc analyses slow DSUR/PSUR, and signal evaluation. A curated database enables near real-time refreshes, and faster turnaround for regulatory queries and inspections.

Consistency across studies: Study-specific definitions and code lead to inconsistent outputs. A safety database enforces standardized definitions, terminologies (MedDRA/WHO Drug), and derivations (TEAE, CTCAE grades), improving comparability and auditability.

Data quality at scale: Automated data profiling, unit normalization, dictionary governance, and rule-based checks catch issues early. Manual workflows are error-prone and hard to reproduce.

Reusability and cost efficiency: Rebuilding analyses for every study wastes time. Reusable ADaM-like structures, shared code libraries, and dashboards reduce effort and improve reliability over the product's lifetime.

Cross-functional collaboration: Biostats, Clinical, and Quality need the same facts. A shared platform with role-based access aligns decisions and reduces back-and-forth over competing extracts.

4. Challenges reviewing Safety of a drug profile

Systems though in place but are mostly limited for individual safety case reporting, not for aggregated analysis that includes multiple studies across the product.

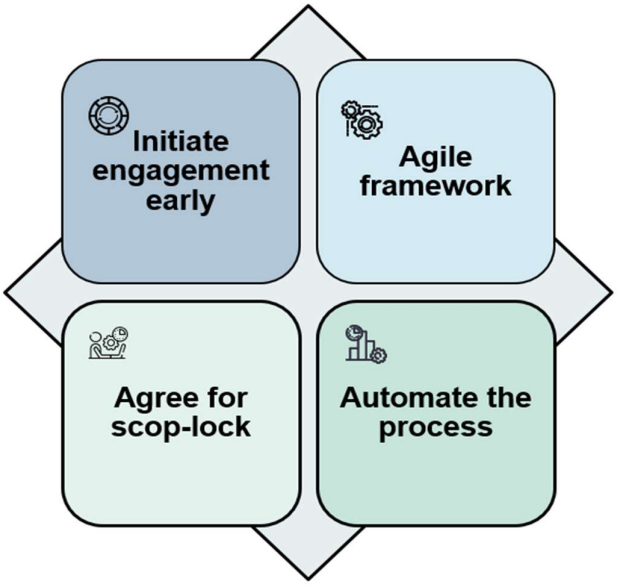
Tools readily available are ideal for study level safety data visualizations, lack of project level information that provides holistic insight into the product.

Clinical Database can only boast study level data, with limited access. Any additional scope or ad hoc requests relevant to FDA/HAQs can demand increased time.

Scope variability driven by evolving outcomes and safety data

Above challenges could delay in assessment of safety of drug profile, eventually leading to risk.

5. Framework for safety database & reporting



Initiate engagement early: Iteratively refine requirements throughout scope definition.

Implement Agile framework: Engage stakeholders regularly to validate priorities and outcomes, Deliver working increments early and often. Adjust scope and plans based on feedback

Agree for scope lock: Define the timeline and so no new features or major changes are added without a controlled change process

Automate the process: Data transfer based on periodical needs, batch run, stabilize the programming workflow and deliver efficiently.

6. Objective of safety database

Build a real time project level safety database that can support safety monitoring

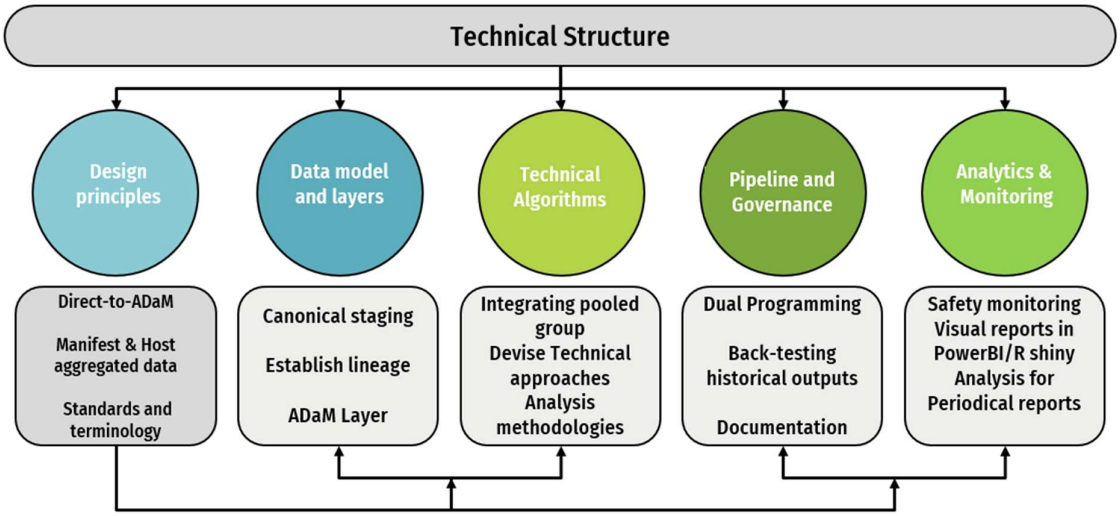
Develop a program for delivery of regular standard safety data reports at defined intervals (weekly, monthly etc) to support routine safety surveillance

Enhance patient safety: Provide timely, reliable safety insights to investigators to support informed clinical decision-making and risk mitigation at sites.

Single source of truth: Centralize study safety data (AEs, exposure, labs, concomitants) into a governed repository with consistent definitions and traceability.

Regulatory readiness: Ensure inspection-ready evidence (lineage, audit trails, version pinning) and support aggregate reporting obligations.

7. Design workflow:



Design workflow of a Safety Database using direct-to-ADaM curation from raw data across multiple studies of the same product.

a) Design principles

Direct to ADaM: Minimize transformations; curate only what's needed for analysis while preserving full traceability to raw. Standards and terminology: Enforce MedDRA for AEs and WHO Drug for concomitants; pin dictionary versions per data cut; use ISO 8601 dates. Reproducibility: Version control specs, code, and configs; generate run manifests for each data snapshot/cut.

b) Data model and layers

Raw layer: Secure, study-partitioned storage of vendor CRFs/exports. Harmonize names, units, and controlled terms into canonical datasets (e.g., CA_AE, CA_EX, CA_CM, CA_LB), plus study metadata (arms, epochs, visit schedule). ADaM layer: Produce ADAE, ADEX, ADCM, ADLB, ADMH with consistent keys (STUDYID, USUBJID, TRT, EPOCH, analysis windows). Aggregated views for dashboards and standard outputs (ISS/ISE tables, DSUR/PSUR summaries).

c) Technical Algorithms

Anchor on first exposure; define wash-in/out; handle partial dates; derive treatment-emergent flags consistently. Exposure: ADEX with dose intensity, interruptions. Labs: ADLB with CTCAE grades, baseline flags, change/shift tables; normalize units and reference ranges. Concomitants: ADCM with standardized dose/route/frequency; ATC class mapping for interaction analyses. Medical history: ADMH for stratification; harmonize terms and systems. Sorting/traceability:

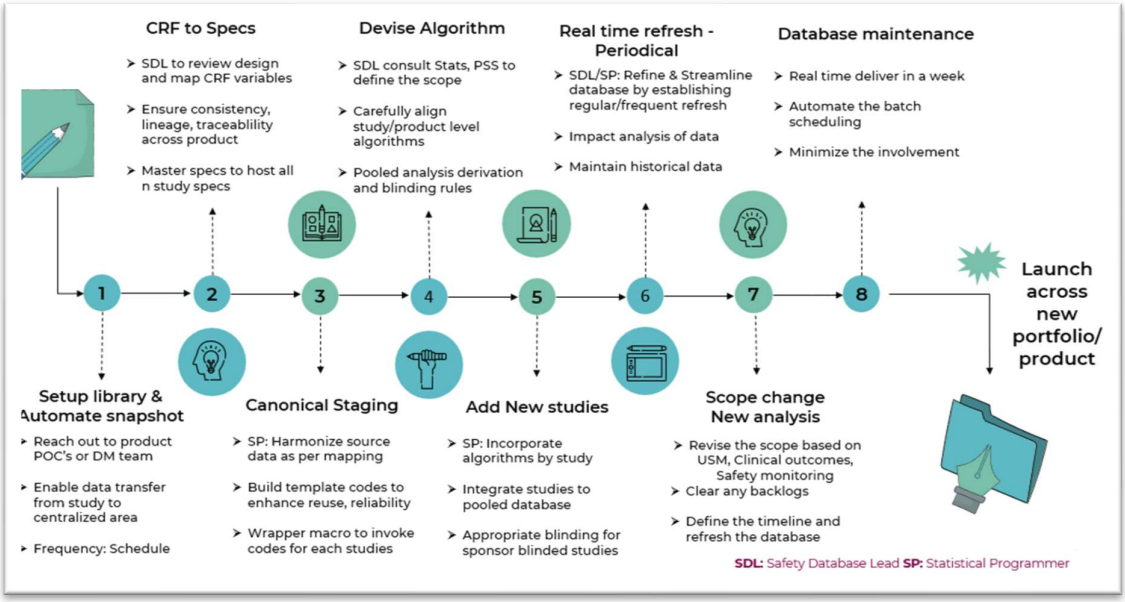
d) Pipelines and governance

Validation: Risk-based testing; dual programming for critical flags, back-testing against historical outputs; data quality dashboards. Change control: Document derivations in Specs; pin dictionary versions; impact reports when versions change., lineage examples, and validation summary. Operating playbook: Onboarding new studies, dictionary updates, and release process.

e) Analytics and monitoring

Dashboards: SOC/PT trends, seriousness/severity, AESI tracking, EAIR, lab shifts; drill-down by study, dose, time. Alerts: Threshold-based for emerging signals; and exposure context for interpretability. Cut management: Routine snapshots with run manifests to ensure reproducibility and inspection readiness.

8. Development and Implementation



9. Data Integrity:

Protecting the integrity of product outcomes is important. Blinding is very essential which strengthens the credibility of the evidence and confidentiality of safety and RECIST/efficacy data.
Let us look at some of the challenges while blinding and the rules that are to be framed.

Challenges: When integrating multiple studies for a drug, obtaining source data can be challenging, following factors include:
➢ For platform studies with similar combination therapies across different indications, access should be endorsed by the relevant product leads before proceeding.
➢ If the study is sponsor-blinded, ensure global safety team does not receive unblinded data. the dataset should be appropriately blinded before sharing with the Safety team.
➢ Late phase study data used for analysis and visualization are appropriately blinded and aligned with the Statistical Integration Plan (SIP).

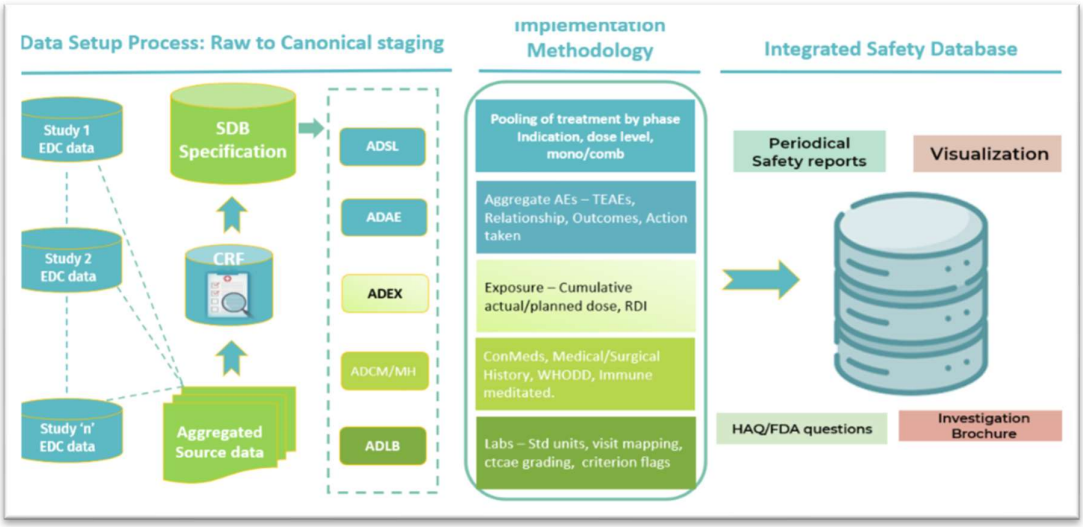
The Safety Database team should continue to oversee data within the product safety database and evaluate whether additional blinding is needed for any new studies.

Blinding Rules: Devise the rules by identifying the studies to be blinded and list down the variables

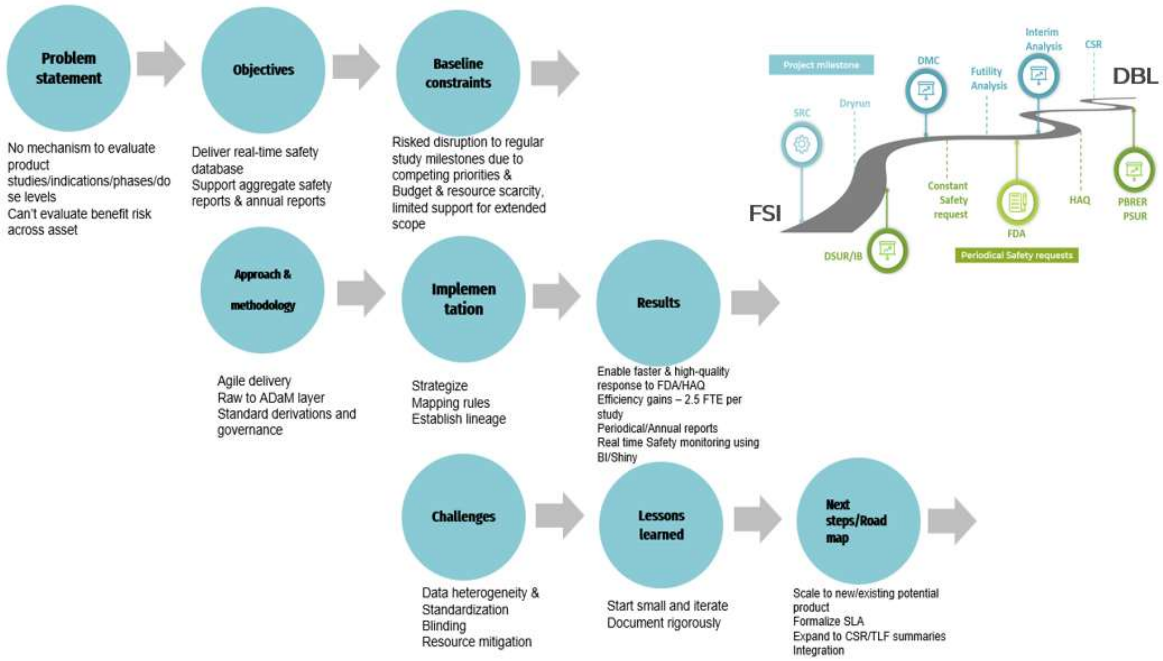
Blinding scope	➢ Prepare the blinding rule for key studies by identifying which variables reveal treatment (e.g., treatment arm, dose level, randomization group, visit/epoch labels). ➢ Mask direct treatment variables or exclude variables that can unblind. ➢ Replace arm-specific names with neutral labels (e.g., "Visit 1/2/3," "Period 1/2")
Exposure and Adverse Events are the Key:	➢ While Adverse event terms and MedDRA coding terms are retained but masking variable that reveal event information like Action taken/relationship to active study drug or combination therapies like AE leading dose reduction/increase/modification/interruption) is utmost important. ➢ AE Severity, seriousness, outcomes, and timing could be included, but ensure no fields reference arm names or investigational product identifiers that differ by arm. ➢ Exclude exposure records with active drug, similarly exposure summary parameters like Intended, total exposure cumulative are to be dropped or masked appropriately etc and drop Lab records.

Governance and access control:	➢ Ensure that a centralized area where blinded datasets are placed and are restricted so that safety team access only to blinded data. ➢ Unblinded data should be fully restricted except for study team members.

10. Process map



11. Case study: Implementing a Safety Database for an Oncology Asset: Direct-to-ADaM Approach



Overview: Build a governed Safety Database that aggregates raw clinical data across multiple studies of a single oncology asset and curates ADaM-ready datasets directly (no SDTM layer). Using an Agile, iterative delivery model, the platform provides near real-time safety analytics (weekly/monthly refreshes), standardized TEAE logic and CTCAE grading, and inspection-ready traceability. This gives us the liberty and privilege to refresh safety data for entire studies of product at any given time with enhanced consistency across studies, and enabled faster, higher-quality responses to FDA and global health authority queries, while reviewing actionable insights through Power BI and R Shiny dashboards.

a) **Problem Statement**: An investigational oncology product typically spans numerous studies across indications, phases, and dose levels. Without a unified mechanism to integrate safety data, teams cannot reliably evaluate the evolving benefit–risk profile across assets, nor deliver real-time safety views. This gap hampers timely responses to FDA and global health authority questions. The lack of integrated, timely safety insights introduced risk to dose strategy optimization and patient selection decisions, with potential downstream impacts on trial outcomes and regulatory interactions.

b) **Objectives**: Deliver near real-time safety analyses on a weekly/monthly cadence, ensuring key safety parameters and derivations are consistently produced (TEAE, CTCAE, exposure-adjusted rates). Support aggregate safety reports and reference documents (PBRER, DSUR, IB, PSUR).

c) **Baseline and Constraints:** Limited resources and competing priorities created bottlenecks and risked disruption to regular study milestones. Budget constraints and limited support for extended scope and standardization. Heterogeneous raw data structures and practices across vendors/studies.

d) **Approach and Methodology: Agile delivery:** Start with a focused MVP scope, then iteratively expand coverage and capabilities. Direct-to-ADaM curation: Map raw CRF/exports directly to ADaM-ready structures, removing the SDTM layer to reduce latency and rework. Maintain lineage and run manifests for traceability. Tools and frameworks: reusable macros for derivations, and validated routines for data quality and blinding where required.

e) **Approach and Methodology: Agile delivery:** Start with a focused MVP scope, then iteratively expand coverage and capabilities. Direct-to-ADaM curation: Map raw CRF/exports directly to ADaM-ready structures, removing the SDTM layer to reduce latency and rework. Maintain lineage and run manifests for traceability. Tools and frameworks: reusable macros for derivations, and validated routines for data quality and blinding where required.

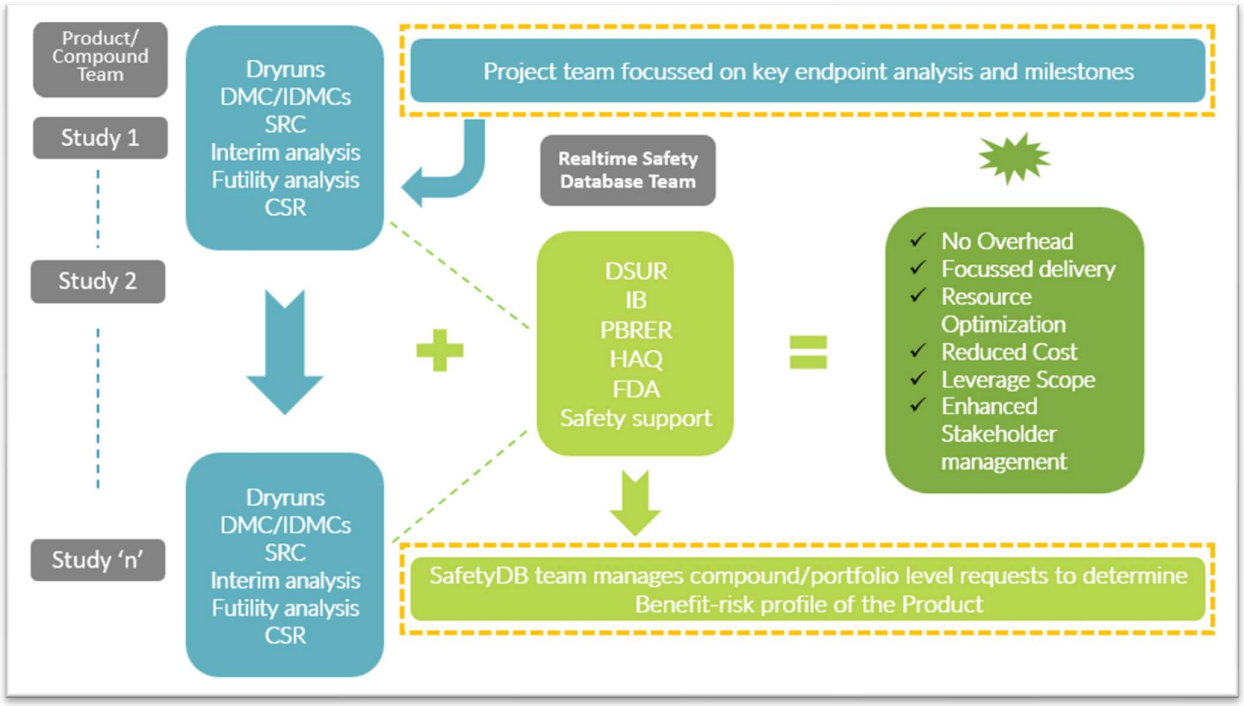
f) **Implementation Details: Project-level centralized database:** Houses all source data with controlled, regular access. Canonical staging and specs: Standardize variable names, units, and codelists; maintain mapping catalogs and derivation specifications. Raw-to-ADaM pipeline: Produce ADAE, ADEX, ADLB, ADCM, ADMH with consistent keys (study, subject, treatment, epochs, analysis windows). Derivations: TEAE logic anchored on first exposure with wash-in/out; CTCAE lab grading; partial-date handling; ongoing event rules. Refresh cadence: Aim for weekly data cuts with automated validation and export feeds for analytics.

g) **Results (Quantitative and Qualitative): Responsiveness:** Enabled faster, higher-quality responses to FDA/HAQs by providing integrated, traceable safety views. Direct all questions to safety team once than reaching out to individual studies. Adoption and usability: Actionable dashboards in Power BI and R Shiny; standardized outputs reusable across “n” studies. Efficiency gains: Saved approximately the equivalent of two FTEs per study across preparation and reviewer time, with cumulative impact approaching ~2.5 FTEs per major request for IB (similarly DSUR, PSUR, PBRER, HAQs, FDA safety inquiries), depending on scope and study count.

h) **Challenges Addressed:** Data heterogeneity and unit normalization via canonical staging and conversion services. Blinding requirements met with masked datasets and inference checks for sponsor-blinded studies. Resource constraints mitigated by reusable code assets, automation, and phased onboarding.

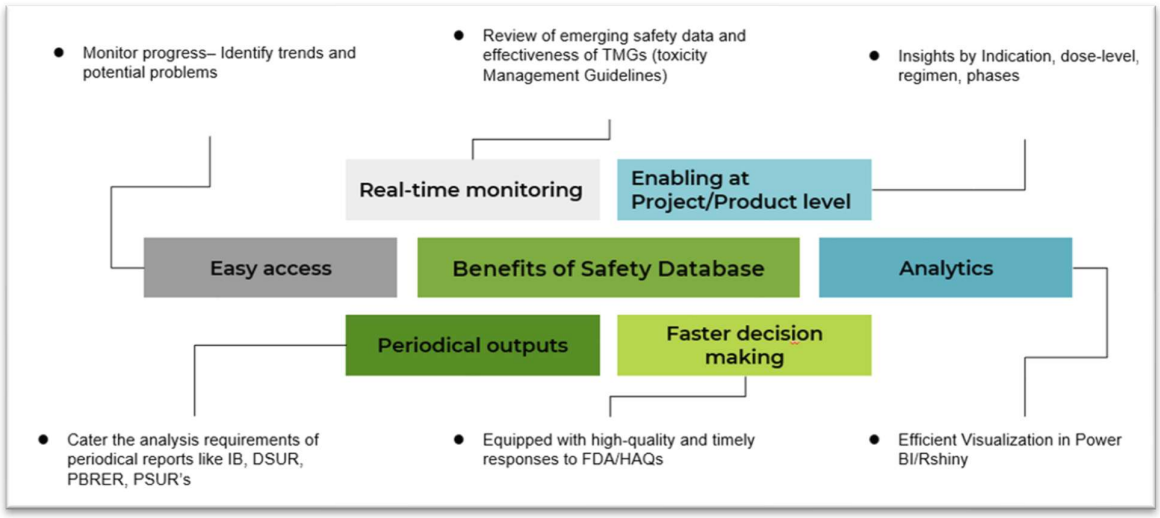
i) **Lessons Learned:** Start small and iterate: Early MVPs build momentum and reveal practical mapping rules. Document rigorously: Reviewer-friendly lineage. Pin dictionary versions and track changes: Prevent silent shifts that can alter results.

11.1 Collaborating Project and Safety Database



12. Benefits and Optimization

Having a well-designed Integrated safety database helps us to better understand product safety profile which would assist in optimizing the dose strategy and patient selection. This is vital as we could enhance patient monitoring approach and build effective toxicity management strategy. Such database evolves toward successful outcomes, promote continuous improvement and enabling strategic decision making. Facilitates with rapid, high-quality responses to FDA and other global authority queries. The database supports proactive management and communication of hepatotoxicity findings, enabling a robust, evidence-based FDA submission. It also equips the safety team with efficient surveillance and streamlined reporting/visualization tools, improving their understanding of the investigator product's safety profile.



13. Conclusion

Selecting a safety database is a lengthy process that requires considering multiple perspectives. It's important to involve a diverse, cross-functional team so that expertise and insights from various disciplines can be leveraged to take informed decisions. The Safety Database establishes a single, governed source of truth for product safety, delivering standardized, near real-time insights across studies. By curating direct-to-ADaM datasets with robust traceability, terminology governance, and validation, it accelerates regulatory reporting, strengthens signal detection, and improves cross-functional decision-making. Early outcomes include faster turnaround and actionable dashboards that enhance patient-centric risk assessment.

Contact Information

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

Ravichandra Hugar

AstraZeneca India Private Limited

Block D3, 2nd floor, Manyata embassy business park,

Rachenahalli, outer ring road

Bangalore, India

Email - Ravichandra.hugar@astrazeneca.com