

Non-Clinical Safety Data Integration Platform



Alain Nanzer*, Alessandro Brigo, Richard Maloney, Silvio Albertini, and Manfred Kansy

F. Hoffmann-La Roche Ltd., Pharma Research and Early Development, Grenzacherstrasse 124, 4070 Basel, Switzerland

*Corresponding author e-mail: alain.nanzer@roche.com

1 – Background

Today, drug discovery and development programs involve the generation of large amounts of *in vitro* and *in vivo* data, with the goal to ensure that potential drug candidates meet all the requirements needed to become new medicines. Many studies are conducted aiming at assessing that selected clinical candidates possess a suitable safety profile. Fast and easy access to information is essential for data analysis, knowledge generation/exchange and interpretation of results. Thereby, an integrated information environment enables the extraction of maximum value from research data in an efficient manner and allows faster and better informed decision making.

In vivo non-clinical Safety data generated from in house studies are captured and stored in different, independent Laboratory Information Management Systems (LIMS). Additionally, data from Contract Research Organizations (CROs) are obtained as documents and stored as PDF files in Document Management Systems (DMS), making it very time consuming - or even impossible - to query and mine study data across LIMS and document repositories. The Safety Data Integration (SDI) project was initiated with the objective to provide **non-clinical safety experts** with an application permitting rapid and easy electronic access to *in vivo* non-clinical study data, seamlessly integrated from active and decommissioned LIMS applications and study data from selected CROs.

Key objectives of the SDI project:

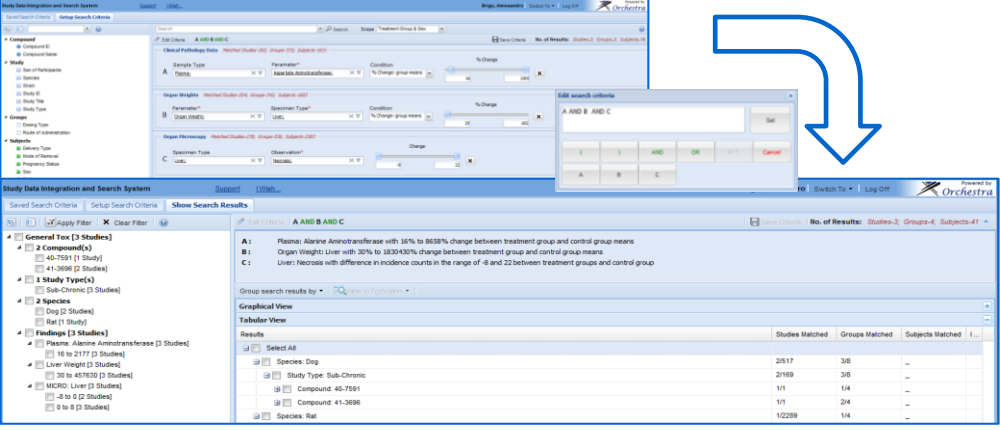
- Integrating all relevant in-vivo non-clinical safety data from in-house LIMS & selected CROs
- Allowing simple text searches and complex structured queries across all imported data
- Providing tabular and graphical representation of selected data and export data into defined structured formats for further analysis
- Interfacing with internal chemical compound registration system, allowing searches by chemical structures or substructures and the visualization of the chemical structures
- Interfacing with document repository providing direct access to the final study report

3 – The SDI Application

The SDI application is using ontologies, automated metadata extraction, units conversion and harmonization of terms to normalize and integrate study data. Mathematical indexing of structured data is combined with specialized text and parametric meta-data extraction, indexing, meta-analysis tools and search technology to support and guide scientific decision making.

Due to seamless data integration, the SDI tool allows both simple **“Google-like” queries** as well as **highly complex structured searches** (qualitative and quantitative) using full Boolean logic. The **SDI application** offers access to **about 4000 non-clinical studies** for scientific data querying and can answer questions like:

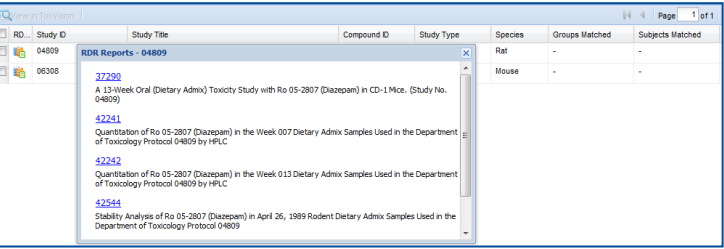
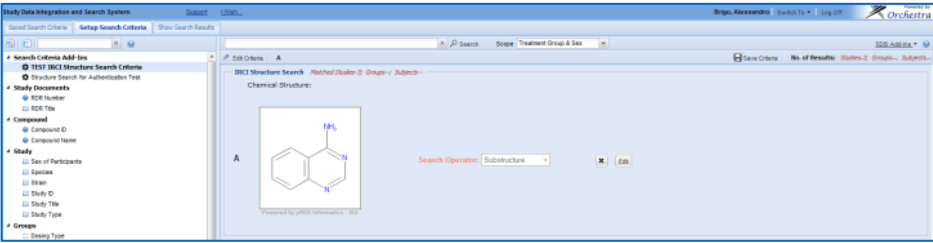
- “Which compounds showed an increase of >16% in Aspartate Aminotransferase AND an increase of > 30% in liver weight AND necrosis in liver?”**



5 – SDI Interfaces

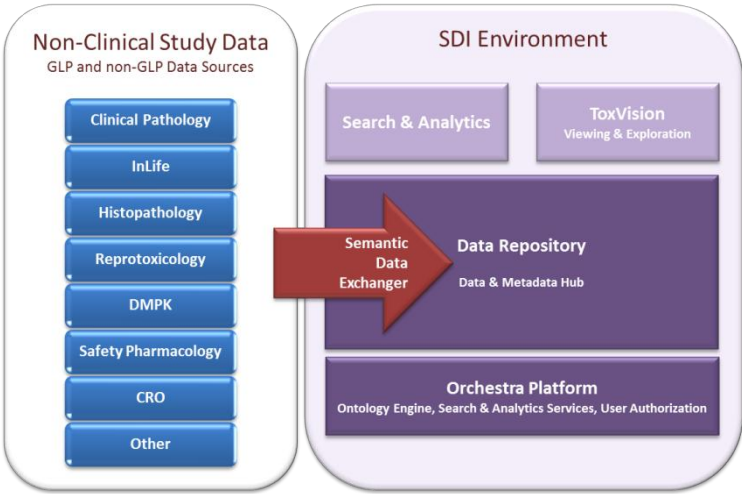
To increase the value of SDI and to streamline scientist workflows the application is interfaced with two important in-house applications:

- Hyperlink to corresponding **study reports** stored as PDF files in the research document repository
- Interface with the **Integrated Roche Chemistry Information (IRCI)** allowing searches by chemical structures, substructures, similarity, etc. and the visualization of the chemical structures of the test items included in SDI



2 – Data Sources and SDI Environment

Historically **non-clinical safety data** were captured in different applications across domains and sites. An analysis of source systems to be integrated identified a total of **13 in-house LIMS** and various other internal and external sources. After a careful evaluation **PointCross’ Orchestra platform** was selected as the integration platform. It offers a set of applications for representing and managing all types of structured and unstructured data, a data repository, a semantic data exchanger, the ToxVision™ viewer, an extended analytics engine and a powerful search engine. The application is web based and globally accessible to authorized Roche scientists using a web browser. The chart below summarizes the data types loaded and the high level architecture of the SDI environment:



4 – Hits and Data Visualization

When performing a search, the system will return, if present, all the studies matching the criteria defined in the submitted query. A list of studies will be displayed and each of them can be explored by drilling down to the individual animal raw data using the implemented ToxVision™ feature. Data can be visualized by **‘Group or Subject Summary’** and at an **‘All Subject’** level.

After selecting the parameter of interest, several automated plotting capabilities are available, allowing quick graphical visualization of selected data:
Bar Charts, X-Bar Charts, Box & Whisker Plots and Line Charts.



6 – Discussion & Perspectives

Based on the PointCross Orchestra™ platform, Roche and its partners have created a system which is providing scientist with unique query and visualization functionality in the non-clinical safety framework. SDI enables **rapid and easy access to complex raw data** allowing:

- The identification of **trends across studies**, compounds, species and data types
- Data analysis/exchange and **interpretation**
- The identification of **complex toxicological patterns** of findings and their relevance
- The identification of tool compounds for **assays validations**
- The development of **predictive models**
- Knowledge** generation and consolidation
- Rapid response to **queries** coming from **Health Authorities**

Currently the team is extending the application by adding more non-clinical data types and discussions started to integrate clinical data into the SDI platform. Additionally, interfacing SDI to the early discovery data warehouse is planned, resulting in a long-awaited full integration of discovery, pre-clinical and clinical data, representing a significant milestone in the Translational Medicine arena.

7 – Acknowledgments

Project Sponsors Thomas Singer Bryn Roberts	Steering Committee Thomas Weiser Rob Tudor Margret Assfalg Michael Braxenthaler Conney Berger Silvio Albertini Manfred Kansy Alain Nanzer	SDI Core Team Richard Maloney (IT Project Lead) Alessandro Brigo (Safety Project Lead) Michael Pignatello SDI Extended Core Team Matthias Festag John Patterson Jürgen Funk Igor Mikaelian	Vendor PointCross, Inc. Safety Data Consultant Phil Drew (PDS)
--	--	--	---