

Integrating Clinical Data with the Oracle Life Sciences Data Hub

An Oracle White Paper
November 2007

Integrating Clinical Data with the Oracle Life Sciences Data Hub

The Oracle Life Sciences Data Hub empowers and automates a life science company's ability to combine clinical and related data at lower overall cost, with fewer staff and less risk than with other commercial or custom solutions.

EXECUTIVE OVERVIEW

In the life sciences industry, the process of clinical development requires that patient data collected from multiple sources be merged and combined for the purpose of understanding the safety and efficacy of a compound. Quick and accurate access to this combined data can:

- inform management decisions to alter, accelerate, or cancel the development of a particular compound and thus increase profits
- deliver data for formal structured analysis to accelerate regulatory submissions
- minimize the risk of introducing an unsafe compound to the market

The Oracle Life Sciences Data Hub empowers and automates a life science company's ability to combine clinical and related data at lower overall cost, with fewer staff, and with less risk than with other commercial or custom solutions.

INTRODUCTION

Today's life sciences companies must gain rapid insight into the safety and efficacy profile of any compound or device under development. Any delay in pulling together all the safety data for a compound postpones safety analysis, which can expose patients unnecessarily to dangerous compounds and even cause the approval of a product that is not as safe as originally perceived. In addition, rapidly pulling together all available patient data for a particular compound or family of compounds, allows organizations to make stop-or-go decisions sooner, optimizing the company's investment strategy.

Every life sciences company must prepare its patient data for structured analysis and regulatory reporting – a process that can be very labor-intensive.

While all companies managed to combine and integrate their patient data for formal analysis to satisfy regulatory requirements, most only use this vital information for the sole purpose of regulated analysis. Reasons for this limited use of data are many and include:

- Most companies do not have sufficient metadata to enable personnel outside of the project team (who may not be familiar with the meaning of the data) to use patient information. For example, the protocol may state that a blood pressure measurement is to be taken on the arm when the patient is in the supine position. However, the fact that the blood pressure was taken in the supine position may not actually be contained in the data, but only in the text of the protocol.
- The data is usually contained in SAS® data sets on file systems and is not generally accessible in a secure manner to the wider clinical development staff due to security restrictions and lack of specialized skills.
- In many cases, the team that responsible for consolidating the data for a submission is under tremendous time pressure to prepare the submission and is not able to prepare standard, reusable components with rich metadata so that that data can be shared in a secure manner with the rest of the clinical development team.
- While on the surface the industry standard CDISC SDTM is helpful in the process of compiling patient data, the source systems in which the data is collected do not produce CDISC SDTM-compliant data, and the activity of preparing compliant data falls to the group preparing the submission.
- Data that is not part of the formal submission is usually maintained in systems siloed from those that contain data used for formal analysis. Genomics data, pre-clinical data, and pharmaco-economic data that may be useful for determining the safety and efficacy, as well as the economic potential of a compound are often difficult or impossible to query alongside the formally collected patient data

THE ORACLE LIFE SCIENCES DATA HUB IS AN APPLICATION FOR INTEGRATING CLINICAL DATA

The Oracle Life Sciences Data Hub is an application that has been designed specifically for the life sciences industry to integrate, transform, and report on data in a secure and regulatory compliant manner.

Adapters allow loading data from multiple technologies

As part of integrating data for formal analysis, ad-hoc exploration, data mining, and patient data reporting, the Oracle Life Sciences Data Hub allows easy loading of data from multiple technologies. Oracle Life Sciences Data Hub comes with adapters preconfigured to accept data contained in SAS data sets, SAS transport files, text files (including XML files), and any system based on an Oracle database. If the source system is based on an Oracle database, it is also possible to use a pass-through view to the remote system to virtually load the source data. Pulling the data together from any data source based on these adapted technologies is as simple as

The Oracle Life Sciences Data Hub allows easy loading of data from multiple technologies. It is preconfigured to accept data from systems based on an Oracle database, SAS data sets and transport files, and text files (including XML files).

filling out a form, dramatically reducing the overall cost of pulling raw data together when compared to traditional techniques.

Integration with Oracle Clinical and Remote Data Capture

Oracle Life Sciences Data Hub comes preconfigured and integrated with Oracle Clinical/Oracle Remote Data Capture so that all the information and metadata in the Oracle Clinical global library and the Oracle Clinical data extract system is automatically synchronized and available for use within the Oracle Life Sciences Data Hub. This means that all the corporate data standards contained within Oracle Clinical can be used for data integration, analysis, and reporting with no additional cost or effort. The Oracle Clinical adapter also brings across patient data, discrepancy data, lab data, study metadata (including sites and investigators), and randomization data and loads it into Oracle Life Sciences Data Hub.

Supported engines transform data to standard formats

Once data has been loaded into Oracle Life Sciences Data Hub, the data can easily be transformed to standard structures using the transformation and reporting engines. Oracle Life Sciences Data Hub comes preconfigured for integration with PL/SQL and SAS. Any manipulations required to make data conform to one or more company or industry standards, including CDISC SDTM, can be done completely within the Oracle Life Sciences Data Hub using programs developed by the life sciences company.

Libraries and version control optimize reuse and traceability

All programs, reports, table definitions, and other metadata used to transform and report loaded data are maintained in libraries under version control. These libraries are organized according to customer needs. Users can classify any metadata and make it available for reuse as a standard. Oracle Life Sciences Data Hub also allows users to trace the output of any program back to the programs and tables that have touched the data. As a result it is very easy to reuse programs and tables to conform to standards.

Open APIs allow metadata-driven data transformations

Oracle Life Sciences Data Hub includes a set of published and supported APIs that can be used to automatically build data manipulation and transformation programs based on metadata. By calling the APIs, users can write master programs that build other programs based on user-supplied mappings or other metadata. The result is that standardizing and pooling programs that traditionally have been handcrafted can be automatically generated based on metadata, enabling substantial resources savings.

Execution module controls running programs and delivering results

Oracle Life Sciences Data Hub includes an execution module that controls all the loading, transforming, and reporting of data in a secure and efficient manner. All

By calling the APIs, users can create master programs that build other programs based on user-supplied mappings or other metadata. The result is that standardizing and pooling programs that traditionally have been handcrafted can be automatically generated based on metadata, enabling substantial resource savings.

programs, table structures, and other metadata are automatically deployed to the database with a push of a button. This saves significant resources and risk over classical data warehousing and analysis applications since there is no need to manually create, validate, copy, or run installation scripts.

All programs, table structures, and other metadata are automatically deployed to the database with a push of a button. This saves significant resources and risk over classical data warehousing and analysis applications since there is no need to manually create, validate, copy, or run installation scripts.

Loads, transformations, and reports are all submitted from a common interface independent of the underlying technology. Jobs may be submitted on demand, scheduled, repeated, or triggered by a message placed on a message queue.

Oracle Life Sciences Data Hub comes with special features that allow running jobs against snapshots of data or on the most current data without any need to create special copies of the data. This saves significant time and risk when attempting to get a consistent view of the data. When running against the most current data, Oracle Life Sciences Data Hub executes only required programs and loads in the correct order in order to give the fastest results possible. This back-chaining feature yields significantly faster results at lower risk than classic applications that require all programs be run to be sure the results are up to date.

Design promotes and assists regulatory compliance

Oracle Life Sciences Data Hub has been designed for regulatory compliance. It includes a number of features to assist in complying with regulations when integrating clinical data.

Blinded and sensitive data

Oracle Life Sciences Data Hub supports the blinding and un-blinding of real data. Programs and reports can be developed and validated against dummy, or fake, data and then executed without any special user programming in a secure manner against the real data by users with special blind-break security privileges. The real data remains blinded unless it is explicitly un-blinded to give access to a wider community, for example when a trial completes and the treatment codes are made more generally available. Each access to the blinded data is recorded and audited to maintain regulatory compliance. Since the handling of sensitive data is built into Oracle Life Sciences Data Hub, no special infrastructure is required to handle the often complex business problem of handling data delivery for interim analysis, safety data monitoring boards, adaptive trial decisions, or Pharmacokinetic (PK) and Pharmacodynamic (PD) data.

Since the handling of sensitive data is built into the Oracle Life Sciences Data Hub, no special infrastructure is required to handle the often complex business problem of managing data delivery for interim analysis, safety data monitoring boards, adaptive trial decisions, or PK and PD data.

Validation assistance

As programs, tables, and reports are developed, all the metadata is tagged with a validation status. As users test and validate programs and other objects, users with a special security privilege can promote the objects from a validation status of Development to Quality Control and from Quality Control to Production. It is also possible to store documents such as requirements documents alongside the programs and tables. If users run jobs that produce outputs for testing, they can link those outputs to the program or table as validation evidence. In addition,

special controls are in place to prevent permanently deleting data from production areas.

Data audit trails

Oracle Life Sciences Data Hub automatically includes audit trails on all data tables. The user who is writing programs and reports to integrate and conform the data to standards does not need to put in the extra effort to write to the audit. As a result of the audit trail, it is possible to create downstream snapshots of the data in any table, as it existed in a point in time.

Open architecture fits into enterprise architecture

In order to integrate data from many sources, Oracle Life Sciences Data Hub comes with an adapter toolkit that a customer can use to adapt additional technologies and applications to work with Oracle Life Sciences Data Hub to load data, transform data, report on data, visualize data, and prepare data for export. Since Oracle Life Sciences Data Hub is designed to use adapters to extend its power, it remains fully supported as users add new technologies.

In addition, Oracle Life Sciences Data Hub works with inbound and outbound messaging. Authorized processes from outside Oracle Life Sciences Data Hub can trigger processing within the hub. In addition, Oracle Life Sciences Data Hub allows outbound messaging as a way to communicate with external applications.

Oracle Life Sciences Data Hub is built on state-of-the-art Oracle Fusion Middleware and Oracle Database technology, which further enables Oracle Life Sciences Data Hub to fit within an enterprise architecture for integration of clinical data.

CONCLUSION

The clinical development process requires that patient data collected from multiple sources be merged and combined for the purpose of understanding the safety and efficacy of a compound. Oracle Life Sciences Data Hub is specifically designed to help organizations in the life sciences industry integrate clinical and non-clinical data, and make all data available for formal analysis and reporting as well as ad hoc exploration, in a secure manner.

Because Oracle Life Sciences Data Hub comes with many features built in, there is no need to spend resources on building many features that are normally required for traditional integration applications. It allows integrating and conforming data to standards at a lower cost and risk than traditional applications with:

- Adapters for loading Oracle, SAS, and text data
- Multiple engines for transforming data, including PL/SQL and SAS
- Multiple libraries to maximize re-use
- Built-in version control

The Oracle Life Sciences Data Hub is built on state-of-the-art Oracle Fusion Middleware and Oracle Database technology, which further enables Oracle Life Sciences Data Hub to fit within enterprise architecture for integration of clinical data.

Oracle Life Sciences Data Hub empowers and automates a life science company's ability to combine clinical and related data at lower overall cost, with fewer staff, and with less risk than with other commercial or custom solutions.

- Built-in data auditing and snapshotting
- Traceability to answer what data came from where and how it came to that state
- Blinding and un-blinding of sensitive data
- Validation assistance
- Open architecture

The Oracle Life Sciences Data Hub enables life sciences organizations to automate the process of standardizing and combining clinical and related data at lower overall cost, with fewer staff, and with less risk than other commercial or custom solutions.



Integrating Clinical Data with the Oracle Life Sciences Data Hub
November 2007
Author: Mike Grossman

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Copyright © 2007, Oracle Corporation and/or its affiliates. All rights reserved.
This document is provided for information purposes only and the contents hereof are subject to change without notice.
This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.
Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.