

Jennifer Desai is a Senior Technical Data Architect at Janssen Pharmaceuticals. In this role, she is responsible for supporting Data Engineering activities for Infectious Diseases and Vaccines (ID&V) clinical trials, as well as contributing to the Data Ingestion Capability shaping team. Jennifer joined Johnson & Johnson in 2016.



DATA LOAD AUTOMATION

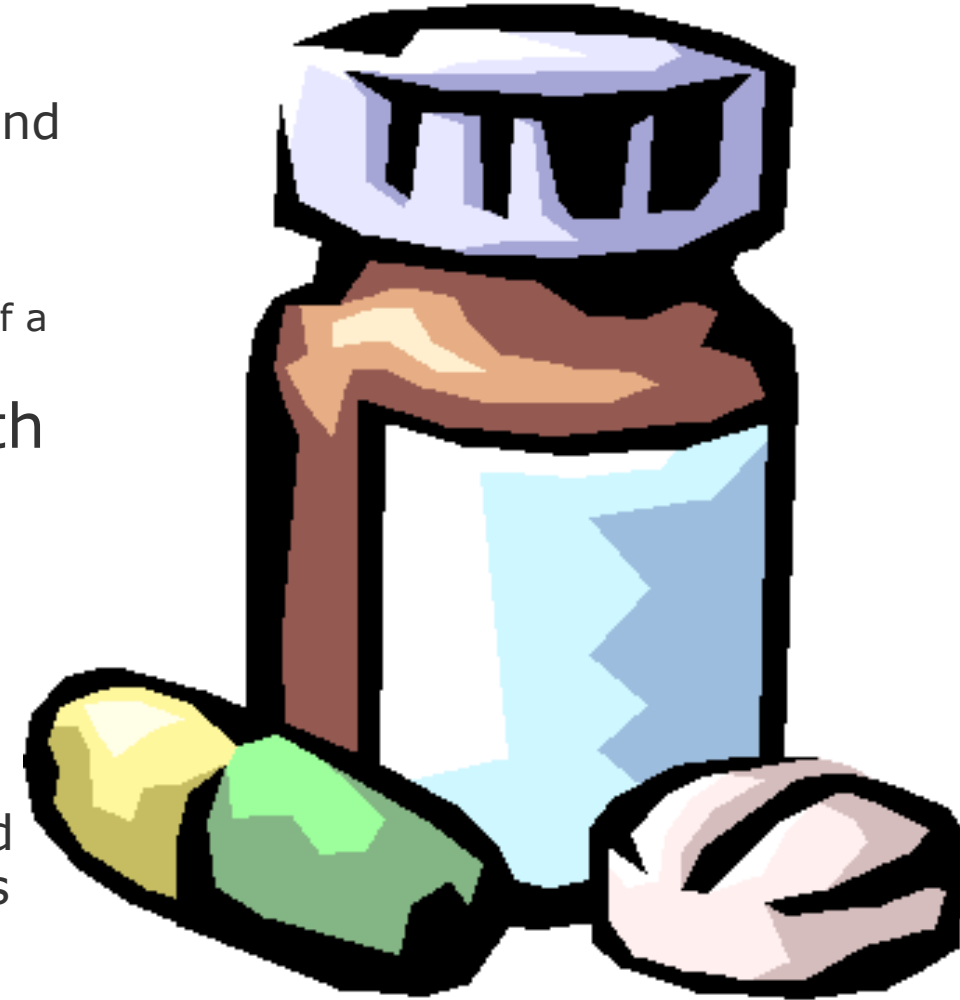
Presented By: Jennifer H. Desai, MPH
Janssen R&D, Raritan, NJ, USA

Table of Contents

- Introduction
- Data Transfer Agreement
- SAS Life Science Analytics Framework (LSAF)
- Current Manual Data Transfer Process
- New Automated Data Transfer Process
- Set Up Process For Data Load Automation
- Limitations
- Conclusion
- References
- Contact Information

Introduction

- Data is the backbone of medicine
 - Pills, vaccines, etc. -> exist due to clinical research and the collected and examined data from clinical trials
 - Important to produce clean data from clinical trials in order to advance in the medical world
 - Data transfers are essential aspects throughout the duration of a clinical trial
- Current Janssen process of data transfers with its external vendors:
 - Manual
 - Time consuming
 - Hassel
- Solution -> data load automation
 - Data load automation - when source data is delivered through an automated process from external vendors directly into Janssen's SAS data repository platform, called Life Science Analytics Framework (LSAF)



Data Transfer Agreements



- Study start-up phase
 - Data streams and vendors that will be utilized in the study are identified
 - Once the database is set up and the study goes live, it begins the conduct phase
- Conduct phase
 - Trial specific data transfer agreements (tsDTAs) are created between Janssen, the study's contracting research organization (CRO), and the vendor

Data Transfer Agreements

- Data transfer agreements (DTAs) are specifications that outline the expectations of the file transfers between Janssen and each of its vendors, per study
 - Janssen and its study CRO work together with the vendor to gather the necessary data needed to complete the tsDTA template
 - The tsDTAs are created per each data stream and each vendor in the study
 - These tsDTAs consist of trial specific information
 - Visit/visitnum combinations, test and test codes, transfer file type, frequency of transfer, etc.
 - The tsDTAs are living documents and can be updated throughout the study as needed
- Method of transfer section
 - Specifies the folder path and locations of where the vendor will need to manually drop the data file in LSAF
 - States the frequency of the data transfers, which can be daily, once a week, biweekly, once a month, etc





- SAS Life Science Analytics Framework (LSAF) is Janssen's data repository tool
 - Used to share data between our CROs and vendors
 - LSAF provides the ability to integrate analytic applications and other SAS solutions into the framework to deliver the power of analytics to more roles within the organization
- LSAF is worldwide accessible and can be used at any time and any location
 - Uniform platform for various departments and vendors that all contain the same configuration, same user interface, same central repository for data, SAS programs, output files, etc.
- LSAF simplifies sharing & transferring of data in a secure way
 - Typical types of data sharing and transferring include when CROs upload data packages to Janssen Data Management (DM) via LSAF, when external vendors share their data with Janssen DM and other vendors via LSAF, and when Janssen DM shares the data with statistics and programming for analysis via LSAF
- LSAF helps streamline and automate processes
 - The tool contains automated process flows that consist of automated job running, it has central security management (users, groups, folders), it can process large volumes of data, and it has an integrated SAS environment (statistics analysis, SDTM QC, scheduling jobs, etc.)

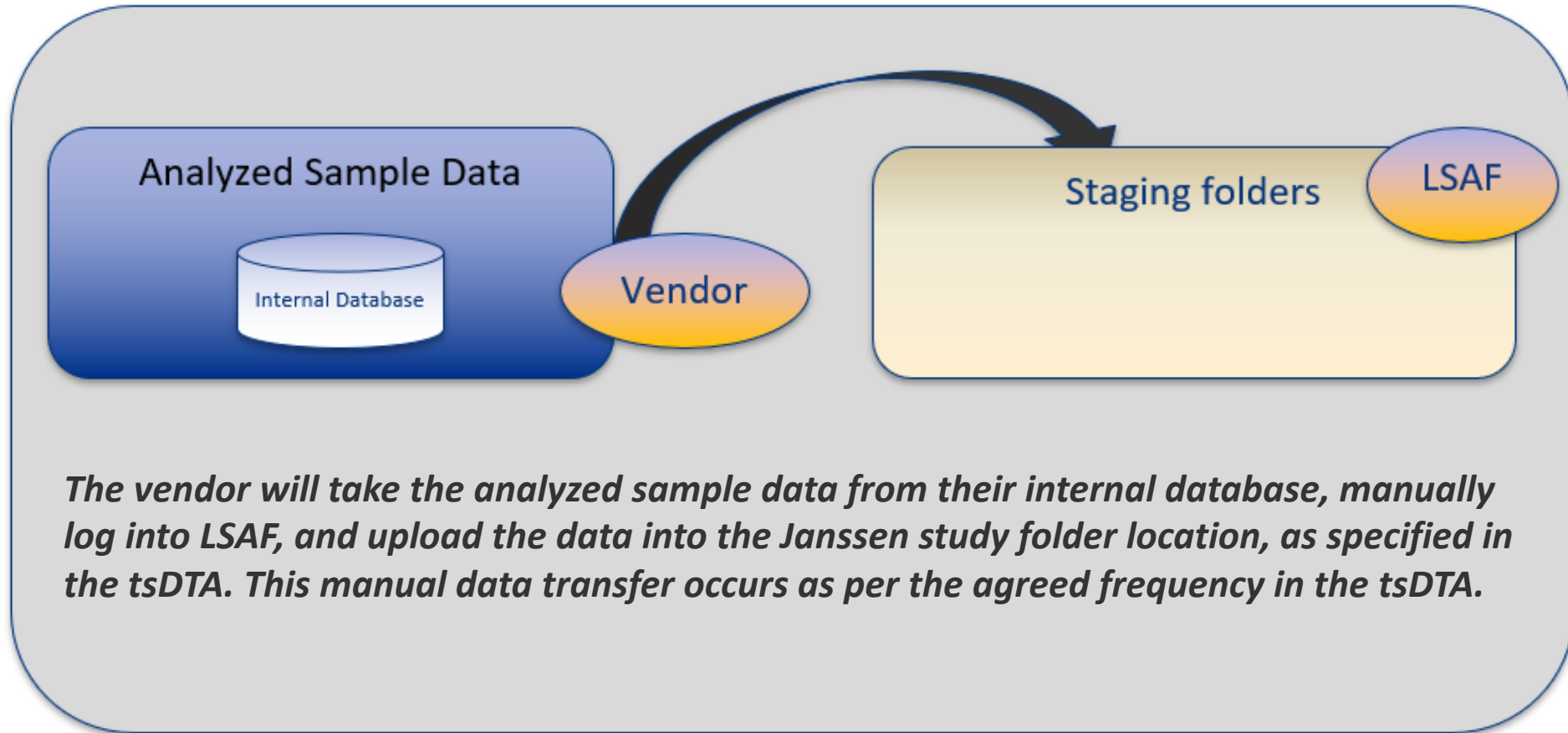
Current Manual Data Transfer Process

- Once a tsDTA is signed and finalized, the data transfers can begin, but before they can begin, the logistics must comply:
 - Janssen DM will have to individually grant access to their study in LSAF for each vendor and the vendor's individual team members
 - Before any vendor personnel can receive access, they need to complete the appropriate trainings
 - When team members at the vendor changes, Janssen DM will need to revoke/grant access as necessary
- The process is time consuming and the coordination can become frustrating which is why manual data transfers remain to be a bothersome experience for both Janssen and its vendors

Vendor Data Current Process – High Level Overview



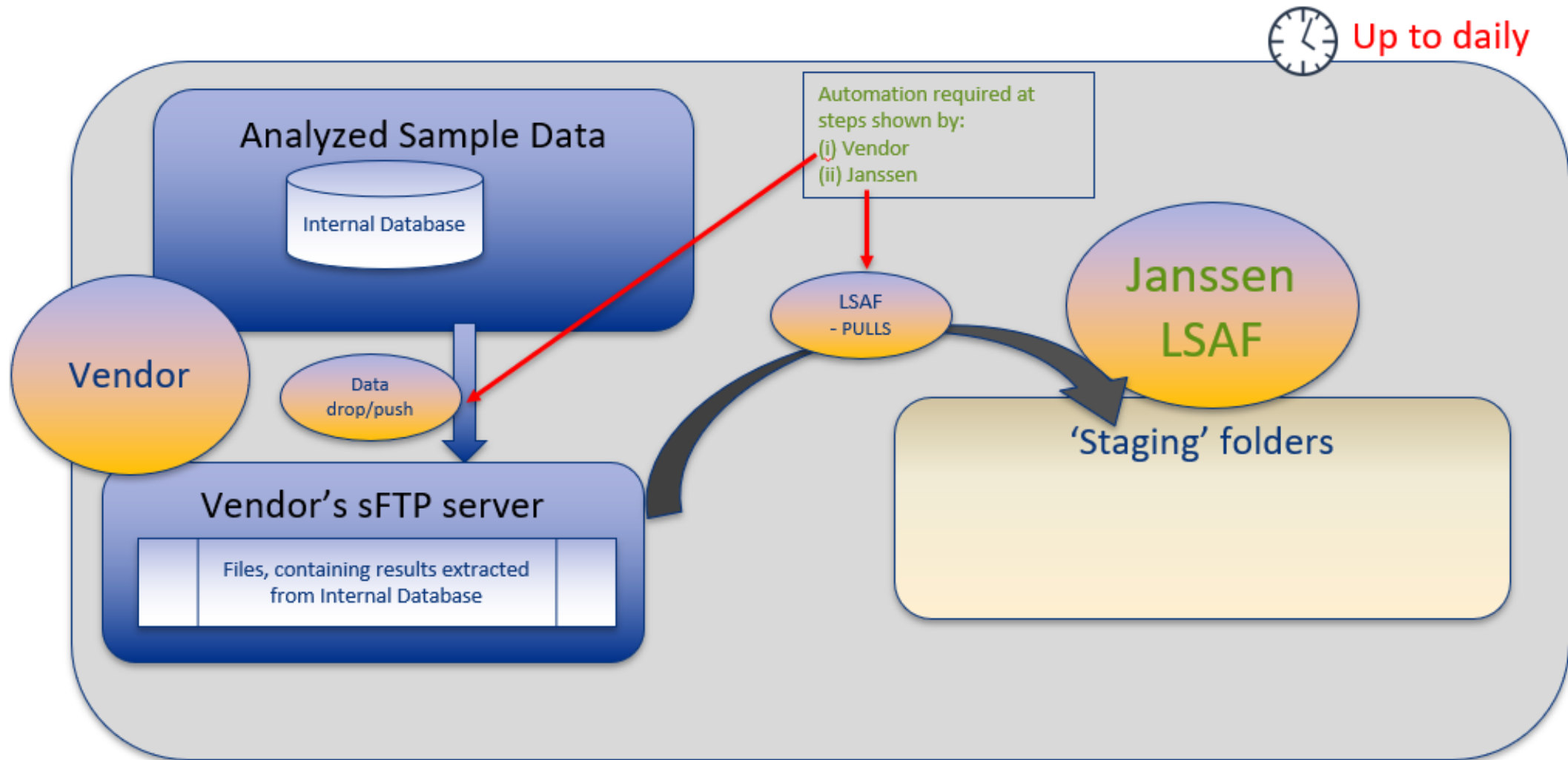
As per DTA frequency



New Automated Data Transfer Process

- With automated data transfers there is no need for the vendors to have each team member trained nor granted access to LSAF
- Vendors will drop their data onto their servers and allow Janssen access to their servers
 - Once the vendors drop their data onto their servers, Janssen can automatically go into them and pull the data into LSAF
- This process is time efficient for the vendors as they do not have to manually upload the data into our data repository, and, it is time efficient for Janssen as we can program the data pulls to have an automatic daily pull schedule
 - This removes the extra hassle of granting the vendors individual access to the specific folders in our data repositories

Autoload – for Vendor Data



Set Up Process for Data Load Automation

- One-time implementation process per each vendor
- The vendor and their individual team members will not all need access to LSAF nor will need to take the LSAF trainings
 - It is recommended that at least one or two members from each vendor still have access to LSAF in case of any automation issues, so they can resort to a manual upload if needed as a backup procedure



Vendor Eligibility Criteria

■ Yes

- Can automate data transfers
- Has an appropriate server (sFTP, MBOX, etc.)
- Meets all technical requirements (allowing traffic via port 22, landing at root folders, etc.)
- Allows Janssen read access to their servers with a username and non-expiring password
- Has multiple studies with Janssen and planned future studies

■ No

- Cannot automate data transfers
- They do not generate their own data (CROs)
- Do not have appropriate servers
- Do not allow sponsors access to their servers
- Janssen only utilizes vendor's services for 1 or 2 studies and/or has no future plans to work with the vendor

****For vendors who cannot automate data pulls, we are looking into trying increase their data push frequency****

Set Up Process for Data Load Automation

- The method section on the tsDTAs will say automated data transfers instead of manual data transfers and it will include the vendor's sFTP server location and folder details
 - The structure, format, naming conventions, etc. will remain the same, as the only change is the method of transfer, from manual to automated
 - The LSAF pulls are programmed to automatically run daily, even if the vendor does not push data daily
 - The data will only be pulled into LSAF if there is a new file in the vendor's server

Method and frequency

Method of subject data (test or production) transfer (choose only one):

☐ Automated transfer:

The (External) Data Provider will automatically post the data according to the agreed frequency at the following sFTP location: [\[specify\]](#)

The “pull” module will automatically pull data from the sFTP server to the following Janssen Clinical Data Repository location: [\[specify\]](#)

☐ Manual transfer:

The (External) Data Provider will manually post the data according to the agreed frequency to the following Janssen Clinical Data Repository location: [\[specify\]](#)

Frequency of (production data) transfer

During Trial (check all that apply):

- | | | |
|--------------------------|--------------------------|---------------------------|
| Once | ☐ | |
| Daily | ☐ | |
| Monthly (once per month) | ☐ | |
| Weekly (once per week) | ☐ | |
| Other, specify | ☐ | [Specify] |

Prior to Database (DB) Lock (check all that apply):

- | | | |
|------------------------|--------------------------|--------------------------------------|
| Weekly (once per week) | ☐ | |
| Other, specify | ☐ | [example: As needed] |

If non-automated transfer date falls on weekend or holiday:

- | | |
|-----------------------|--------------------------|
| Previous Business Day | ☐ |
| Next Business Day | ☐ |

The frequency of transfers may increase before interim and final DB lock.

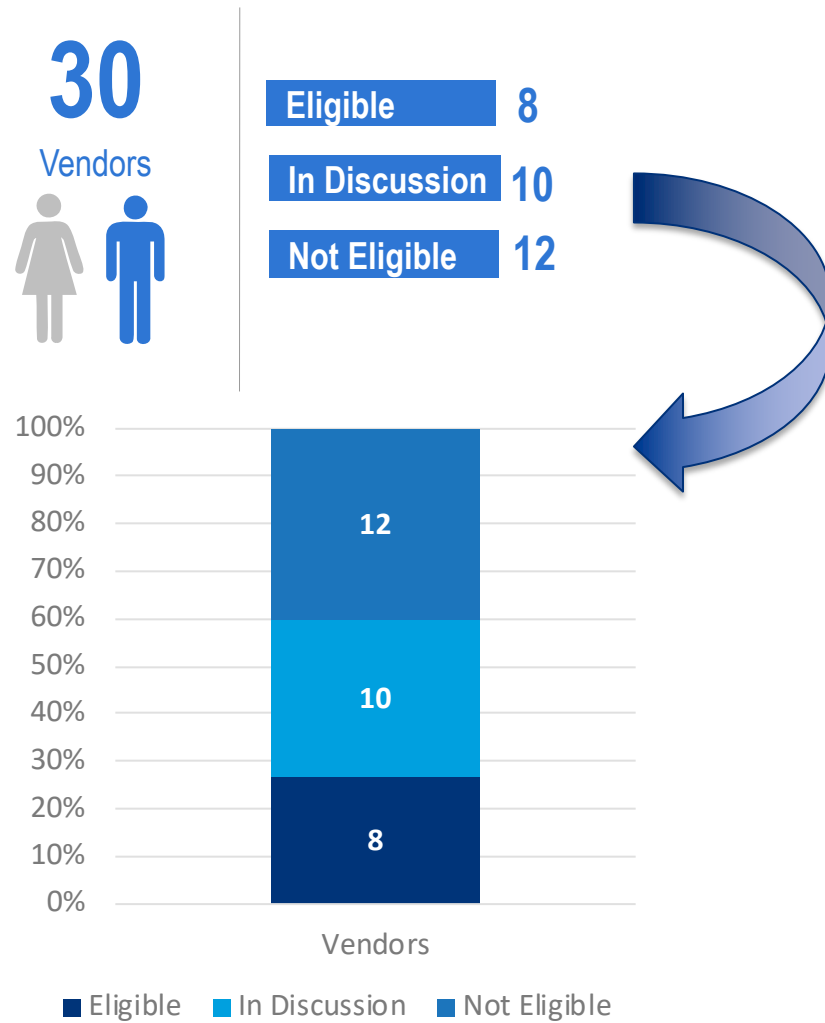
Set Up Process for Data Load Automation

Set Up Process for Data Load Automation

- Data load automation can be set up with all data types
 - Laboratory, electrocardiography (ecg), medical imaging, etc.
- Because of data load automation, source data is now readily available more frequently
 - Supports Real Time Learning and Decision Making (RTLDM), as well as improving Study Data Tabulation Model (SDTM) conversion timelines



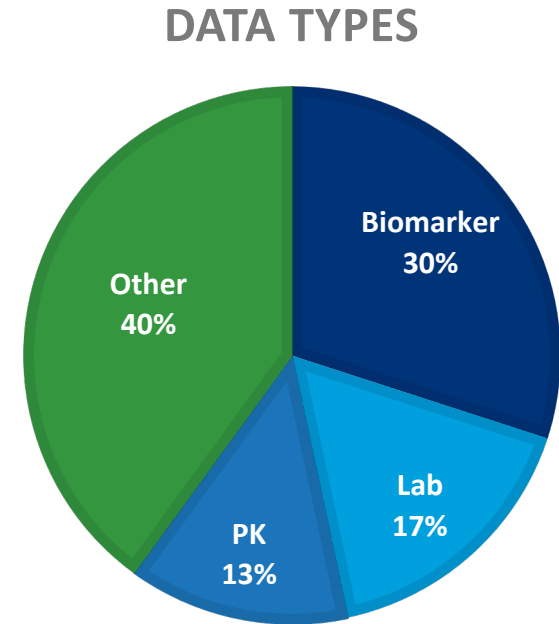
Autoload Progress at Janssen



Data types

9 Biomarkers
5 Lab
4 PK

12 Other (cognitive tests, ecg, imaging, etc.)



Limitations

- Not all vendors can implement automated data transfers
 - Smaller vendors such as wet labs or universities
 - Counter solution -> increase their manual data transfer push frequency
 - For example, for the smaller vendors who manually transfer their data into LSAF once a month, we will like to increase it to twice a month, weekly, etc.
- Potential contractual or financial impact
 - However, because data load automation benefits the vendors themselves as well, they have been supportive in its implementation



Conclusion

- Data load automation is a solution to enhance vendor data transfers
- It helps deliver data on a more frequent schedule
 - Continually looked at and reviewed by all parties involved in the clinical trial
 - Data managers, statisticians, programmers, study responsible physicians (SRP), study responsible scientists (SRS), etc.
 - Aids in delivering cleaner data for submissions
 - More efficient and effective data transfer process throughout the duration of a clinical trial

References

SAS Life Science Analytics Framework. A singular analytics foundation for clinical research to increase efficiency. Fact Sheet. (2019). Retrieved from <https://www.sas.com/content/dam/SAS/documents/product-collateral/fact-sheets/en/life-science-analytics-framework-107108.pdf>

Contact Information



Jennifer H. Desai, MPH
Janssen Research and Development
920 US 202 South
Raritan, NJ, 08869
908 927 4994
jdesai12@its.jnj.com
<http://linkedin.com/in/jenniferdesai>