RxGPS
THE ALLIANCE FOR GLOBAL
PHARMACEUTICAL SERIALIZATION

Data Sharing and Reporting: An Overview

The success of a product tracing system for pharmaceutical supply chain security hinges on secure and interoperable data sharing and data reporting across the entire pharmaceutical supply chain. However, there is no single or uniform solution that will work across systems and between markets. The foundational decisions made by regulators and industry around how to construct a system for using serialization of pharmaceuticals have important implications for how serialized data are shared and communicated.

As outlined in the RxGPS Model Regulation, there are three fundamental, foundational questions that any market seeking to implement a serialization model needs to answer:

1. How will individual pharmaceutical packages be identified?

   Serialization of pharmaceuticals involves assigning a unique identifier to a product for identification purposes and encoding that data in a barcode that is printed on the product package. The corresponding data record for the product is then retained and shared as required. For more information on the format, data elements, and location of a unit identifier, see the RxGPS Position Statement on Unit Identifier.

2. How will serialization data be shared to enhance supply chain security?

   While serialization forms the foundation of a supply chain security system, the use of serialized data among trading partners to verify, authenticate, track, or trace products at the unit level, can be a major step to secure the supply chain.[1] Depending on the specific needs of a given market, there are different models for data sharing. For more information on why RxGPS recommends implementing point-of-dispense verification prior to consideration of a full traceability system, see the RxGPS Position Statement on the Benefits and Complexities of Common Serialization Models.

3. What data architecture will be used to implement the selected approach?

   The RxGPS Model Regulation and Implementation Plan provides a framework for a market to select and embark on implementation of pharmaceutical serialization and models for sharing and storing serialization information. The Model Regulation also includes a side by side comparison of the two different approaches to data architecture (storage): centralized and distributed.

While answering each of these questions will help a market chart a path toward a system for supply chain security, successful implementation of such systems requires careful consideration of the mechanisms for communicating serialized data within a pre-determined database architecture, between trading partners, and with regulators as appropriate. This document outlines how to leverage serialization data, through standardized communication and data reporting protocols and infrastructure, to advance supply chain security and other market-specific goals. Depending on the answers to the

[1] https://www.rxgpsalliance.org/serialization/

questions above, markets will have unique challenges and considerations with regard to data sharing and reporting. As such, this document elaborates on the 7 key considerations for the utilization of serialized data and how they vary across the chosen model for data sharing (*i.e.*, verification or traceability) or across a chosen data architecture (*i.e.*, centralized database/repository or distributed, company-owned databases).

1. Data Storage

When it comes to data storage, the chosen data architecture model largely dictates how data are stored and where data are located. However, key considerations for data storage are who stores the data and for how long.

Who stores the data?

Centralized Database Architecture

With a centralized database model, all data are stored and then accessed from one central storage repository. All data are mapped to the central location, and all supply chain participants replicate their data to the location for storage and query.

Typically, for global markets implementing centralized databases, a government agency has taken on the responsibility to store data within a government-owned repository and to manage data storage, including development of the necessary technical specifications. Most centralized databases are built, maintained, and at least partially funded by a government agency. However, responsibility for setting up and/or funding construction of the database has sometimes resided with industry.

Once a centralized database is constructed, trading partners transmit all serialization data to the central database for storage and query. Any supply chain security functionality that utilizes product serialization data would be routed to the central database. However, trading partners may still choose to hold/store their data while the centralized database structure is being built and even over time for their own internal use.

It is common for regulators to call for implementation of a single, centralized database for an entire market (*e.g.*, Turkey). However, for markets with more federated governance, such as the European Union, it is possible to create a network of centralized databases. The European Medicines Verification System (EMVS) is made up of National Medicines Verification Organizations (NMVOs) that operate national, centralized data repositories for individual EU member states. The EU Hub is a centralized database for the entire EU market, which stores and transmits manufacturer product data to the relevant national systems for the use of trading partners in that market.[2]

[2] https://emvo-medicines.eu/mission/emvs/

Distributed Databases Architecture

With a distributed database model, each company owns and maintains its data in its own repository. These repositories must adhere to standards for interoperability and can therefore be queried to retrieve or share the data, as appropriate.

Unlike a centralized database, distributed systems spread responsibility for data storage across trading partners. Each trading partner is responsible for maintaining the security, storage, access, and governance for their own data in company-owned servers. Cloud storage is a common and acceptable solution, and many trading partners outsource management of their proprietary databases to third-party solutions providers. The dispersed responsibilities lead to a range of data retention requirements and solutions based on a company's business and regulatory needs, maturity, and infrastructure but requires companies to have an over-arching policy to ensure global regulatory compliance and consistent procedures that adhere to industry standards for interoperable data exchange. Unlike in a centralized database architecture responsibility for data storage, management, security, and integrity in a distributed architecture lies not with the governing authority, but with individual companies.

How long are data stored?

Regardless of the data architecture chosen, data "residence time" and "retention time" requirements should be in place both for individual trading partners' databases as well as the centralized database(s). Residence time requirements will dictate how long data must be available in a "live system" or database (*i.e.*, retrievable via query). RxGPS believes that data should be retained in a live system/database for, at minimum, 1 year after the expiry date for a product. Data retention requirements refer to how long data must be kept or archived for potential audit or investigation. Global markets tend to require data to be retained for approximately 6-10 years. Together, these requirements will ensure data are available throughout the life of the product, plus some minimum additional time period.

2. Data Security & Access

Each data architecture brings unique considerations for how to manage data access and balance the need for access with the need for security. Regulators and industry must collaborate to agree upon who will need access to serialized data, which data will be able to be accessed and in what context, and how to protect against unauthorized access.

Who has access to stored data?

Centralized Database Architecture

Centralized storage provides benefits and challenges for security and access. A single point of access provides greater opportunities to enforce standardization, to secure and manage individual access, and to define the circumstances under which individuals have permission to access the data. However, a single point of access is also considered more vulnerable to a data breach or other nefarious activity.

Access controls for stored data is equal parts challenging and important. Data access should be reserved for authenticated supply chain trading partners and regulators, and only for specific activities such as verification and tracing to investigate spurious product or to facilitate a recall. Patients or end users should not have access to the central database (though potential options exist for limited patient access to company-owned medication information for patients without routing requests through the central database used by industry and regulators).

Patient-level verification can create significant security concerns because authentication by patients would necessitate a database that is accessible by any person in a country. This would open these secure databases to significant risk of unauthorized access, which would completely undermine supply chain security. Further, it is the position of RxGPS that the ultimate value of serialization is the ability to verify the authenticity of packages before they are dispensed to patients.[3]

Further, trading partners should not be able to see data from other trading partners that has been submitted to the database. Data should only be available from the database for verification or tracing purposes.

Distributed Databases Architecture

Storing data within distributed databases increases the amount of secure access points to supply chain data and limits the amount of data that may be accessed at each point. As compared to a centralized database, there is less risk that a single security breach would threaten the entire supply chain. However, the tradeoff is that each access point must be appropriately secured and managed.

Unlike with a centralized database, distributed databases allow for limited database access, often reserved only for company employees or contracted solutions providers. Other trading partners or regulators would not access data on individual company databases, but rather must rely on a data query to receive data from a distributed database. Given that data queries are limited to verification or tracing requests, distributed databases protect against unfettered access to additional serialized or product data.

How is data access managed?

Centralized Database Architecture

With a centralized database, the general responsibility for data security falls to the agency or entity governing the database and to the registered users uploading their own data. It is critical that all entities that receive access to a central database are properly credentialed (*i.e.*, authenticated). Access to data must be managed from the central governing body responsible for the database by requiring entities to follow a clear credentialing policy.

Further, credentials must be in place at the individual trading partner level for the specific individuals who will require data access. Secure, unique access (*i.e.*, usernames and passwords) can establish user

[3] https://www.rxgpsalliance.org/principles/

access to appropriate areas only and can vary based on the user's level of access and level of credentialing.

Distributed Databases Architecture

Individual trading partners manage their data and govern the access to that data. Security of distributed data also varies by company. Some users may need to be limited to read-only data, while others will need full authoring access; this can be managed through unique usernames and passwords for each authorized user. Ultimately, however, data security and governance are the responsibility of individual trading partners.

Data should only be shared with legitimate supply chain trading partners and regulators, and only in response to a verification and tracing request as required by market regulators. Therefore, it is essential that a market implement requirements to credentialing of legitimate supply chain entities so that trading partners can leverage such credentials to ensure that they are only responding to data requests from authorized, legitimate trading partners. It is also appropriate for regulators to be able to request data from individual trading partners.

Though individual trading partners are responsible for the data in their databases, an overarching data retention policy should be put in place.

3. Data Integrity

Preservation of the integrity of serialized data is essential to the functioning of any system for serialization, verification, or traceability. The commissioned data from the manufacturer (*i.e.*, the identification of units within their production lot) is the "source of truth" for serialization. Any time the commissioned data are transmitted to another database, or when serialization data are derived from another source such as scanning of packages, there are risks for data errors. Data errors, data mismatches, or missing data can lead to false product alerts for legitimate product and potentially result in delays for patient access or destruction of legitimate product. As such, maintaining a standardized level of data integrity should be a priority for industry and regulators alike.

Depending on the data architecture chosen, a system may need to protect against potential errors due to the duplication of data. Further, the frequency of data transmission, and therefore risk of errors, varies between a verification system and a traceability system.

Are data duplicated?

Centralized Database Architecture

A centralized database provides a single location for uploading, storing, and accessing serialization data. In a centralized database system, commissioned data are uploaded directly to the central database by the manufacturer at the time those units are placed in inventory and therefore available for sale. These data could then be used by other supply chain entities and/or regulators, directly, via the same database. Processes like verification and tracing utilize the central database. Data are not duplicated and stored in multiple places. Rather than risking errors or modifications as data are transferred between trading partners, a central database maintains the integrity of the supply chain data that are uploaded. Therefore, it is incumbent upon trading partners to put in place the proper controls to preserve data integrity as data are uploaded.

Distributed Databases Architecture

Distributed database systems increase the complexity of maintaining data integrity since there is no commonly used data source and data must be passed/transmitted between trading partners. As such, maintaining distributed databases necessitates that the same data set is duplicated for storage across entities. Transmitted data between databases increase the chance of errors and mismatches between datasets, which could impact verification and tracing. It is essential that the manufacturer's commissioned data is relied upon as the source of truth.

How frequently are data transmitted?

Traceability System

Within a traceability system, data are frequently passed between supply chain entities and all trading partners must maintain databases of serialized data they own or process. With every product transaction, information about that product must be sent from the seller to the buyer. This increases the likelihood of errors in data capture or transmission for each node in the supply chain. It is therefore all the more critical that systems are in place to handle data errors and exceptions in a way that can quickly ascertain data errors for true product legitimacy concerns and continue to allow legitimate product to flow to patients.

Further, foundational discussions for traceability systems must determine whether traceability will follow a change of ownership or a change of possession model. Capture and/or reporting of traceability data is required for every product transaction (*i.e.*, a change of ownership model) or for every product movement (*i.e.*, a change of possession model), including those within the same company. In a change of ownership model, product movements within entities or by entities that do not take ownership of the product are not required to be captured. As such, a change of possession model requires greater process changes to integrate entities that take possession of a product but not ownership and increases the amount and frequency of serialized data capture, which has important implications for data integrity.

Verification System

Systems that rely on verification reduce the challenge of data integrity given that verification requests should only be routed back to the product manufacturer's database or the manufacturer's data in the central repository. While all dispensers and hospitals would need to be able to make requests back to each product manufacturer's data, a verification system does not require connectivity and data duplication across all supply chain entities.

4. Data Sharing & Communication

The use of serialized data hinges on the communication of these data in a standardized way so that all supply chain entities can interoperably share and efficiently utilize product data. The need for and utility of a common standard for data sharing and communication does not vary across data architectures, however the utility of a standard for visibility and transaction data can vary depending on the supply chain security model chosen.

What is the standard for data sharing and communication?

Traceability System

Should a market choose to pursue traceability across the supply chain, sharing of event and transaction data is essential and must leverage a common "language" and structure for such sharing. Global pharmaceutical industry members have aligned around the use of EPCIS for consistency and efficiency.

EPCIS is a standard that provides a common language for supply chain entities and regulators within and across global markets. EPCIS is a GS1 standard that enables trading partners to share information about the physical movement and status of products as they travel throughout the supply chain (*i.e.*, *what*, *where*, *when*, and *why*).[4] Therefore, leveraging EPCIS creates data consistency across transactions. In addition, EPCIS also defines open, standardized interfaces that allow seamless integration of services. The capture interface allows trading partners to easily capture data from scanned product for upload in a standardized way and the query interface defines how data are requested and delivered.[5]

The EPCIS standard has been widely recognized and utilized by global companies and is the current standard form of communication within various functioning supply chain security systems (*e.g.*, EU, US, India). Given the vast experience of global companies utilizing EPCIS, harmonization of requirements across markets to utilize the EPCIS standard provides helpful efficiencies in implementation, allowing companies to draw upon their prior experiences.

It is important that regulators require the use of a common communication standard to support interoperability. Industry should be consulted in determining which standard to use and should be given ample time to prepare for implementation. As noted above, without a common standard for communication, serialization loses its utility.

Verification System

A system for verification does not require information about product location, movement, possession, or other visibility and transaction data that the EPCIS standard enables. A system for verification is focused on the dispenser (*i.e.*, pharmacy) checking back to the manufacturer's commissioned data to ensure that the unit identifier information on the package is consistent with commissioned data for that

[4] https://www.gs1.org/standards/epcis

[5] https://www.gs1.org/standards/epcis-and-cbv-implementation-guideline/12#2-Overview-of-EPCIS+2-1-What's-in-the-EPCIS-and-CBV-standards?

product package. As such, while a standard approach is still needed, the common information shared or communicated within a verification system can be limited to a subset of the data elements retained.

5. Data Capture & Reporting

Data capture and/or reporting varies according to both the data model chosen and the data architecture.

Centralized Database Architecture, Traceability System

In a centralized database architecture, all data are reported (*i.e.*, sent, communicated) to the central database by supply chain trading partners. For a traceability system, transaction information and visibility data are reported (via EPCIS as described above) by every trading partner in the supply chain at the time of data capture and when a product is shipped (or within some short time period after). When leveraging the reported information for traceability, no additional measures must be taken to gather product information as all data is already housed in one central location. It is up to the body that governs the central database to determine whether the analysis of serialized data (*i.e.*, event data under EPCIS) is performed automatically by the database or whether entities may query the system for reported data to perform their own analyses.

Centralized Database Architecture, Verification System

For a system that relies solely on verification of data within a central database, such as the EU system, reporting of EPCIS data is less necessary. An end-user verification system does not trace product through its supply chain journey, but instead leverages the manufacturer's commissioned data that is created when products are packaged, and a serial number is affixed. As such, only the manufacturer must report data to the central database. Therefore, it may be appropriate, given internal market dynamics, for a regulator to decide to limit the product information reported to include only the information which is necessary to perform verification (*e.g.*, lot number, unit serial number, expiry date, and perhaps a trade identifier). The EU FMD system has determined and required reporting of such a set of data elements. What is critical, however, is that the data elements to be reported are pre-determined well in advance of implementation of any data reporting.

Distributed Databases Architecture, Traceability System

Unlike with a centralized database system where data are reported to a centralized location, in a distributed database system, individual trading partners capture product information and store that information within their own systems.

For a traceability system, transaction information and visibility data are captured by every trading partner in the supply chain. Without a centralized location for product information, creation of a tracing report would require polling multiple companies across the supply chain for information on a particular product. In order to effectively gather such information, there must be a method to determine which individual data repositories must be queried for a specific tracing request (*i.e.*, to determine the chain of ownership for a particular unit).

Distributed Database Architecture, Verification System

For a system that relies solely on verification of data within distributed databases, data capture is predominately required on the part of manufacturers that commission and store serialized data in their individual databases.

6. Master Data Management

Master data refers to descriptive information about a product (*e.g.*, product trade item number, product name, manufacturer name). Unlike EPCIS event data, master data is static. Therefore, it is not necessary to capture/report such data with every transaction if that data can be properly stored and accessed. However, master data management will vary based on the data architecture chosen.

Central Database Architecture

A central database enables master data to be uploaded in standard way and uploaded a single time. Markets utilizing a central database could also choose to leverage the GS1 Global Data Synchronization Network (GDSN) for upload, maintenance, and sharing of product master data.[6]

Distributed Databases Architecture

Without the need to upload master data to a centralized database, decisions about master data management can be left to individual trading partners. However, standardization of master data can allow for supply chain efficiencies and reduce the amount of data that needs to be transmitted with every transaction. Should master data formats be dictated by each individual entity, master data will need to be transmitted, and reformatted, with every transaction, which increases the risk for data errors and adds unnecessary complexity to the supply chain. Alternatively, with a consistent master data structure, master data can be transmitted once (and/or provided via a link rather than necessitating the completion of a full data record). It is also possible to maintain a centralized database of master data within a distributed database system. However, any innovative architectures will require a standardized master data format.

7. Data Accuracy & Alerts

As noted above, it is incumbent upon trading partners to put in place the proper controls to preserve data integrity as data are commissioned and uploaded.

[6] https://www.gs1.org/services/gdsn/how-gdsn-works

Centralized Database Architecture

As a component of the controls to preserve data access and integrity, centralized databases should be constructed to alert trading partners in the event of: data mismatch errors (*i.e.*, product data from a scan does not match the product data in the database), missing data, and duplicated data. It is also possible to enable a central database to produce an alert if a product owned by a distributor or a dispenser was not previously owned by another supply chain entity. Further, centralized databases should be designed to maintain appropriate data access controls and requirements. A component of such controls is ensuring that a central database confirms that serial numbers and GTINs are only uploaded by entities with responsibility for those products.

Should trading partners receive such alerts, requirements should be in place for the quarantine and inspection of product suspected to be counterfeit, diverted, substandard, or otherwise falsified. However, in the start-up phase of any system for supply chain security, there are likely to be errors and alerts that are due to system glitches rather than true product or data discrepancies. As such, regulators should consider appropriate measures to maintain the flow of legitimate products through the supply chain during the start up phase.

Distributed Databases Architecture

It is the responsibility of each company to construct their database systems to manage data errors. Within a distributed database architecture, data checks and alerts are managed on a company-by-company basis.

Akin to the controls in a centralized database system, should trading partners encounter data mismatches, missing data, etc. requirements should be in place for the quarantine and inspection of product suspected to be counterfeit, diverted, substandard, or otherwise falsified. As with any system, there are likely to be errors and alerts that are due to system glitches rather than true product or data discrepancies, and these errors are more likely during the start-up phase. As such, trading partners should implement appropriate controls to maintain the flow of legitimate product through the supply chain and to patients.

Conclusion

There are various foundational considerations for regulators looking to implement a system for pharmaceutical supply chain security that leverages serialization data. While choosing a data architecture (centralized versus distributed) and a model (verification versus traceability) impacts the specific of storage, sharing, capturing, reporting, and communicating data, the need for a standardized approach to data sharing remains constant. Understanding the intricacies of data storage and communication for the architecture and model chosen can allow decisions to be made up front that will increase efficiency, reliability, and interoperability within the supply chain.