

---

## Commentary

Theme: Ligand Binding Assays in the 21st Century Laboratory  
Guest Editors: William Nowatzke, Ago Ahene and Chad Ray

---

# Ligand-Binding Assays in the 21st Century Laboratory: Recommendations for an Automated Data Interchange Process

Sheldon S. Leung,<sup>1,7</sup> Joel Usansky,<sup>2</sup> Robert Lynde,<sup>3</sup> Theingi Thway,<sup>4</sup> Robert Hendricks,<sup>5</sup> and David Rusnak<sup>6</sup>

Received 10 October 2011; accepted 22 December 2011; published online 14 January 2012

**KEY WORDS:** 21 CFR Part 11; automated data interchange; common data standard; ligand-binding assays; metadata.

---

## INTRODUCTION

With the ever increasing speed with which data are generated and the continual implementation of new instrumentation, maximizing the efficiency of data management for ligand-binding assays (LBAs) remains an increasing challenge. At the American Association of Pharmaceutical Scientists Workshop on the 21st Century Bioanalytical Laboratory: Maximizing Quality and Efficiency through Innovation, the attendees recognized the need to address the issue of data management. The eSolutions team, made up of end users and vendors, was formed with this challenge in mind, under the 21st Century Laboratory initiative.

The eSolutions team has identified the need for a fully automated data interchange process as the first step to optimizing data management. This will require at least two major advances. First, for instruments capable of capturing raw data and metadata, a common open-source data standard is

critical. Software independent of instruments will need to be compatible with the common data standard. Secondly, vendors must ensure that instruments and data analysis systems that process LBA data are enabled for direct bidirectional, file-less transfer of data between laboratory information management systems (LIMS), and instrument systems.

In many laboratories conducting assays, LBA data systems and laboratory instruments are often islands of information, separated by an ocean of non-communication. Multiple applications within the laboratories generate reams of data that are stored in separate, non-connected silos of files. Data translation is required to share the data stored within these files with a LIMS or other analytical software. At best, this can be resolved using information technology (IT) resources. However, in the worst case scenario, this translation is performed manually leading to tedious error-prone tasks that require additional quality control (QC) oversight. In order to meet US FDA 21 CFR Part 11 (Electronic Records Electronic Signatures, ERES) regulations, data transfer through a file-based process requires assurance that the data imported into the LIMS are the same data exported from the data acquisition instrument. Therefore, one way to increase laboratory productivity would be through the use of a file-less automated process for data interchange tasks. Automation offers the performance of tasks in a secure, repeatable, and consistent manner without human intervention.

Additional benefits of implementing an automated data interchange include:

1. Facilitation of results and metadata transfer from source laboratory instruments (e.g., plate reader) to a data processing system (e.g., a LIMS)
2. The ability to automate laboratory business processes where the format-consistent data can be programmed to be parsed and read by computer software
3. Consistency of data formats between versions of the same software

---

Guest Editors: William Nowatzke, Ago Ahene and Chad Ray

<sup>1</sup> Pfizer Inc., 1 Burt Rd, Andover, Massachusetts 01810, USA.

<sup>2</sup> Thermo Fisher Scientific, Philadelphia, Pennsylvania 19102, USA.

<sup>3</sup> Hitachi Solutions America, Ltd., South San Francisco, California 94080, USA.

<sup>4</sup> Amgen Inc., Thousand Oaks, California 93012, USA.

<sup>5</sup> Genentech Inc., South San Francisco, California 94080, USA.

<sup>6</sup> BioAgilytix Labs, Inc., Durham, North Carolina 27713, USA.

<sup>7</sup> To whom correspondence should be addressed. (e-mail: sheldon.leung@pfizer.com)

**ABBREVIATIONS:** API, Application programming interface; ERES, Electronic Records and Electronic Signatures regulations (US FDA 21 CFR Part 11); GLP, Good Laboratory Practice regulations (21 CFR Part 58); GUID, Globally unique identifier; IT, Information technology; LBA, Ligand-binding assay; LC/MS, Liquid chromatography/mass spectrometry; LIMS, Laboratory information management system; SDK, Software development kit; USB, Universal serial bus.

4. Data comparability between software applications even from different vendors avoiding the business risk of having to maintain legacy systems in order to retain the ability to read stored proprietary raw data in the future

While many LBA laboratories clearly see a need for an automated data interchange process, these laboratories must raise awareness of this requirement by supporting the eSolutions team in establishing this process as well as supporting those vendors who choose to participate in the effort. It is absolutely necessary to have active vendor participation to be successful. There is a compelling reason for instrument and software vendors to participate in the creation of an automated data interchange process and to support its adoption. Innovative vendors of LIMS and analytical software will benefit from not having to keep up with the ever changing landscape of data formats. The automated data interchange process also allow new instrument vendors to more easily comply with data management requirements of LBA laboratories, thereby increasing the likelihood that innovative technologies will be adopted. On the other hand, LBA laboratories should be willing to accept an increase in cost for applications to account for costs incurred during adoption of the process and to patronize vendors that adopt the automated data interchange process. In this article, we provide perspectives on the benefit to the LBA community (vendors and users) and what is required to establish an automated data interchange process for LBA data.

## END USER PERSPECTIVE

### **Innovation: Adoption of New Technology in the LBA Laboratory**

Continuous improvement in LBA laboratories can involve the assessment of new technology early in instrument and/or software product life cycle. Once an instrument has been evaluated and it is determined that it fulfills a scientific need of a laboratory, it must be integrated into the laboratory's workflow. Integration requires a system evaluation from an IT perspective. When a product is introduced by the vendor, the evaluation often reveals several business risks. The highest risk is the lack of features that enable compliance with 21 CFR Part 11. Noncompliant instruments cannot be used to support Good Laboratory Practice (GLP) studies, making it unlikely to be integrated into either a GLP laboratory or a non-GLP laboratory, which must transfer assays to a GLP laboratory. Therefore, adoption of such an instrument can be contingent on significant resources to ensure that the correct metadata are collected (such as user ID, date and time stamps, as well as critical instrument settings). In addition, many instruments produce data in a proprietary data format that requires translation before importing it into another application for analysis, such as a LIMS. If resources are available, a translation tool can be built, but this requires validation of the tool to support regulated studies. If this approach is not an option, then a manual process is used, which often requires cutting and pasting of data and manual verification of the data. Clearly, a

manual data transfer process is an undesirable situation for regulated support. Finally, proprietary data formats pose a risk with respect to regulated data retention, especially when the data are encoded in a binary or nonhuman readable format, thereby requiring maintenance of obsolete instruments and software applications, or converting data file into another format or paper copy.

These issues could be resolved by an automated data interchange process. A common data standard would establish what information would need to be collected for compliance requirements. With respect to data retention, a common data standard would ensure that newer software would be able to import historical data, eliminating concerns with version updates.

## Increasing the Value of LBA Data

Effective collaboration between groups (internal or external) necessitates the exchange of data. As discussed above, having to add a translation step is not an efficient process unless it can be automated. In most cases, a translator does not exist to reformat data into the desired format, requiring IT support, as described in the previous case, which presents another obstacle to efficiency.

An often overlooked risk is that data translations can lead to the loss of metadata. This can arise if the translation focuses only on raw data or the importing application does not have the ability to accept incoming metadata. Analyzed data in proprietary data formats combined with the loss of metadata impede the ability to retrospectively mine data and will reduce the value of the information generated. Automated data interchange processes would solve these issues by enabling different software applications to seamlessly exchange data and establish the appropriate metadata to associate and store with the raw data.

## VENDOR PERSPECTIVE

### **Benefits of a Common Data Standard**

This section describes the benefits of a common data interchange format from the perspective of a laboratory instrument vendor. Vendors of instruments used in the LBA laboratory not only have to manufacture and commercialize the instrument hardware but also have to provide software that can acquire the instrument's readings, store the raw data, provide compliance with GLP and ERES requirements (where appropriate), and can interface with external data systems.

In late 2007, Hitachi Software Ltd. began developing a multi-platform analysis and reporting tool for absolute quantitative assays called MasterPlex®. The goal was to allow laboratories to standardize on one analysis and reporting tool regardless of which platform or technology generated the raw data. The objectives of the tool were to reduce license, training, and support costs for laboratories that used the MasterPlex® Platform.

Developing MasterPlex® to seamlessly integrate with multiple platforms proved and continues to be a daunting task. At the start, the diversity in LBA instruments mirrored the diversity in output file formats: multiple versions of

output file formats may exist, even for a single application, and the format of the data file may be varied by the instrument vendor as their needs evolve over time. It must be realized that changes even as small as addition or removal of a single comma or space character can adversely affect parsing algorithms which can prevent the file's data contents from being read and interpreted correctly. Therefore, a library of file output converters had to be created for MasterPlex®. For this tool to be successful, the development and release of up-to-date instrument interfaces is a continual and ongoing effort. Platform vendors commonly release upgrades to instrument control software that include changes to the raw data output file not to mention new platforms enter the market with their own variety of file formats. Consequently, the MasterPlex® development team is constantly reacting to the market to release updates to ensure compatibility with these new versions and/or products. Considering there are multiple platforms with multiple versions in use and new versions being released all the time, ensuring compatibility represents a significant cost to the vendor to develop and maintain a product like MasterPlex®. Since the pharmaceutical/biotechnology industries are highly regulated fields, end users and contract research organization (CRO) businesses require that software undergoes a predefined software development life cycle involving project initiation, requirements definition, system analysis, development of code, testing, and release cycles together with appropriate documentation of the software and archival and escrow of the codebase. These costs must ultimately be passed to the end users.

From the perspective of an end user and a multi-platform solution provider like Hitachi Software Ltd., the benefits of a common data standard are clear. However, these benefits may not necessarily be recognized by all vendors. There are a number of reasons why vendors continue to use proprietary or nonstandardized data storage formats and data interchange methods.

First of all, a widely adopted common data standard does not exist. Without the market pressure to deliver products that adhere to an accepted common data standard, software product development will naturally follow the preferences of the development team. For example, the data format used for a new product may be similar to but different from formats used in previous products; as described above, different formats typically require adjustments to the data parsing

algorithm. Developers may use formats they are comfortable with rather than thinking about the impact to the end user or how the data are to be utilized in the next step.

There are also positive and negative economic drivers that contribute to vendors' use of proprietary or nonstandard formats. Developing on a format that meets the minimum requirements of the product may be more cost-effective than trying to adopt a common data standard; a data standard may be complex to understand and implement, and the common specification may be unsuited to the needs of the product. Additionally, the use of proprietary formats can be viewed as a competitive advantage by creating product dependency and increasing the likelihood of future revenue.

Conversely, the cost to shift to a common data standard could be quite large both from a development and support aspect. Firstly, vendors would have to invest manpower to familiarize themselves with a new common data standard and know how to develop software to meet the specifications of that data standard. New versions or entire new products may have to be developed. Software support services would also have to be created to help customers migrate to new versions and new standards.

From a vendor's perspective, the initial economic impact of migrating to a common data format is likely to be significant; however, the long-term benefits of adopting a standardized data transfer mechanism far outweigh the short-term costs. Vendors who are involved from the beginning in such an innovative movement will benefit from having a voice in developing and establishing the functional specifications and scope of the common data standard. Collaborating with consumers will increase brand visibility and credibility for those vendors who choose to lead rather than reluctantly follow.

It is not hard to see that an adopted standard will lead to an improved end user and customer experience. Vendors who are early to the market will enjoy the benefits of creating an improved customer experience leading to brand loyalty and future revenue potential. A common data standard could also facilitate better integration across their own and strategic partner product lines offering cross-selling opportunities. Development costs could be lowered as one development project could be shared across multiple products.

It is interesting to note that the definition of some types of data standards has already been established and are in use in the pharmaceutical industry (Table I). This article

**Table I.** Organizations and Forums that Promote the Standardization of Data Interchange in the Pharmaceutical Industry

Organization	Acronym	Description	Web site
Clinical Data Interchange Standards Consortium	CDISC	Standards for the interchange of clinical, non-clinical, laboratory, and statistical data	<a href="http://www.cdisc.org">www.cdisc.org</a>
Analytical Information Markup Language	AnIML	Data format for analytical chemistry data	<a href="http://animl.sourceforge.net/">http://animl.sourceforge.net/</a>
Generalized Analytical Markup Language	GAML	General purpose data format that can encompass spectrometry data	<a href="http://www.gaml.org">www.gaml.org</a>
Standardization in Laboratory Automation	SILA	Device and data interface standard related to laboratory automation	<a href="http://www.sila.coop/">http://www.sila.coop/</a>
Health Level Seven International	HL7	Messaging standards for recording and transferring specific medical data	<a href="http://www.hl7.org">www.hl7.org</a>
ASTM International	ASTM	netCDF, a standard designed for mass spectrometry data	<a href="http://www.astm.org">www.astm.org</a>

advocates the creation and widespread adoption of analogous standards for instruments and data systems in the area of LBA. While none of the existing data standards were designed specifically to address the needs of LBA end users, further work will reveal whether we may be able to leverage off of these existing data interchange standards.

### File-Less Data Interchange

The solution to the problem of manual file transfer is the usage of modern computer technologies and applications to enable automated data interchange. Several technologies have already been developed and are in widespread use for the internet and commercial applications. Technologies available such as (a) software development kits (SDKs) and application programming interfaces (APIs), (b) automated connectivity tools, and (c) web services. However, in the LBA bioanalytical laboratory, the software controlling laboratory instruments is often not enabled with these advanced interfacing technologies. Therefore, fully automated file-less transfer processes from one system to another cannot be created, and the laboratory users have to rely on the creation of data files or even copy-and-paste for the transfer of data. File-based interfaces can also require extensive and time-consuming manual data verification steps.

Due to the need to adhere to GLP and ERES regulations, file-less instrument interfaces are advantageous because the need to control, securely store, and archive/de-archive a data file is obviated. In contrast to the above

description of file-based instrument interfaces, laboratories that employ chromatographic data systems (CDS) and liquid chromatography/mass spectrometry (LC/MS) instruments typically use a secured transfer of sequence (run and sample) information from the LIMS to the data acquisition system, followed by the secured transfer of results data (*e.g.*, peak area, retention time, *etc.*) back to the LIMS. In the majority of cases, the process starts at the LIMS side because the receipt and login of samples are performed in the LIMS. When the samples are to be analyzed, a set of samples comprising standards, QCs, unknowns, and other samples is assembled into a sequence (an analytical run). Sometimes, these samples are also mapped out onto positions on one or more plates. The analytical run and sample data can then be transmitted directly into the CDS or LC/MS software using a file-less interface. This is possible because these data systems are enabled with an SDK or API. Following data acquisition, the instrument software can transmit the measured response data and its metadata back to the LIMS in a file-less manner.

The types of data and metadata that are useful to be transferred from LIMS to instrument software are shown in Table II. Metadata are information about data, providing context and important ancillary information, and in this situation, it is information regarding the measured responses. Metadata are critical to the laboratory automation process as these provide context around data acquisition, permit the tracking of instrument performance, and provide an audit trail of the data.

**Table II.** Types of Data and Metadata to be Transferred Between LIMS and Lab Instruments

Direction of transfer	Mandatory	Optional
Transfer of worklist information from LIMS to laboratory instrument or liquid handler	Unique sample ID: •Barcode ID, or •Sample ID, or •GUID Replicate number Position on plate •Plate ID •Position on plate (row, column) or position in sequence or Tube ID	Type of sample ( <i>e.g.</i> , standard, quality control, unknowns, blank, <i>etc.</i> ) Project ID and study ID Analytical run number Sample volume Dilution factor Assay type Assay name Plate barcode ID
Transfer of results information from laboratory instrument or liquid handler to LIMS	Unique sample ID: •Barcode ID, or •Sample ID, or •GUID Response measurements	Replicate number Plate ID or plate barcode ID Position on plate Instrument operator Read start date Read finish date Instrument serial number Instrument model number Wavelength <sup>a</sup> Gain <sup>a</sup> OD subtraction <sup>a</sup> Detector ID [multiplexing] <sup>a</sup> Over the limit flag QC state Error conditions Comments

GUID globally unique identifier, ID identifier, LIMS laboratory information management system, QC quality control

<sup>a</sup>Detection parameters are dependent on platform and instrument technology

A critical piece of information to transfer from the LIMS worklist to the instrument is the sample's unique identifier. This same unique identifier can be used on the return trip to transfer the measured responses back to the LIMS. It is important that all sample types, including standards and QCs, have a unique identifier. The unique sample identifier (ID) can take a number or forms based on the originating data system and the target instrument. Typically, the LIMS generates a unique sample ID which can be a Barcode ID, or it could be a composite string of characters describing the subject's characteristics in the study (such as study ID, study's subject ID, treatment, group, visit, biological matrix, time point, *etc.*). Alternatively, a sample in an analytical run can be identified by a combination of information such as Study ID and Run ID and Position in Sequence, or alternatively for 96-well or 384-well plates by a combination of Plate ID and Position in Plate (*e.g.*, A1 and B6). For interfacing to robotic liquid handlers, parameters such as the plate ID and position on the plate are critical for automation.

Some of the instrument parameters are universal, that is, are independent of the physicochemical properties that are measured (*e.g.*, date and time of data acquisition). Other parameters are dependent on the physicochemical or biological properties of the assay methodology (*e.g.*, absorption wavelength for ultraviolet assays, absorption, and emission wavelengths for fluorescent assays).

### Software Development Kits

A SDK is a set of programming tools and specifications that allows software engineers to connect to and interact with an external software application. SDKs provide a controlled "doorway" for external software to interchange data with a target third-party application in an approved, reliable manner. The vendor who creates the instrument software system also creates and defines the SDK for that software. The user of the SDK does not have to be knowledgeable in the inner workings of the target software itself or how the data are stored internally. A user of the SDK usually receives it either directly from the target system developer, or it can sometimes be downloaded from the internet. Vendors often create SDKs to encourage the use of their instruments and software. SDKs also frequently include sample code and supporting technical notes or other supporting documentation to help clarify points from the primary reference material.

The SDK may be an API in the form of some files to interface to a particular programming language. The SDK may also be implemented as web services which allow connection of systems over wide area networks.

It is advantageous if the SDK is constructed in a modular manner. Modularity means that the SDK can be changed when the need arises without changing the main application. This occurs primarily when a link to new third-party software is needed or when the current SDK needs to be enhanced to accommodate new requirements for data transfer (*e.g.*, where parameters that were not specified in the original SDK need to be transferred or when additional features are needed). Modularity reduces implementation and validation costs and effort for both the application vendor and the end user.

It is preferable, but not an absolute requirement, that the SDK specification is published in an open and public manner.

The SDK specifications need to be made available to the people who are performing the system integration. This allows integrators to create interfaces in an independent manner. However, if some vendors decide not to openly publish their interfacing specifications, then an SDK-based interface can be created provided that both parties agree to share information between themselves by virtue of a contractual agreement for licensing and/or a non-disclosure agreement.

### Case Study: Automated Data Interchange

This section describes how an interfacing tool can be used to effect direct, application-to-application data interchange in a file-less manner. This case study uses Thermo Scientific's Integration Manager application as an example of such a tool. Integration Manager (IM) is a flexible data translation and transmission tool that can transfer data from point to point across diverse types of laboratory instruments and software systems. IM operates by collecting/importing data from any data source, converting it internally into an XML data stream, then transmitting it over the network to the target where the XML data stream get converted into a format that the target can accept. The flexibility of IM stems from the two separate transformation agents: one at the source end that converts data from the source's structure into XML and one at the target end that converts the XML to the data format required by the target. Sources and targets can be any type of data in any format such as a data file of any type, LIMS, CDS, web service, SDK, API, serial devices such as a balance, or a computerized system or database. Sources and targets can be connected directly over a local or wide area network, or they can be folders on centralized servers. Sources and targets can be either local or remote to IM.

Most importantly, IM has the capability to perform direct, system-to-system data interchange without human intervention and without the creation of files. Obviously, IM's ability to perform in a file-less manner is dependent on the availability of an SDK for each system to which it should be connected. Once SDKs are established, the benefits of a file-less transfer mechanism become realized.

### Globally Unique Identifiers

One useful concept that is aligned with the idea of an automated data interchange process is the Globally Unique ID (GUID), also known as a Universally Unique Identifier. These special purpose identifiers are 32-character hexadecimal strings which enable computerized systems to uniquely identify a single physical entity. In the context of LBAs, GUIDs could be used to describe samples, child samples, and sample-replicates as the physical samples and/or tubes progress through clinical study conduct from central laboratory, to doctor's office, to bioanalytical laboratory, to clinical data management group, to biostatistics, and finally to submission to regulatory authorities. GUIDs do not require central coordination and could be used to identify samples uniquely. This has practical utility because the bioanalytical laboratory receives samples from multiple sample sources such as pharmaceutical companies, other departments in the same company, universities conducting clinical trials, CROs, and central laboratories, each of which may have their own

computerized system with its own barcode numbering schemes. It can be visualized that systems may have to deal with multiple GUIDs: one generated by itself as well as other (s) imported during a prior sample transfer process. Thus, the ability to exchange multiple GUIDs from one system to another represents a potential breakthrough in developing a global, universal sample exchange mechanism.

### CORRELATION WITH ELECTRONIC RECORDS AND ELECTRONIC SIGNATURES REGULATIONS

In order to provide data to the regulatory authorities such as the US FDA and EMA, all companies operating in the pharmaceutical industry need to be in compliance with GLP 21 CFR Part 58 regulations and Good Manufacturing Practices and Good Clinical Practices. Computerized systems that are used to generate and process data for use in GxP must be compliant with the ERES (US FDA 21 CFR Part 11) regulations. These regulations stipulate the requirements for control over electronic records and cover such aspects such as security, login, independent time-stamped audit trail, raw data storage, electronic signatures, archiving, and retrieval. In order to be compliant with ERES regulations, all stages of

the data acquisition, transfer, storage, and archiving have to be under control. Editing and deleting of raw data must be controlled and secured. Any portions of the process that cannot be secured and controlled must have workarounds created. Developing, validating, maintaining, and using computerized systems for compliance with ERES has been described in many articles and is beyond the scope of this article.

There are three methods that can be used to transfer data from one system to another, to adhere to ERES requirements: (a) direct system-to-system electronic interchange without human intervention, (b) using paper records with manual verification of the data in both systems, or (c) using electronic records with manual verification of the data in both systems. In methods (b) and (c), the process requires human intervention and is therefore slow and labor-intensive and potentially error-prone. Because of the lack of manual or inefficient processes, only method (a) provides streamlined operations and simplified workflows and allows data exchange in a controlled and secure manner that can be audited and verified. Therefore, the ability to enable direct system-to-system data interchange is important to improving laboratory productivity and assuring compliance with ERES. Data can be transferred from one system to another by a number of different techniques and technologies.

**Table III.** A Comparison of Different File Formats Used for the Transfer of LBA Data

Transfer mechanism	File type	Advantages	Disadvantages
File-based	Text file ( <i>e.g.</i> , ASCII file)	<ul style="list-style-type: none"> <li>• Simple to create</li> <li>• Human readable</li> <li>• Easy to parse and read</li> </ul>	<ul style="list-style-type: none"> <li>• Easy to change</li> <li>• Contents can be inadvertently or deliberately changed with no evidence of tampering</li> <li>• Files can be lost, misplaced, or renamed</li> <li>• Files need to be stored securely and archived</li> </ul>
	Binary file	<ul style="list-style-type: none"> <li>• Possible to incorporate evidence of tampering (<i>e.g.</i>, cryptographic hash function)</li> </ul>	<ul style="list-style-type: none"> <li>• Not human readable</li> <li>• Hard to create and read</li> <li>• Files can be lost, misplaced, or renamed</li> <li>• Files need to be stored securely and archived</li> </ul>
	Application file format ( <i>e.g.</i> , Microsoft Excel)	<ul style="list-style-type: none"> <li>• Can be secured with a password</li> </ul>	<ul style="list-style-type: none"> <li>• Easy to change unless password-protected. But a password-protected file may be opened without the knowledge of the password</li> <li>• Dependency on a third party for the definition of the file format</li> <li>• The file format is subject to change at the whim of the application's owner. For example, Microsoft Excel's file type and content structure changed from *.XLS to *.XLSX with Office 2007 suite, thus any interfaces written for XLS would need to be updated for the XLSX file type</li> <li>• Files can be lost, misplaced, or renamed</li> <li>• Files need to be stored securely and archived</li> </ul>
Direct transfer	XML file format	<ul style="list-style-type: none"> <li>• Highly structured</li> <li>• Human readable</li> <li>• Easy to parse and read</li> <li>• XML can be validated against a schema (XSD)</li> </ul>	<ul style="list-style-type: none"> <li>• Files can be lost, misplaced, or renamed</li> <li>• Files need to be stored securely and archived</li> <li>• Files can be lost, misplaced, or renamed</li> <li>• Files need to be stored securely and archived</li> <li>• Verbose</li> </ul>
	Interfacing tool or web service	<ul style="list-style-type: none"> <li>• File-less transfer</li> <li>• No manual intervention</li> <li>• Can be made 21 CFR Part 11 compliant</li> <li>• No files to be lost, misplaced, or renamed</li> <li>• No files to be stored securely or archived</li> </ul>	<ul style="list-style-type: none"> <li>• More costly for vendor to develop</li> <li>• More costly for end user to purchase, deploy, and validate</li> </ul>

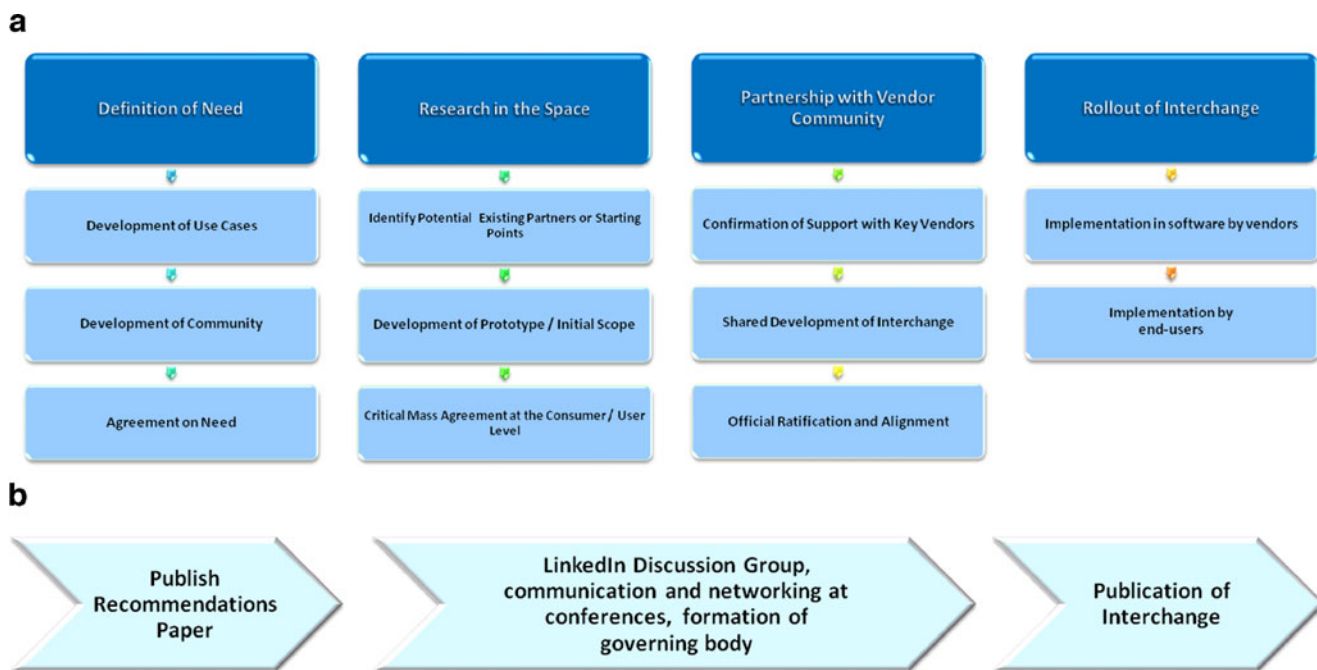


Fig. 1. a Process map and b deliverables for automated interchange development and implementation

The use of a data standard alone is insufficient to completely address laboratory efficiency or ERES requirements. The key to successful automation, and also compliance with ERES, relies on the ability to transfer data and metadata from one system to another in a controlled, secured, and auditable manner. This is because of the weaknesses inherent in any file-based transfer method. Currently, many types of files are used for transfer of data between systems, they may be ASCII text files, binary files, files in a proprietary file format (e.g., Microsoft Excel), or XML. A comparison of these different file formats is shown in Table III. With all file formats, the content and structure of the file is determined by the software vendor and it is usually in a vendor-specific format; thus the receiver needs to know the format of the file's contents in order to be able to parse and read the contents. For compliance with 21 CFR Part 11, the audit trail of electronic data begins at the instant the data hits the durable media (i.e., local computer hard disk or server) (1). Therefore, once a data file has been exported to a local computer hard disk or server and is transferred to a LIMS, verification that the data exported by the data acquisition instrument is the same as the data imported into the LIMS or other data processing system must be performed. Due to the

use of proprietary file formats, the verification of the data is necessarily a manual process.

More importantly, from a GLP perspective, there are major limitations that any file-based transfer method suffers from: alteration of the content, deletion of the file, storage, and long-term archiving and retrieval. Steps can be taken to mitigate some of these issues such as inclusion of a cryptographic hash function in order to detect (but not prevent) tampering, saving of the file onto a secured location (e.g., password-protected server directory that allows users to place files there but not edit or delete them, i.e., a drop folder), and not allowing user to save transfer files into unprotected locations. The aforementioned mitigation steps are dependent on implementation by the end user, not the software vendor.

**DEVELOPMENT OF AN AUTOMATED DATA INTERCHANGE PROCESS**

Any path forward for the industry to adopt a data standard will require the participation of both consumers and vendors. Consumers have a responsibility to create market pressure that encourages vendors to engage in this discussion. Vendors have

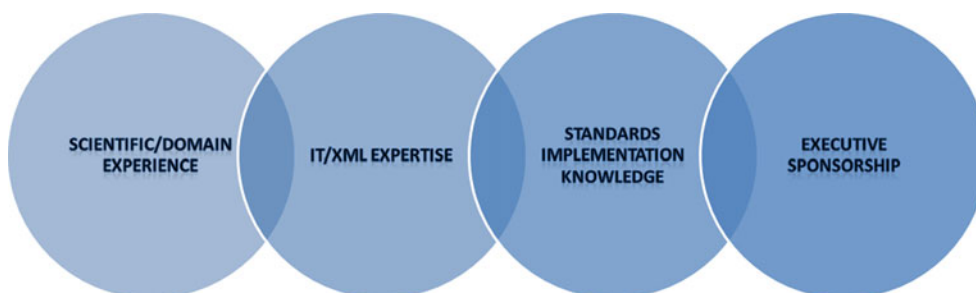


Fig. 2. Team membership needs

the responsibility of listening to their customers and delivering products and services that meet the market needs.

Certainly building the requirements for a universal standard that meet the needs of consumers and vendors alike is no small task and the economic impact would not only be felt by vendors. However, the call for standards is not unique to LBA laboratories and there are many examples both within and outside of the life science market where standards were effectively developed and widely adopted. Consumers and vendors alike would do well to look at similar initiatives such as the adoption of data standards for flow cytometry or the Proteomics Standards Initiative (2,3).

There are many examples of competitive companies collaborating in highly complex and technical industries. One example is the Universal Serial Bus (USB) specification (4,5) which allows external devices such as computer mouse devices, printers, external hard drives, and digital cameras to be easily and readily plugged into computers using “plug-and-play” simplicity. By creating a universal standard such as USB has enabled widespread global adoption of interchangeable devices these collaborating companies have benefitted from increased adoption and reduced effort in maintaining their own proprietary standards. For USB, the standard involved the specification of hardware as well as data interchange protocols. In the pharmaceutical laboratory environment, there is the SILA collaboration among companies that manufacture automated laboratory equipment such as plate and liquid handlers, robotics, and other components that are used in bioanalytical assays. In the case of interfacing LIMS with laboratory instruments, only a data standard and an interchange protocol would need to be developed because these components are typically already connected to the laboratory’s network.

With the advent of the internet, open standards have become a commonplace. However, some of the most successful standards remain proprietary, *e.g.*, USB. The USB Implementers Forum, Inc. (5) is a nonprofit corporation founded by the group of companies that developed the Universal Serial Bus specification. The USB-IF was formed to provide a support organization and forum for the advancement and adoption of Universal Serial Bus technology.

The stages to developing and implementing an automated interchange are depicted in Fig. 1. We propose to establish a new forum to facilitate the development of a data interchange format and to encourage vendors to develop and publish SDKs. Our initial effort to establish a forum will involve supporting discussions within the LinkedIn 21st Century Bioanalytical Laboratory group. A part of the discussion on the forum will be to consider exploring existing mechanisms and data standards such as CDISC, AnIML, GAML, HL7, ASTM, or SILA (Table I) to establish if they are suitable for interfacing laboratory instruments that analyze ligand-binding assays in a bidirectional, file-less manner.

Ultimately, a governing body would have to be created that would act as a liaison between vendors and consumers to build specifications and ultimately release the data standard and interchange protocol. In order to be successful, we will need to recruit volunteers with a wide range of skill sets (Fig. 2). This body would also have the responsibility of creating documentation to assist vendors and end users in adopting the automated data interchange process. Maintaining the process and adapting to changes in

technology and the industry would also be the responsibility of this governing body.

## CONCLUSION

For end users, effective laboratory data management is critical to an efficient LBA laboratory in the twenty-first century. An initial step in this direction is to establish an automated data interchange process, resulting in benefits to the entire LBA community. An automated data interchange process will address many issues with respect to efficiency, including adoption of new technology, compliance, data retention, and creation of relationships within the community during its development. For the vendors, an automated data interchange process will provide clear requirements from the end user enabling vendors to create products aligned with customer needs.

Through this paper, the eSolutions team is working to engage the LBA community to participate in the development of an automated data interchange process. We will look to take advantage of social networking avenues to educate, discuss, and recruit the necessary team members required to develop an automated data interchange process. If you are interested in participating in this effort, please look to join the 21st Century Bioanalytical Laboratory group on LinkedIn ([www.linkedin.com](http://www.linkedin.com)) where the eSolutions team will be hosting discussions on the process, contact any of the authors, or email the corresponding author.

## ACKNOWLEDGMENTS

The authors would like to acknowledge Jim King’s efforts in initiating this effort, Chad Briscoe’s participation on this subcommittee, and Ago Ahene, Marian Kelley, Denise O’Hara, Jean Lee, and Philip Oldfield for their review and comments. This manuscript was prepared by members of the 21st Century Bioanalytical Laboratory: eSolutions Subcommittee of the Ligand Binding Assay Bioanalytical Focus Group (LBABFG).

**Conflict of interest** This article describes two commercial systems: MasterPlex® from Hitachi Solutions America and Integration Manager from Thermo Scientific. Robert Lynde and Joel Usansky are employees of Hitachi Solutions America and Thermo Scientific, respectively. These two systems were provided as examples of commercially available systems that do not provide a complete solution. To the authors’ best knowledge, due to the lack of a universally accepted automated data interchange process, there are currently no commercially available systems that provide the complete solution that we seek.

## REFERENCES

1. [www.21cfrpart11.com/pages/faq/index.htm](http://www.21cfrpart11.com/pages/faq/index.htm)
2. Spidlen J, Gentleman RC, Haaland PD, Langille M, Le Meur N, Ochs MF, *et al.* Data standards for flow cytometry. *OMICS*. 2006;10:209–14.
3. Orchard S, Hermjakob H, Apweiler R. The proteomics standards initiative. *Proteomics*. 2003;3:1374–6.
4. Universal Serial Bus Implementers Forum, Inc.: Universal Serial Bus 3.0 Specification (2011). Hewlett-Packard, Intel, Microsoft, NEC, ST\_Ericsson, Texas Instruments. Accessed 16 Dec 2011
5. Universal Serial Bus Implementers Forum, Inc.: <http://www.usb.org/about>. Accessed 16 Dec 2011