Best Practices in Data Standards Implementation Governance

phuse.global

Contents

Revision History

Version	Date	Summary
1.0	02 Oct 2024	This is the first version.

Introduction

Purpose of this Document

The Best Practices in Data Standards Implementation Governance project, under the PHUSE Optimizing the Use of Data Standards Working Group, was initiated in 2021 to address common industry challenges in data standards implementation governance. The project team started with an assessment of current practices across industry via a survey, designed to gather insight into what level of governance exists, the scope of standards governed, how governance is organised, and what challenges and pain points organisations are experiencing when implementing and governing the standards.

This white paper presents the results of the survey along with feedback from informal discussions at PHUSE's Computational Science Symposium (CSS) sessions in 2022 and 2023. The data and observations shared are intended to inform industry and identify further projects for exploration, while also encouraging industry collaboration for areas which are found to be the most challenging.

Problem Statement

Currently, there is limited sharing of how data standards are developed and implemented across and within organisations. Anecdotally, this is an area of great frustration within companies but is also a roadblock to gaining the efficiencies that data standards are meant to deliver, including automation, which is highly dependent on standardisation.

By discussing and developing cross-industry best practices, it should be possible to address misconceptions about data standards and enable more efficient operation of all areas of clinical and nonclinical data.

Background

Data standards are defined by standards development organisations (SDOs) (e.g. CDISC), though cannot be expected to cover all scenarios, and will generally focus on the areas of largest impact. Additionally, standards development can be a long process, and SDOs do not dictate how standards are implemented within a company's systems and processes. Hence, companies must determine how, when and where in their processes they will apply standards and find ways to address the gaps between what is needed versus what is currently available in published standards. This necessitates some degree of sponsor-specific extensions and customisation.

As a result, many companies have developed their own practices and data standard libraries/repositories, which define their specific implementation of SDO-defined standards (and any extensions to them) and how they are to be implemented within the organisation. Once standards are established, governance is beneficial to enforce compliance and address new situations that arise. Furthermore, sponsor-defined standards must be maintained and updated whenever requirements change, e.g. due to updates from SDOs or regulators, and when new standards are needed due to new study needs, new data sources, or evolving science.

With each company making its own decisions on standards implementation, the use of standards can be inconsistent across companies, and, without proper governance, there may even be inconsistency within a company. Despite all the effort that goes into defining and using standards, the full benefits of standards (content and implementation) may never be achieved without cross-industry collaboration.

This project strives to get a sense of how companies implement and govern data standards (protocol through reporting/submission) and to identify potential best practices across the industry.

Scope

The following major areas were assessed in relation to clinical and/or nonclinical data standards:

- Standards governance organisational structures
- Components of data standards and types of trials governed
- Developing, using and managing data standards
- Building and using data standards libraries/repositories
- Internal compliance of standards
- Managing vendors/partners
- External and non-CRF data
- Managing updates and adoption of new versions
- Providing feedback to SDOs

Topics that are out of scope include how SDOs develop data standards, and company organisational structures.

Overview of Survey

Survey Design and Distribution

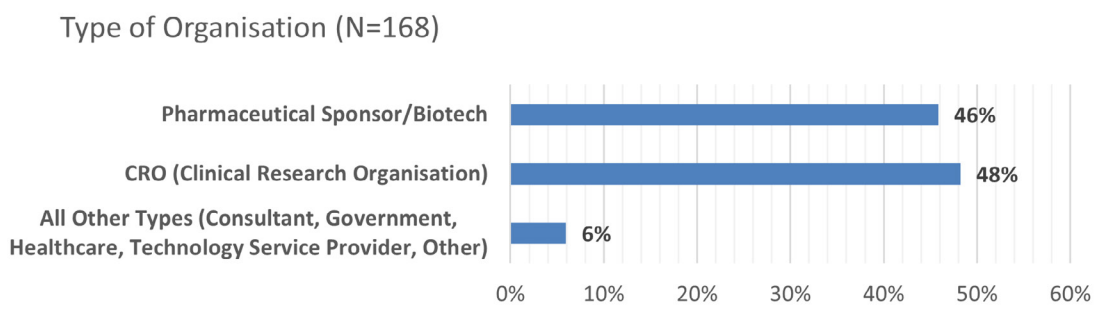
The survey was built using SurveyMonkey and was distributed on 14 February 2022, with entries accepted through to 15 March 2022. There was a total of 33 questions, with five demographic questions, 27 questions on standards implementation and governance (split amongst nine sections), and one free text field asking for any other items the respondent would like to see addressed concerning best practices for data standards governance within an organisation. Of the 27 standards and governance questions, there were 15 single-response type questions, 10 "select all that apply", one "select top 3" and one "select up to 3".

During the open survey period, the survey was promoted via PHUSE social media and by the project team. Responses were requested from anyone involved in data standards governance or from users of data standards, from any functional role. One hundred and seventy-one responses were received.

Respondent Demographics

Of those providing demographics, roughly 94% (excluding blanks) were pharmaceutical sponsors/biotech and clinical research organisations (CROs), with others from government, healthcare, technology service providers, consultants, and other non-categorised organisations (Figure 1). At least 21 organisations provided responses (organisation name was optional and was not provided by all respondents). Clinical data was the primary area of focus for about 98% of respondents (nonclinical focus for the remainder).

Figure 1:



In terms of functional roles represented, most respondents were statistical/clinical programmers or statisticians, with other respondents representing standards management, data management, clinical research scientists/associates, regulatory, clinical operations, or other non-categorised roles (Figure 2). Of the functional role respondents, about 68% are involved with standards governance within their organisations, with about 42% of that subset dedicated full time to standards (Figure 3).

Figure 2:

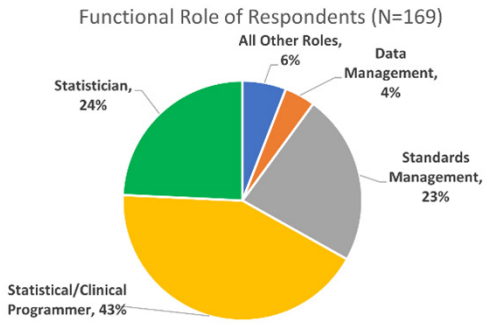
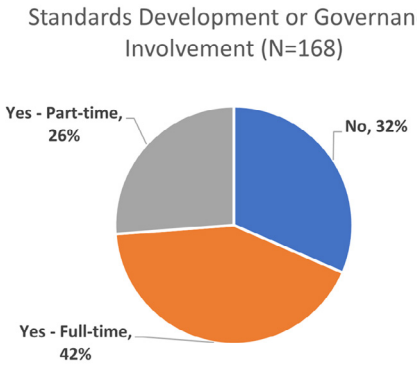


Figure 3:



Analytical Methods

Survey results were exported in Excel format, and data was imported to SAS with some reformatting to support analysis. Note that some data was recategorised as follows:

- Organisation Name (optional field) was collected as free text. Values which were entered in different ways but determined to be the same organisation (e.g. “Company 1” versus “Company 1, Inc.”) were consolidated to a single name.
- Organisation Type: Due to inconsistencies in how pharmaceutical sponsor and biotech were selected within the same organisation, these values were combined into a single type of “Pharmaceutical Sponsor/Biotech”. Clinical Research Organisation (CRO) was not changed, but the remaining types were combined due to low number of respondents in each category (Government, Consultant, Healthcare, Technology Service Provider, Other) and analysed as “All Other Types”.
- Functional Roles of Respondents: Several types which had low numbers of respondents (Regulatory, Clinical Research Scientist/Associate, Clinical Operations, Other) were combined for analysis as “All Other Roles”.

There are some interesting potential trends seen in the data, but with some questions having smaller numbers of responses, it is hard to draw definitive conclusions. Data is generally presented here “as is” (apart from the few exceptions noted above), with denominators noted alongside each analysis. Any analyses based on a subset of respondents (excluding blank responses, for instance) are described accordingly. While there are notable observations included throughout the summary, the number of respondents (Ns) should be taken into consideration relative to the overall number of survey respondents. Certain questions with lower response rates may indicate areas of less confidence, experience, and/or awareness, which may provide valuable insights into the overall state of standards and governance within the industry.

Summary and Observations

Organisational Structures

Within an organisation, governance groups may be "dedicated", "federated", or a combination. For this survey, a dedicated group was defined as being comprised of employees who exclusively work on data standards and a federated group as being comprised of employees who support data standards along with other functional responsibilities. No structural best practice emerged. Forty-six per cent of respondents indicated they had no official governance group, did not know if one existed, or did not respond. Of the remainder, the majority employ a combination of both dedicated and federated (28% of respondents), with others using a dedicated-only (16%) or federated-only (10%) approach (Figure 4).

Figure 4:

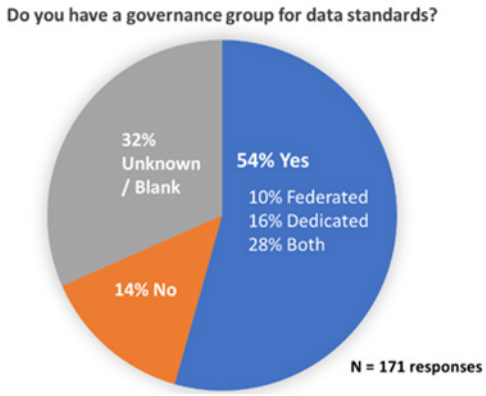
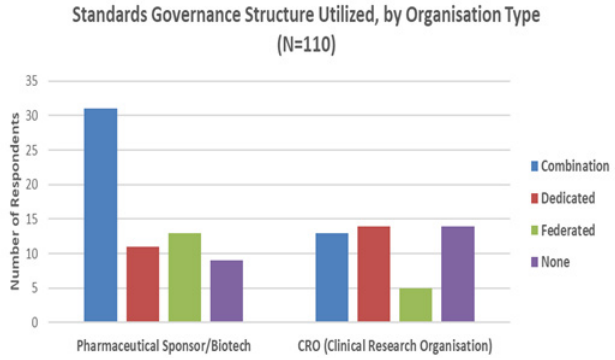


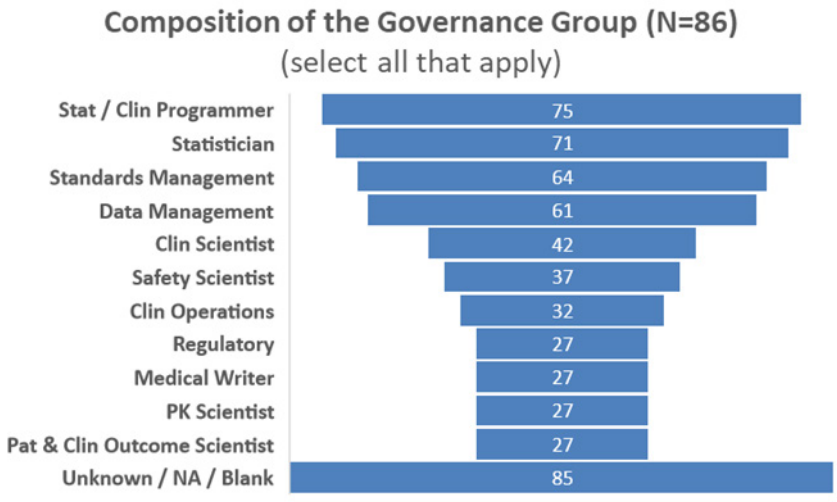
Figure 5:



Notable Observation: Pharma/biotech companies are most likely to use a combined (dedicated plus federated) approach.

Standards governance groups appear to have diverse functional representation, with many comprised of several different roles. The most common responses included statistical/clinical programmers, statisticians, standards managers and data managers (Figure 6). Many CSS attendees agreed that it was beneficial to have representation from different functional lines and areas of expertise, but this may also make it more challenging to make decisions (also see Figure 9, Table 3).

Figure 6:



Notable Observation: Average # roles participating in governance group = 5.6 (for 86 respondents who answered with role details).

While the survey did not assess the practice and impact of using a multi-level governance structure, consideration should be given to clearly defining roles and responsibilities as well as an escalation path in case decisions cannot be agreed within the responsible governance group.

Scope of Governance

Standards governance may be focused on only core/global standards (standards which are used by most studies, e.g. demographics, adverse events, ECG) or may also include therapeutic area (TA) focused standards. Forty-eight per cent of respondents indicated some level of centralised standards governance, with 52% of respondents answering "Other", "Not applicable", "Unknown" or blank (Table 1).

Table 1:

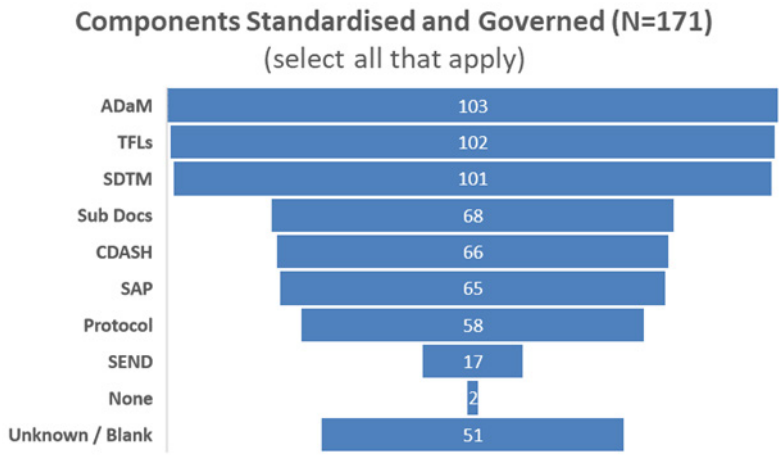
% Respondents (N=171)	Data Standards Governed
29%	Only core/global and therapeutic area standards are governed
13%	Only core/global standards are governed for common forms
6%	Core/global governed by centralised group and therapeutic area by individual function/molecule
52%	Other type of organisation, Unknown/Not applicable, or did not respond

Notable Observation: Excluding "Other", "Not applicable", "Unknown" and blank responses, most respondents centrally govern both core/global and therapeutic area standards.

Some companies noted in CSS discussions that new TA- or indication-focused data points may initially be developed and implemented by the TA or study team and later promoted into company-wide standards if it is determined there would be re-use.

SDTM/ADaM datasets, and Tables, Figures and Listings (TFLs), are generally standardised and governed by most organisations, with a fair amount also governing protocol, statistical analysis plans (SAPs), data collection/CDASH, and submission document standards (Figure 7). The vast majority govern both early-phase (63% of respondents) and late-phase (62% of respondents) interventional studies, with some including non-interventional (20%), low-interventional (9%), and/or nonclinical studies (7%). Note that for some of these components (i.e. protocol, SAP) the "standard" may just be a document template.

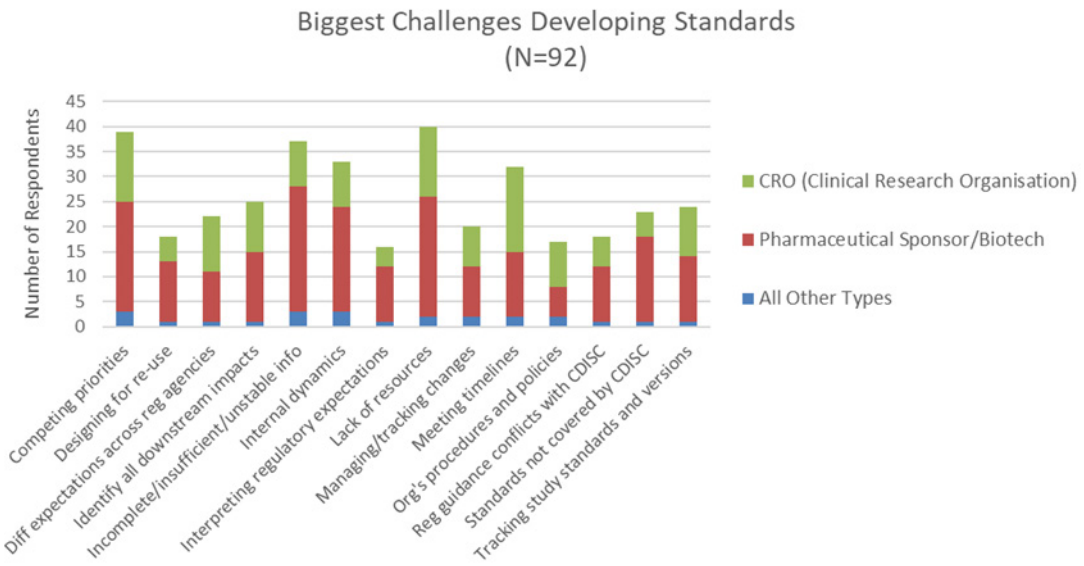
Figure 7:



Developing and Using Standards

When asked about the top challenges encountered in developing and implementing standards, the top selections (Figure 8, Table 2) were lack of resources, competing priorities, incomplete/insufficient or unstable information, internal dynamics (e.g. lack of alignment across groups), and meeting timelines. Additional comments provided by respondents suggest that use of a functional metadata repository (MDR) and automation could help resolve some of these challenges.

Figure 8:



Note that while the question asked respondents to select the "top 3", respondents were able to check all that apply. All responses have been included in the analysis.

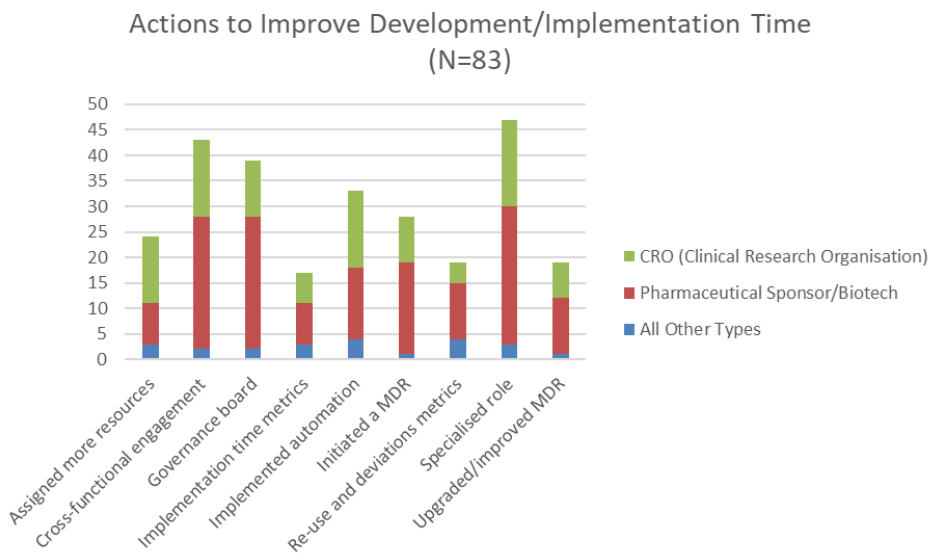
Table 2:

	Top 3 Challenges	
	Pharmaceutical Sponsor/Biotech	**CRO (Clinical Research Organisation)**
1	Incomplete/insufficient or unstable information	Meeting timelines
2	Lack of resources	Lack of resources
3	Competing priorities	Competing priorities

Notable Observation: While all types of organisations experience challenges due to lack of resources and competing priorities, the top challenge for pharma/biotech was incomplete/unstable information, while the top challenge for CROs was meeting timelines.

The top actions to improve standards development and implementation time (Figure 9, Table 3) include use of a specialised role to develop/implement standards, cross-functional engagement, having a governance board to ensure and maximise re-use of existing standards, and implementing automation based on data standards content. While cross-functional engagement was found to improve standards implementation (it also helps to call in experts from other teams and involve all relevant therapeutic areas), it was noted in subsequent discussions that it was easier to make decisions when a smaller group was involved. There needs to be a careful balance between including diverse functional representatives and limiting the number of decision-makers. Education is also a challenging area for many companies but is an important consideration to ensure users outside of those involved in standards decisions have awareness of the standards and related tools to use them.

Figure 9:



Note that while the question asked respondents to select the "up to 3", respondents were able to check all that apply. All responses have been included in the analysis.

Table 3:

	Top 3 Actions	
	Pharmaceutical Sponsor/Biotech	**CRO (Clinical Research Organisation)**
1	Specialised role to develop/implement standards	Specialised role to develop/implement standards
2	Governance board to ensure and maximise re-use of existing standards	Implemented automation based on data standards content
3	Cross-functional engagement	Cross-functional engagement

Notable Observation: The top action indicated by pharma/biotech and CROs was the use of a specialised standards development role within the organisation. Cross-functional engagement was also within the top 3 selected actions; however, for the second choice, pharma/biotech favoured a governance board, while CROs favoured automation.

In discussions at the CSS, it was noted that many companies use some type of tracking for change requests; however, it was mainly a manual effort for some organisations. It was also challenging to go back to determine why a change was made and to track related changes for a standard across different requests. A suggested best practice is to try to automate workflows for change requests and document the full history of discussions that take place, including impact assessments and rationale. Attendees also pointed out that all requested changes should initially be assessed to determine if they are truly needed prior to moving forward for full impact assessment and implementation.

Many companies allow some degree of customisation without requiring approval from the standards governance board (Figure 10). Most (61% of those who answered, excluding "not applicable" responses) allow for optional fields or terminology to be removed. Forty-eight per cent indicated that studies may add new fields or terms without needing governance approval, although in later discussions it was determined that this is mainly the allowance of new terms to be added to extensible code lists. Smaller numbers indicated that labels can be modified (28%), or mandatory fields can be removed (9%). A majority (76%) indicated that any study-level deviations from standards must be reviewed and approved by the standards governance group or a designee (Figure 11).

Figure 10:

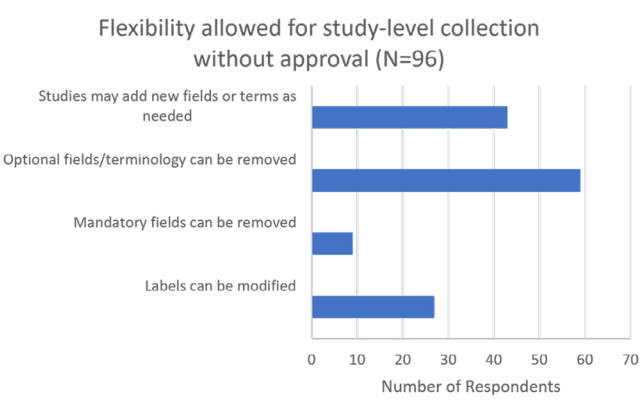
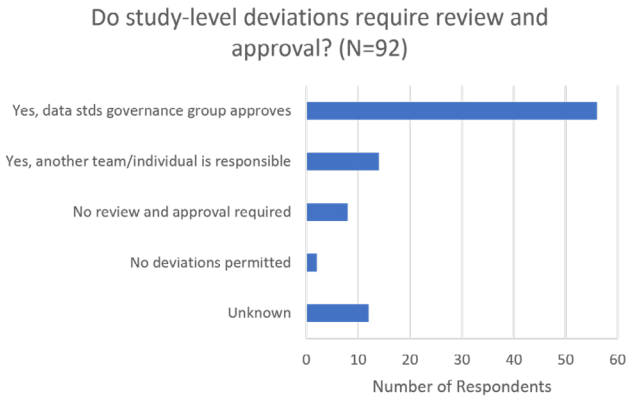


Figure 11:



One of the challenges noted in discussions at the CSS was how to maximise re-use of standards while accommodating study needs across therapeutic areas and indications. Allowing some degree of flexibility in the standards can help reduce the need for multiple variations of standards (and maintenance of those standards) and reduce requests for changes or deviations (and thus the burden on data standards governance groups, particularly for larger organisations). Greater flexibility can present a challenge for downstream automation, and organisations must still use caution to ensure compliance and quality.

Standards Metadata Libraries/Repositories

In terms of storing and maintaining standards metadata (respondents were asked how they primarily store data standards metadata and were allowed to select one response), the majority (32% of those who answered the question) primarily use spreadsheets or some other document-based approach. Around 25% are using an off-the-shelf or customised commercial MDR (Figure 12). Pharmaceutical sponsors/biotech companies indicated higher usage of commercial MDRs compared to CROs (42% of pharma/biotech responses versus 7% of CRO responses, Table 4).

Figure 12:

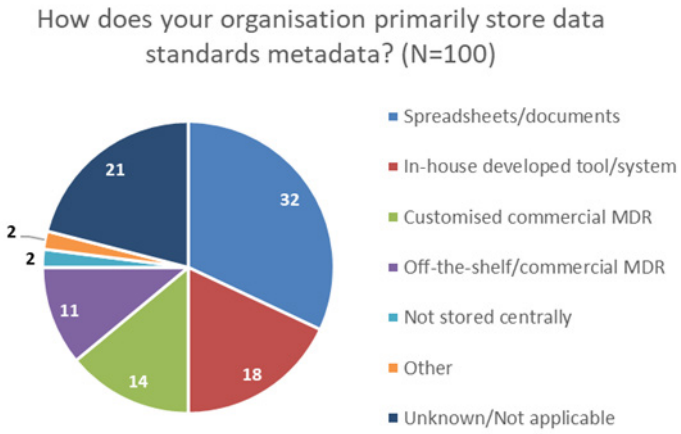


Table 4:

Primary Standards Metadata Storage Method	N	
	CRO	Pharma/Biotech
Spreadsheets/documents	12	18
In-house developed tool/system	10	8
Customised commercial MDR	1	13
Off-the-shelf/commercial MDR	2	9
Other	2	0
Not stored centrally	0	2
Unknown/not applicable	16	2

In discussions at the CSS, many attendees acknowledged that they use spreadsheets because it is easier than going through their MDR (end users were generally more comfortable working with documents) or because they do not have an MDR. Anecdotally, this was an area of much frustration. One survey respondent commented: "State of MDR solutions leaves a lot to be desired in terms of functionality." Commercial MDRs were generally found to be difficult to navigate and use in support of automation, were large investments in time and money, and did not seem to be designed with all end users in mind. Several companies using MDRs noted they export the metadata into documents for easier/quick reference.

While this question focused on the primary tool for storing metadata, discussions acknowledged that some companies may actually be using a combination or hybrid approach where different components are stored in different systems.

Despite the challenges with establishing an effective standards metadata library, companies can realise benefits from their efforts. Out of nine pre-specified potential benefits of using a standards library (respondents could select all that applied), most respondents (60% or more of those who responded) noted benefits of reduced cycle time to set up studies as well as to generate outputs, automation, improved quality, and improved traceability/documentation (Figure 13). These benefits can result in increased productivity, efficiency and quality, while reducing time and costs.

Figure 13:

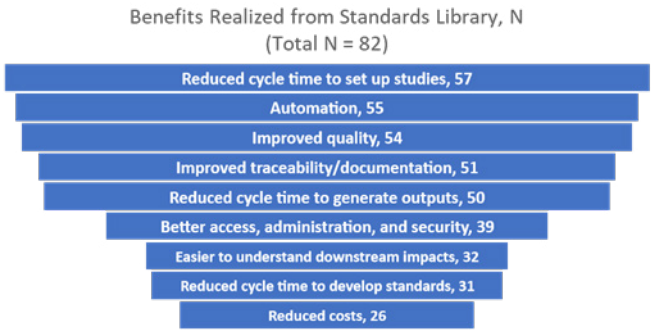


Figure 14:

Number of Benefits by Primary Metadata Storage (N=82)

Primary Metadata Storage	Average Number Benefits Selected
In-house developed tool/system	5.4
Customised commercial MDR	4.6
Spreadsheets/documents	4.5
Off-the-shelf/ commercial MDR	4.4
Not stored centrally	3

Notable Observation: Those who selected "In-house development tool/system" for their primary standards metadata storage indicated the most benefits on average, while those who selected "Not stored centrally" indicated the fewest benefits.

Based on the survey results and CSS discussions, recommended best practices are to make standards easily accessible, centrally located and user-friendly. Optimally, standards metadata should be searchable, filterable and consolidated. It is important users can get to the information they need easily and quickly and understand how to use it. Education is key to enabling this.

Internal Compliance

When it comes to validating a study against organisation (or sponsor) standards, the most commonly used method selected (Figure 15) was in-house automated tools and macros (86% of the respondents who selected at least one type of validation, excluding unknown and no validation responses) followed by manual review (55%) and use of commercial automated tools (42%). While many companies are still conducting some level of manual review, 38 of the 41 who selected manual review are doing so in combination with automated checks (Figure 16, Table 5).

Figure 15:

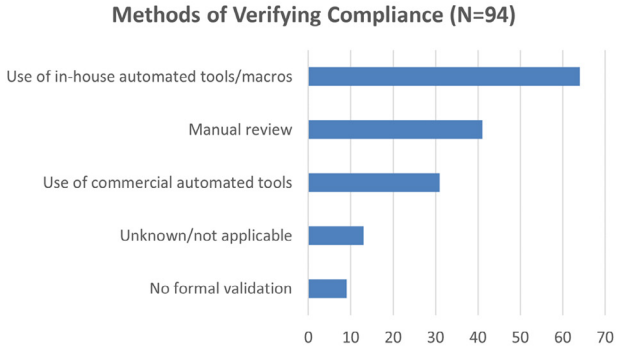


Figure 16:

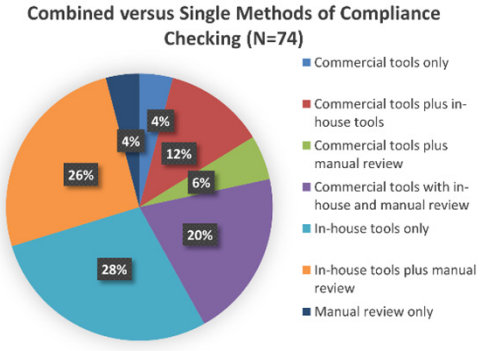
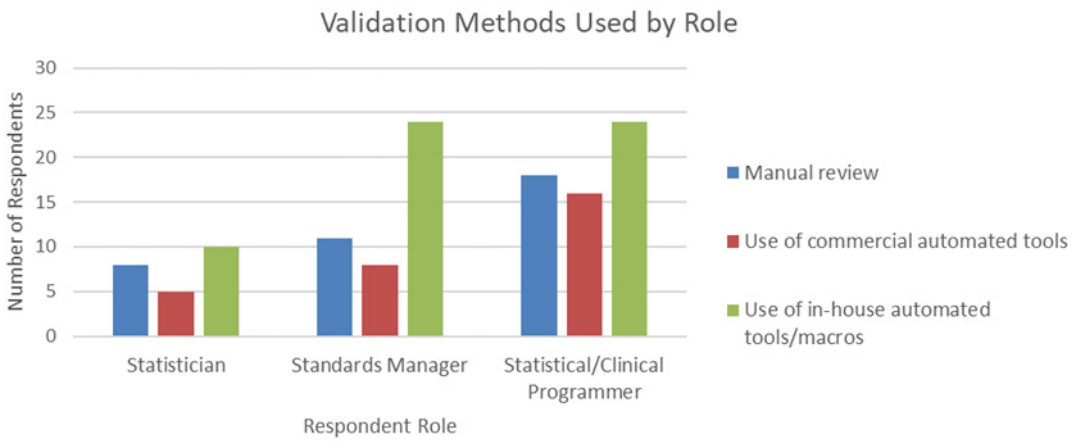


Table 5:

Type of Organisation (N=74)	Number of Methods Used for Compliance Checking (N)		
	1 Method	2 Methods	3 Methods
Pharmaceutical Sponsor/Biotech	20	14	5
CRO (Clinical Research Organisation)	6	16	9
All Other Types	1	2	1

Notable Observation: More than one compliance-checking method could be selected, and 64% indicated using a combination of commercial tools, in-house tools and/or manual review – with CROs usually employing a variety of methods. While it is unclear if "more is better", it does seem to indicate that automated tools alone are not sufficient.

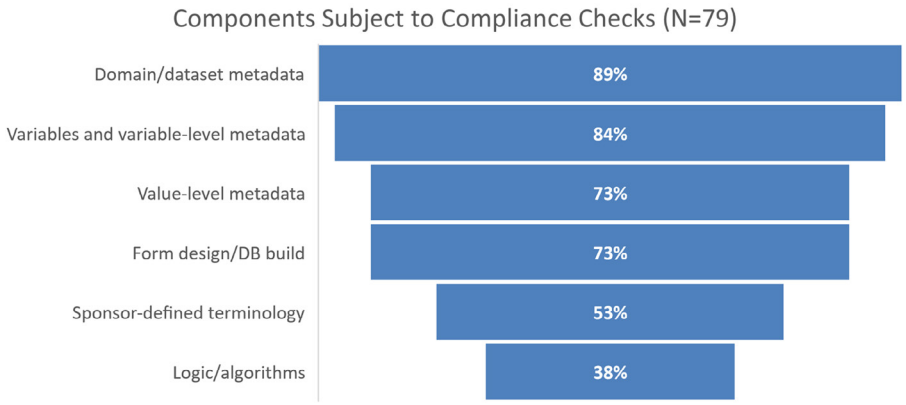
Figure 17:



Notable Observation: Standards managers are most likely to use in-house validation tools whereas programmers and statisticians use a mix of commercial tools, in-house tools and manual review. These role-based differences may be due to differences in scope of validation/compliance checking.

Most respondents (60% or higher) noted that domain/dataset metadata, variable-level metadata, value-level metadata, and form design/database build are subject to internal compliance checks. More than half of respondents also include sponsor-defined terminology and at least one third include logic/algorithms.

Figure 18:



Notable Observation: As expected, submission data components (for SDTM/ADaM datasets) are most likely to be checked for compliance; however, a high amount also check form design and database specifications.

The majority (42%) of those surveyed indicated that an individual study is typically in the range of >75–90% compliant with standards, with 27% noting >90% compliance (Figure 19).

Figure 19:

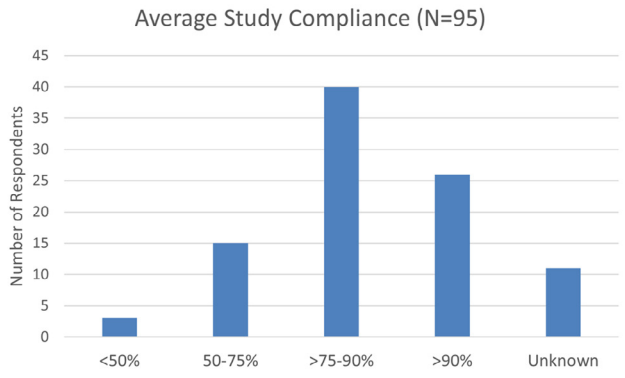


Table 6:

Estimated Study Compliance				
Role	**<50%**	**50-75%**	**>75-90%**	**>90%**
Standards Manager	1	8	19	1
Statistical/Clinical Programmer	2	4	13	16
Statistician	0	1	6	7
All Other Roles	0	2	2	2
Involvement in Standards Governance	**<50%**	**50-75%**	**>75-90%**	**>90%**
Yes, full-time	0	0	26	8
Yes, part-time	3	7	8	4
No	0	2	6	14

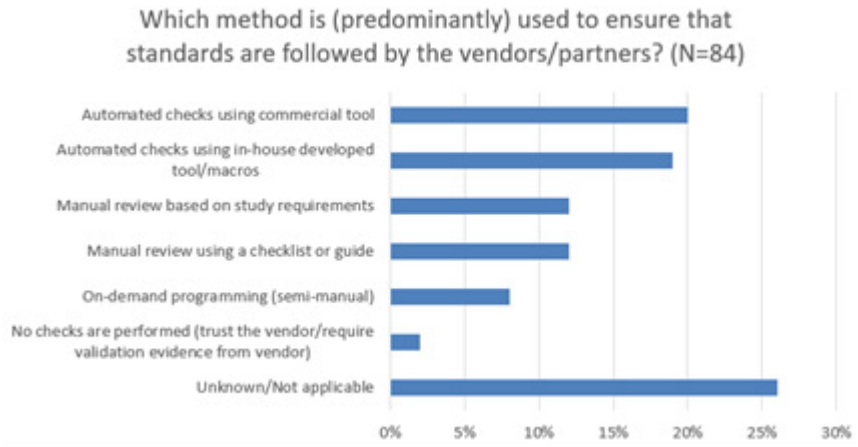
Notable Observation: Those involved in standards management tended to rate study compliance in the >75–90% range, while those who are not involved tended to rate it towards >90%.

Enforcement of standards was discussed at the CSS. Study teams are expected to follow the organisation's (or sponsor's) standards, and some companies have tools to check the level of compliance to some degree. Tools may provide a compliance score or identify non-standard elements, but generally they verify whether standards have been used and cannot necessarily tell if they have been used correctly. Teams are expected to take action to address findings/low scores; however, there is no active "policing" in most cases to ensure studies are at an acceptable level of compliance with appropriate usage. There are generally no penalties for non-compliance apart from having to document general CDISC compliance findings in the data reviewer's guides for regulatory submissions. Assessing compliance is also challenging when there are multiple active versions of standards in an organisation's library. To be most effective, organisations cannot just govern the development of standards, but need to also govern implementation within studies. While it is helpful to have tools in place to check compliance to the organisation's standards, there is no clear best practice on how to check and enforce correct use of the standards within a study.

Managing Vendors/Partners

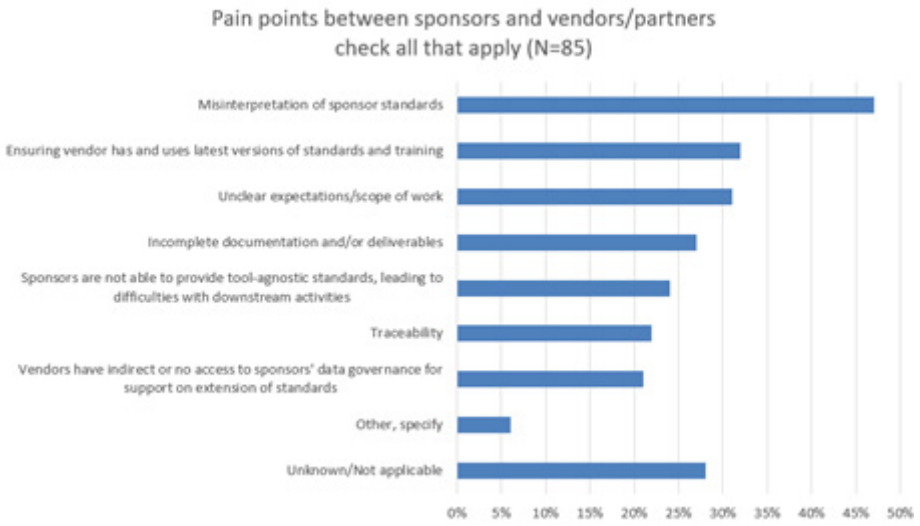
As many sponsors are working with vendors/partners to support study work, we asked several questions on how incoming data from vendors is checked for compliance. Most respondents (~40% of those who use vendors/partners) indicate automated checking through commercial or in-house tools is the primary means to check data provided by vendors. Roughly 25% use manual methods of review (Figure 20).

Figure 20:



The top 3 pain points for implementing and governing standards between sponsors and vendors/partners selected by respondents (Figure 21) are misinterpretation of sponsor standards, ensuring vendors/partners use the latest versions of standards (including training), and unclear expectations/scope of work, with each receiving at least 30% of respondents selecting. Additional pain points selected include incomplete documentation and/or deliverables, not providing tool-agnostic standards, traceability, and indirect or no access to sponsors' data governance, which received 20% or more in each category. There were several comments concerning high employee turnover within vendors/partners and lack of sufficient experience/training, which are additional pain points.

Figure 21:

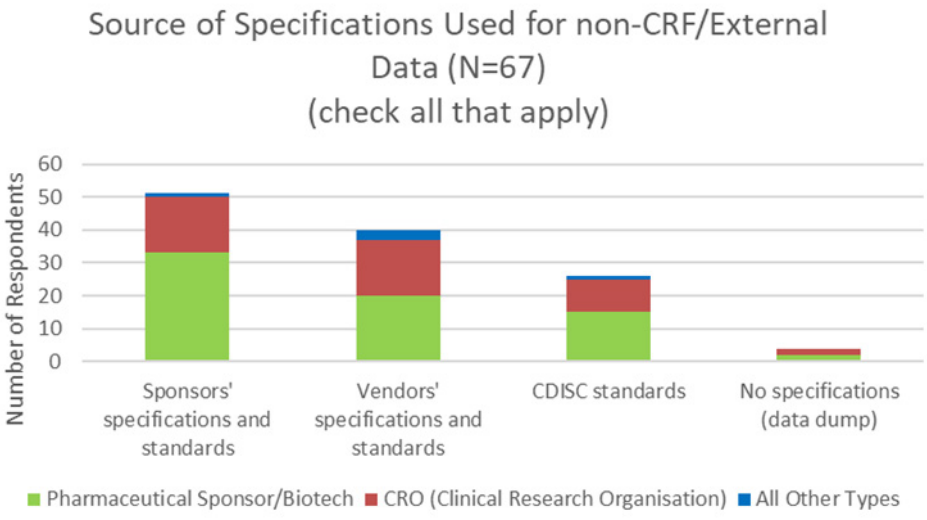


Best practices recommended by attendees at the CSS included sharing lessons learned after major deliveries at the vendor level and study level, having open lines of dialogue and regular check-ins, and working as a team (versus just handing off the work and awaiting delivery). It was also noted that there are benefits to long-term relationships with vendors.

External (Non-CRF) Data

For organisations which use vendors to provide non-CRF/external data, 67% of those responding (excluding unknown/not applicable) indicated that vendors are expected to use the sponsors' specifications/standards for transferring data. Sixty per cent indicated that the vendors' specifications/standards are used, and 39% use CDISC standards (Figure 22). Note that more than one source could be selected, and more than half of the respondents (54%) selected two or three sources of specifications.

Figure 22:



Notable Observation: With 54% of respondents selecting multiple sources of standards, this suggests that different sources of standards may be used across different programmes or studies, and the source used may depend on the type of vendor/the vendors' capabilities.

Figure 23:

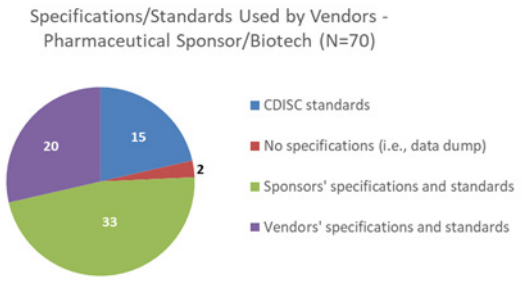
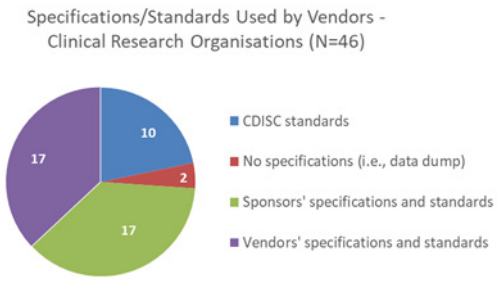


Figure 24:



Notable Observation: Pharmaceutical sponsors were most likely to use sponsors' specifications/standards, while CROs were evenly split between sponsors' and vendors' specifications/standards. Both use general CDISC standards less frequently.

Most of those responding (65%) check both contents and structure when non-CRF data is received from a vendor; for a few respondents, only structure or only contents are checked (Figure 25). The majority use a combination of automated, on-demand and manual checks (Figure 26).

CSS attendees noted that certain data collection such as eCOA/ePRO (electronic clinical outcome assessments/electronic patient-reported outcomes) may be challenging to align with standards and that some vendors have challenges complying with controlled terminology.

Figure 25:

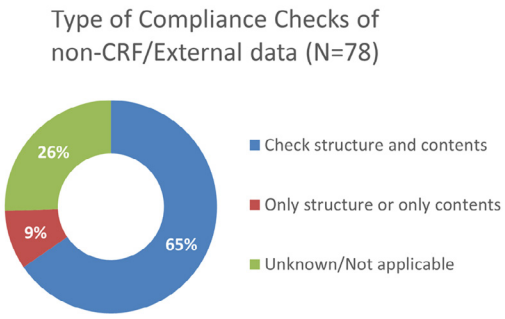
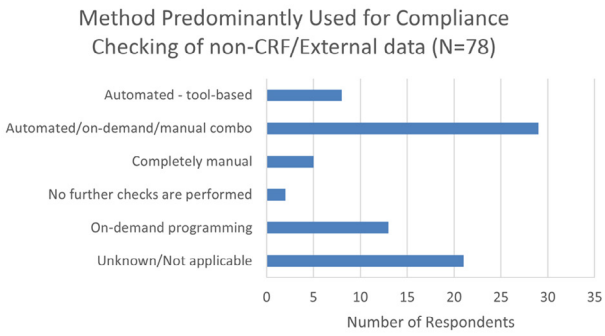


Figure 26:



Note that one blank response for type of compliance checks was assigned to the Unknown/Not applicable category (Figure 25) for analysis due to the response to the subsequent question identifying a method of compliance checking (Figure 26) so that this pair of questions could be assessed using the same denominator.

Notable Observation: While there is still a lot of manual review being performed (44% completely manual or in combination with other checks), it is not typically the only method used for checking compliance. Most often, organisations use a combination of automated, on-demand, and/or manual compliance checks.

Managing Updates/Versions

The primary driver for initiating an update to an organisation's standards library (Figure 27) was the release of a new standard from an SDO (i.e. CDISC), but also appears to be influenced by programme/study needs. The survey questions summarised below and discussion at the CSS revealed that updating a standards library is often a complex activity, depending on multiple factors such as the scope and complexity of the changes, the size of the organisation and the breadth of the therapeutic areas.

Figure 27:

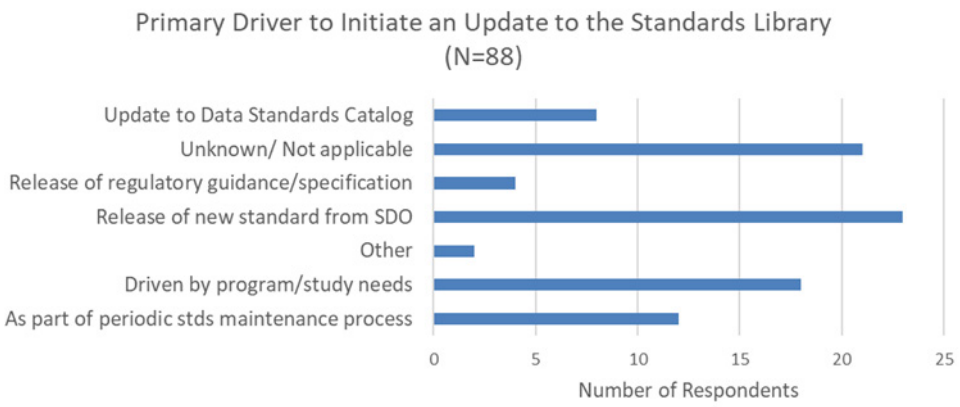
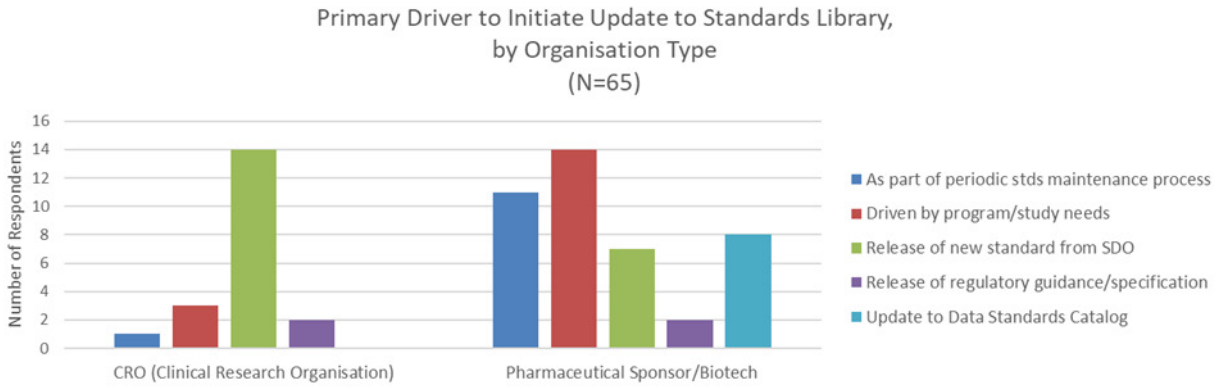


Figure 28:



Notable Observation: Excluding Unknown/Other/NA/blank, CROs tend to be driven by release of a new standard from an SDO whereas pharma/biotech companies are primarily driven by programme/study needs or as part of a regular standards maintenance process.

Companies which provide services to pharmaceutical sponsor/biotech organisations may be more proactive in preparing to adopt new standards, while sponsor companies may be more reactive (implementing as needed/required) or may roll changes into their regular maintenance schedules. Sponsors may also be incorporating components from newer versions while officially using older versions of a standard.

When deciding to move to new standards (whether by choice or requirement), companies must assess for impact on ongoing studies. If multiple versions of standards are required to support the portfolio (due to different studies being in different versions), the organisation must determine a process for managing the different versions via its library/repository or other method. This has been an area of challenge for most companies. Furthermore, when deciding what the company's active standards versions will be, consideration needs to be given to compounds/clinical programmes which may want to maintain a consistent version across studies. Programmes which have studies in different versions may have to upversion one or more studies for integrated analyses, which adds challenge to submission preparations and can result in discrepancies between integrated data and the individual studies.

Regarding timing to start implementing a new standards version, there is no clear best practice/common strategy. Of those responding, there is an almost even split between three of the responses ("Once health authority announces plans to start accepting", "Following release of the standard by the SDO", and "Between Date Support Begins and Date Requirement Begins"), as shown in Figure 29. This is likely because all three responses contribute to the decision. Note that definitions of each time point were intended to align with the FDA's Data Standards Catalog (DSC),[1] but were not specifically defined in the survey.

Figure 29:

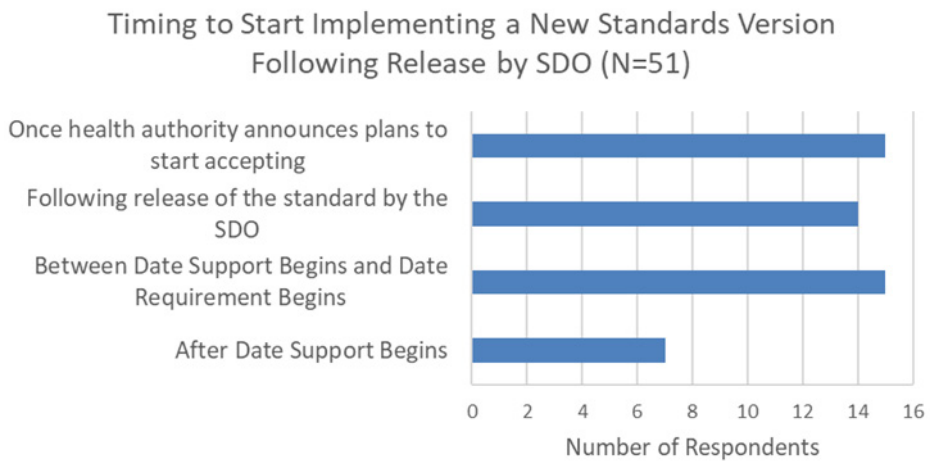
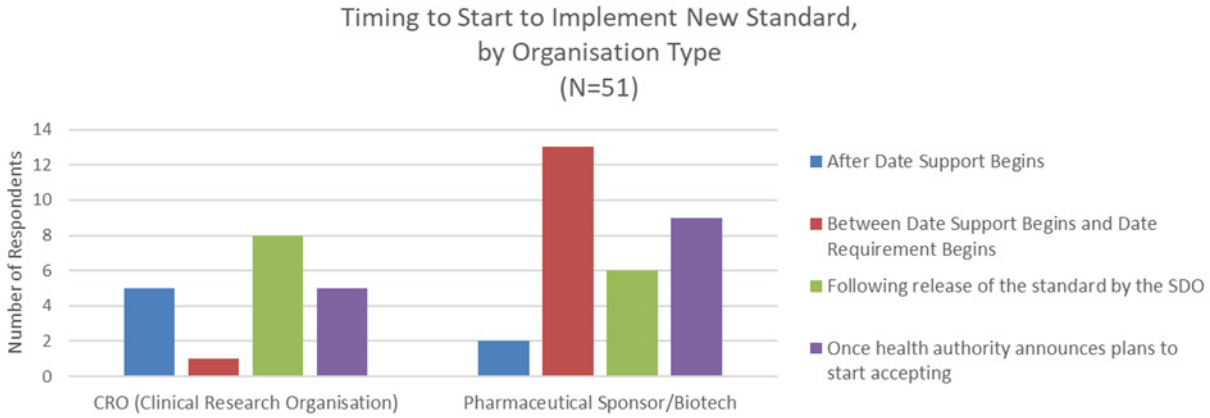


Figure 30:



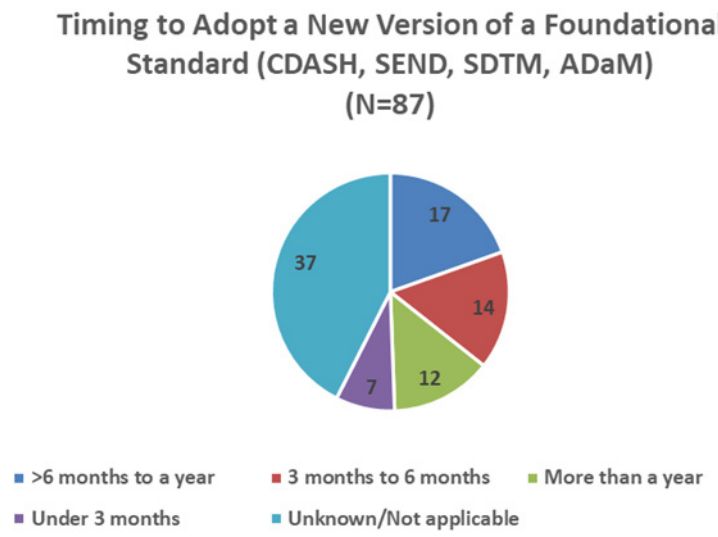
Notable Observation: Excluding Unknown/Other/NA/blank, CROs tend to start following the release of a new standard from an SDO whereas pharma/biotech companies tend to be driven by regulatory dates.

Discussions at the CSS revealed several considerations contribute towards the time required to implement a new standards version. Companies must assess the extent of the changes and whether standards impact would apply end to end or could be managed on the back end. When standards are applied end to end, metadata must drive the update, and all associated standards must be updated at once due to dependencies (this cannot be done in pieces, i.e. SDTM now, ADaM later). Another consideration is backwards compatibility, which can be a challenge. Other factors which may affect this decision are the size of an organisation and the breadth of therapeutic areas involved.

The timing needed to adopt a new version of a foundational standard (such as CDASH, SEND, SDTM, ADaM) generally seems to take >6 months to a year (Figure 31), but this may be dependent on the scope of changes and systems impacted. For many (43% of respondents), the duration was unknown or perhaps too variable to select. Based on CSS discussions, this can be a major undertaking (some of the challenges have already been mentioned above). An additional factor for some organisations is the complexity of outsourcing the work of standards implementation.

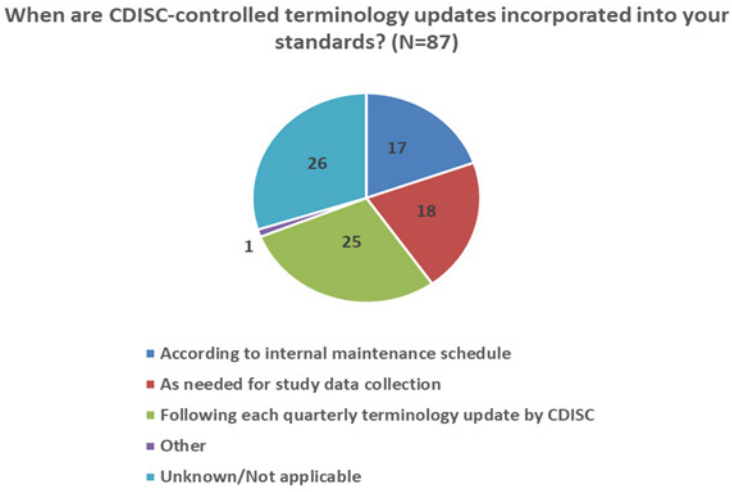
A complete end-to-end impact assessment should be performed following the release of a new standards version by an SDO, to understand the extent of the changes and the scope of the work involved. Companies should be mindful of when a health authority plans to start accepting a new standard, but ultimately need to be prepared to implement a new version in time for the “Date Requirement Begins” for a standard in the DSC. Most companies begin this initial effort following a release from the SDO but prior to the new standard being added in the DSC. Once this assessment is completed to determine scope and effort, companies should work backwards from the “Date Requirement Begins” to build out development and implementation timelines, building in plenty of extra time for unknowns/unanticipated challenges.

Figure 31:



Terminology updates are generally less burdensome, but still require assessment for impact on any ongoing or completed studies. Updates are applied following each quarterly update from CDISC at some companies, while others update as needed for study data collection or based on an internal maintenance schedule (Figure 32).

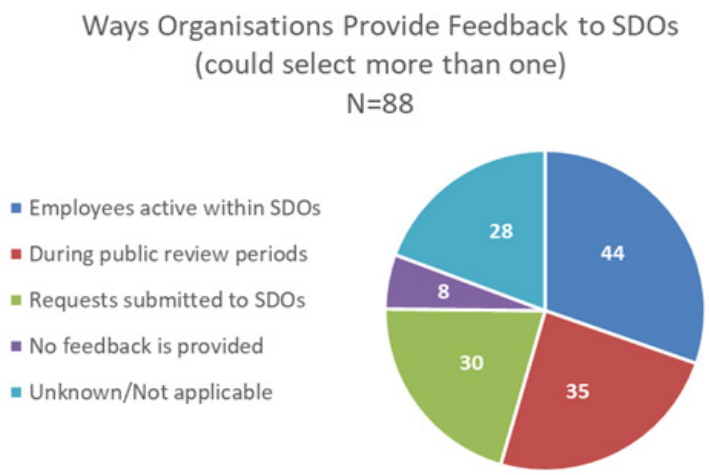
Figure 32:



Providing Feedback to Standards Development Organisations

Of the 89 responses, 53 indicated there was some level of company participation within SDOs (Figure 33): through active participation in SDO teams and activities (83%), during public review periods (66%), or by submitting requests to SDOs (57%). Most of those respondents selected multiple types of participation (68% responding with at least two types). There was no distinct difference in the level of or type of participation between pharmaceutical/biotech companies and CROs.

Figure 33:



For organisations submitting requests to SDOs (e.g. new term requests to the National Cancer Institute (NCI)), decisions must be made for how to proceed while waiting for a response on the request (Figure 34). Of the 83 respondents, 30% indicated using existing standards in the meantime, and 24% indicated using the new proposal/internal standard at risk that the request may be denied. There may not be a clear singular strategy, as many factors could impact the decision, such as timelines and whether there are similar existing standards that may be leveraged. For those using existing standards, 64% of the respondents are CROs and Others, while 36% are pharmaceutical/biotechnology sponsors. For those using the new proposal/internal standard at risk, 95% of the respondents are pharmaceutical sponsors/biotechnology companies (Figure 35).

Figure 34:

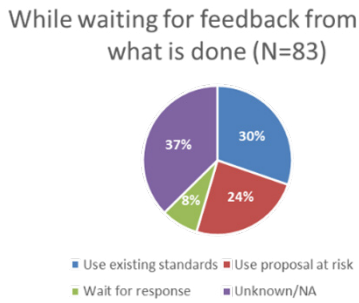
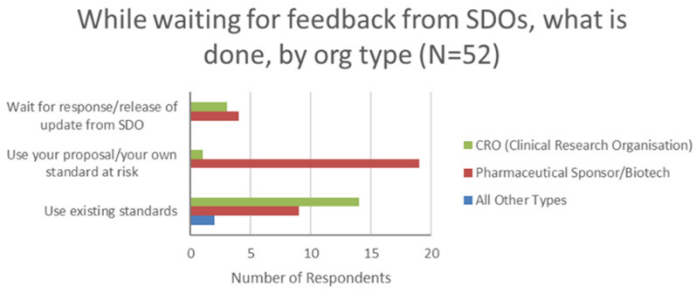


Figure 35:



Notable Observation: Pharmaceutical/biotech companies are more likely to use their own proposal/standard at risk (may be due to having pre-submission discussions with regulatory agencies), while CROs appear to be more conservative and tend to use existing standards.

When using new proposals at risk, it is recommended to review existing standards/terminology and create new standards which align as closely as possible. Requests for new standards should be submitted to SDOs, and/or new terminology should be submitted to the NCI; however, there was feedback that it is challenging to obtain resources to submit requests. If a request is rejected or another standard or term is recommended by the SDO, companies generally apply the recommendation to new studies. Older studies may be remapped (collection remains the same, but values are mapped over to the recommended standard in SDTM) or companies may decide to continue using the proposed standard based on timing, analysis requirements, and submission and integration plans.

For those trying to use existing standards when no standard currently exists, CSS attendees responded that they are generally making use of extensible code lists (sponsor extensions to terminology) or using supplemental qualifiers or custom domains.

Conclusion

This paper summarises the results of a cross-industry survey on current practices in developing, implementing and governing data standards. The main conclusion is that data standards are very beneficial towards gaining efficiencies in study set-up time, data quality, and automation, yet adoption of standards by individual companies (including sponsor extensions which are dependent on study needs) can present multiple challenges. While there is no single best practice or process flow that can be recommended based on the survey results, a few highlights can be made. The keys to success might include having specialised standards developers/implementers and a centralised standards governance team with a robust, cross-functional governance process, along with maintaining a user-friendly standards library and thorough user training. Due to the complexity of standards implementation and expertise required to comprehend the standards, having subject matter experts devoted to standards implementation, governance and training certainly aids in overcoming major challenges with standards adoption, thus increasing studies' standards compliance. Equally important is storing standards in one central location with user-friendly functionalities: searchable, filterable and consolidated. As a result of these functional requirements, the most common form of a library identified in the survey is a spreadsheet or some other document-based approach. Of note, implementation of off-the-shelf or customised commercial MDRs has been a notable trend, yet companies' journeys towards a functional MDR have proven difficult and are not overly successful to date.

Potential areas identified for collaboration include upversioning standards, standards libraries/repositories (MDRs), internal standards compliance/enforcement, and automation/AI for data standards. Improving, expanding and integrating MDRs may be an area of great opportunity in the future as organisations seek to incorporate artificial intelligence and automation. Furthermore, as clinical trial design, along with data collection and analysis requirements, continues to evolve over time with advancements in science and technology, it is critical to provide support to SDOs so that the standards can keep up with the industry's needs and reduce the gaps that result in sponsor extensions and customisation.

Disclaimer

The opinions expressed in this document are those of the authors and should not be construed to represent the opinions of PHUSE members, respective companies/organisations or regulators' views or policies. The content in this document should not be interpreted as a data standard and/or information required by regulatory authorities.

References

1. U.S. Food & Drug Administration (FDA) Data Standards Catalog. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/data-standards-catalog

Recommended Reading

- "Data Standards Governance: Current State and Challenges" (PP07), PHUSE CSS 2022: https://phuse.s3.eu-central-1.amazonaws.com/Archive/2022/CSS/US/Silver%20Spring/POS_PP07.pdf
- "Data Standards Governance and Implementation Across Industry: Key Challenges and Benefits" (DS02), PHUSE US Connect 2023: https://phuse.s3.eu-central-1.amazonaws.com/Archive/2023/Connect/US/Florida/PRE_DS02.pdf
- "Data Standards White Paper", PHUSE Data Standards Community (2024): https://phuse.s3.eu-central-1.amazonaws.com/Deliverables/Special+Projects/WP086.pdf
- "Best Practices in Data Standards Governance Survey: Insights and Next Steps" (PP16), PHUSE CSS 2024: https://phuse.s3.eu-central-1.amazonaws.com/Archive/2024/CSS/US/Silver%20Spring/POS_PP16.pdf

Project Contact Information

Project Lead:
Sandra VanPelt Nguyen
Email: sandra.vanpeltnguyen@pfizer.com

PHUSE Working Groups:
Email: workinggroups@phuse.global

Acknowledgements

A special note of gratitude to Steve Pearce from Amgen, who helped import the data and organise it in SAS, and to Ken Getz from the Tufts Center for the Study of Drug Development, who provided guidance on statistical considerations and data interpretation.

We would also like to thank Jane (Lozano) Owens and Janet Low, Co-Leads of the PHUSE Optimizing the Use of Data Standards Working Group, for their input and support.

Authors and editors include Sandra VanPelt Nguyen, Ellen Asam, Matt Baldwin, Monika Buda, Janet Low, Jennifer Manzi, Priyanka Pollarine, Susan Takacs and Tingting Tian.

Other individuals who have contributed to the survey design and data summarisation plans include Kapil Anand, Todd Bazin, Titania Dumas-Roberson, Patrick Genyn, Brooke Hinkson, Rie Ichihashi, Jennifer Jarvis, Gayathri Mahadevan, Nicholas Pemble, Torsten Petsching, Carlo Radovsky, Ranvir Singh, Gina Wood, Cathy Yang, Rong Ye and Aatiya Zaidi.