

# The Roadmap for Nonclinical Data Standards and Elements to Improve Data Access



## Table of Contents

### Information on data standards implementation

- Considerations for Implementing Standards
- Setting the right foundation for communication

### Considerations for data standardization

- Future Standardization

#### How to Move Forward with a Standardization Effort

Creating new standards and new conventions are by no means easy tasks. However there are some tricks that should be considered when looking into standardizing a new study type or new endpoints within studies where some standards already apply. As your organization becomes SEND-aware and develops its go-forward strategy, it must be acknowledged that SEND is being used for both operational as well as submission purposes. Understanding how your organization intends to respond to the standard is important. Perhaps you will use it only operationally or only to participate in pilot activity until it becomes a requirement. Or, you may intend to submit one package for one study in your next update. Alternatively, you may plan to submit a package for the majority of the studies in your next NDA or IND. For a vendor or a CRO, the need is to be as ready as your first client request.

#### How to Attack New Study Types Using Established Standards

When considering a standard for a study type, it is important to first identify the 'low hanging fruit' (i.e., there may already be standards or conventions developed that could be applied to the study type). The CDISC SEND standards built on SDTM offer a wealth of predefined concepts and terminology around in vivo animal studies, and these are not specific to any one study type. The endpoints specific for general toxicology studies, defined in SEND version 3.0 (**Stakeholders details below**), will apply for many other study types as well. The SEND standard offers a platform that should be leveraged when trying to come up with new standards for nonclinical data. Knowledge of SDTM and SEND will enable definitions of specific concepts and terminology around them.

- Prioritization of Nonclinical Data
  - Figure 1. Priorities for Nonclinical Study Groups.
  - Figure 2A. Priorities for Nonclinical Study Types
  - Figure 2B. Priorities for Miscellaneous Studies
  - Figure 3. Priorities for Data Elements

## Introduction

Advances in informatics and data standards are helping to optimize nonclinical data utility to meet demands for more effective and efficient drug development. The CDISC Standard for Exchange of Nonclinical Data (SEND) team is currently working to standardize nonclinical study types. The FDA, as well, is establishing regulatory guidance for the electronic submission of these nonclinical studies. Collectively, these efforts make this a formative time for stakeholders of these standards.

As a result, the FDA's Center for Drug Evaluation and Research (CDER) in conjunction with representatives across the industry have formed a collaboration aimed at addressing common issues with regards to nonclinical data management, data standards, and standards implementation. A Nonclinical Working Group within the collaboration identified several areas of interest and formed teams to work on each one. A summary of the Nonclinical Working Group and its various work-group teams is provided in the resource list at the end of this document.

One of the working groups, the Nonclinical Standards Roadmap Team (Roadmap Team), identified a need to orient newcomers to standards, set priorities for nonclinical data, and consider ways to facilitate future standardization efforts. Similar to what has already been done for clinical data by CDISC, the group realized a need to develop a strategy on how to prioritize, maximize, and revamp the standardization efforts for nonclinical data. The team aimed to create a nonclinical standards roadmap which could be useful to all, whether one is a novice or working on future steps in nonclinical data standardization.

The Roadmap Team focused on the following objectives:

- Identify points to consider for implementing a standard with a goal to minimize redundancy and maximize collaboration
- Highlight important steps of the data standardization process
- Prioritize the nonclinical data types
- Design a survey and obtain results from members of industry to further understand the priorities of the many stakeholders of nonclinical data
- Compile list of key standards resources (Links for Stakeholders below) for implementing the best approach for future standardization efforts
- Inform stakeholders about the types of standards and standards organizations that are available and the potential role that stakeholders can play in standards development

## Considerations for Implementing Standards

Standards implementation poses many challenges for organizations of all sizes. The PHUSE Nonclinical Working Group prioritized it as a major challenge and formed an Implementation Team devoted to addressing questions about standards utilization. Their work culminated in a Frequently Asked Questions (FAQ) resource which is posted at the following link ([SEND Implementation FAQ](#)). The Roadmap Team considered that many organizations are in early stages of implementation, and therefore also documented questions to consider during the process.

At the outset, it is important to realize what a standard can and cannot provide. A standard provides a language, but you need to decide the message you want that language to convey. Quite often in the clinical world, it is stated that there are company-specific flavours of a standard, meaning there are company-specific implementations and uses of that standard. The various 'flavours' or interpretations of a standard will usually be less flexible than the overall standard itself, which is acceptable. However, a standard must not be so rigid that it can only be used in one setting, as no one will ever adopt the standard. As such, it must be flexible enough to allow for customized use, but with the caveat that it will no longer be a standard if it becomes too flexible.

Examples of such interpretation using the SEND standard could be:

- Some sponsors may decide that certain permissible variables in SEND are required for their data.
- Sponsors can include Parameters for the Trial Summary for their specific operational use (such as the name and contact details for a Principle Investigator in a study)

So, what are the necessary steps to become compliant with requirements for electronic data? Where should you begin? This may seem obvious, but it cannot be repeated too often: Start with what you know, and accept that the first iteration of your standardization exercise will not be 100%. If it takes you to 80%, at least you are only 20% from your target instead of where you started.

### 2.1 Planning Phase: What Should be Considered

The following is a list of tasks that should be considered as you develop a plan for implementing standards. During the Planning Phase, you are not necessarily trying to answer all of the questions described below. Rather, you are planning for the time and resources required to address these topics and ensuring you have a complete list of tasks.

It is critical to inventory all of your data sources and consider how the data from each source fits into the standard.

- Plan for assessing how the standard will be implemented for each system (more on this in the next section).
- Understand where and when a single study may have data derived from multiple sources.
- When data is derived from multiple sources consider the following:
  1. What are all of the sources (e.g., in-house systems, CRO systems, both)?
  2. At what frequency will data be provided: at the end of the study, monthly, weekly?
  3. What study designs will be addressed now and what will (or must) be deferred until a later time?
  4. Is the format of the study identifiers the same across sources? If not, plan for assessing how this will be addressed.
  5. Is the format of the animal/pool/method identifier the same across sources? If not, plan for assessing how this will be addressed.
  6. Will data sources be providing duplicate information (e.g., define.xml file, trial domains, etc)? If so, plan to assess how that duplication will be handled and merged to create a final version for submission.
  7. Identify an authoritative source for a data verification QC process. For example, against what are you going to verify duplicate information?
  8. In the plan, consider how you will ensure that keys throughout the dataset (across sources) are aligned and pointing to the appropriate records. For example, consider a situation where in-life data is in one system and pathology data is in another system. How will you ensure that RELREC records are correctly aligned across all domains?
  9. Ensure that your process provides traceability from creation of the data in the source system throughout the final version of an electronic dataset.
  10. Consider how datasets will be versioned and stored, what will this include (original datasets from the source system, merged datasets, multiple versions, etc.).

## 2.2 Implementation Phase

The earlier section on Planning Phase formulates a list of "what" needs to be done to implement a standard. With that list in place, the next step is to formulate a plan for "how" the standard will be implemented. There are several ways to approach this, and the strategy will be company and department specific.

How do you get your source data into the standard format?

- Enforce the entire standard on existing systems thus setting an expectation that each source system will provide electronic data in the applicable standard.

This approach would require a moderate analysis, and result in an extensive implementation impact. For example, many LIMS systems would require some degree of customization for this approach to work.

- Enforce a portion of the standard or relevant parts of the standard on existing systems and create a new system for dealing with the new parts of the standard.

This approach would require an extensive analysis, and result in a moderate implementation impact. For example, it may be possible to enforce the use of controlled terminology, but not the final output format of the electronic data.

- Consider a new system to convert data and enforce the standard.

This approach would require a moderate analysis, and carries with it the potential for extensive implementation impact. For example, if the data standard is only enforced at the end when creating the submission data, the collected data face a conversion process that likely ends up being a manual exercise requiring specialist knowledge.

A single approach may suffice. However, once your analysis is complete, a combination of the above approaches may be required. Once the approach(es) is identified, the following must be considered and/or planned for:

- For any implementation, it is extremely important to have well-defined and documented processes that are thoroughly validated including data workflow; how data will be stored, backed-up, and archived; if/when/how Quality Assurance reviews will be completed; and who will have access to the data and when.
- Consider required scientific resources to properly assess how data will be converted. During the assessment, include representatives from each business unit so that the entire life-cycle of the data and the processes that it must go through are well thought-out.
- When developing the process, ensure that the same scientific conclusion will be met when analyzing data from either the source system or from the standardized data. For example, check group summary information for various data points, tumor reports, etc., in the original data and standardized data. Ensure that there are no compliance risks or that risks are identified and have a proper mitigation process defined.
- From a submission perspective, it is important to look forward to the date when electronic datasets will be required for submissions. If you have certain study types or data types which are still collected manually, or processes which preclude the transfer of data from one system to another, now is the time to implement change management.
- Automation of a process reduces both the risk of human error and the amount of QA review that must be performed on an on-going basis.

While there is a lot that must be considered, planned, and accomplished to implement a standard, there are extensive benefits to data standardization and overall increases in the quality of the data. Data discernment, reproducibility, exchange, analysis, cross-study analysis, storage, etc., will significantly improve as more standards are developed and implemented across the industry.

### Setting the right foundation for communication

SME engagement is essential to the success of the standards implementation process. In order to capture your audience, you must address them at their level of awareness. Below you can find some questions and answers that can help you when delivering the standardization message to the subject matter experts (SME) in your organization. What does your audience know about SEND? SEND, the Standard for Exchange of Nonclinical Data, is an implementation of the CDISC Standard Data Tabulation Model (SDTM) which specifies a way to present nonclinical data in a consistent format.

2. What does your audience know about PDUFA V and the draft guidance that will allow the FDA to make requiring SEND output a regulation?

"When is SEND mandatory" response on the PHUSE Implementation Advance Hub. (**Stakeholders details below**)

3. What does your audience know about SEND history and process for developing new domains?

The SEND Consortium has been in existence for about a decade. Previous iterations of the implementation guide, pilot data, and lessons learned were used as a baseline during the development of the current SEND IG. The team worked to rebaseline the model and bring it up to SDTM v1.2 standards from 2007 through 2009 when a draft was published for use in the FDA pilot. Subsequent adjustments were made and public comment received before the current SENDIG v3.0 was published in May 2011.

The creation of new domains is an ongoing activity of both the CDISC Submission Data Standards (SDS) and SEND teams. The current SEND IG (SEND Implementation Guide) covers multiple domains which support single-dose and repeat-dose general toxicity and carcinogenicity studies. Domains for respiratory and cardiovascular safety pharmacology studies are currently being piloted. Embryonic and fetal development domains for reproductive studies are awaiting pilot status. CNS safety pharmacology domains are under development.

4. What does your audience know about FDA projects like the Janus Data Warehouse? (**Stakeholders details below**)

5. Is your audience siloed (focusing on their area of expertise) or big picture (responsible for nonclinical data and/or the submission as a whole)? Do you have the right audience?

The ideal of electronic data submissions is shared across clinical and nonclinical domains, and both areas have used the same baseline standard in the SDTM. Yet, there are some significant differences between clinical and nonclinical data, and there is certainly a difference in the stage of development for standards. Take the time to ensure understanding throughout the organization; across, within, and between business units; and definitely within the regulatory structure. Include scientists in your discussion as they need to understand the evolution and may become your biggest advocate in moving the data visualization concept forward.

6. Are there other data/IT initiatives related to a given study type which would cause someone to either raise or lower the ranking in the hopes of less interference/more information?

Perhaps your organization just purchased new instruments for a specific data type or you are in the middle of a multi-year upgrade or perhaps you have budgeted for changes next year and incorporating SEND requirements now might throw the projects off track. Understanding and addressing ongoing/planned project timelines in the context of submission readiness is critical to gaining support from key stakeholders.

## Future Standardization

### How to Move Forward with a Standardization Effort

Creating new standards and new conventions are by no means easy tasks. However there are some tricks that should be considered when looking into standardizing a new study type or new endpoints within studies where some standards already apply. As your organization becomes SEND-aware and develops its go-forward strategy, it must be acknowledged that SEND is being used for both operational as well as submission purposes. Understanding how your organization intends to respond to the standard is important. Perhaps you will use it only operationally or only to participate in pilot activity until it becomes a requirement. Or, you may intend to submit one package for one study in your next update. Alternatively, you may plan to submit a package for the majority of the studies in your next NDA or IND. For a vendor or a CRO, the need is to be as ready as your first client request.

### How to Attack New Study Types Using Established Standards

When considering a standard for a study type, it is important to first identify the 'low hanging fruit' (i.e., there may already be standards or conventions developed that could be applied to the study type). The CDISC SEND standards built on SDTM offer a wealth of predefined concepts and terminology around in vivo animal studies, and these are not specific to any one study type. The endpoints specific for general toxicology studies, defined in SEND version 3.0 (**Stakeholders details below**), will apply for many other study types as well. The SEND standard offers a platform that should be leveraged when trying to come up with new standards for nonclinical data. Knowledge of SDTM and SEND will enable definitions of specific concepts and terminology around them.

## Prioritization of Nonclinical Data

The Standards Roadmap Team created a survey for members of industry to understand the priorities of the many stakeholders of nonclinical data. We asked them to consider the entire expanse of nonclinical studies listed for Module 4 of the eCTD. Responders rated the study types with a scale of high, medium, or low priority or had the option of leaving it blank (indicative of no preference). A spot to comment on rationale for the preference was provided. Finally, responders were asked to indicate whether they were responding as an individual or on behalf of their organization. Results Respondents consisted of both companies/institutions (n=4) and individuals (n=11) throughout industry. Overall, there was no significant difference between company and individual respondents.

The results of the survey indicated that Carcinogenicity and General Toxicology were the highest priority study groups (Figure 1). This was not surprising and supports previous decisions by the SEND team to standardize these first. Developmental and Reproductive Toxicology (DART) and Safety Pharmacology were rated the next highest priorities. Indeed, standards for these two study groups are currently being developed. Pharmacokinetics studies were considered to be a medium priority, along with Genetic Toxicology studies. Respondents differed on the utility of pharmacology studies, which received both high and low priority status, hence it is labeled as lower priority in Figure 1.

## Priorities for Nonclinical Study Groups

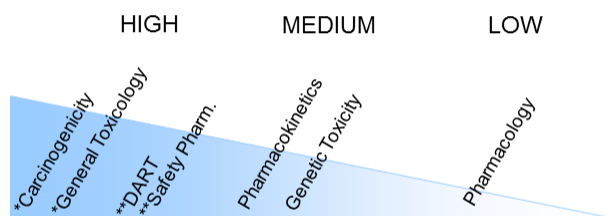


Figure 1. Priorities for Nonclinical Study Groups. Study groups are listed in order of their priority based on the results of the survey. Single asterisks (\*) indicate study groups which have been standardized. Double asterisks (\*\*) represent those study groups currently being standardized

The Roadmap team recognized value in keeping study types grouped since conventions of standardization may be applied to several study types within a study grouping. Study types for developmental and reproductive toxicity (DART studies: Segment I, II, III) were similar in their perceived high priority. Safety Pharmacology had both high and low priorities among their study types (Figure 2A). Note that Respiratory and Cardiovascular Safety Pharmacology study types received the highest priority and are presently being standardized. CNS Safety Pharmacology also rank high, while Renal, Autonomic Nervous System and Gastro-Intestinal systems are lower. Pharmacokinetic and Genetic Toxicity study groups received high /medium scores across their respective study types.

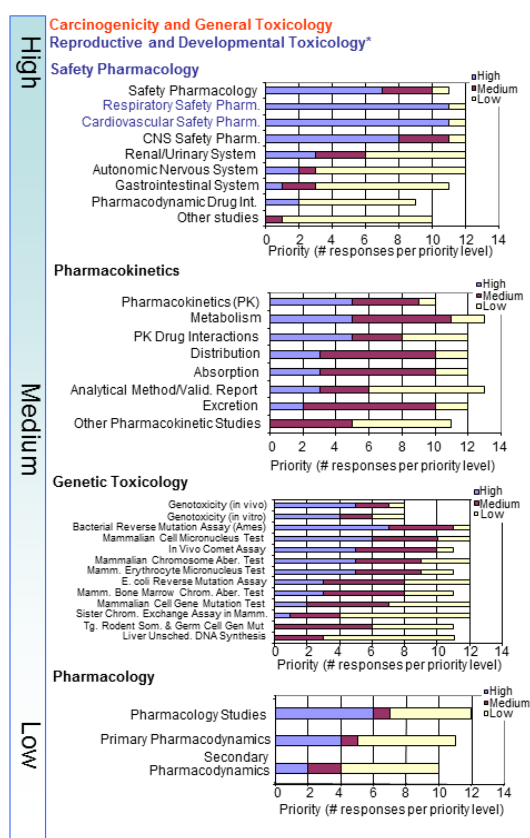


Figure 2A. Priorities for Nonclinical Study Types. Individual study types were rated as high, medium, or low priority by survey respondents. Red Font indicates study groups that have been standardized. Blue font denotes study types that are currently being standardized. Asterisk (\*) indicates that all study types within the Reproductive and Developmental Toxicology were of a similar high priority based on survey results.

Figure 2B captures prioritization of miscellaneous studies. Immunotoxicity and Receptor Screens were ranked highest, while all other studies take medium to lower importance.

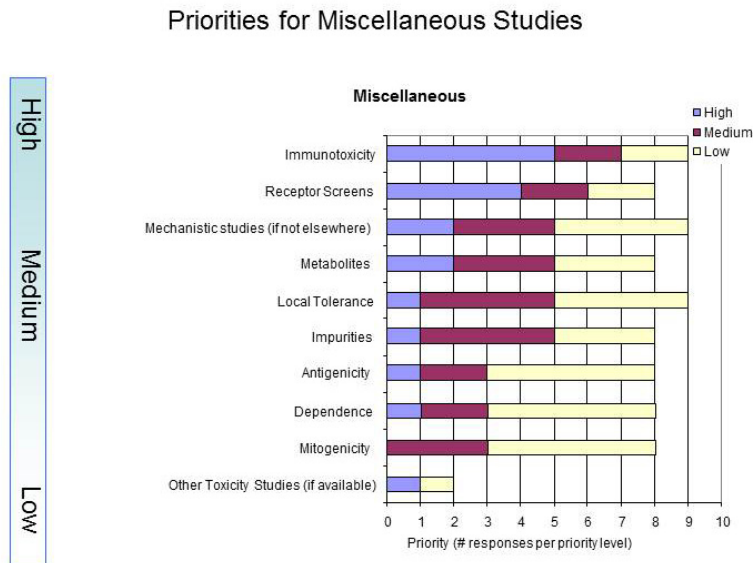


Figure 2B. Priorities for Miscellaneous Studies. Study types were rated as high, medium, or low priority by survey respondents.

A prioritization of several data elements which can be applied across study types is also provided in Figure 3. Respondents considered creating standards for Historical Control data, Study Design/Logistics, and Clinical Signs of highest importance. Formulation Data, Biomarker Incorporation, and Metadata content also were important for standard implementation.

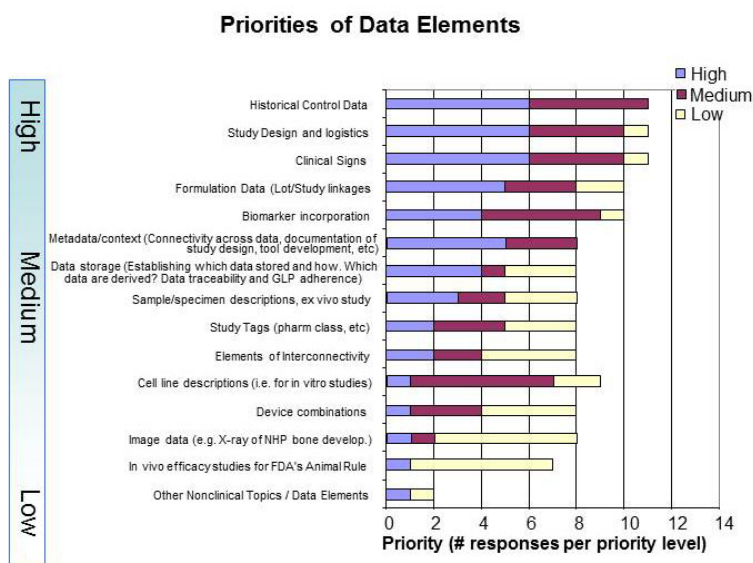


Figure 3. Priorities for Data Elements. Study types were rated as high, medium, or low priority by survey respondents.



## Survey Conclusion

Overall, Carcinogenicity, General Toxicology, DART, and Safety Pharmacology were ranked highest among study groups for creating standards. Work is already in progress for standard creation in each of these study groups. Additional study types that are high priority for standards development include CNS Safety Pharmacology, Pharmacokinetics, Metabolism, and Bacterial Reverse Mutation Assays. Further, there is a desire for historical control data, study design, clinical signs, formulation data, biomarkers, and metadata to be standardized sooner rather than later. These will be more challenging and difficult to implement.

## Summary

### Future Directions

Large volumes of nonclinical data are being produced to provide evidence of the safety of investigational drugs. Currently, these data are recorded, compiled, and transferred in various forms. This is not an impediment to data access as long as our disparate systems allow the visualization to be electronic and standardized. By taking the many and varied processes, systems, lexicons, and formats used in the collection and reporting of drug development data across our industry and collaborating to create a common standard for visualizing it, we will improve our access to data. By basing both clinical and nonclinical data on the same data tabulation model, the science of predictive toxicology can be furthered, proving hypotheses and moving theory to reality. Flexible standards for collecting, conveying, and submitting data provides access to a vast amount of information. The Roadmap Standards team is now undertaking an exercise to identify the possibilities that exist with such a model. Ultimately, the goal is to improve the collection and transmission of nonclinical information and to consider existing models which may give us access to more data. The benefits and challenges to such model are being considered by many stakeholders as they plan to test such a model in a pilot.

### Conclusion

The nonclinical scientific community has been harmonizing data collection methods and data analysis for many years. Some of these efforts have resulted in powerful research and predictive tools and others in publications of harmonized nomenclature and diagnostic criteria for nonclinical lesions. These efforts should be leveraged when looking into developing a standard. The challenge can be that some of these harmonized ontologies ('scientific synonym dictionaries') are considered proprietary and therefore not easily accessible. Ideally the companies owning these can find an appropriate business incentive to participate in a standardization effort, such as becoming a recognized standards training and implementation partner for the industry or as a valued software vendor. A harmonization effort is the first step on the way to developing a new standard. One can quite easily get lost in translation, so to speak. Many people believe that once you have harmonized the nomenclature, it is a standard, but it takes quite a bit more. A standard should include both uniquely defined concepts that ideally can be applied throughout the nonclinical (and potentially the clinical) world as well as recommended or preferred terminology for every concept. Whether concepts or terminology come first is the 'chicken and egg' discussion. In truth, concepts can help define terminology and vice versa. A standard also requires adoption, use, maintenance, and support. The nonclinical scientific community should strive to have broadly adopted standards to enable a common understanding of the science. In the end, everyone is dependent on others understanding our data. This, however, requires up front "buy in" from multiple stakeholders, and there should be incentive to adopt the standard such as a regulatory push, an unmet scientific need, a measurable saving in time to reporting, man hours per study, etc. A standard must evolve according to science and business needs. This necessitates ownership to its continuous maintenance and support as can be provided by a Standard Development Organization (SDO). Various SDOs exist as organizations or industrial consortia, build on expertise and interest from the people involved, with an established infrastructure allowing informed decision-making about the development of a standard. A standard can exist on many levels. You may have your own personal standard for how you do things or a department or company could have a standard. Incentive may help the tricky task of advancing a standard to the next level of adoption. Why should a company adopt a global standard or why should you adopt a colleague's standard for that matter? It is about convincing the right people, in the right way, that having well defined standards increases the effectiveness, scientific quality and, ultimately, the regulatory compliance of your work. The Roadmap Team believes that improved access to data through standards provides benefits on many levels.

## Links for Stakeholders

### **SEND Implementation guide**

- <https://www.cdisc.org/standards/foundational/send>

CDISC terminology: NCI Thesaurus

- <http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/cdisc>

### **Regulatory information**

Study Data Standards for Submission to FDA webpage:

- <http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm>

Draft guidance: Providing Regulatory Submissions in Electronic Format - Standardized Study Data

- <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292334.pdf>

Study Data Specifications v2.0

- <http://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM312964.pdf>

Prescription Drug User Fee Act (PDUFA) V

- <http://www.fda.gov/downloads/forindustry/userfees/prescriptiondruguserfee/ucm270412.pdf>

### **Standards implementation**

[SEND Implementation User Group](#) team (PHUSE Nonclinical Working Group)