ABSTRACT

The CDISC SEND standard benefits tremendously from the highly regulated environment. Since the FDA is requiring data in the SEND format, industry is creating SEND datasets for numerous nonclinical studies. Pharmaceutical companies are not only including them in regulatory submissions but also leveraging them to create internal data warehouses. Consortia are also forming to facilitate sharing across the industry hoping this information will enable us to learn to develop better drugs and get them through the pipeline and delivered to patients faster. Unfortunately, many challenges must be overcome to reach the full potential. SEND currently supports an important sub-set of nonclinical studies, but expanding its scope and improving its quality will enable even greater value to be derived. We present the objectives and anticipated delivery dates of the active SEND development work. We also present the truth about a number of common misunderstandings related to the SEND standard.

INTRODUCTION

The SEND Implementation Guide (SENDIG) is one of the CDISC global foundational platform-independent data standards (11) that is required for certain submissions to the FDA as described in regulatory guidance (17). This paper will provide an understanding of how the CDISC SEND standard development has benefited from the highly regulated environment by sharing historical progress, recent deliverables, and the current development work leading to the largest number of deliverables ever from the SEND team this year. In addition to the SEND team progress, the broadening use of SEND in the industry is mentioned to underscore the benefits of SEND. The paper will also share the SEND team’s additional steps in standards development (pilots and example datasets) that are intended to improve the quality of each version of a standard and avoid the costs associated with errors. The cadence we set for delivering versions based on this development life cycle is intended to be managed in close collaboration with our stakeholders, including and especially the FDA, to provide benefits for all stakeholders (SEND producers and SEND consumers).

In an attempt to make reading through text with a high volume of acronyms easier, some of the more frequently used ones are listed at the end of this paper.

HISTORICAL PROGRESS OF SEND DEVELOPMENT AND THE INFLUENCE OF THE REGULATED ENVIRONMENT

This section contains some information about the history of the SEND standard development as it relates to the benefits of collaborating closely with our key stakeholder, FDA. For a more detailed historical background of SEND, see Wood and Kramer (19).

The SEND team first began with fewer than 20 members in the first quarter of 2003 as an industry consortium focused on FDA submission. By November 2005, they had completed a pilot of manually created data and produced the SENDIG v2.3. At this time, the SEND efforts stalled and very nearly ended with many companies unwilling to continue to commit resources without some formal statement from the FDA regarding the longevity of SEND (19,20). Many companies were choosing to “wait and see” if SEND would ever become implementable and useful across the industry.

A resurgence of SEND activity began in mid-2007 when a few pharmaceutical sponsor companies saw a significant need for standardized data operationally (for transfer of data from CRO to sponsor) and volunteered resources to lead and support efforts with the FDA (7,19). At the same time, the FDA had a renewed interest in SEND, not only from CDER but from the Office of the Commissioner for its usefulness in submissions (19,20). It is the collaboration between the FDA and industry on SEND development that remains the hallmark of the SEND team successes today.
Since 2007, the SEND team has held two to four 4- to 5-day face-to-face meetings each year. They have published SENDIG v3.0 (2011) and SENDIG v3.1 (2016), both of which are on the FDA Data Standards Catalog as supported submission standards today.

The time following the 2014 FDA publication of the guidance documents (17) in December of 2014, commonly known as guidance that mandated standardized data in submissions, has clearly been the period of the most growth for the SEND standard, including significant gains in SEND team membership as well as actual SEND implementations and successful submissions of SEND data packages. Since 2014, the SEND team has jumped to over 100 people with over 19 active sub-teams and more than 70 attendees at the week-long face-to-face meetings which are now hosted by the FDA at their White Oak facility. The years since 2015 have also led to growth in implementation in the industry. In the annual Industry SEND Progress Survey conducted by the PhUSE Nonclinical Topics team, the number of responding organizations who had included a SEND dataset(s) in an actual IND or NDA submission jumped from 5% of responders in 2015 to 28% of responders in 2017(8). And most recently, data shows that 43 studies (NDAs from 18 Dec 2016 to 31 August 2018) and 134 studies (INDs from 18 December 2017 to 31 August 2018) have been submitted to the FDA. (5) Of the 15 original IND submissions with nonclinical study data, there was a critical error rate of 15.94%, where a critical error is defined as not compliant with Technical Rejection Criteria #1734 (regarding trial summary datasets) or #1736 (regarding demographics datasets.) (5,14)

The co-authors believe that the FDA need for this data in submission drove the support that turned SEND into a successful standard for submission and beyond. Given the significant cost of standards implementation, we all benefit from working in this highly regulated environment because the regulations are driving adoption and a broad adoption of the standard enables serious efforts to build efficient processes and tools, data sharing consortiums, and data analytics that are not focused purely on regulatory submission. While it is still too early to know, we suspect that FDA participation and submission mandates will be seen as the turning point that moved the SEND efforts towards the ultimate broader vision for data across all nonclinical research activities.

CURRENT ACTIVITIES AND DELIVERY SCHEDULE OF THE SEND DEVELOPMENT TEAM

The activities of the CDISC SEND team are organized into multiple sub-teams. The SEND team is faced with the challenge of prioritizing work across many competing priorities. Today, several teams are providing clarity on the previously published standards, others are overcoming challenges within the existing scope, and still others are expanding SEND’s scope of studies and scientific domains. As a result of the adoption of SEND for regulatory reasons, all of these SEND team activities are enabling the industry to exchange nonclinical study data in a common structure leading to broader use and easier re-use of the valuable data collected on nonclinical studies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Latin Square POC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DART v1.1 POC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controlled Terminology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENDIG-AR v 1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENDIG v3.1.1 (PC, PP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCG feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoDEx for SENDIG v3.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENDIG v3.0 Conformance Rules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENDIG v3.1 Conformance Rules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genotoxicity IG with POC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dermal/Ocular IG with POC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENDIG V3.1 FFU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENDIG V3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Last year, the SEND team delivered more items than in any prior year. We have delivered proof-of-concept data for DART v1.1, an embryo-fetal development study, and proof-of-concept data for a Latin Square study design using a cardiovascular safety pharmacology study in addition to our usual quarterly controlled terminology updates. This year, again, we expect there will be even more publications delivered by or with significant support from the
SEND team than in any prior year. The most prominent of these is the SENDIG-AR v1.0, expanding the scope of the SEND family of standards to cover studies conducted under the FDA’s animal rule (16). This new IG was developed by the CPATH and CDISC organizations in close collaboration with the FDA and SEND team members. We will also be releasing SENDIG v3.1.1 to more completely standardize pharmacokinetic concentration data.

Many of the other publications for this year will improve clarity on the previously published standards. Not only will we continue to release quarterly updates to the controlled terminology, we will also be releasing the Confirmed Data Endpoints for Exchange (CoDEx) for SENDIG v3.1. This CoDEx will bring clarity to the types of measurements that can be confidently exchanged using the standard; so, when organizations are considering the use of SEND to exchange data they can use this document to understand which measurements will require special coordination and which don’t. In addition, we will be publishing the SENDIG 3.0 Conformance Rules. These are a critical tool to help implementers measure how well they are conforming to the standard. This set of rules will form the basis for preparing and releasing conformance rules for the subsequent releases of the SEND standards.

Looking forward to 2020-2021, we expect to release proof-of-concept data as well as standards covering four types of Genetic Toxicology studies, multiple types of dermal and ocular irritation studies and safety pharmacology studies evaluating effects on the central nervous system.

We are planning for the next major release of SENDIG to be published in 2022-2024. Major activities this year are intended to prepare the scope for this release. A fit-for-use pilot of SENDIG v3.1 with the FDA. Since studies with this version of the standard will begin flowing to the FDA early in 2020, this pilot will be important to help us see where improvements in the standard are needed and will enable the FDA to provide meaningful guidance to the industry on how to make datasets that are useful for regulatory reviews. In concert with this pilot we invite and encourage others to similarly perform pilot operations exchanging and reviewing SENDIG v3.1 data. Please provide feedback to CDISC with the things you learn through our JIRA comment tracker: https://wiki.cdisc.org/pages/viewpage.action?pageId=68949529

Also this year we are working on improvements in the areas of anatomic pathology, tumor findings, and enhancements to support the needs of FDA CBER in addition to efforts striving to establish more consistency across the CDISC standards.

One of the challenges associated with this highly regulated environment is that the practices described in the regulations and/or standards are widely adopted even if the practices are flawed initially. Additionally, the costs and effort to change practices to make improvements is often higher than it was to implement the practices in the first place. This elevates the importance of establishing good standards the first time. To meet this challenge, the CDISC SEND team, in collaboration with the FDA OCS has established a standards development process that ensures the standards are tested as they are developed and proven to be useful before they are adopted.

The following diagram (Figure 1) shows the major steps throughout the standard development life cycle (6). The first step is for the CDISC SEND team to develop the standard with POC datasets, conformance rules, and a listing of the confirmed data end points for exchange. Once the standard is released, software is developed and deployed. At this point we have the opportunity to see how well the standard functions as implemented in the production-ready software by conducting a FFU pilot. We found this to be extremely valuable with SENDIG v3.0 and enabled the FDA to provide meaningful guidance in the Study Data Technical Conformance Guide to enable the industry to prepare datasets that would be useful to them. This also identified areas requiring improvements in the standard itself. As a result, the SEND team has started identifying long-term solutions to these challenges and are making plans to include these changes in a future release. Given the value of this FFU, the CDISC organization has now added the concept of FFU piloting (optionally) into their standards development process, COP-001 (1), for all CDISC teams to consider.

Figure 1. The Standards Development Life Cycle
With the release of a standard accompanied by the example datasets, regulators have the information to determine when they can begin to support the standard; however, it is unknown how well the implementations actually support the regulator’s needs until the full usability testing in the FFU pilot. At the conclusion of the FFU pilot, regulators can confidently determine if the standard will improve their operations and if it would be appropriate to require this standard in submissions. This could be the trigger for determining the date this standard is mandated to be used in submissions.

As previously noted, the FFU identifies key areas for improvement. Armed with this information, the standards development team can identify near- and long-term solutions to these challenges and earmark these changes for future releases. When considering the best cadence to aim for in the industry (see Figure 2), the start of development for a subsequent major release should depend upon having completed a FFU pilot. Staggering the development cycles including a reasonable length of overlap in time that old standards are supported while new ones are introduced will allow industry to react in a consistent manner, giving vendors time to adapt software and users time to re-tool their environments.

**Figure 2. Considering the Best Cadence**

**QUALITY GAINS FROM THE COMBINATION OF EXAMPLE DATASETS AND USABILITY PILOTS**

Both the FFU pilots and POC datasets help to establish confidence in a version of a standard, but they also serve distinct functions. The example datasets are released twice. First as draft with the request for public comment on a draft standard. The draft example datasets ensure the standard is more completely considered at this early point in development and enables the standard to be more completely reviewed. This is a key opportunity for the world to give advice to CDISC on how to adjust the standard to more effectively enable the exchange of meaningful data before the standard is released.

The second time the example datasets are released is with the final release of the standard. At this point the datasets promote the efficient adoption of the standard. One of the key distinctions between the FFU pilots and the POC example datasets is that the example datasets are publicly sharable and represent “good practice”. The FFU pilot is used to test the standard by specific stake-holders such as the FDA or consortia. Sponsor organizations provide data from real studies that are transformed into the standard using production-ready software. Generally, this means the datasets cannot be shared outside the FDA or outside the consortia, but they do provide a good representation of how the standard was interpreted and implemented. As a result, they mimic production exchanges. Feedback from the FFU pilots are critical to ensure successful production data exchanges.

For additional details of the CDISC SEND teams improvement efforts, including quality improvement pilots, to meet today's challenges, see Houser and Kramer (6).
MYTHS AND MISCONCEPTIONS

With the increased regulatory interest and use of SEND and, of course also the mandate for submissions, the SEND team members have often been quite challenged to work on developing new versions of the standard while gathering and dispositioning information on the existing published versions. Heightening this balancing act is the pressure for experienced SEND members expected to assist a greater number of new learners within their own organizations as well as with their external business partners and collaborators. This creates a struggle for many SEND members to maintain a pace on their volunteer activities while implementing and assisting others in real time. However, in the midst of this challenging environment, the team has learned so much from this real world experience. The teams learning has also been enhanced by their completion of multiple piloting efforts recently. From all of these experiences, it is apparent that there are some common misconceptions or misunderstandings in the interpretation of the standard. The following paragraphs describe common misconceptions that are seen repeatedly by SEND members and so, they are the myths we’d like to dispel.

MYTH: ALL I NEED FROM THE IG ARE THE EXAMPLES

Examples are presented throughout the SEND Implementation Guide (IG) for each version of the standard. Examples are key to demonstrating the principles and assumptions of the model and many people find them the best way to begin to learn about the standard. However, a common practice that we see in the industry today is in taking the examples as the literal definition of the standard. The examples are just that, an example of how data can be modeled. Examples are considered informative material in any CDISC guide and are therefore, not the definition of the standard. Instead the user should understand the normative material in each section, including the paragraph explanations, domain specifications, CDISC notes, assumptions, and core values.

The SENDIG recommends the following sequence of learning to ensure an understanding and to better implement SEND for each use case: the reader should first read the SDTM model (13) to understand the principles and concepts that apply to all of the IGs in the SDTM family of standards (including SENDIG, SDTMIG, etc.) Then the reader should read the SENDIG (11) in order by section, specifically Sections 1-3 (key concepts), 4 (assumptions), 5-6 (domain models), 7 (trial design datasets), 8 (relationships between datasets and records, adding variables not specifically defined in the models, and pooled data) and then review SEND Controlled Terminology. It is best to read the SDTM model and SENDIGs online so you can make use of the many hyperlinks (11).

MYTH: ONLY VARIABLES THAT APPEAR IN THE DOMAIN SPECIFICATION TABLE CAN BE USED IN THAT DOMAIN DATASET

The SEND standard is based upon and meant to be fully aligned with the SDTM model. The SDTM model is the authoritative source for which variables may be used in each domain class. For example, identifiers and timing variables can be used in any domain based on the three general observation classes: Interventions, Events, or Findings. The domain specification tables have only the most commonly used variables for that domain.

MYTH: NO NEED TO LOOK UP CT, IT’S IN THE EXAMPLES AND CDISC NOTES

To meet the requirements for Controlled Terminology (4) it is best to understand the terminology structure and use. The examples and values listed in CDISC notes should never be considered the full CT list. In addition, the CT is updated frequently. as a result, you can’t expect the examples in the SENDIG to be complete or current.

MYTH: THE SEND STANDARD AND THE BASE MODEL ARE ALIGNED AND SO, ALWAYS AGREE

As mentioned above, the SEND standard is based upon and meant to be fully aligned with the SDTM model. With each new version of a SEND IG, the team takes great care to review and adjust to align with the most current appropriate SDTM model version. However, as the standard develops over time new content can sometimes include errors (misspelled variable names, variables not in correct order). These are inadvertent ways that the standard may become misaligned from the SDTM model. In that case, as these errors become known to the CDISC SEND development team, they are either published as errata (3) on the CDISC public facing wiki space (18) for that version of the standard or lead to changes in a future version of the standard and are noted in the team’s JIRA project as a future change. If a reader finds an error in the SENDIG at any time, they should inform the CDISC SEND team by entering a description of the error in the JIRA comment tracker:

In addition, there are times when the differences between the SEND standard and the SDTM model are intended and correct. In the case of the “Core” attribute of a variable, the values for Core can differ between the IG and the model and even between IGs. The value will depend upon the use cases represented in that IG. For example, due to the specific use cases related to timing and reproductive events in Developmental and Reproductive Toxicity studies in SEND, the Core value for most—NOMDY (Nominal Study Day for Tabulations) variables is Expected in SENDIGv3.0 but is Permissible in SENDIG-DARTv1.1, while the Core value for RXFSTDTC (Date/Time of First Study Treatment) and RXFENDTDC (Date/Time of Last Study Treatment) variables in Demographics is Expected in DART studies but Permissible in SENDIGv3.1.
MYTH: SEND PREPARES ALL MY DATA FOR ANALYTICS
SEND is an exchange standard; it is not meant to produce datasets that are fully ready for analysis. The ideal dataset for any specific analysis need is different from the ideal dataset for another analysis need. To accommodate these diverse needs, CDISC provides an analysis standard, ADaM (Analysis Data Model), however, to date the SEND team has not found a compelling need to create ADaM datasets for nonclinical data. The consistency provided by the SEND standard has been effectively leveraged by many software packages to transform the data and present useful analytical visualizations. The SEND standard should be expected to be used along with such analysis tools to produce the analytics needed by an organization.

MYTH: SEND IS ONLY NEEDED FOR GLP STUDIES IN SUBMISSIONS
SEND is expected to be used for any study, regardless of GLP status, that falls within the FDA requirements for study submission as outlined in regulatory guidance (17) and also as described in the Technical Rejection Criteria for Study Data (14). The Technical Rejection Criteria provides concise information on the scope of studies relative to sections of the eCTD that require SEND data packages; however, as of the writing of this paper, these criteria have not become effective yet.

MYTH: SEND IS ONLY FOR SUBMISSION PURPOSES
While we must recognize the critical importance that regulatory use and regulator involvement has had in the successful implementation of the SEND standard, it is also important to note that the scope of the SEND standard was never meant to be limited to FDA submission. The original vision for having this data standardized included a very important benefit of having truly sharable nonclinical data within companies and across the industry for the first time. The FDA requesting, supporting and mandating these data have led to broad adoption across the industry making this a goal that now seems attainable. Pharmaceutical companies have begun to use SEND to create internal data warehouses and to consistently communicate with business partners, like CROs, leading to data warehouses with searchable, shareable data from multiple sources.

Industry consortia have also begun to leverage SEND data and the experiences of the SEND users to meet their goals. For example, the BioCelerate organization plans to leverage SEND data for shareable toxicology and background control data. Their vision is to “enable member companies to make data-driven decisions on compound progression based on an increased understanding of on-target and off-target toxicity” (15).

MYTH: SEND IS ONLY USEFUL IN THE U.S.
As with the comments about scope beyond submissions, it is also important to note that the scope of the SEND standard was never meant to be limited to U.S. submissions either. We do not yet have another region that has defined a mandate for nonclinical datasets in submission, but interest is building in multiple countries.

HOW TO GET YOUR SEND QUESTIONS ANSWERED?
There are so many different implementations of SEND in progress in the industry today and there will always be varying levels of experience among those who are beginning to create or consume SEND data; therefore, we highlight here the best ways to get your technical SEND questions answered.

First, the Nonclinical Topics Working Group of the PhUSE organization has a project entitled “SEND Implementation User Group.”(9) This project team provides Frequently Asked Questions and other tips for newer SEND users. In addition, there is a “SEND Implementation Forum”(10) that new and experienced SEND users will find a valuable place for asking any questions regarding SEND. A group of experience SEND developers meet regularly to provide answers to your posted questions in a timely way.

Also, the CDISC SEND team leaders can be contacted through the CDISC organization contact form from the CDISC home page, www.cdisc.org. (2)

And, of course, the official answer for anything related to FDA regulatory submissions must come from the FDA themselves. Their Study Data Standards Resources page (12) has a great deal of information on requirements for submission, including email addresses for asking additional specific questions of CDER or CBER.

CONCLUSION
Throughout this paper we’ve presented many challenges of the SEND team, from the lack of sponsor company participation historically, to the many competing priorities of the SEND developers and competing needs for the SEND standard, from the challenges and costs associated with publishing standards that have not been tested, to the balancing act most experienced SEND team members are in between volunteering to develop SEND and assisting their own organization or their business collaborators with their implementation issues. But with all of these challenges the SEND team is more and more successful each year, has plans to beat their record for the number of deliverables in a year again this year, and we are even beginning to see creative SEND use in the nonclinical
research space that goes beyond the regulatory requirement for submission. We credit many of the advances and successes to the close collaboration with key stakeholders, including and especially the FDA, so operating in a highly regulated environment has been of great benefit to the SEND standard. We believe we will see continued benefits to all stakeholders by the requirement to pilot (POC and FFU) at key points in the development process whereby all stakeholders, including FDA, can be more prepared much earlier for smoother adoption of successive standards and versions.

REFERENCES
ACKNOWLEDGMENTS
The authors would like to thank the following:

- The members of industry who participated in or arranged for the initial, very small meetings in April/May of 2007 at the Pfizer office in D.C. to resurrect the SEND effort with renewed interest in operational benefits. They had the vision and the personal drive to lead and influence the critical pieces of work along with our FDA liaisons. This includes: Peggy Zorn, Mary Jo Brucker, Fred Wood, Louis Norton, Gitte Frausing, and Lou Ann Kramer, with support from Bill Rosen and Steve Ward behind the scenes. If we’ve omitted anyone, please accept our apologies.

- The current SEND leadership team (aka SEND Extended Leaders) who have the leadership skills and the personal commitment to continue to drive our many sub-teams towards a better future of standardized nonclinical research data.

- The current SEND development team who keep coming back year after year bringing their unique experience and skills, commitment, and collective enthusiasm to the work.

- The current FDA liaisons and leadership for maintaining a level of involvement across so many important team efforts and for the many FDA personnel who actively participate in our Spring and Fall Face-to-Face meetings.

- CDISC leadership for supporting the SEND team with a CDISC liaison, IT tools and support and the many CDISC staff members who answer the daily needs for collaboration on solutions, communication needs, and governance.

- Senior leadership of Bristol-Myers Squibb for their support and advocacy of CDISC standards.

- Leadership of Lilly Research Laboratories for their vision and early support of the SEND efforts, especially David S. Negilski, Michael Masnyk, and Andrew M. Dahlem

SOME OF THE ACRONYMS USED IN THIS PAPER

CDER          Center for Drug Evaluation and Research (FDA)
CDISC         Clinical Data Interchange Standards Consortium
CoDEx         Confirmed Data Endpoints for Exchange
CRO           Contract Research Organizations
FDA           Food and Drug Administration (U.S.)
FFU           Fit-For-Use
IG            Implementation Guide
JIRA          Jira is a proprietary issue tracking software product developed by Atlassian, used by CDISC
PhUSE         Pharmaceutical Users Software Exchange
POC           Proof-Of-Concept
SDTM          Study Data Tabulation Model
SDTMIG        SDTM Implementation Guide
SEND          Standard for Exchange of Nonclinical Data
SENDIG        SEND Implementation Guide

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Mr. William Houser  Ms. Lou Ann Kramer
Principal Scientist I  Senior Director, Standards Development
Bristol-Myers Squibb  CDISC
1 Squibb Dr  401 W 15th Street
New Brunswick, NJ  Suite 800
732-227-3795  Austin, TX 78701
william.houser@bms.com  +1-317-498-1527
磷l.kramer@cdisc.org  www.cdisc.org/send

Brand and product names are trademarks of their respective companies.