PAPER DS04

# A practical exploration of SEND for CBER submissions

Jack Baker, Labcorp, Alconbury, Cambridgeshire, UK

## Abstract

Now that the Center for Biologics Evaluation and Research (CBER) division of the FDA has initiated support to receive SEND datasets in future submissions, sponsors are turning their attention to the requirement date of 15 March 2023. Industry understanding of how to report and leverage immunogenicity, biodistribution, and/or tumorigenicity data endpoints will be critical to the advancement of gene therapies, vaccines, and/or cell products in the fight against Cancer and other diseases. At Labcorp, we too are focused on our SEND readiness for submissions made to the OVRR, OTAT, and OBRR offices of CBER, and we would like to share our findings as we prepare SEND datasets for studies conducted by our Nonclinical Toxicology and Preclinical Oncology teams.

## Introduction

Preparing for new data types in SEND can provide many challenges, and data in relation to studies that will be submitted to CBER is no different. CBER will be expecting these data endpoints as early as the requirement date of 15$^{th}$ of March 2023 however it has been the case that CBER are already accepting some of these data endpoints now. This is why at Labcorp we believe it is imperative to start preparing for these submissions sooner rather than later. We currently have an established network for CDER data i.e. communications with scientists, systems set up with SEND in mind, however this is yet to be established for CBER, setting up networks and establishing communication channels is a key part of ensuring the SEND generation process is as accurate and efficient as possible. Many current biologics data collection systems are not prepared to output flat files, let alone files in SEND format, and therefore acknowledging this and working with the right teams to get systems caught up and outputting in a SEND friendly format is essential. The current implementation of the SENDIG 3.1 is not suitable for many of the data types that are introduced with the CBER submissions. I have used the latest version of the SENDIG 3.2 on the CDISC Wiki as well as the latest version of the SDTMIG (v3.4) in my research/modelling to create practical ways to integrate the data. The SENDIG 3.2 is currently still in development and its projected date for when it will be locked for editing and ready for the publishing phase is towards the end of 2023. I am hoping my findings will be able to impact choices made in future SENDIG versions such as what domains are included, what variables from the SDTM are used and updated examples of data typically seen on studies that will be submitted to CBER.

## Discussion

To be as prepared as possible it's important to recognize the challenges we face. One example of this was lack of previous knowledge and understanding of the data. In general, a lot of the data in CBER studies are more complex than the typical data endpoints of general toxicology studies submitted to CDER. The additional data required much more niche scientific knowledge and so a bit more research and investigation was needed for certain data endpoints. In many cases it was necessary to set up meetings with the scientists and principle investigators to get a deeper understanding of the data that was being included in the SEND domains. In some cases, it was the first time the scientists or principle investigators were introduced to the concept of SEND which added further challenge. Often it was necessary to preface questions with an explanation as to why we were asking and what the information was being used for. Another challenge was learning the new domain structures themselves. New domains introduced to the SENDIG 3.2 include the CP (Cell Phenotype Findings) and IS (Immunogenicity Specimen Assessments) domains which are used for many data types introduced with CBER submissions, and so researching and understanding these new domain models was essential. Lastly, the format of data in general was a challenge due to collection systems not being set up with SEND in mind. A significant percentage of the data was not collected or provided in SEND format or even a flat file format and so required a much more manual approach to manipulate the data into SEND format, primarily using SAS.

During this project I learned many new skills and developed a deeper understanding of many new concepts. This included not only learning the new or updated domain specifications, but familiarizing myself with the new controlled terminology that will be introduced with these domains, as well as additional terminology used for existing domains that may not have applied to data commonly used for datasets that would have been submitted to CDER. I also needed to learn some of the background of the science behind what data we are inputting and formatting into SEND to ensure I had the relevant knowledge to map tests, test details, categories etc. to the correct SEND terms. I expanded my knowledge on the SDTMIG; although this is specific to human clinical trials I found a lot of the information could be transferable to the data endpoints seen in nonclinical CBER data, and so familiarizing myself with the SDTMIG was very valuable in my interpretation of the data and modelling of the SEND Datasets. This is all knowledge I have now that I wish I knew at the start of the process.

In order to model CBER data effectively in SEND format a number of updates and changes are necessary to current implementation guides. The current 3.2 version of the SENDIG on the CDISC Wiki does address many of these changes however there are still more concepts from the SDTMIG that could be brought into the SEND implementation guide that would allow for a much better representation of the data. These include additional variables within domains; additional domains themselves; additional assumptions and examples of domain models included in implementation guides. The next section includes specific domain examples including updates/changes to existing SENDIG implementations to model the data effectively:

## MI Domain Examples

The first data I explored was Immunohistochemistry data, Immunohistochemistry (IHC) is a method that involves treating tissue with a stain that adheres to very specific substances. In this example IHC was used to verify the graft survival of intrastriatal human embryonic stem cell derived progenitor cells and to confirm differentiation with human-specific neural cell adhesion molecule (hNCAM), human-specific anti-nuclear antibody (KU80) and assess for proliferating cells with the Ki67 antibody. There are currently no references to this data in the SENDIG, and therefore in order to model this data in SEND format I utilized the guidance and examples in the SDTMIG. My initial instinct was that this data may belong in the new CP domain, however the SDTMIG 3.4 states the following:

"The CP domain is not intended to supplant use of the LB domain for routine lab hematology (e.g., blood cell differentials), nor is it intended for findings originating from microscopic assessment of cells, including those employing immunohistochemical (IHC) techniques."

This lead me to find appropriate examples and assumptions in SDTMIG that concluded this data belongs in the MI domain. Assumption 1 of the MI assumptions in the SDTMIG states:

"This domain holds findings resulting from the microscopic examination of tissue samples. These examinations are performed on a specimen, usually one that has been prepared with some type of stain. Some examinations of cells in fluid specimens (e.g., blood, urine) are classified as lab tests and should be stored in the Laboratory Test Results (LB) domain. Biomarkers assessed by histologic or histopathological examination (by employing cytochemical/immunocytochemical stains) are stored in the MI domain."

One key difference to the MI domain in the SDTMIG in comparison to the SENDIG is the additional variables of MITSTDTL (Microscopic Examination Detail). MITSTDTL is used to further describe the test within the MITEST/MITESTCD variables. It is used to represent specific attributes, for example in this case it is used to represent the intensity score of the test. The MITEST and MITESTCD terminology in the SDTMIG is a lot more expansive, whereas the SEND terminology only contains tests for General Histopathologic Exams and Sexual Maturity Microscopic exams. For this example, the controlled terminology for MITEST and MITESTCD had to be extended for the hNCAM and KU80 tests however the Ki67 test was taken from the latest SDTM Terminology.

Another key difference to the MI domain in the SDTMIG in comparison to the SENDIG is the use of the MISPEC variable as well as the additional variable MILOC that is present in the SDTMIG. In the SDTMIG, MISPEC (Specimen Material Type) is used to represent the type of specimen used e.g. "BLOOD", "TISSUE", "BONE" etc. MILOC (Specimen Collection Location) is used to represent the location relevant to the collection of the specimen. For this example, where brain tissue was examined, a combination of the MISPEC and MILOC was used i.e. an MISPEC of "TISSUE" with an MILOC of "BRAIN". In the SENDIG, MILOC is not present and therefore typically you would only use the MISPEC (along with MIANTREG, MILAT and MIDIR) to identify the examined tissue, therefore if

we were to model the example in the current SENDIG format we would have an MISPEC of "BRAIN" with no MILOC present.

Example Dataset:

| Row | STUDYID | DOMAIN | USUBJID | MISEQ | MITESTCD | MITEST | MITSTDTL | MIORRES |
|---|---|---|---|---|---|---|---|---|
| 1 | ABC | MI | ABC-001 | 1 | HNCAM | hNCAM | STAINING INTENSITY | 0 |
| 2 | ABC | MI | ABC-001 | 2 | KI67 | Ki-67 | STAINING INTENSITY | 1+ |
| 3 | ABC | MI | ABC-001 | 3 | KU80 | KU80 | STAINING INTENSITY | 2+ |
| 4 | ABC | MI | ABC-002 | 4 | HNCAM | hNCAM | STAINING INTENSITY | 4+ |
| 5 | ABC | MI | ABC-002 | 5 | KI67 | Ki-67 | STAINING INTENSITY | 3+ |
| 6 | ABC | MI | ABC-002 | 5 | KU80 | KU80 | STAINING INTENSITY | 3+ |

| Row | MISTRESC | MIRESCAT | MISPEC | MILOC | MIMETHOD | MIDTC | MIDY |
|---|---|---|---|---|---|---|---|
| 1 (cont) | 0 | NEGATIVE | TISSUE | BRAIN | IHC | 2020-02-26 | 29 |
| 2 (cont) | 1+ | POSITIVE | TISSUE | BRAIN | IHC | 2020-02-26 | 29 |
| 3 (cont) | 2+ | POSITIVE | TISSUE | BRAIN | IHC | 2020-02-28 | 29 |
| 4 (cont) | 4+ | POSITIVE | TISSUE | BRAIN | IHC | 2020-07-22 | 183 |
| 5 (cont) | 3+ | POSITIVE | TISSUE | BRAIN | IHC | 2020-07-29 | 183 |
| 6 (cont) | 3+ | POSITIVE | TISSUE | BRAIN | IHC | 2020-07-29 | 183 |

In this example, IHC assessment of samples in brain tissue yielded staining intensity on a scale of 0 to 4+. Staining intensity values of 0 were categorized as negative; values of 1+, 2+, 3+ and 4+ were categorized as positive.

## EX Domain Examples

CBER submitted studies also introduces new/ unfamiliar dosing methods. This is an example of an Exposure dataset for selected animals from a stereotactic Intra-Striatal Toxicity, Tumorigenicity and Biodistribution Study. Subjects were randomized to one of six treatment groups, where they were administered either Human Embryonic Stem Cell Derived Progenitor Cells (HESC-PC), Human Embryonic Stem Cells (100% HESC), or Human Embryonic Stem Cell Derived Progenitor Cells (HESC-PC) spiked with Human Embryonic Stem Cells (HESC) at 3 different levels (0.001% HESC, 0.1% HESC and 1% HESC), one group also served as the vehicle control group. The animals dose consisted of 4 injections of 1.5 uL to the right striatum, one after the other, leading to a single total dose of 6 uL.

In this example the data was able to be modelled using the SENDIG 3.1 however there was some items of note. For EXDOSU the UNIT codelist was extended to allow for the dosing units of 10^5 cells, these types of units are not often seen in conjunction with dosing in typical CDER SEND studies and so it may be that new units will need to be added to controlled terminology. For records 4-6 the EXTRT was used to also represent the different levels that the HESC-PC was spiked with HESC, it may be worth considering whether additional variables should be introduced to the domain to represent this information separately.

Example Dataset:

| Row | STUDYID | DOMAIN | USUBJID | EXSEQ | EXTRT | EXDOSE | EXDOSU | EXDOSFRM |
|---|---|---|---|---|---|---|---|---|
| 1 | ABC | EX | ABC-001 | 1 | HESC-PC | 0 | 10^5 cells | SUSPENSION |
| 2 | ABC | EX | ABC-002 | 2 | HESC-PC | 7 | 10^5 cells | SUSPENSION |
| 3 | ABC | EX | ABC-003 | 3 | 100% HESC | 7 | 10^5 cells | SUSPENSION |
| 4 | ABC | EX | ABC-004 | 4 | HESC-PC\|0.001% HESC | 7 | 10^5 cells | SUSPENSION |
| 5 | ABC | EX | ABC-005 | 5 | HESC-PC\|0.1% HESC | 7 | 10^5 cells | SUSPENSION |
| 6 | ABC | EX | ABC-006 | 6 | HESC-PC\|1% HESC | 7 | 10^5 cells | SUSPENSION |

| Row | EXDOSFRQ | EXROUTE | EXLOT | EXLOC | EXMETHOD | EXTRTV |
|---|---|---|---|---|---|---|
| 1 (cont) | ONCE | INTRACEREBRAL | | Right Striatum | INTRASTRIATAL INJECTION | Vehicle A |
| 2 (cont) | ONCE | INTRACEREBRAL | ABC1 | Right Striatum | INTRASTRIATAL INJECTION | Vehicle A |
| 3 (cont) | ONCE | INTRACEREBRAL | DEF1 | Right Striatum | INTRASTRIATAL INJECTION | Vehicle A |
| 4 (cont) | ONCE | INTRACEREBRAL | ABCD1 | Right Striatum | INTRASTRIATAL INJECTION | Vehicle A |
| 5 (cont) | ONCE | INTRACEREBRAL | ABCD2 | Right Striatum | INTRASTRIATAL INJECTION | Vehicle A |
| 6 (cont) | ONCE | INTRACEREBRAL | ABCD3 | Right Striatum | INTRASTRIATAL INJECTION | Vehicle A |

| Row | EXVAMT | EXVAMTU | EXSTDTC | EXENDTC | EXSTDY | EXENDY | EXDUR |
|---|---|---|---|---|---|---|---|
| 1 (cont) | 6 | uL | 2022-01-31 | 2022-01-31 | 1 | 1 | PT10M |
| 2 (cont) | 6 | uL | 2022-01-31 | 2022-01-31 | 1 | 1 | PT10M |
| 3 (cont) | 6 | uL | 2022-01-31 | 2022-01-31 | 1 | 1 | PT10M |
| 4 (cont) | 6 | uL | 2022-01-31 | 2022-01-31 | 1 | 1 | PT10M |
| 5 (cont) | 6 | uL | 2022-01-31 | 2022-01-31 | 1 | 1 | PT10M |
| 6 (cont) | 6 | uL | 2022-01-31 | 2022-01-31 | 1 | 1 | PT10M |

## CP Domain Examples

The CP domain is a new domain to SEND that will be introduced with the SENDIG 3.2 that contains information for cell phenotype findings. In this example I have modeled Immunophenotyping in this new domain. Immunophenotyping is a method that is used to identify cells based on the types of markers or antigens present on the cell's surface, nucleus, or cytoplasm. In this example Immunophenotyping was used to evaluate phenotypes of Total T cells, Helper T Cells, Cytotoxic T cells, Natural Killer cells and B cells as well as activated subsets of the phenotypes.

I was able to model this data using the SENDIG 3.2 examples from the CDISC wiki. I used the latest SDTM controlled terminology to map the controlled terms, however the terminology was extended for the THLYHSTLY test code and Tly Help Sub/Tlym Help test name. In this example I felt like I had all the resources available to model the data into the CP domain however in order to understand the domain model I had to do much more research into the background of the data itself. Developing a deeper understanding and knowledge of the data was essential for me to model the data accurately. Understanding the different tests and how they relate to the marker strings, as well as understanding how subsets of cell populations can be identified in different ways was key to modelling this data in a meaningful way. I found that communicating with the scientists and principle investigators behind the data to better understand the science itself helped when modeling this data into SEND format. This domain also has scope for much more complicated tests than those seen in this example and so in those cases it is vital to understand the background of the data to ensure the SEND dataset represents the data in the most meaningful way.

Example Dataset:

| Row | STUDYID | DOMAIN | USUBJID | CPSEQ | CPTESTCD | CPTEST | CPCELSTA |
|---|---|---|---|---|---|---|---|
| 1 | ABC | CP | ABC-001 | 1 | BLYCE | B-Lymphocytes | |
| 2 | ABC | CP | ABC-001 | 2 | BLYCELY | B-Lymphocytes/Lymphocytes | |
| 3 | ABC | CP | ABC-001 | 3 | NKCE | Natural Killer Cells | |
| 4 | ABC | CP | ABC-001 | 4 | NKLY | NK Cells/Lym | |
| 5 | ABC | CP | ABC-001 | 5 | NKSNK | NK Cells Sub/NK Cells | ACTIVATED |
| 6 | ABC | CP | ABC-001 | 6 | TLC | TLym Cytx | |
| 7 | ABC | CP | ABC-001 | 7 | TLCSTLC | TLym Cytx Sub/TLym Cytx | ACTIVATED |
| 8 | ABC | CP | ABC-001 | 8 | TLSTLY | TLym Sub/TLym | ACTIVATED |
| 9 | ABC | CP | ABC-001 | 9 | TLYCE | T-Lymphocytes | |
| 10 | ABC | CP | ABC-001 | 10 | TLYCTLY | TLym Cytx/TLym | |

| Row | STUDYID | DOMAIN | USUBJID | CPSEQ | CPTESTCD | CPTEST | CPCELSTA |
|---|---|---|---|---|---|---|---|
| 11 | ABC | CP | ABC-001 | 11 | TLYH | TLym Help | |
| 12 | ABC | CP | ABC-001 | 12 | TLYHSTLY | TLym Help Sub/TLym Help | ACTIVATED |
| 13 | ABC | CP | ABC-001 | 13 | TLYHTLY | TLym Help/TLym | |
| 14 | ABC | CP | ABC-001 | 14 | TLYHTLYC | TLym Help/TLym Cytx | |
| 15 | ABC | CP | ABC-001 | 15 | TLYLY | TLym/Lym | |

| Row | CPCSMRKS | CPMRKSTR | CPCAT | CPORRES | CPORRESU |
|---|---|---|---|---|---|
| 1 (cont) | | CD3-CD20+ | IMMUNOPHENOTYPING | 0.4 | 10^9/L |
| 2 (cont) | | CD3-CD20+ | IMMUNOPHENOTYPING | 11 | % |
| 3 (cont) | | CD3-CD159a+ | IMMUNOPHENOTYPING | 22 | % |
| 4 (cont) | | CD3-CD159a+ | IMMUNOPHENOTYPING | 1.5 | 10^9/L |
| 5 (cont) | CD69+ | CD3-CD159a+CD69+ | IMMUNOPHENOTYPING | 50 | % |
| 6 (cont) | | CD3+CD4-CD8+ | IMMUNOPHENOTYPING | 1.2 | 10^9/L |
| 7 (cont) | CD69+ | CD3+CD4-CD8+CD69+ | IMMUNOPHENOTYPING | 4.6 | % |
| 8 (cont) | CD69+ | CD3+CD69+ | IMMUNOPHENOTYPING | 1.7 | % |
| 9 (cont) | | CD3+ | IMMUNOPHENOTYPING | 3.4 | 10^9/L |
| 10 (cont) | | CD3+CD4-CD8+ | IMMUNOPHENOTYPING | 33.5 | % |
| 11 (cont) | | CD3+CD4+CD8- | IMMUNOPHENOTYPING | 2.1 | 10^9/L |
| 12 (cont) | CD69+ | CD3+CD4+CD8-CD69+ | IMMUNOPHENOTYPING | 1 | % |
| 13 (cont) | | CD3+CD4+CD8- | IMMUNOPHENOTYPING | 62.4 | % |
| 14 (cont) | | CD3+CD4+CD8-/CD3+CD4-CD8+ | IMMUNOPHENOTYPING | 1.9 | RATIO |
| 15 (cont) | | CD3+ | IMMUNOPHENOTYPING | 68.4 | % |

| Row | CPRESTYP | CPSTRESC | CPSTRESN | CPSTRESU | CPSPEC | CPMETHOD |
|---|---|---|---|---|---|---|
| 1 (cont) | NUMBER CONCENTRATION | 0.4 | 0.4 | 10^9/L | WHOLE BLOOD | FLOW CYTOMETRY |
| 2 (cont) | NUMBER FRACTION | 11 | 11 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 3 (cont) | NUMBER FRACTION | 22 | 22 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 4 (cont) | NUMBER CONCENTRATION | 1.5 | 1.5 | 10^9/L | WHOLE BLOOD | FLOW CYTOMETRY |
| 5 (cont) | NUMBER FRACTION | 50 | 50 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 6 (cont) | NUMBER CONCENTRATION | 1.2 | 1.2 | 10^9/L | WHOLE BLOOD | FLOW CYTOMETRY |
| 7 (cont) | NUMBER FRACTION | 4.6 | 4.6 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 8 (cont) | NUMBER FRACTION | 1.7 | 1.7 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 9 (cont) | NUMBER CONCENTRATION | 3.4 | 3.4 | 10^9/L | WHOLE BLOOD | FLOW CYTOMETRY |
| 10 (cont) | NUMBER FRACTION | 33.5 | 33.5 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 11 (cont) | NUMBER CONCENTRATION | 2.1 | 2.1 | 10^9/L | WHOLE BLOOD | FLOW CYTOMETRY |
| 12 (cont) | NUMBER FRACTION | 1 | 1 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 13 (cont) | NUMBER FRACTION | 62.4 | 62.4 | % | WHOLE BLOOD | FLOW CYTOMETRY |
| 14 (cont) | RATIO | 1.9 | 1.9 | RATIO | WHOLE BLOOD | FLOW CYTOMETRY |
| 15 (cont) | NUMBER FRACTION | 68.4 | 68.4 | % | WHOLE BLOOD | FLOW CYTOMETRY |

| Row | CPDTC | CPDY | CPNOMDY | CPNOMLBL | CPTPT | CPTPTNUM | CPELTM | CPTPTREF | CPRFTDTC |
|---|---|---|---|---|---|---|---|---|---|
| 1 (cont) | 2021-07-29 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-07-29 |
| 2 (cont) | 2021-07-30 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-07-30 |
| 3 (cont) | 2021-07-31 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-07-31 |
| 4 (cont) | 2021-08-01 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-01 |
| 5 (cont) | 2021-08-02 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-02 |
| 6 (cont) | 2021-08-03 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-03 |
| 7 (cont) | 2021-08-04 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-04 |
| 8 (cont) | 2021-08-05 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-05 |
| 9 (cont) | 2021-08-06 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-06 |
| 10 (cont) | 2021-08-07 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-07 |
| 11 (cont) | 2021-08-08 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-08 |
| 12 (cont) | 2021-08-09 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-09 |
| 13 (cont) | 2021-08-10 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-10 |
| 14 (cont) | 2021-08-11 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-11 |
| 15 (cont) | 2021-08-12 | 1 | 1 | Day 1 | 6 Hours PD | 1 | PT6H | Day 1 Dose | 2021-08-12 |

## IS Domain Examples

The IS domain is a new domain to SEND that will be introduced with the SENDIG 3.2 that contains information for Immunogenicity Specimen Assessments.

Example 1:

In the first example for this domain I have modeled Antidrug Antibody data in this new domain. This type of data is commonly assessed on general toxicology studies we have submitted previously to CDER however this data has always either been considered out of scope for SEND or it has been modeled into the most appropriate domain following SENDIG 3.1, typically the LB. Once the SENDIG 3.2 is released, this data will then have a home in the IS domain. In this example the detection of Binding Antidrug Antibodies was assessed via Meso Scale Discovery.

I was able to model this data using the SENDIG 3.2 examples from the CDISC wiki, and it gives a nice fairly straightforward example of the CP domain in use. I found I was able to model this data with very few complications or effort.

Example Dataset:

| Row | STUDYID | DOMAIN | USUBJID | ISSEQ | ISTESTCD | ISTEST | ISBDAGNT | ISTSTOPO |
|---|---|---|---|---|---|---|---|---|
| 1 | ABC | IS | ABC-001 | 1 | ADA_BAB | Binding Antidrug Antibody | DrugA | SCREEN |
| 2 | ABC | IS | ABC-001 | 2 | ADA_BAB | Binding Antidrug Antibody | DrugA | CONFIRM |

| Row | ISCAT | ISSCAT | ISORRES | ISORRESU | ISSTRESC | ISSTRESN | ISSTRESU |
|---|---|---|---|---|---|---|---|
| 1 (cont) | Antidrugantibodies | Humoral immunity | Positive | | Positive | | |
| 2 (cont) | Antidrugantibodies | Humoral immunity | Negative | | Negative | | |

| Row | ISSPEC | ISMETHOD | ISDTC | ISDY | ISNOMDY | ISNOMLBL |
|---|---|---|---|---|---|---|
| 1 (cont) | SERUM | Meso Scale Discovery | 2017-09-08 | 1 | 1 | Day 1 |
| 2 (cont) | SERUM | Meso Scale Discovery | 2017-09-08 | 1 | 1 | Day 1 |

Example 2:

In the second example for this domain the presence of Neutralizing Binding Antidrug Antibodies was assessed via luminescence output. This is an example of data that may be slightly more unfamiliar to individuals who have only worked with CDER submitted data. I was able to model this data using examples from the SENDIG 3.2 on the CDISC Wiki. With this example I found that the data was provided in not the most desirable format for SEND conversion, and so utilizing SAS to transpose and manipulate the data first into a flat file format and then convert to SEND was necessary to get this data in the desired format.

Example Dataset:

| Row | STUDYID | DOMAIN | USUBJID | ISSEQ | ISTESTCD | ISTEST | ISBDAGNT |
|---|---|---|---|---|---|---|---|
| 1 | ABC | IS | ABC-001 | 1 | ADA_NAB | Neutralizing Binding Antidrug Antibody | AAV VirusA |
| 2 | ABC | IS | ABC-002 | 2 | ADA_NAB | Neutralizing Binding Antidrug Antibody | AAV VirusA |

| Row | ISTSTDTL | ISCAT | ISSCAT | ISORRES | ISORRESU | ISSTRESC |
|---|---|---|---|---|---|---|
| 1 (cont) | NEUTRALIZING TITER 50% | IMMUNOGENICITY | HUMORAL IMMUNITY | 1 in 135 | titer | 1 in 135 |
| 2 (cont) | NEUTRALIZING TITER 50% | IMMUNOGENICITY | HUMORAL IMMUNITY | 1 in 405 | titer | 1 in 405 |

| Row | ISSTRESN | ISSTRESU | ISSPEC | ISMETHOD |
|---|---|---|---|---|
| 1 (cont) | | titer | SERUM | Neutralizing Antibody assessment by luminescence assay |
| 2 (cont) | | titer | SERUM | Neutralizing Antibody assessment by luminescence assay |

| Row | ISDTC | ISDY | ISNOMDY | ISNOMLBL |
|---|---|---|---|---|
| 1 (cont) | 2020-01-08 | -6 | -6 | Pretreatment |
| 2 (cont) | 2020-01-08 | -6 | -6 | Pretreatment |

## PC Domain Examples

Quantitative polymerase chain reaction (QPCR) is a common laboratory method used when evaluating biodistribution. In this example QPCR was used to measure tissue and shedding samples. When first investigating this data it was hard to establish a domain where it belonged. There are currently no examples of this data type in either SENDIG 3.1 or on the SENDIG 3.2 wiki.

I decided to model the data in the PC domain, although this is not quite the right domain to model the data I found it was the most appropriate fit to model the information I felt needed to be included in the domain.
There are some key parts of the PC domain where I felt like I could have used to model the data more accurately. For example, a --LAT, --ANTRAG and --DIR variable would all be useful to model the specimen information more meaningfully, currently I have extended the terminology for the specimen to include this information all in the PCSPEC variable.

I also found that a “Test Detail” variable could be useful for data such as this. I have currently included test details into the PCSCAT variable however I feel like this information is much better placed in a new variable which is specific for adding additional details to further describe the test.

Another note I had on this particular data is that it had reference to a lower limit of detection as well as a lower limit of quantification. Many domains include --LLOQ and --ULLOQ but there are currently no variables that can be used to populate a --LLOD. If this is something that we commonly see in data types and this information is deemed important and relevant enough to include in the SEND dataset, then it may be worth considering introducing a new variable to capture this information.

Something to consider would be whether this data could be modeled in the GF domain (Genomics Findings) from the SDTMIG 3.4, there are no current examples of quantitative polymerase chain reaction in the SDTMIG however it could certainly be worth considering whether this data belongs here or whether an entirely new domain is necessary.

Example Dataset:

| Row | STUDYID | DOMAIN | USUBJID | PCSEQ | PCTESTCD | PCTEST | PCCAT |
|---|---|---|---|---|---|---|---|
| 1 | ABC | PC | ABC-001 | 1 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 2 | ABC | PC | ABC-001 | 2 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 3 | ABC | PC | ABC-001 | 3 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 4 | ABC | PC | ABC-001 | 4 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 5 | ABC | PC | ABC-001 | 5 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 6 | ABC | PC | ABC-001 | 6 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 7 | ABC | PC | ABC-001 | 7 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 8 | ABC | PC | ABC-001 | 8 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 9 | ABC | PC | ABC-001 | 9 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 10 | ABC | PC | ABC-001 | 10 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |

| Row | STUDYID | DOMAIN | USUBJID | PCSEQ | PCTESTCD | PCTEST | PCCAT |
|---|---|---|---|---|---|---|---|
| 11 | ABC | PC | ABC-001 | 11 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 12 | ABC | PC | ABC-001 | 12 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 13 | ABC | PC | ABC-001 | 13 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 14 | ABC | PC | ABC-001 | 14 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 15 | ABC | PC | ABC-001 | 15 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |
| 16 | ABC | PC | ABC-001 | 16 | VIRUSAVG | VirusA vector genome | Adeno-Associated Virus |

| Row | PCSCAT | PCORRES | PCORRESU | PCSTRESC | PCSTRESN | PCSTRESU |
|---|---|---|---|---|---|---|
| 1 (cont) | Shedding Sample Results | N/D | DNA copies/ug | N/D | N/D | DNA copies/ug |
| 2 (cont) | Shedding Sample Results | 700 | DNA copies/5u | 700 | 700 | DNA copies/5u |
| 3 (cont) | Shedding Sample Results | N/D | DNA copies/5u | N/D | N/D | DNA copies/5u |
| 4 (cont) | Shedding Sample Results | 5000 | DNA copies/5u | 5000 | 5000 | DNA copies/5u |
| 5 (cont) | Shedding Sample Results | <=LOQ | DNA copies/5u | <=LOQ | <=LOQ | DNA copies/5u |
| 6 (cont) | Tissue Sample Results | 950 | DNA copies/ug | 950 | 950 | DNA copies/ug |
| 7 (cont) | Tissue Sample Results | 100000 | DNA copies/ug | 100000 | 100000 | DNA copies/ug |
| 8 (cont) | Tissue Sample Results | 88000 | DNA copies/ug | 88000 | 88000 | DNA copies/ug |
| 9 (cont) | Tissue Sample Results | 400 | DNA copies/ug | 400 | 400 | DNA copies/ug |
| 10 (cont) | Tissue Sample Results | 600 | DNA copies/ug | 600 | 600 | DNA copies/ug |
| 11 (cont) | Tissue Sample Results | 2000 | DNA copies/ug | 2000 | 2000 | DNA copies/ug |
| 12 (cont) | Tissue Sample Results | 37000 | DNA copies/ug | 37000 | 37000 | DNA copies/ug |
| 13 (cont) | Tissue Sample Results | 700000 | DNA copies/ug | 700000 | 700000 | DNA copies/ug |
| 14 (cont) | Tissue Sample Results | 12000 | DNA copies/ug | 12000 | 12000 | DNA copies/ug |
| 15 (cont) | Tissue Sample Results | 200 | DNA copies/ug | 200 | 200 | DNA copies/ug |
| 16 (cont) | Tissue Sample Results | <=LOD | DNA copies/ug | <=LOD | <=LOD | DNA copies/ug |

| Row | PCSPEC | PCMETHOD | PCBLFL | PCLLOQ |
|---|---|---|---|---|
| 1 (cont) | FECES | Quantitative polymerase chain reaction | | 50 |
| 2 (cont) | FLUID, NASAL | Quantitative polymerase chain reaction | | 50 |
| 3 (cont) | PLASMA | Quantitative polymerase chain reaction | | 50 |
| 4 (cont) | SALIVA | Quantitative polymerase chain reaction | | 50 |
| 5 (cont) | URINE | Quantitative polymerase chain reaction | | 50 |
| 6 (cont) | ESOPHAGUS | Quantitative polymerase chain reaction | | 50 |
| 7 (cont) | GALLBLADDER | Quantitative polymerase chain reaction | | 50 |
| 8 (cont) | GLAND, ADRENAL, LEFT | Quantitative polymerase chain reaction | | 50 |
| 9 (cont) | GLAND, SALIVARY, PAROTID | Quantitative polymerase chain reaction | | 50 |
| 10 (cont) | GUT-ASSOCIATED LYMPHOID TISSUE | Quantitative polymerase chain reaction | | 50 |
| 11 (cont) | HEART | Quantitative polymerase chain reaction | | 50 |
| 12 (cont) | LIVER, LEFT MEDIAN LOBE LOWER HALF | Quantitative polymerase chain reaction | | 50 |
| 13 (cont) | LIVER, RIGHT MEDIAN LOBE UPPER HALF | Quantitative polymerase chain reaction | | 50 |
| 14 (cont) | LYMPH NODE, MESENTERIC | Quantitative polymerase chain reaction | | 50 |
| 15 (cont) | PANCREAS | Quantitative polymerase chain reaction | | 50 |
| 16 (cont) | URINARY BLADDER | Quantitative polymerase chain reaction | | 50 |

| Row | PCDTC | PCDY | PCNOMDY | PCNOMLBL |
|---|---|---|---|---|
| 1 (cont) | 2020-01-15 | -6 | -6 | Pretreatment |
| 2 (cont) | 2020-01-15 | -6 | -6 | Pretreatment |
| 3 (cont) | 2020-01-15 | -6 | -6 | Pretreatment |
| 4 (cont) | 2020-01-15 | -6 | -6 | Pretreatment |
| 5 (cont) | 2020-01-15 | -6 | -6 | Pretreatment |
| 6 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 7 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 8 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 9 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 10 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 11 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 12 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 13 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 14 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 15 (cont) | 2020-07-22 | 184 | 183 | 26 week |
| 16 (cont) | 2020-07-22 | 184 | 183 | 26 week |

## Future Domain Examples

Other data types I have considered and would like to investigate further as a part of this project are:

- Data in relation to T Cell Dependent Antibody Response (TDAR)
- Data in relation to Delayed-Type Hypersensitivity Studies
- Data in relation to Host Resistance Studies
- Data in relation to Hypersensitivity Studies
- Data in relation to Autoimmunity Studies

Currently these types of studies are not in scope for SENDIG 3.2 however data in relation to these could be included as supportive data endpoints for general toxicology studies and so there is still a precedent to investigate and explore this data.

## Conclusion

In order to be prepared as possible for when SEND for CBER is required for submission, it's important that we focus on what can be done to ensure a smooth and effortless transition to these studies and data types. Updating future SEND implementation guides with CBER data types in mind is key to ensuring the data can be modeled in a clear, concise and meaningful way. While it is currently possible to add variables from SDTMIG to a SEND domain dataset, if the variables were actually present in the SENDIG they are far more likely to be used and understood. Working with scientists and system operators to update processes and data collections systems with SEND in mind is key to ensuring the data manipulation process is as efficient as possible. Finally, educating ourselves on the new domains we will be using and new data types will be valuable to modelling this data. These are all actions the industry can and should be taking now, in order to ensure we are prepared for future CBER SEND submissions.

## References

SEND Implementation Guide v3.2, available on the CDISC Wiki
CDISC SENDIG v3.1
CDISC SDTMIG v3.4

## Acknowledgements

I would like to acknowledge and thank William Moseley, BSc (Labcorp, Supervisor and SEND subject matter expert, Global Data Management Solutions) and Jason Pitzer (Labcorp, North America Regional Manager, Global Data Management Solutions) for the expert review of the content in this paper and their guidance and contribution towards the project as a whole. I would also like to acknowledge and thank the space contributors who contributed towards the content on the SEND Implementation Guide v3.2 on the CDISC Wiki, the information available on this Wiki was essential in this project. A list of space contributors is available on the SEND Implementation Guide v3.2 Home page.

## Contact Information

Your comments and questions are valued and encouraged. Contact the author at:
Jack Baker
Labcorp, Woolley Rd, Alconbury, Huntingdon PE28 4HS, UK
Email: Jack.Baker@labcorp.com