TREAT-NMD
SMA Dataset Implementation
Workshop 2 Report

27th and 30th October 2020
Virtual

Contents

1. DISCLOSURE AND CONFLICT OF INTEREST

This workshop was part of the TREAT-NMD SMA Dataset Implementation Project supported by Biogen; a pharmaceutical company with an approved therapy (Spinraza) for spinal muscular atrophy (SMA). This report provides an overview of the discussions and recommendations made during the workshop; it does not necessarily represent the full perspectives of any individual attendees, Biogen, or TREAT-NMD.

This report has been prepared by Joanna Das (SMA Dataset Project Coordinator) and Jo Bullivant (SMA Dataset Project Manager) with review and input from the SMA Registry leads Miriam Rodrigues and Victoria Hodgkinson and from IT consultant and registry representative Marcel Heidemann.

2. EXECUTIVE SUMMARY

TREAT-NMD is an international neuromuscular network, which coordinates a global network of SMA registries who all collect a common core dataset. The TREAT-NMD SMA Dataset Project is funded by Biogen and led by Project Manager Jo Bullivant at the John Walton Muscular Dystrophy Research Centre, Newcastle University, United Kingdom. The three-year implementation project (2019-2022) is supporting SMA registries in the TREAT-NMD network to collect an agreed expanded dataset for SMA patients.

A virtual workshop was held on Tuesday 27th October and repeated on Friday 30th October 2020, to provide an overview of version 2 of the TREAT-NMD SMA Core Dataset to registry curators participating in the project. This workshop report provides a summary of the key discussion points, questions raised and any agreed next steps.

3. ABOUT THE WORKSHOP

3a. Project context

The SMA core dataset supports standardised data collection across all registries in the TREAT-NMD network. This model has been in place for 10 years. The original dataset was minimal, designed to identify patients eligible for recruitment into clinical trials. However, with disease modifying therapies (DMT's) now available in many countries, TREAT-NMD expanded the core dataset with a view to providing real-world data from patient registries to support post marketing surveillance.

The TREAT-NMD SMA Core Dataset was established in 2009 (version 0), expanded in September 2018 (version 1) and revised in October 2020 (version 2). It supports data collection to inform on the natural history of SMA and provides data to support post-marketing surveillance (safety and effectiveness) of new treatments. Due to the number of disease-modifying therapies (DMTs) now authorised or approaching authorisation - and in order to ensure the dataset is kept relevant - a formal revision plan is in place. Consultation for and preparation of version 2 took place between March and October 2020.

3b. Workshop aims and objectives

The target audience for this workshop was registry curators participating in the SMA Dataset Implementation Project.

The workshop focused on providing training and guidance on version 2 of the Core SMA Dataset. Version 2 looks radically different however, the most significant changes lie in the presentation and description of the data. The content (that is, the data to be collected) has not been drastically altered.

Workshop Objectives:

- Provide an overview of the work done during the revision of v1 to v2 of the dataset
- Provide an overview of v2 structure and general principles
- Update on information not contained within the v2 specification
- Discuss the headline changes, future considerations, and possible research topics
- Provide an outline of patient-focussed work planned for 2021

3c. Workshop agenda and structure

The workshop aimed to provide accurate and clear information and to ensure all participants had opportunities to:

- voice their opinion and ask questions;
- participate in discussions and share their experience;
- learn from other registries facing similar challenges.

Table 1: Workshop Agenda

11:00	**Welcome**	Anna Ambrosini (TGDOC Chair Elect)
11:05	**Introduction**	Jo Bullivant (Project Manager)
11:15	**Work done for this revision**	
11:25	**Patient-focussed work 2021**	
11:30	**V2 Structure: General principles**	Marcel Heidemann (IT/Data Expert)
12:05	**Guidance Appendix**	Joanna Das (Project Coordinator)
12:10	**Questions**	All
12:25	**Comfort break (10 minutes)**	
12:35	**V2 Contents** • **Headline changes** • **Future considerations** • **Research questions**	Miriam Rodrigues and/or Victoria Hodgkinson (TGDOC SMA Subgroup Leads)
12:55	**Discussion groups**	

13:15	**Feedback from groups**	
13:30	**Next steps**	Jo Bullivant (Project Manager)
13:35	**Questions and discussion**	All
13:55	**Close**	Anna Ambrosini (TGDOC Chair Elect)

3d. Workshop participants

Chairs / Organisers:

1. **Anna Ambrosini, TGDOC Chair Elect**
2. **Jo Bullivant, Project Manager**
3. **Joanna Das, Project Coordinator**

In attendance:

4. Aya Ahmed
5. Anna Ambrosini
6. Poorani Anandakrishnan
7. Davit Babikyan
8. Nina Barisic
9. Dominique Baumann
10. Anna Bedoshvili
11. Vesna Brankovic
12. Numan Bulut
13. Maria Grazia Cattinari
14. Teodora Chamova
15. Marjan Cosyns
16. Rasha El Sherif
17. Robin Forbes
18. İpek Gürbüz
19. Jana Haberlová
20. Sahar Hassanein
21. **Marcel Heidemann**
22. **Victoria Hodgkinson**
23. Marlène Jagut
24. Veronika Karcagi
25. Teik-Beng Khoo
26. Andrea Klein
27. Anna Lusakowska
28. Vitaliy Matyushenko
29. Vedrana Mlic Rasic
30. Lenka Mokra
31. Soledad Monges
32. Lindsay Murphy
33. Kostandyan Natella
34. Tatishvili Nino
35. Anna O'Malley
36. Damjan Osredkar
37. Beatrix Pálmafy
38. Gabriela Plosnic
39. **Miriam Rodrigues**
40. Sureshkumar Sankaran
41. Sonia Segovia
42. Lip Yuen Teng
43. Simone Thiele
44. Venkataraman Viswanathan
45. Vana Vukic

TREAT-NMD Secretariat

46. Cathy Turner, DMD Programme Coordinator
47. Ben Watling, TREAT-NMD CEO
48. Janet Wilkins, DMD Project administrator

Key

Bold: Speakers

4. WORKSHOP DISCUSSIONS (all slides available)

4a. Welcome

Anna Ambrosini welcomed everyone to the workshop and thanked everyone for attending. Anna highlighted that a lot of work had been done in the last 12 months by the project team, SMA Leads and Marcel Heidemann on preparing version 2 of the dataset. Everyone who provided feedback and comments during the revision process was also thanked.

4b. Introduction

The dataset implementation project started in May 2019 and is now in its second year. It supports registries through a phased adoption of the Core SMA Dataset. The year 1 report is available on the TREAT-NMD website.

The project implementation plan included the development of a formal revision process for the dataset. This demonstrates TREAT-NMD's commitment to responding to the needs of the SMA community, continuous improvement of data standards and quality, and harmonisation with other initiatives. The formal revision process also aims to manage the burden placed on registries and participants from changes to the core dataset.

The v1 to v2 revision process launched in March 2020. It involved two rounds of extensive stakeholder consultation, and over 700 individual items of feedback were received and considered. Version 2 was released on 15th of October 2020. Version 2 looks and feels drastically different to version 1, however the data that registries are asked to collect are largely the same and has actually reduced in quantity.

Table 1: Version 2 data item count in comparison to version 1:

	v1	v2
Mandatory CR	113	117
Mandatory PR	92	91
Not mandatory	50	37
TOTAL	**167**	**154**

4c. Work done for this revision

The dataset must provide enough guidance for registries to collect the right data in the right way, whilst being flexible enough to be applicable in the wide range of registries in the network.

Examples of issues identified in v1:

- Dataset presented as a data collection form (items appeared as questions, focusing only on what to collect, not on how/when to collect it)
- Mixed data types
- Too many free text fields, useful for registries but not TREAT-NMD
- Numbered items which were not unique or stable
- Included items that were not in remit of a core dataset

Key improvements made in version 2 of the dataset:

- A clearly defined and structured technical dataset specification document which is machine readable
- Clear guidance on how to collect required data supported by example questions where relevant
- A dataset based on the FAIR data principles where possible
- Everything contained within one source file from which the specification website as well as PDF and Excel documents are generated
- Stable and unique identifiers for each data item
- Clearly defined data types with minimal free text fields

All of this will support more standardised data collection by the registries, and more efficient central aggregation and analysis of data by TREAT-NMD.

4d. Patient-Focussed work 2021

Through the revision process consultation patient organisations provided a lot of feedback. Along with SMA Europe we have identified some priority areas for 2021:

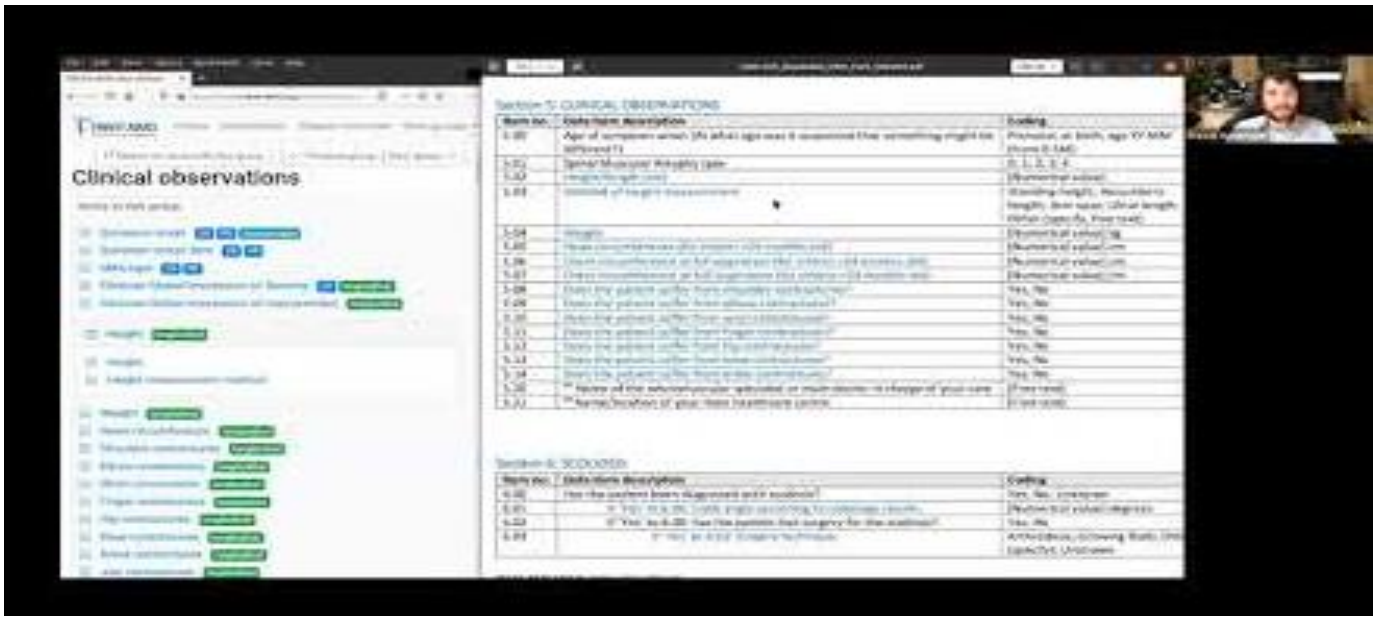
- Patient centricity of both patient-reported and clinician-reported data in the dataset
- Patient Reported Outcome Measures – gap analysis, consensus building
- Influencing pharma on the value of patient input
- Training workshop for patients on the importance and practicalities of patient input in the development of registries and patient reported outcome measures (PROMS).

4e. V2 structure: general principles

Marcel Heidemann provided a comprehensive demonstration and explanation of the key changes from v1 to v2, covering the following topics:

- Version 1 and version 2 differences to naming conventions of data groups and items
- Use of unique identifiers in v2
- Data descriptions
- Data types
- Differences between longitudinal and datestamped items
- Explanation of 'reference periods' and 'episodes'
- Interactive example questions and example data representation

A full overview of is available via the video below:



4f. Guidance Appendix

The version 1 Dataset Manual contained a lot of important 'supplementary' information to support usage of the core dataset. Most of this has been merged into the new v2 dataset where it was appropriate.
Not all the supporting information belongs in a core dataset specification but is nevertheless important. This has been compiled into a support document called the Dataset Appendix. This document covers:

- Project context and background
- Information on feedback, harmonisation, and revisions
- Data sharing and publications information
- Suggested wording for consent and ethics
- Data items not included in v2 but should be collected locally as best practice

4g. Session 6: V2 Contents

Headline changes

Headline changes were outlined to the attendees. These included changes to:

- Stable demographic data used to generate a **PPRL (Privacy Preserving Record Linkage) is now mandatory** in the core dataset. Collecting the data to support this tool is important to reduce and/or link duplicate records across the TGDOC registries and externally with other groups. However, this data will always remain with the local registry and not be submitted to TREAT-NMD.
- **Genetic Diagnosis & SMA type** now matches what is listed in the Standards of Care. The value "Undetermined 5q SMA" has also been included to meet the needs to the community where there is a genetic diagnosis but no clinical diagnosis of the SMA subtype.
- **Ambulatory status** has replaced 'wheelchair use' and is more clearly defined.
- **Spirometry** now includes Peak Cough Flow (PCF) with 'type of airway clearance' removed
- **SMA Disease Modifying Therapies (DMT's)** replaces 'Therapies and Medications' from version 1. Has been expanded to include the new therapies now available.
- **Allopathic drugs & therapy interventions** have moved to co-morbidities group and is further defined as managing symptoms.
- **Patient Global Impression of Severity/Improvement**, PGI-S (baseline) and PGI-I (follow-up) are mandatory in v2
- **Clinician Global Impression of Severity/Improvement** CGI-S and CGI-I have moved to clinical outcomes section
- **General language** was ambiguous in places and now much more directive and specific throughout

Requests for TREAT-NMD to provide guidance on which Motor Measures and Patient Reported Outcomes to collect have been addressed through the publication of the Outcome Measure Library (OML). The OML provides a list of validated motor measures and PROM's included in the dataset, however, registries are free to collect other validated motor measures if they wish. It is felt at this time it would be outside of scope for TREAT-NMD to define what should be collected. There is a lack of validated PROMS and we have identified the patient perspective is not well captured currently.

Research Questions and Future Considerations

All feedback received was valuable, but we have been unable to implement every suggestion. Each item was carefully considered and discussed. The decision for each feedback item can be broadly separated into:

- applied with modifications
- not applied (out of scope or not in line with project objectives)
- not applicable (already changed or removed)
- possibly in scope for future versions (Future Considerations)
- not in scope of core dataset but could be an interesting research question (Research Questions)

The feedback items identified as possible future considerations and research questions were outlined to the participants. During the Tuesday session the group considered the Future Considerations and Research Questions as one large group due to reduced numbers. On Friday, the attendees were split into 4 groups and asked to focus on a specific section. Below are the key points raised from both sessions.

Future Considerations

Reasons for not choosing therapy: good idea to be included to provide some concrete data for payers (e.g. social security) to discuss pricing etc.; good to identify funding issues in various countries.

PROMs: it would be nice to analyse the first 2 years of data on PROMS and motor measures to identify what has been used and maybe give some guidance. Seems difficult to impose one choice of PROM but there could be an option to reduce the list.

Motor Outcome Measures: see PROMs. Difficult for TREAT-NMD to be prescriptive as what to collect is sometimes driven by payers and reimbursement agencies. Would be good if TREAT-NMD could provide some training (such as Train the Trainer sessions).

Sleep Study: not available in all countries but would be interesting to have data were available and possibly consider for future inclusion on a non-mandatory basis.

Bulbar Function: consider a feasibility questionnaire to identify how many registries collect this data already. Good to have test and measures included instead of just asking 'does the patient have….' in order to get meaningful data.
Heart Function: see bulbar function

Independence: understand it is especially important to patients and would be interesting but how could it be included unless in an independence related PROM. Would be difficult to standardise.

Biomarkers: more and more companies are asking for this information but not all registries will have it. Possible not all biomarkers could/should be included. This would need more exploration.

Research Questions

We received a lot of feedback identifying data that would be interesting to collect for assorted reasons. Many of these suggestions were judged as not in scope for the core dataset but could pose an excellent opportunity for registries to engage in collaborative research on topics relevant or interesting to them.

For clinician-reported registries many of these research topics are already being looked at in some way but not necessarily as part of the registry itself. PROMS are important for patients and should cover items like bulbar function, fatigue, impact of scoliosis surgery and meaningful outcomes for patients and would be good to collect this information directly from patients without the need to use paper forms. Another item that was discussed was the intervals at which data is collected for different patients and across different registries, and the difficulties and barriers around standardising this.

Scoliosis: important to have an understanding of the meaningful outcome for patients from surgery.

Respiratory physiotherapy/orthoses: consensus reached that collecting further detail on these areas is outside the scope of the core dataset.

Therapeutic (rehabilitative) interventions: considered essential for the majority due to the potential vastness of the research question and is particularly important to understand.

Outpatient/Care level/Education and work: all interesting and potentially valuable questions in particular understanding the levels of standards of care that patients themselves would like, versus what they are receiving and what their clinicians feel able to provide. This would make an interesting and ongoing research question.

DMT funding: who pushed for the administration of each DMT and who decided on the selection of treatment and why, would make interesting global research and could highlight success factors and barriers.

5. NEXT STEPS

1. Any registries currently conforming to v1 have 6 months to identify relevant changes and make necessary changes to conform with v2 of the SMA Core Dataset. Registries not yet conforming to v1 should work towards v2 using existing timeline
2. Work on 2021 priorities will get underway in the new year
3. Plan to publish the Core SMA Dataset as a technical document.
4. Send OML link with Report and slides

6. APPENDICES

Appendix 1: Q&As from Tuesday 27th October 2020

Do the data type descriptions, such as yes/no; integer etc correlate to any data standard (e.g., CDISC) descriptions?
Not Known, we are not familiar with the CDISC descriptions

What is the definition of SMA4?
This has been added to the dataset

We collect non 5q SMA in Australia
This is a really good example the difference between the core data set and a data collection form. You can add these options into your data collection form but TREAT-NMD would never request this data as part of the core dataset. d to undefined.

Have patient groups approved the use of 'suffer from,' e.g., 'suffer from scoliosis' rather than, 'have scoliosis.' I know it can be a sensitive term, depending on opinion and DMD is looking at this.
We received the same feedback. In v2 we say, 'diagnosed with'. However, registries may of course phrase it however they like on their data collection forms

This example form is very clear and intuitive, it might be very helpful if we could collect data in the same programme (like this example form) to make the Core Dataset more uniform and accurate. For the purposes of national registries, we could collect also different type of information.
We are glad you find it useful. Registries are welcome to use this format if it suits their circumstances. However, many registries prefer to collect differently (and have already been doing so for many years). So as long as their data can still be mapped to the core dataset specification, they do not need to adopt this version of the example form.

How often the data should be updated?
This tends to vary across different registries and hospitals (and depends on age, type, treatment status etc) so we try not to dictate that, as long as everything is date stamped. We would recommend however at least once a year as a minimum

Undetermined type 5q is very different in SMA. Why it is interesting to be collected in this type of registry?
We are presenting the data TREAT-NMD is being collected. It is not always known what type of SMA they typically have due to not knowing point of symptom onset. For those patients it is not always possible to collect which type e.g., 3a or 3b. therefore 5q SMA can be used in these cases.

Cathy: Data like wheelchair use/walking cane etc and type of airway clearance is useful for looking at SoC but I guess this is deemed outside the scope/purpose of the registries? We currently have them IN for DMD, but I imagine will follow your lead
We received an overwhelming amount of feedback for the importance of type of wheelchairs in use. And there are many factors including socio economic factors like access, size of house etc... and an interesting research question. At a county level there is a lot of data that can be collected and useful but to look at it at an international is very difficult and is out of scope.

Some centres store CSF and blood samples before starting Nusinersen therapy and during each subsequent application. Is it good to know which centres are doing this for future collaboration? (*This discussion will be continued offline).*

Are you coding/collecting reasons for not collecting any of the core items?
We ask this for motor measures only (whether a motor measure was taken for that patient at that visit, and if not why).

Appendix 2 Q&As from Friday 30th October 2020

Many times, parents do not remember the exact date when a particular milestone was achieved or lost but may be roughly the month. There is no way to state that. How do we do that?
We specifically ask for partial dates, rather than exact dates, so you can submit partial dates which would be the month and year.

When entering the data but it is not available can you save the entry and come back to it later.
That is something you can provide as a feature of your own registry. The online dataset is only a list of items you should collect as part of your registry, it does not replace your current way of collecting and saving your data. You will need to continue to do this as you currently are until the URP is ready.

Some data is collected by clinicians and some by patients in my registry, is there a way of organising the dataset in such a way that makes it clear who we need to enter each data item?
Theoretically all data items can be collected from a patient or clinician. This is something that would need to be done/built in at a local level. From a TREAT-NMD perspective We just need we would need to know who was provided, whether its patient reported, or clinician reported. We would not mandate this centrally.

The category 'Unknown' has been removed can we use the category undetermined?
You can include whichever categories are most appropriate for you in your data collection forms as long as you can map them to response options in the dataset.

The start stop date completion would be difficult, imprecise and a lot of work. For immunisations, the stop date is not relevant as they are often a onetime shot e.g., annual influenza, vaccine. Also, would be difficult to collect for seasonal vitamin D supplements and bisphosphonates which can be given in a single dose or as a continuous treatment.
The start and stop date would be the same date. We will make sure the establishing conventions is explicit and we will update the data item descriptions to reflect this. Additionally, the reference period (period between follow ups or 12months at baseline) is mandatory for these items but the episodes would be optional. Therefore, if the episode approach does not make sense in any situation curators can choose not to enter it.

When is the next revision of the SMA dataset likely to be?
We do not know for sure; we may make some minor or patch revisions next year if there is something urgent that comes up. There will probably be another full revision before the end of the project in May 2022.