Experimental Neurology 385 (2025) 115100

Contents lists available at ScienceDirect

Experimental Neurology

journal homepage: www.elsevier.com/locate/yexnr

Research paper

Data reporting quality and semantic interoperability increase with community-based data elements (CoDEs). Analysis of the open data commons for spinal cord injury (ODC-SCI)

Anushka Sheoran[a], Kenneth A. Fond[b], Lex Maliga Davis[b], J. Russell Huie[b], Romana Vavrek[c], P.J. Axtman[b], Vance Lemmon[d], John L. Bixby[d], Ubbo Visser[e], John C. Gensel[f], Karim Fouad[c], Adam R. Ferguson[b,g], Jeffrey S. Grethe[h], Anita Bandrowski[a,i], Maryann E. Martone[a,g,h,*], Abel Torres-Espin[b,c,j,**]

[a] *Department of Neuroscience, University of California San Diego, San Diego, CA, USA*
[b] *Department of Neurological Surgery, Weill Institute for Neurosciences, Brain and Spinal Injury Center, University of California San Francisco, San Francisco, CA, USA*
[c] *Department of Physical Therapy, Faculty of Rehabilitation Medicine, Institute for Neuroscience and Mental Health, University of Alberta, Edmonton, Alberta, Canada*
[d] *Miami Project to Cure Paralysis, Department of Neurological Surgery and Department of Molecular and Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, FL, USA*
[e] *Department of Computer Science, University of Miami, Coral Gables, FL, USA*
[f] *Department of Physiology, Spinal Cord and Brain Injury Research Center, University of Kentucky College of Medicine, Lexington, KY, USA*
[g] *San Francisco Veterans Affairs Health Care System, San Francisco, CA, USA*
[h] *Center for Research in Biological Systems, University of California San Diego, San Diego, CA, USA*
[i] *SciCrunch Inc., San Diego, CA, USA*
[j] *School of Public Health Sciences, Faculty of Health, University of Waterloo, Waterloo, Ontario, Canada*

ARTICLE INFO

Keywords:
Data sharing
ODC-SCI
Community-based data elements (CoDEs)
Standards
Spinal cord injury
FAIR

ABSTRACT

Data interoperability is crucial for effectively combining data for scientific inquiry. To facilitate interoperability, data standards such as a common definition of variables are often developed. The Open Data Commons for Spinal Cord Injury (odc-sci.org) has established an initial set of community-based data elements (CoDEs)—a minimal set of variables for sharing—to promote data interoperability in SCI research, aligning with FAIR (Findable, Accessible, Interoperable, and Reusable) data principles. We sought to understand the use of CoDEs by the SCI community to inform current standards adherence and future standards development. We systematically analyzed 39 public datasets in relation to 17 required CoDEs and found variations between reported data and the structure specified by the CoDEs. Overall, we found that the enforcement of data standards improved reporting rates of CoDEs variables. Notably, different variables were found to require different levels of curation to ensure semantic equivalence among datasets. We also uncovered specific reporting habits of researchers such as formatting and naming patterns. A need for different data standards based on the nature of the study (e.g., human study, derivative study) was realized alongside a detailed list of issues that should be addressed when implementing such standards. Among the various approaches to developing data standards, ODC-SCI adopted a semi-formal approach by creating standards that are easy to adopt by the user. Our data-driven evaluation of actual reporting behavior shows that this flexibility can lead to subsequent problems in harmonization. This study serves as a baseline analysis of reporting behaviors for shaping and facilitating data standards.

1. Introduction

The realization of the full value of biomedical data—which is dependent on individual researchers actively sharing their data—holds major implications for biomedical science including increased transparency and credibility, meta-analysis based on individual subject data,

* Corresponding author at: Department of Neuroscience, University of California San Diego, San Diego, CA, USA.
** Corresponding author at: School of Public Health Sciences, Faculty of Health, University of Waterloo, Waterloo, Ontario, Canada.
E-mail addresses: maryann@ncmir.ucsd.edu (M.E. Martone), abel.torresespin@uwaterloo.ca (A. Torres-Espin).

https://doi.org/10.1016/j.expneurol.2024.115100
Received 25 June 2024; Received in revised form 2 December 2024; Accepted 4 December 2024
Available online 7 December 2024

informed experimental design (e.g., sample size estimation from shared data), and the integration of big data assets for knowledge discovery and hypothesis generation (Bandrowski and Martone, 2016; Begley and Ioannidis, 2015; Chan et al., 2014; Collins and Tabak, 2014; Ferguson et al., 2014; Levesque, 2017; Nielson et al., 2015). This belief that data sharing will benefit scientific research has gained prevalence in recent years, resulting in many regulatory organizations requiring data management best practices and mandating the sharing of research data. In the US, the NIH (National Institutes of Health) recently issued the Data Management and Sharing (DMS) policy (effective January 25, 2023) in which researchers submitting for funding must develop a DMS plan, including data sharing strategies (National Institutes of Health, 2023). As efforts of data sharing increase, new challenges have presented themselves, such as the need for establishing data standards across members of a research community to facilitate the sharing and integration of data.

In 2014, the FAIR data-sharing principles (Findable, Accessible, Interoperable, and Reusable) were developed as a framework to facilitate the development of efficient and effective data-sharing strategies (Wilkinson et al., 2016). One of the core recommendations of FAIR is following certain data standards which can be defined as a set of rules on various dataset characteristics, such as format, structure, metadata, and definitions (e.g., data elements) that have been agreed upon by a group of experts in the field. Notably, even if the data is shared using the same file formats and structures, the semantic interoperability of shared variables, i.e. that variables have the same meaning and refer to the same concept, is a primary choke point for data integration and harmonization. The FAIR principle of interoperability discusses the use of controlled and recognized vocabulary to ensure clear semantics during data integration.

The Open Data Commons for Spinal Cord Injury (ODC-SCI) is a community-governed data repository for the management and sharing of SCI data built following FAIR principles (Callahan et al., 2017; Fouad et al., 2020; Torres-Espín et al., 2022). The ultimate goal of the ODC-SCI is not just the sharing of single data sets, but the ability to compare and aggregate data across studies to facilitate data reuse for different scientific purposes (Nielson et al., 2015; Almeida et al., 2022; Chou et al., 2022a; Torres-Espín et al., 2021). To move toward interoperability and reusability, the ODC-SCI community board developed a list of minimal data elements that should be included in all animal SCI studies shared through the ODC-SCI. We refer to these data elements as Community-based Data Elements (CoDEs), since their definitions have been agreed upon by representatives of the ODC-SCI community. We view CoDEs as a precursor to the development of more formal Common Data Elements through the process being developed by the NIH and hope our work informs that process.

In this study we conducted a systematic analysis of data elements reported in public datasets before and after CoDEs were required for public data release in the ODC-SCI to answer two major questions: 1) How are researchers implementing and interpreting the CoDEs? 2) What is the impact of soft-enforcing CoDEs in the data reporting behaviors? Both lenses provide insight into the feasibility of building a more harmonizable data repository based on the semantic interoperability of the CoDEs. Better harmonizability across datasets is an important step toward more transparent and repeatable research, and to enable automatic, faster, and more accurate data integration for secondary data use by researchers within the laboratory and beyond. All of these benefits will ultimately accelerate research and translation. Furthermore, implementing CoDEs encourages standardization among lab members and can also streamline collaboration efforts.

2. Methods

2.1. CoDEs development and implementation

The initial draft of the minimal set of variables or CoDEs was extracted from common elements from the Minimum Information about a Spinal Cord Injury Experiment (MIASCI) model (Lemmon et al., 2014) by the ODC-SCI team, including some of the authors of this work. The draft was submitted for internal evaluation by additional experts within the ODC-SCI team, resulting in a list of 17 elements. This list was provided to the ODC-SCI Community Board, an independent group of stakeholders from the SCI community (i.e., researchers, clinicians, funders, and people with lived experience) for review and endorsement. The Community Board approved the minimal set of variables, deemed the standard necessary for publicly releasing data from animal studies through the ODC-SCI. The resulting elements were branded CoDEs, to distinguish them from officially recognized Common Data Elements listed by the NIH. Starting in 2021, the CoDEs were introduced in the ODC-SCI documentation and tutorials describing the publication process, the information was provided to users submitting for publication, the template data dictionary was updated with the CoDEs, and incorporated in the procedures on data checks that all datasets must undergo for publication (ODC-SCI documentation link: https://fdilab.gitbook.io/open-data-commons/guides/odc-standards/common-terminology/odc-sci-codes). To facilitate the data curation team reviewing datasets for publication, an automatic validation check for the presence of these CoDEs was also developed. To facilitate user uptake and compliance, the implementation was progressive; initially made as a recommendation, followed by a soft requirement guided by the ODC-SCI curation team. Prior to the implementation of the CoDEs, the curational data primarily helped researchers structure their data to comply with ODC-SCI formatting. Once this minimal set of variables had been established, the curatorial team further helped researchers navigate and curate their data in alignment with the CoDEs. The definition of these variables are provided in Table 1, and the template data dictionary can be found in the protocol adopted in this analysis (dx.doi.org/10.17504/protocols.io.5qpvo3wodv4o/v1) or directly on ODC-SCI documentation.

2.2. Analyzed data

All datasets and accompanying data dictionaries up to May 2022 - the time at which this work was conducted- that were publicly available in the ODC-SCI (https://odc-sci.org, RRID:SCR_016673) (n = 39) (Ferguson et al., 2018; Liu et al., 2019; Torres Espín et al., 2021a; Torres Espín et al., 2021b; Schmidt et al., n.d.; Kyritsis et al., 2021; Puko and McTigue, 2020; Aceves et al., 2020; Schmidt et al., 2020a; Schmidt et al., 2020b; Batty et al., 2020; Schmidt et al., 2021a; Keller et al., 2021a; Ehsanian et al., 2020; Fenrich et al., 2021; Brennan et al., 2021; Batty et al., 2021; Nielson et al., 2021; Schmidt et al., 2021b; Almeida et al., 2021; Parhizi et al., 2021; Mah et al., 2022a; Stehlik et al., 2021; Keller et al., 2021b; Stewart et al., 2021; Mah et al., 2022b; Brennan et al., 2022; Cerqueira et al., 2022; Mirkiani et al., 2022; Chou et al., 2022b; Stewart et al., 2022; Madalena et al., 2022; Mifflin et al., 2022; Raposo et al., 2022; FRAUSSEN et al., 2022; Sydney-Smith et al., 2022; Metz et al., 2022; Liu et al., 2022; Metz et al., 2023) were downloaded and used for analysis. The link and reference citation for all datasets are found in Table S1. Although CoDEs were designed for standardizing reporting in data from animal studies, datasets from human studies were also analyzed with the secondary goal of understanding how we can serve those studies better in the future. Of all analyzed datasets, 18 datasets were published before the CoDEs implementation in April 2021, and 21 datasets after. Table 2 shows key characteristics between datasets in the pre and post CoDEs implementation groups.

2.3. Systematic annotation of shared variables

To ensure effective organization, management, and a systematic annotation and analysis of shared variables in all datasets, we established a protocol with appropriate documentation, naming conventions, and clear annotation guidelines. To maintain consistency in data extraction and annotation decisions, a spreadsheet template was created

Table 1
List of community-based data elements with their definitions according to the ODC-SCI documentation.

CoDE	Definition
Subject_ID	Unique identifiers for each subject in the dataset
Species	Species of the subject
Strain	Strain of the subject
Animal_origin	Vendor or origin of the animal
Age	Age of the subject at start of experiment. If age is available at different timepoints, age is provided at the corresponding time in a corresponding time/timepoint variable
Weight	Weight of the subject at start of experiment. If weight is available at different timepoints, weight is provided at the corresponding time in a corresponding time/ timepoint variable
Sex	Sex of the subject
Group	Name or identifier of the experimental group in which the subject was included, if any
Laboratory	Name of laboratory, usually the PI
StudyLeader	Name of person responsible for overseeing project
Exclusion_in_origin_study	Whether the subject was included in the study that originated the data. "Total exclusion" if excluded from the entire study, otherwise, specify experiment or measures of which the animal was excluded if any. For example: animals that were run in behavior but maybe tissue is loss and excluded from histological analyses. Reasons for exclusion might be specify in the exclusion_reason variable.
Exclusion_reason	Reason by which the subject was excluded from the study that originated the data as specified in the Exclusion_in_origin_study variable
Cause_of_Death	Cause of death (e.g. perfusion/necropsy, died during surgery, euthanized for health reasons, etc)
Injury_type	Type or model of injury used in the subject (e.g. contusion, complete transaction, partial section)
Injury_device	Name of the device used for the injury
Injury_level	Spinal cord level at which the injury was performed including segment (e.g. cervical; C) and number (e.g. C5)
Injury_details	Other details referent to the injury that might be relevant to understand the severity and type of injury performed

(Sheoran et al., 2024), and versioning was used to keep track of different curation runs—defined by their annotation decisions. DOI links to their respective datasets and data validation options (Ex: dropdown-style for different columns) were implemented, whenever applicable. The set of features to be extracted for each CoDE was discussed among members of the team from a preliminary annotation of the first 5 datasets to finalize the annotation protocol for the whole set. To ensure consistency, all data was extracted and annotated by the same researcher. After 50 % of the datasets were annotated, two independent members of the team spot-checked the extraction and ensured protocol conformance. The same two reviewers were consulted to resolve special cases and uncertain annotations. Data extraction and annotation were periodically reviewed by researchers on the team for feedback and resolution of ambiguities. A log file was created and maintained to document, reference, and track specific annotation decisions made during the process. Qualitative observations made during the data extraction and annotation process were also documented in a separate file for each one of the datasets and CoDEs.

The data extraction and annotation process involved reviewing each dataset, and the respective data dictionary and metadata published by the ODC-SCI to extract a set of features that reflected how the data were acquired and documented. For each of the 17 required CoDEs, some extracted features—for example, does the variable exist in the data dictionary—were constant, while other criteria were specific for a given CoDE, e.g., value sets. In cases where a variable was not included in the data dictionary or if the content was missing in the dataset, we consulted the associated paper to determine if more information was available. If

Table 2
Mean characteristics of included datasets.

	Pre CoDEs $n = 18$	Post CoDEs $n = 21$	Total $n = 39$
Species			
Mouse	4 (22)*	7 (33.3)	11 (28.2)
Rat	11 (61.1)	4 (19)	15 (38.4)
Human	3 (16.6)	8 (38.1)	11 (28.2)
Rat and Mouse	0 (0)	1 (4.7)	1 (2.5)
Minipig	0 (0)	1 (4.7)	1 (2.5)
Number of rows	311 [164.2–945.7]**	140 [54–230]	182 [89–404]
Number of unique subjects	46.5 [33.7–148.2]	40 [29–91]	42 [30–115]
Number of variables	80 [48–101]	82 [49–134]	80 [48–124]
Injury Type			
Contusion	6 (33.3)	9 (42.9)	15 (38.5)
Crush	1 (5.6)	0 (0)	1 (2.6)
Dorsolateral Quadrant Section	4 (22.2)	0 (0)	4 (10.3)
Hemisection	1 (5.6)	0 (0)	1 (2.6)
Multiple	2 (11.1)	0 (0)	2 (5.1)
None	1 (5.6)	4 (19)	5 (12.8)
Transection	0 (0)	1 (4.8)	1 (2.6)
Unilateral Pyramidotomy	0 (0)	2 (9.5)	2 (5.1)
Not applicable	3 (16.7)	5 (23.8)	8 (20.5)
Injury level			
Cervical	9 (50)	3 (14.3)	12 (30.8)
Multiple	3 (16.7)	5 (23.8)	8 (20.5)
Not applicable	1 (5.6)	4 (19)	5 (12.8)
Pyramid	0 (0)	2 (9.5)	2 (5.1)
Thoracic	5 (27.8)	7 (33.3)	12 (30.8)
Sex			
Both	6 (33.3)	7 (33.3)	13 (33.3)
Female	10 (55.6)	8 (38.1)	18 (46.2)
Male	1 (5.6)	0 (0)	1 (2.6)
Not applicable	1 (5.6)	6 (28.6)	7 (17.9)
Percent of CoDEs reported	43.4 % (30)***	92.7 % (7.1)	69.9 % (17)
Percent of CoDEs reported following ODC naming	21.8 % (23.7)	83.7 % (7.1)	55 % (12.8)

*Count (%), **Median [Q1 – Q3], *** Mean (SD).

the data was not found or referenced in the paper, we assumed that the data was not recorded during the research process. If a variable was not present in the data dictionary, the corresponding columns in the analysis sheet were filled with "n/a" to indicate its absence. All annotated data can be found in the supplementary documentation (file A). Note that the datasets have been anonymized for sharing.

2.4. Analysis

The data was analyzed from two perspectives: a standards adoption analysis and a qualitative usage analysis. In the standards adoption analysis, we compare how users behaved in reporting variables before versus after the CoDEs were introduced. A proportion test was conducted in R (RRID:SCR_001905) (R Core Team, 2023) using the *prop.test ()* function to test whether introducing reporting recommendations or requirements into the ODC affected the reporting of CoDEs. An adjusted p value <0.05 using the Benjamini and Hochberg method for controlling the false discovery rate was considered significant. In addition, the data was qualitatively analyzed to address the issue of harmonization potential across the 39 datasets. This data was visualized using MS Power BI software (RRID:SCR_025118).

3. Results

3.1. Summary statistics on reporting CoDEs

On average, 70 % of datasets include a given CoDE. However, the reporting rates varied depending on the specific CoDE. Subject_ID was reported 100 % of the time because it is the only variable required for initial data upload to ODC-SCI. Sex was reported in 97 % of the analyzed ODC-SCI datasets, and Species was reported in 90 % of them. Conversely, some CoDEs were frequently not included in the datasets, such as Injury_details, Cause_of_Death, and Animal_origin at 43 %, 46 %, and 54 % of the analyzed datasets, respectively. Specific rates of reporting for each CoDE can be found in Fig. 1. Further, we found that in certain cases the variable was mentioned in the supporting documentation of the dataset (e.g., summary metadata abstract) or the accompanying manuscripts but not included in the dataset itself. For example, 4 datasets had not reported the species variable in their dataset but all 4 had stated the species in the manuscript. Similarly, 8 of the 10 datasets that had not reported Injury_level in the dataset stated this variable in their manuscript. Interestingly, all but one of these 8 cases had induced injury at the same level for all their subjects. When it comes to variable naming, scientists within the SCI community predominantly reported CoDEs using the ODC-SCI naming convention, accounting for 81 % of the instances (Fig. 1), however, the rates varied before and after the implementation of the CoDEs as explained below.

In addition to the presence or absence of CoDEs, we also examined the data type and value sets of the reported variables. We found that the content for any given CoDE was not provided 21 % of the time, i.e. the variable was either left blank or a null indicator such as “not applicable” was provided (Fig. 2). Datasets utilized three primary data types for the 17 CoDEs: string, integer, and decimal. On average, 91 % of all reported data was in string format, 8 % in integer format, and 1 % in decimal format. Even for variables that could be considered numerical, the string format was predominantly used. This situation was observed when presenting numerical data as a range or including the unit in the content of the variable itself. For example, the “Weight” variable was represented in string format 61 % of the time, integer format 21 % of the time, and decimal format 18 % of the time. A breakdown of all CoDEs can be found in Fig. 2.

3.2. Pre and post CoDEs implementation

Of the total 39 datasets that were analyzed, 18 datasets were published prior to the implementation of CoDEs as a requirement for public release, and 21 datasets were published after their implementation. Prior implementation of CoDEs, the average reporting rate for these data elements was 43 % (Fig. 3). The implementation of CoDEs led to a significant improvement in data reporting behavior to an average of 93 % of datasets across all CoDEs (Table 3). The increased reporting rate was accompanied by a decrease in the rate of reporting the actual contents of the CoDEs. Prior to the implementation of data CoDEs, the content filling rate for the CoDEs was 99 %; which means while there were often very few datasets reporting a given CoDE, the content was usually filled if the CoDE was there. After implementation of the CoDEs, the average rate of filling the content of a variable dropped to 71 %. This drop was mainly produced by studies for which specific CoDEs were not applicable, but still included in the dataset for consistency with the standards. Prior to the implementation, on average, only 56 % of the datasets used the exact CoDEs naming convention, while after implementation, this percentage increased to 90 % (Fig. 3). For example, while Subject_ID was reported in 100 % of the cases due to being the only variable required for uploading to the ODC-SCI, the naming of such variable changed from only 12 % of the dataset following the convention prior to implementation of the CoDEs to 100 % after.

3.3. Qualitative description of reporting

Throughout the 39 datasets, the data was reported in different forms, carrying different semantic meanings. Data was often riddled with data quality issues, each of which required different types and complexity levels of computations in order to achieve harmonized data. Moreover, two types of datasets were notable for their lack of or variable use of

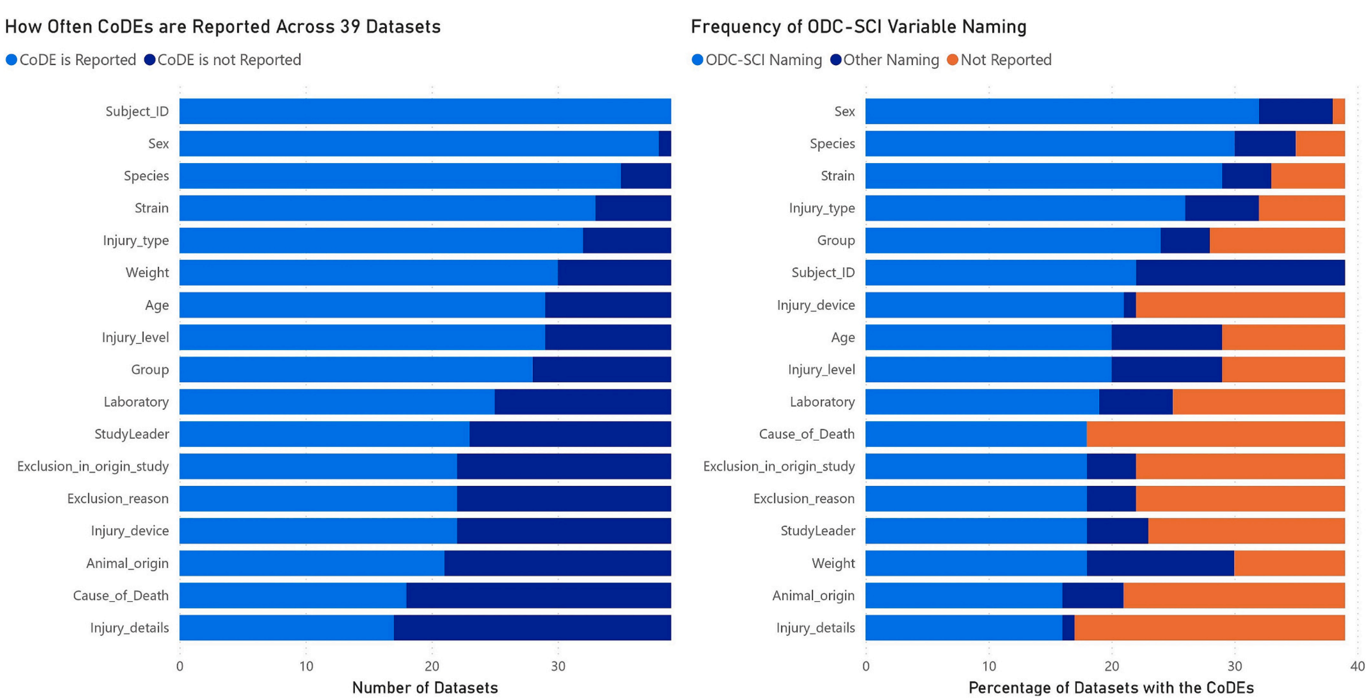

Fig. 1. The reporting rates chart (left-side) shows how often each of the CoDEs was included on average across 39 datasets; shown in decreasing order. The frequency of ODC-SCI variable naming shows how often the reported CoDEs were assigned the same name as detailed by the ODC-SCI, along with a inclusion of datasets that did not report the CoDE (in orange).

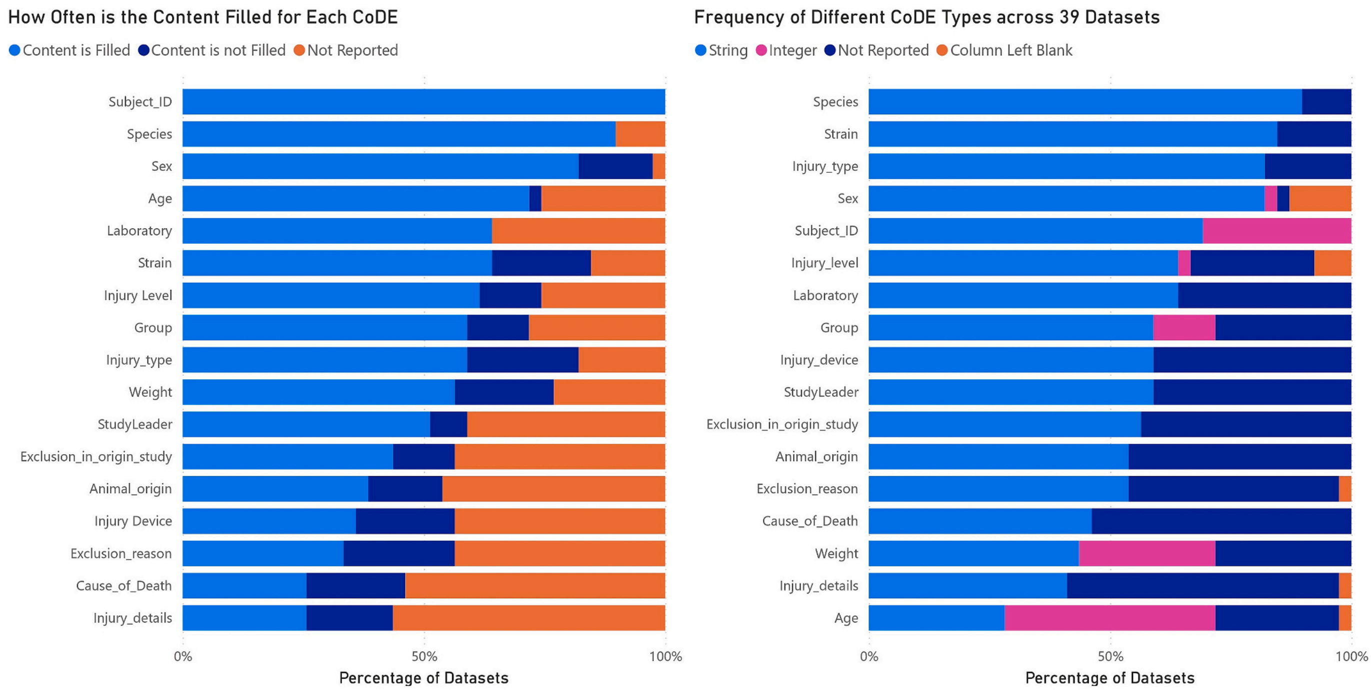

Fig. 2. The two charts together describe the content of data submitted for each CoDE. The left-most chart portrays how often the content of a variable was filled when the variable is reported. The right chart demonstrates the occurrence of data reported as strings, integers, or decimals, while also indicating the percentage of datasets that did not report the variable or have blank columns for that variable (while still including it in their data dictionary/dataset).

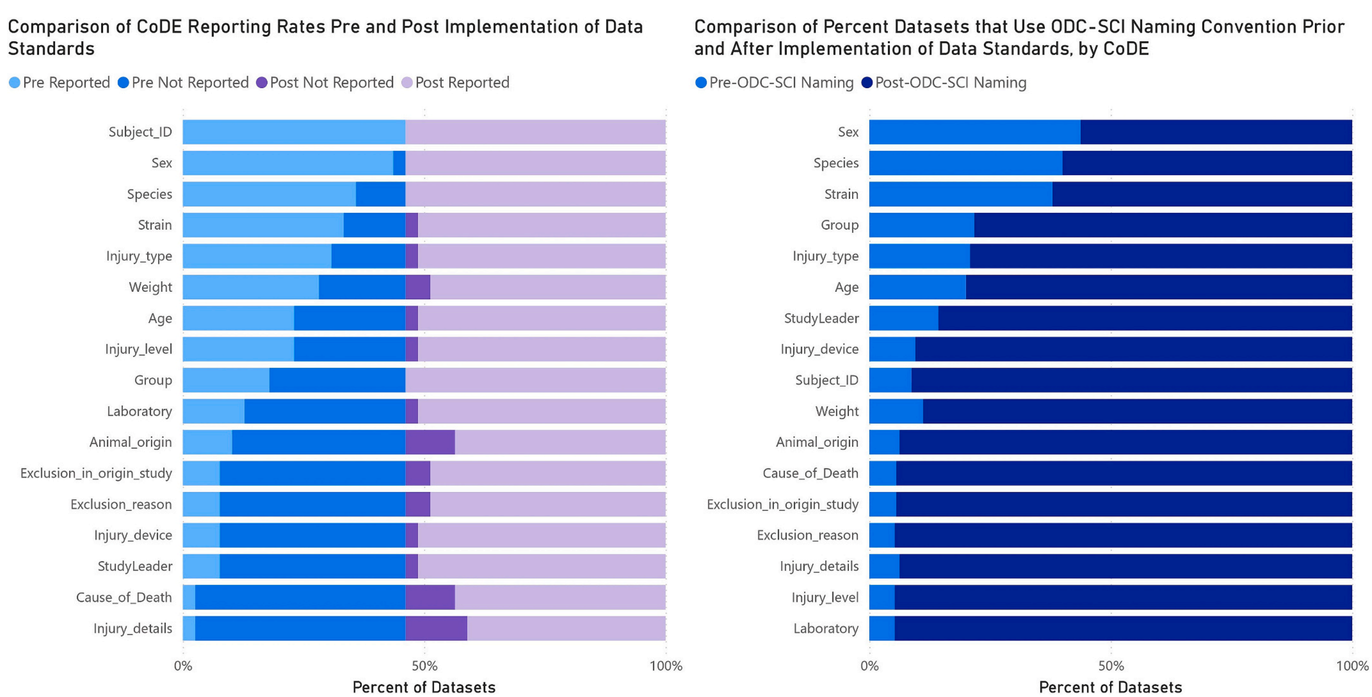

Fig. 3. These charts portray the shift in reporting behaviors pre and post standards implementation. The left-hand chart shows a dramatic increase in reporting rates after the standardization (purple). The right hand chart shows of all the datasets that had named their CoDEs in alignment with ODC-SCI, the percent that had been published prior in comparison to after the data standards had been implemented (showing a much larger percent of after/dark-blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CoDEs studies; these are derivatives from previous datasets and human studies. It was found that on average, 39 % of datasets included a given CoDE in secondary studies which is significantly lower than the 70 % rate when analyzing all 39 datasets. It is important to note that only 3 of the 39 datasets in our study were derivative studies and thus, we must be cautious generalizing that secondary studies exhibit lower CoDE reporting rates. However, one of the datasets included two different subject ID's, one from the current study and one mapping it to the original dataset. These kinds of variations may amplify when harmonizing a higher number of derivative studies. Human studies presented

Table 3
This table shows the results of a proportion test before and after CoDEs implementation.

	Variable Exist			Variable follows naming convention		
CoDEs	Pre	Post	Adj P value	Pre	Post	Adj P value
Subject_ID	1	1	1	0.11	1	**<0.001**
Species	0.77	1	0.09	0.66	0.85	0.3
Strain	0.72	0.952	0.13	0.61	0.85	0.18
Animal_origin	0.22	0.81	**0.001**	0.05	0.71	**<0.001**
Age	0.5	0.952	**<0.01**	0.22	0.76	**<0.01**
Weight	0.61	0.904	0.09	0.11	0.76	**<0.001**
Sex	0.94	1	0.93	0.77	0.85	0.82
Group	0.38	1	**<0.001**	0.27	0.85	**<0.001**
Laboratory	0.27	0.952	**<0.001**	0.05	0.85	**<0.001**
StudyLeader	0.16	0.952	**<0.001**	0.16	0.85	**<0.001**
Exclusion_in_origin_study	0.16	0.904	**<0.001**	0.05	0.80	**<0.001**
Exclusion_reason	0.16	0.904	**<0.001**	0.05	0.85	**<0.001**
Cause_of_Death	0.05	0.809	**<0.001**	0.05	0.80	**<0.001**
Injury_type	0.66	0.952	0.076	0.27	0.90	**<0.001**
Injury_device	0.16	0.952	**<0.001**	0.11	0.90	**<0.001**
Injury_level	0.5	0.952	**<0.01**	0.05	0.85	**<0.001**
Injury_details	0.05	0.762	**<0.001**	0.05	0.71	**<0.001**

data differently due to compliance with HIPAA regulations and the inapplicability of certain CoDEs, which were originally designed for animal studies (e.g., 'Strain' and 'Animal_Origin'). In such cases, variables were either not reported or filled with synonyms of 'not available.' Of all human studies, 64 % of the time, the studies either did not have a CoDE reported or the content filled out. In the case of Strain, Animal_-Origin, Weight, Cause_of_Death, Injury_Device and Injury_level, all human studies either did not report the variable or did not fill out the content of the CoDE. 74 % of cases where the content was not filled, and 34 % of cases where the CoDE was not reported, were human studies (Fig. 4). Along with these special cases, there were three key types of data issues that were discovered in the qualitative analysis of these datasets: coding differences, data quality issues, and semantic differences (Table 4).

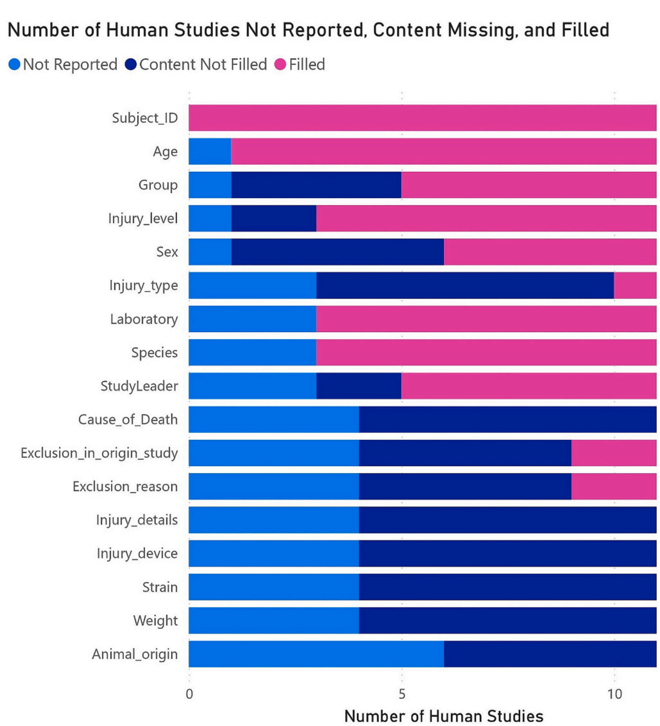

Fig. 4. This figure shows the number of datasets (out of the total number of human studies, $n = 11$) that had the CoDEs not reported, their content not filled out even when the CoDE was present (content missing), or the CoDE was present and with content.

4. Discussion

We provide evidence of the positive impact of implementing a minimal set of standard data elements and their potential effect on increasing the interoperability of shared data. The substantial increase in the overall reporting of standard variables following their implementations highlights the effectiveness of such initiatives in promoting data consistency and enhancing data usability. Nonetheless, we also highlight that simple inclusion of a CoDE (or by extension, a CDE) in a dataset does not mean that the data are automatically interoperable. In particular, we show that there is considerable variation in how a given CoDE can be interpreted, often deviating to such an extent that the data cannot be easily combined. Additionally, our results provide estimates of how implementation of CoDEs reduces unstructured reporting of variables often manifested through actions like implicit mentions (i.e. excluding variables in their dataset but including them in your manuscript). This implicit mentioning makes operations such as cohort selection across multiple datasets much more difficult and increases downstream tasks such as searching and pooling when attempting to harmonize across datasets. An example is the benefit of including the CoDEs even when they might not apply to a given dataset. For example, the CoDE "Exclusion_reason" might not be meaningful for a dataset that did not exclude any animal. Yet, the CoDE can be reported and filled with values pre-specified values such as "No exclusion", which will be important for harmonizing across datasets. Although the increase in standardization is not a surprise as it was the goal for the CoDEs and these have been required for publishing, having measures of the magnitude of this impact provides a starting point for benchmarking future initiatives. Our analysis provides key insights on reporting behavior and gaps on the CoDEs itself that should be considered for further development. In addition, we provide lessons learned during the process of implementing and evaluating the use of these data elements, which should serve as a starting point for discussing semantic standardization in SCI. Finally, our study underscores the value of merging manual CDE creation, rooted in ontological principles, with a data-driven methodology. We apply this hybrid approach by enforcing CoDEs semi-formally and evaluating their impact. Our study contributes to the iterative improvement of CoDEs, while also reflecting the ongoing evolution of this process (Fig. 5).

4.1. Recommendations on improving ODC-SCI standards

In examining the complexities of data standardization within our selected datasets, it is apparent that while data standards are relevant to improving interoperability, there are inherent trade-offs between system

Table 4
Qualitative reporting behaviors observed throughout the 39 datasets and 17 CoDEs.

Data Issue	Explanation	Subtype	Example cases	Quantification
Coding Differences	Variations in data formats, structures	Data Formats	Split between using "F/M "and "male/female" for 'Sex'.	11 datasets used 'F/M' and 20 datasets used 'male/female'
		Units	Unit inconsistencies in 'Age' and 'Weight, such as some datasets using weeks and others using months etc.	1, 13, 4, and 10 datasets used days, weeks, months and years respectively
		Synonyms/ Naming Conventions	'AnimalID' and 'OrigSubNum' were both utilized to refer to the 'Subject_ID' CoDE	22, 2, 6, and 4 datasets used Subject_ID, SubjectID, Animal_ID and ID; 5 datasets used other variations
Data Quality	Different types of errors affect data quality	Missing Data	Leaving certain cells or columns blank; definition different notions of missing data	2 datasets where the column header was included in the datasets, (e.g. 'organism_sex') however, the entire column is left blank
		Human errors	Apparent copying and pasting mistakes.	2 cases where the definitions in the data dictionary shifted down one row to their respective CoDE
Semantic Differences	Deviations were found in the interpretation of the CoDEs' meaning, and nuanced definitions	Multiple similar Variables	Temporal measurements like weight at different time points, where different labels such as 'weight_w1,' 'weight_w2,' 'weight_w3,' and 'weight_w0'	6 datasets had multiple temporal measurements. Specifically, these datasets had 5–8 different CoDEs for weight
		Vague definitions	Injury_type' CoDE, exhibited variability in reported components; some said C5 contusion Spinal Cord Injury; and others just said contusion. In 'Group' CoDE, where the values assigned to this CoDE were often unrelated and lacked standard meaning across different studies.	6 datasets included the injury level (e.g. cervical or C5) in their injury_type value despite there being another CoDE called 'Injury_level'
		Granularity	Reporting variables like 'Age' and 'Weight' as ranges (e.g. 2–11 weeks old). Further there is granularity beyond units when some datasets only included the group in 'Injury Level' (e.g., cervical, thoracic, lumbar, sacrum, and coccyx); others included the specific number of the vertebrae (e.g., reporting Cervical versus C2)	Of the 29 datasets that reported 'Age', 5 datasets reported it as a range as opposed to a single value

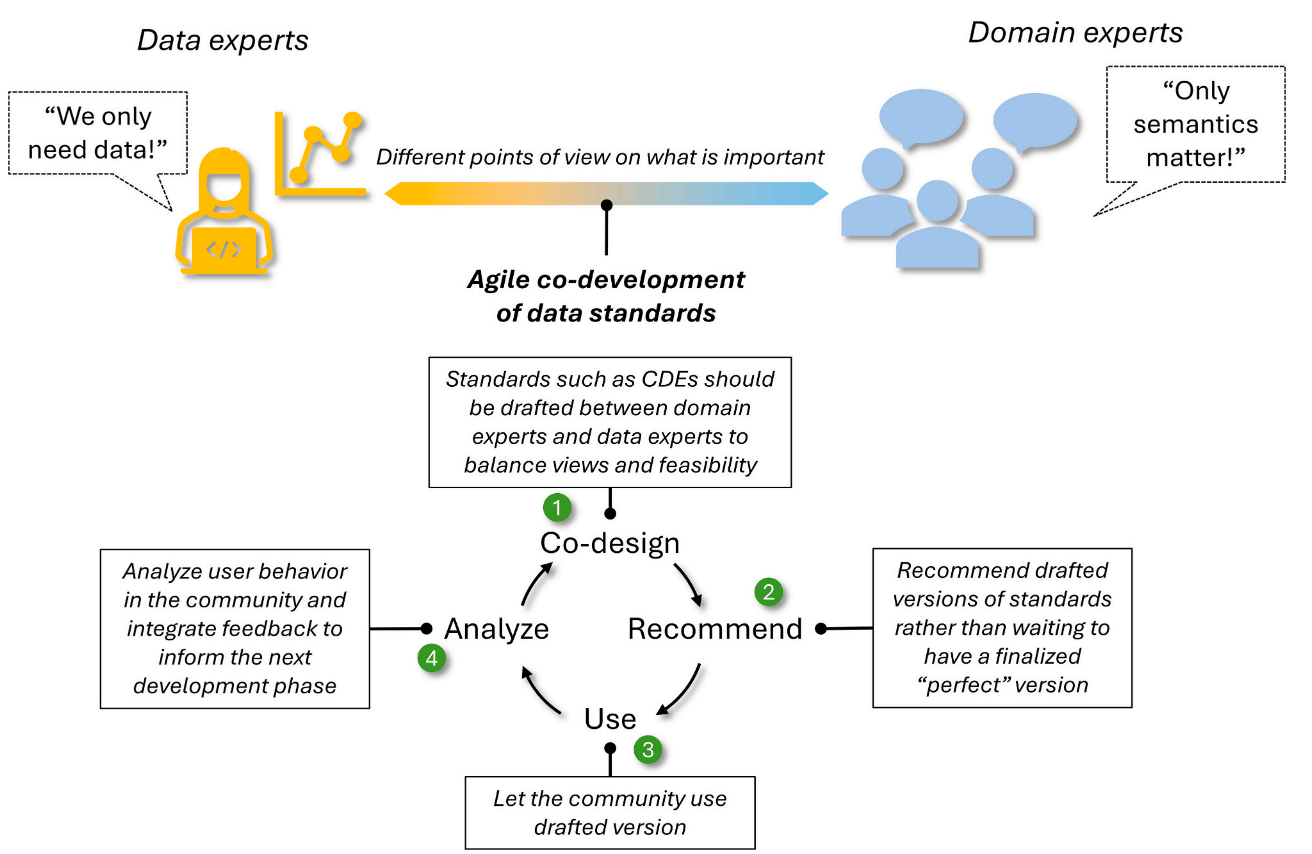

Fig. 5. Conceptual framework for the co-development and evaluation of data standards using agile design principles.

automation and user workload/burden. Further, data sharing repositories like the ODC-SCI operate under limited resources including the time and effort required to harmonize data and thus resources must be optimized to critically promote adherence to best sharing practices. The semantic differences and variations in naming conventions for CoDEs demonstrate the challenges associated with achieving semantic interoperability across diverse research contexts and highlight the need for clearer and more precise definitions of data elements and their permitted values and structure. Along with clear specifications of the CoDEs, practical examples should be included to promote consistent interpretation across datasets. Examples should also be included for cases such as missing data. The implementation of these specifications must be done in creative and effective ways but will depend on the context of the CoDEs themselves. For instance, when addressing variables such as age, establishing guidelines that address units while avoiding reporting discrepancies, such as ranges or strings, requires solutions like template-driven data collection. This is something ODC-SCI is potentially going to create. As repositories can enforce certain validation checks, it is important to provide complementary resources such as documentation, data dictionaries, protocols, manuscripts, and even other mentioned/related datasets. Moreover, it is vital to engage researchers in discussions that define standards, guidelines, ontologies, or controlled vocabularies to consider their needs and align standards to their practice.

The type of analysis performed here clearly shows areas where CoDEs can be improved. ODC-SCI must identify a different set of data standards for human studies, which could be derived or adopted from current international standard initiatives (Biering-Sørensen et al., 2023). It may also be valuable to consider the guidelines when using derivative studies which may be missing some variables and thus would not be able to comply with our CoDEs standards. Furthermore, our analysis indicated that the definition of CoDEs 'Group' and 'Injury_Type' must be improved to be more specific, while the inclusion of 'Injury_details' and 'Cause_of_death' in the required CoDEs must be reconsidered to make it optional as many datasets had substituted these columns for 'n/a'. For a complete set of recommendations, consult Table 5.

Limitations.

This study has several limitations. First, the analysis was based on a small number of datasets, which does not fully represent the entire range of SCI research data. Future studies should aim to include a larger and potentially more diverse sample to capture a more accurate image of data reporting behaviors. Additionally, the study focused primarily on analyzing reporting behavior and characteristics without delving into the underlying reasons behind these behaviors. Future research could consider conducting surveys or interviews with members of the SCI community to gain a deeper understanding of their motivations, challenges, and decision-making processes regarding what data is being reported. While this study focused on the influence of data standards on reporting rates and patterns, other factors may also contribute to the observed behaviors. For example, researchers' workload, time constraints, and familiarity with data standards could impact their reporting practices. Additionally, differences in disciplinary backgrounds, research objectives, and study designs could contribute to variations in reporting behaviors. For example, we note that there are differences in the number of studies per species pre and post CoDEs implementation. Exploring these alternative explanations can offer valuable insights into the contextual factors influencing data reporting and addition to standards, and guide the development of targeted interventions and support mechanisms.

5. Conclusion

Establishing the most effective data standards involves navigating a delicate balance. On one hand, if the established/platform standards deviate too much from the data a researcher naturally produces, the likelihood of the data not being reported increases. On the other hand, in the absence of predefined guidelines, the diversity of reported data makes semantic interoperability and harmonization (and thus the comparisons between studies) challenging. Our analysis emphasizes the importance of establishing and requiring data standards in SCI research to enhance data interoperability. We have also included a set of recommendations for researchers to produce and report data, which are available in the supplementary material (Table S2). This study shares key factors to consider when designing or adopting these data standards. As the amounts of data produced exponentially increase, it becomes imperative for the community to collectively collaborate and maximize the interoperability and reusability of the produced and shared data.

Funding

A.R.F: NIH/NINDS (R01NS122888; UH3NS10689; U24NS122732 to A.R.F.), Department of Veterans Affairs (I01RX002245, I01RX002787, I01BX005871, I50BX005878 to A.R.F.), Wings for Life foundation (A.R.F.), Craig H. Neilsen Foundation (A.R.F., KF), UC Noyce Initiative (A.R.F.). K·F: Wings For Life Foundation, Craig H. Neilsen Foundation, Canadian Institute of Health Research (CIHR), Canada Research Chairs (CRC). A.T.E: Wings For Life Foundation, Canadian Institute of Health Research (CIHR).

CRediT authorship contribution statement

Anushka Sheoran: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kenneth A. Fond:** Writing – review & editing. **Lex Maliga Davis:** Writing – review & editing. **J. Russell Huie:** Writing – review & editing. **Romana Vavrek:** Writing –

Table 5
Recommendations for improving the CoDEs in ODC-SCI, divided into general recommendations applicable throughout the CoDEs and specific changes that should be made.

RECOMMENDATIONS for improving CoDEs in ODC-SCI
General Recommendations

- Create a different set of CoDEs for studies sharing data collected from human subjects
- Update example data dictionary with clear examples on the use of each CoDE
- Specify in the definition of CoDEs a standard use for missing information
- Specify in the definition of CoDEs the type of an element, and examples of the valid form of reporting. For example, Age as a numeric variable, with units specified in the data dictionary
- CoDEs should be divided in minimal set if elements required for publication, and supplemental elements
- Improve instructions in the ODC-SCI documentation for clarity on the use of CoDEs
- Align CoDEs with existing CDEs

Recommendations about specific CoDEs

- The CoDE 'Group' must be refined due to its semantic ambiguity. A new name for the CoDE should be considered with higher specificity, for example "Experimental Group". In addition, the new definition should specify what "group" means in different contexts and provide recommendations for when deviations of this definition might occur.
- 'Injury_Type' must be redefined as the components of 'type' are too ambiguous and result in great semantic diversity. The CoDE should provide clear examples of accepted entries of types and suggestions for deviations of the common types of injuries.
- 'Injury_details' should be removed as required by CoDE for publication. Instead should be considered a supplemental element.
- 'Cause_of_death' should be removed as a required CoDE or be conditional when cause of death is unexpected and reported in the data. In addition, the CoDE should be redefined to specify if it is referred to as the cause of death as a natural progression of an experiment or due to circumstances unforeseen in the study design.

review & editing. **P.J. Axtman:** Writing – review & editing. **Vance Lemmon:** Writing – review & editing, Methodology, Conceptualization. **John L. Bixby:** Writing – review & editing, Methodology, Conceptualization. **Ubbo Visser:** Writing – review & editing, Methodology, Conceptualization. **John C. Gensel:** Writing – review & editing, Methodology, Conceptualization. **Karim Fouad:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Adam R. Ferguson:** Writing – review & editing, Funding acquisition, Conceptualization. **Jeffrey S. Grethe:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization. **Anita Bandrowski:** Writing – review & editing, Resources, Methodology, Conceptualization. **Maryann E. Martone:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Abel Torres-Espin:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

A.R.F. discloses Spine X DSMB and the funding provided below. A.B. is a co-founder and serves as CEO of SciCrunch Inc., a company that works with publishers to improve publications. These conflicts have been reviewed by the UCSD COI office. M.E.M. is a founder and has equity in SciCrunch Inc., a tech startup out of UCSD that provides tools and services supporting scientific rigor and reproducibility. ATE: co-owner of YEG scientific, a company producing behavioral research tools for animal models of disease. All other authors have nothing to declare.

Data availability

The data is publicly available at the ODC-SCI

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.expneurol.2024.115100.

References

Aceves, M., Dietz, V.A., Dulin, J., et al., 2020. Repeated Measurements of Hindlimb CatWalk Variables in Normal Rats, vol. 532 KB. https://doi.org/10.34945/F54S3W.
Almeida, C., Torres Espín, A., Huie, J.R., et al., 2021. Hemodynamics, Weight and Overground Locomotion from 1125 Rats with Different Spinal Cord Injury Thoracic Contusion Severities Recovered from the Multicenter Spinal Cord Injury Study (MASCIS). https://doi.org/10.34945/F5QG66.
Almeida, C.A., Torres-Espin, A., Huie, J.R., et al., 2022. Excavating FAIR data: the case of the multicenter animal spinal cord injury study (MASCIS), blood pressure, and neuro-recovery. Neuroinformatics 20 (1), 39–52. https://doi.org/10.1007/s12021-021-09512-z.
Bandrowski, A.E., Martone, M.E., 2016. RRIDs: a simple step toward improving reproducibility through rigor and transparency of experimental methods. Neuron 90 (3), 434–436. https://doi.org/10.1016/j.neuron.2016.04.030.
Batty, N., Vavrek, R., Raposo, P., et al., 2021. Behaviour and Histopathology after Simultaneous Cortical Infusion of PKA and EPAC Agonists and Rehabilitative Forelimb Motor Training after Incomplete Cervical Spinal Cord Injury in Female Rats. https://doi.org/10.34945/F56C7W.
Batty, N.J., Torres-Espin, A., Vavrek, R., et al., 2020. Behavior and Histopathology after Single-Session Cortical Electrical Stimulation and Rehabilitative Forelimb Motor Training in Cervical Spinal Cord Injury in Females Rats. https://doi.org/10.34945/F59885.
Begley, C.G., Ioannidis, J.P.A., 2015. Reproducibility in science: improving the standard for basic and preclinical research. Circ. Res. 116 (1), 116–126. https://doi.org/10.1161/CIRCRESAHA.114.303819.
Biering-Sørensen, F., Charlifue, S., Chen, Y., et al., 2023. International spinal cord injury Core data set (version 3.0)—including standardization of reporting. Spinal Cord 61 (1), 65–68. https://doi.org/10.1038/s41393-022-00862-2.
Brennan, F., Li, Y., Wang, C., et al., 2022. Microglia Coordinate Cellular Interactions during Spinal Cord Repair after Contusion or Crush Spinal Cord Injury in Female Mice. https://doi.org/10.34945/F5B012.
Brennan, F.H., Nobel, B.T., Wang, Y., et al., 2021. Acute Post-Injury Blockade of A2d-1 Calcium Channel Subunits on Pathological Autonomic Plasticity after T3 Crush Spinal Cord Injury in Female Mice. https://doi.org/10.34945/F5V30C.
Callahan, A., Anderson, K.D., Beattie, M.S., et al., 2017. Developing a data sharing community for spinal cord injury research. Exp. Neurol. 295, 135–143. https://doi.org/10.1016/j.expneurol.2017.05.012.
Cerqueira, S., Benavides, S., Lee, H., et al., 2022. Neuroprotection and Bromodomain and Extra-Terminal Domain (BET) Protein Inhibition after T8 Spinal Cord Contusion in Female Rats and Mice. https://doi.org/10.34945/F5ZP4P.
Chan, A.-W., Song, F., Vickers, A., et al., 2014. Increasing value and reducing waste: addressing inaccessible research. Lancet 383 (9913), 257–266. https://doi.org/10.1016/S0140-6736(13)62296-5.
Chou, A., Torres-Espín, A., Huie, J.R., et al., 2022a. Empowering data sharing and analytics through the open data commons for traumatic brain injury research. Neurotra. Rep. 3 (1), 139–157. https://doi.org/10.1089/neur.2021.0061.
Chou, A., Beattie, M., Bresnahan, J., et al., 2022b. AIS Impairment Scale, Injury Characteristics, Intraoperative Data after Spinal Cord Injury in Patients from a TRACK-SCI Cohort. https://doi.org/10.34945/F5KG6Z.
Collins, F.S., Tabak, L.A., 2014. Policy: NIH plans to enhance reproducibility. Nature 505 (7485), 612–613. https://doi.org/10.1038/505612a.
Ehsanian, R., Haefeli, J., Crew, J.D., et al., 2020. Surgical Blood Pressure Management and Expected Motor Recovery in Individuals with Traumatic Spinal Cord Injury. https://doi.org/10.34945/F58C7H.
Fenrich, K., Hallworth, B., Vavrek, R., et al., 2021. Self-Directed Reaching and Grasping Rehabilitation Using Automatic Pellet Presentation System after Cervical Dorsolateral Quadrant Injury in Female Rats. https://doi.org/10.34945/F5ZW20.
Ferguson, A.R., Nielson, J.L., Cragin, M.H., et al., 2014. Big data from small data: data-sharing in the "long tail" of neuroscience. Nat. Neurosci. 17 (11), 1442–1447. https://doi.org/10.1038/nn.3838.
Ferguson, A.R., Irvine, K.-A., Gensel, J.C., et al., 2018. Cervical (C5), unilateral spinal cord injury with diverse injury modalities, multiple behavioral outcomes, and histopathology. Open Data Commons Spinal Cord Injury 95. https://doi.org/10.7295/W9T72FMZ.
Fouad, K., Bixby, J.L., Callahan, A., et al., 2020. FAIR SCI ahead: the evolution of the open data commons for pre-clinical spinal cord injury research. J. Neurotrauma 37 (6), 831–838. https://doi.org/10.1089/neu.2019.6674.
FRAUSSEN, J., BECKERS, L., van Laake-Geelen, C., et al., 2022. Immune Phenotyping SCI Patients versus Healthy Controls. https://doi.org/10.34945/F5588X.
Keller, A., Singh, G., Sommerfeld, J., et al., 2021a. Hemodynamic Stability, Incidence of Pain, Skin Redness, Autonomic Dysreflexia and Other Safety Related Outcome Measures of Transcutaneous Electrical Spinal Cord Stimulation in Children with Cervical and Thoracic Spinal Cord Injuries. https://doi.org/10.34945/F5HP4N.
Keller, A., Singh, G., Joel, Sommerfeld, et al., 2021b. Trunk Kinematic and Center of Pressure Displacement Time Series Data during Transcutaneous Lumbosacral Spinal Cord Stimulation and Passive Pelvic Tilt in Human Pediatric Participants with Trunk Control Impairments Due to Cervical or Thoracic Spinal Cord Injury (Pilot Clinical Trial NCT03975634). https://doi.org/10.34945/F5NC7X.
Kyritsis, N., Torres-Espin, A., Schupp, P.G., et al., 2021. Acute Blood RNA Sequencing from SCI Human Patients. https://doi.org/10.34945/F5QC7J.
Lemmon, V.P., Ferguson, A.R., Popovich, P.G., et al., 2014. Minimum information about a spinal cord injury experiment: a proposed reporting standard for spinal cord injury experiments. J. Neurotrauma 31 (15), 1354–1361. https://doi.org/10.1089/neu.2014.3400.
Levesque, R.J.R., 2017. Data sharing mandates, developmental science, and responsibly supporting authors. J. Youth Adolesc. 46 (12), 2401–2406. https://doi.org/10.1007/s10964-017-0741-1.
Liu, S., Xing, C., Wei, Haitao, et al., 2022. CatWalk Gait Parameters at 8 Weeks after T10 Thoracic Contusion Using the MASCIS Impactor Model III in Male and Female 8- to 10-Week Old Rats. https://doi.org/10.34945/F5VS3D.
Liu, Y., Wang, X., Li, W., et al., 2019. T10 Lateral Hemisection Spinal Cord Injury with Multiple Histological and Behavioral Outcomes, vol. 643. https://doi.org/10.7295/W9HQ3X20.
Madalena, K., Brennan, F., Popovich, P., 2022. Genetic Deletion of the Glucocorticoid Receptor in Cx3cr1+ Myeloid Cells Is Neuroprotective and Improves Motor Recovery in Female Mice after 75 Kdyne T9 Contusion Spinal Cord Injury. https://doi.org/10.34945/F56887.
Mah, K.M., Wu, W., Al-Ali, H., et al., 2022a. Contralateral Axon Sprouting in Male and Female c57Bl/6J Mouse Cervical Spinal Cord by Uninjured Corticospinal Axons Following Unilateral Pyramidotomy and Cortical Injection of Kinase Inhibitors. https://doi.org/10.34945/F5T01D.
Mah, K.M., Lemmon, V., Bixby, J., 2022b. Contralateral sprouting in C57Bl/J6 mouse cervical spinal cord after unilateral Pyramidotomy and cortical injection of RO48 8 weeks following injury. Figure 3. https://doi.org/10.34945/F5JG6N.
Metz, K., Concha-Matos, I., Li, Y., et al., 2022. Facilitation of Sensory Axon Conduction to Motoneurons during Cortical or Sensory Evoked Primary Afferent Depolarization (PAD) in Humans. https://doi.org/10.34945/F5WS3Q.
Metz, K., Matos, I.C., Hari, Krishnapriya, et al., 2023. H-Reflex Modulation from Sensory Conditioning in Non Injured Female and Male Human Participants. https://doi.org/10.34945/F5M88Z.
Mifflin, K., Brennan, F., Guan, Z., et al., 2022. Lung Immunity after Complete Thoracic Spinal Cord Transection in Female Mice. https://doi.org/10.34945/F52G69.
Mirkiani, S., Roszko, D., O'Sullivan, C., et al., 2022. Gait Kinematic Parameters and Onset/Offset of EMG Muscle Activities during Overground Walking in Healthy Female Yucatan Minipigs at Different Speed Ranges from 0.4 m/s to 1.6 m/s. https://doi.org/10.34945/F5V01Q.

National Institutes of Health, 2023. 2023 NIH Data Management and Sharing Policy. Available from: https://oir.nih.gov/sourcebook/intramural-program-oversight/intramural-data-sharing/2023-nih-data-management-sharing-policy.
Nielson, J.L., Paquette, J., Liu, A.W., et al., 2015. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. Nat. Commun. 6 (1), 8581. https://doi.org/10.1038/ncomms9581.
Nielson, J.L., Liu, Aiwen W., Guandique, C.F., et al., 2021. A Comprehensive Set of Hemodynamic Parameters, Blood Gases and Hemoglobin Levels during Spinal Cord Injury Procedure Followed by Weekly Assessments of Locomotor and Bladder Function Recovery with Survival Histology Data from 334 Rats in the MASCIS Trial. https://doi.org/10.34945/F5V884.
Parhizi, B., Barss, T., Mushahwar, V., 2021. Alteration of H-Reflex and MEP Amplitude in the Flexor Carpi Radialis Muscle with Transcutaneous Spinal Cord Stimulation during Static and Leg Cycling Tasks in Neurologically Intact Male and Female Humans. https://doi.org/10.34945/F5B59T.
Puko, N., McTigue, D.M., 2020. Data for Manuscript: Delayed Short-Term Tamoxifen Treatment Does Not Promote Remyelination or Neuron Sparing after Spinal Cord Injury, vol. 28KB. https://doi.org/10.34945/F5QP4H.
R Core Team, 2023. R: A Language and Environment for Statistical Computing.
Raposo, P., Lucas-Osma, A., Schmidt, E., et al., 2022. Comparing the Effect of Skilled Forelimb Rehabilitative Training after Unilateral Cervical Contusion and Dorsal Quadrant Spinal Cord Injury in Female Lewis Rats. https://doi.org/10.34945/F5DS3C.
Schmidt, E., Raposo, P., Vavrek, R., et al., 2021a. Lipopolysaccharide Treatment in the Subacute Stage of Cervical Spinal Cord Injury Enhances Motor Recovery and Increases Anxiety-like Behaviour in Female Rats. https://doi.org/10.34945/F5FW2B.
Schmidt, E., Raposo, P., Madsen, K., et al., 2021b. Effects of a Fecal Transplant from Anxious Donors on Rehabilitative Training, Microbiota Composition, Systemic Inflammation and Behaviour Following a Unilateral Cervical Spinal Contusion (C5, 125kdyn) in Female Lewis Rats. https://doi.org/10.34945/F5XW2P.
Schmidt, E., Torres-Espin, A., Raposo, P., et al., 2019. Data for the Manuscript: Fecal Transplant Prevents Gut Dysbiosis and Anxiety-like Behaviour After Spinal Cord Injury in Rats, p. 792. https://doi.org/10.7295/W97942VQ.
Schmidt, E.K., Raposo, P., Torres-Espin, A., et al., 2020a. Minocycline treatment for acute cervical spinal cord injury in females rats. Microbiota Compos. https://doi.org/10.34945/F5F30N.
Schmidt, E.K., Raposo, P., Torres-Espin, A., et al., 2020b. Minocycline treatment for acute cervical spinal cord injury in female rats: behavioural. Histopathol. Syst. Immune Effects. https://doi.org/10.34945/F5JS3M.
Sheoran, A., Martone, M.E., Torres-Espin, A., 2024. Protocol for Systematic Analysis Data Elements. https://doi.org/10.17504/protocols.io.5qpvo3wodv4o/v1.
Stehlik, K., Pelisch, N., Kurpad, Shekar, et al., 2021. Treatment with Gallium Maltolate after Spinal Cord Thoracic Contusion in Female Mice, Functional Outcome and Inflammatory Markers. https://doi.org/10.34945/F5S59V.
Stewart, A., MacLean, S., Stromberg, A., et al., 2021. Studying Sex as a Biological Variable: Considerations of Weight, Anemia and Mortality Following Thoracic Spinal Cord Contusion Injury in Male and Female Mice. https://doi.org/10.34945/F5D01P.
Stewart, A., Glaser, E., Mott, C., et al., 2022. Assessment of Spinal Cord Glutathione Levels and Regulatory Pathways at 1- and 3-Days Post-Injury after 60 kDyn T9 Spinal-Contusion Using the Infinite Horizons Impactor in Male and Female 4- and 14-Month Old Mice and Evaluating n-Acetylcysteine Amide as a Therapy in Female 4- and 14-Month Old Mice to Restore Spinal Cord Levels of Glutathione and to Protect against Oxidative Stress and Affect Locomotor Outcomes and Thermal Sensitivity by 28-Days after Spinal Cord Injury. https://doi.org/10.34945/F5XS31.
Sydney-Smith, J.D., Koltchev, A.M., Moon, L., et al., 2022. Delayed Viral Vector Mediated Delivery of Neurotrophin-3 Improves Skilled Hindlimb Placement on the Horizontal Ladder, Hindlimb Coordination during Swimming, and Stability Walking in the Open Field after T9 Thoracic Contusion in Female Rats. https://doi.org/10.34945/F51G60.
Torres Espín, A., Haefeli, J., Ehsanian, R., et al., 2021a. ASIA Impairment Scale, Level of Injury, Intraoperative Time Series Mean Arterial Pressure and Heart Rate after Spinal Cord Injury in Patients in a Multi-Site Retrospective TRACK-SCI Cohort: Site 1 of 2. https://doi.org/10.34945/F5R59J.
Torres Espín, A., Haefeli, J., Ehsanian, R., et al., 2021b. ASIA Impairment Scale, Level of Injury, Intraoperative Time Series Mean Arterial Pressure and Heart Rate after Spinal Cord Injury in Patients in a Multi-Site Retrospective TRACK-SCI Cohort: Site 2 of 2. https://doi.org/10.34945/F5MG68.
Torres-Espín, A., Haefeli, J., Ehsanian, R., et al., 2021. Topological network analysis of patient similarity for precision management of acute blood pressure in spinal cord injury. eLife 10, e68015. https://doi.org/10.7554/eLife.68015.
Torres-Espín, A., Almeida, C.A., Chou, A., et al., 2022. Promoting FAIR data through community-driven agile design: the open data commons for spinal cord injury (odc-sci.org). Neuroinformatics 20 (1), 203–219. https://doi.org/10.1007/s12021-021-09533-8.
Wilkinson, M.D., Dumontier, M., IjJ, Aalbersberg, et al., 2016. The FAIR guiding principles for scientific data management and stewardship. Sci Data 3 (1), 160018. https://doi.org/10.1038/sdata.2016.18.