DISS. ETH NO. 30175

Data-driven approaches to maximize clinical impact in spinal cord injury research

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES
(Dr. sc. ETH Zurich)

presented by

LUCIE MANON BOURGUIGNON
MSc ETH, ETH Zurich

born on 12.01.1999

accepted on the recommendation of

Prof. Dr. Catherine R. Jutzeler (examiner)
Prof. Dr. Karsten Borgwardt (co-examiner)
Prof. Dr. Armin Curt (co-examiner)
Prof. Dr. Torsten Hothorn (co-examiner)

2024

Contents

List of Figures

List of Tables

Acronyms

χ^2 chi-squared. 163, 165, 166
10MWT 10-meter walking test. 33, 46, 47
6MWT 6-minute walking test. 33, 46, 47

AANS american association of neurological surgeons. 94
AI artifical intelligence. 4, 10, 34, 204, 297, 298
AIS American spinal injury association (ASIA) impairment scale. 8, 27–29, 35, 45–48, 51–58, 61, 69, 74–77, 79, 81, 82, 85, 97–101, 111, 138, 144, 145, 159–161, 165–174, 180, 182, 183, 186–188, 190, 192, 243, 244, 281, 282
ALAT alanine aminotransferase. 78, 79, 82, 274, 275, 278, 279
ANOVA analysis of variance. 71, 74, 75, 79
ASAT aspartate aminotransferase. 78, 79, 274, 275, 278, 279
ASIA American spinal injury association. 8, 11, 27, 28, 45, 46, 52, 54, 57, 69, 74, 76, 97, 99, 138, 145, 159, 166, 167, 171, 180, 182, 187, 192, 244, 281, 282

BBB Basso Beattie Bresnahan. 123, 124
BBB blood brain barrier. 95, 101, 103
BGU berufsgenossenschaftliche unfallklinik. 74, 140
BMS Basso mouse scale. 123, 124

CBC complete blood count. 72, 77
CCA complete case analysis. 157, 158, 162, 170, 173
CCS central cord syndrome. 9, 184, 188, 193, 195, 215, 295
CI confidence interval. 7, 164, 167, 170, 171, 173
CNS congress of neurological surgeons. 94
CNS central nervous system. 24, 85
COVID-19 coronavirus disease 2019. 32, 212
CRP C-reactive protein. 79, 274, 275, 279
CSF cerebrospinal fluid. 85, 138
CTCAE common terminology criteria for adverse events. 95, 108

DAP deep anal pressure. 27

df degree of freedom. 52, 56, 71, 76, 77, 144, 165, 172

DILI drug-induced liver injuries. 85

EHRs electronic health records. 216

EMG electromyography. 244

EMSCI European multicenter study on human spinal cord injury. 8, 9, 15, 17, 20, 33, 39–41, 43, 45, 47, 49, 51, 52, 54, 56–59, 61–63, 65, 67, 69, 74, 87, 139, 155, 173, 175, 178–181, 183, 186–193, 195, 198, 216, 295

EP electrophysiology. 124

ETH Eidgenössische Technische Hochschule. 88, 118

EU GDPR European general data protection regulation. 75, 87, 102, 143, 176, 185, 197

FDA food and drug administration. 73, 74

FIM functional independence measure. 35, 49, 244

Gamma-GT γ-glutamyl transferase. 78, 79, 274, 275, 278, 279

GBR gradient boosting regressor. 141, 148, 276, 277, 283, 285, 286, 288–294

GFAP glial fibrillary acidic protein. 138, 151

GGT c-glutamyl transferase. 79

GM-1 gangliosidosis 1. 31, 49, 73, 93, 94, 159, 181

GRASSP graded redefined assessment of strength sensibility and prehension. 25

HIV human immunodeficiency virus. 157

ICH international council for harmonisation of technical requirements for pharmaceuticals for human use. 95

INR international normalized ratio. 79, 274, 275, 278, 279

ISNCSCI international standards for neurological classification of spinal cord injury. 6, 25, 26, 28, 33, 35, 41, 46, 47, 49, 51, 74, 139, 181, 243, 244

k-NN k-nearest neighbors. 162, 169, 171

LASSO least absolute shrinkage and selection operator. 141, 146–148, 276, 277, 283–294

LEMS lower extremity motor score. 7, 9, 26, 29, 35, 46, 134, 137–145, 147–151, 156, 158–162, 164–173, 180–183, 186–188, 191–193, 214–217, 279–282, 289–294

LightGBM light gradient boosting machine. 141, 146, 148, 276, 277, 284, 285, 287–294

LME lésion de la moelle épinière. 17–19

LOCF last observation carried forward. 16, 140, 142, 144, 153, 157, 161, 162, 169, 171–174, 183, 184, 187, 189, 191, 192, 214, 281

LR linear regression. 162, 164, 169, 170, 173, 289–294

LT light touch. 26, 27, 243, 244

MAE mean absolute error. 7, 137, 143, 146–148, 164, 169, 170, 283–294
MAR missing at random. 155, 156, 158, 160, 161, 165, 168–171, 173, 174
MCAR missing completely at random. 47, 155, 156, 158, 160, 161, 163, 165, 168–174
MCH mean corpuscular hemoglobin. 78, 79
MCHC mean corpuscular hemoglobin concentration. 77, 79, 82, 274, 275, 278, 279
MCV mean corpuscular volume. 6, 77, 79, 83, 274, 275, 278, 279
MedDRA medical dictionary for regulatory activities. 95
MEPs motor evoked potentials. 124
ML machine learning. 15, 16, 34–36, 134, 143, 204, 212, 214, 216
MNAR missing not at random. 155, 156, 158, 160, 161, 165, 168–171, 173, 174
MP methylprednisolone. 123, 126
MPSS methylprednisolone sodium succinate. 31, 93, 123, 159
MS motor score. 183, 188

NA not available. 145, 166, 167, 281, 282
NASCIS national acute spinal cord injury study. 31, 93, 159
NF-L neurofilament light chain. 138, 151
NIH national institutes of health. 29
NISCI Nogo inhibition in spinal cord injury. 32, 62, 85, 217
NLI neurological level of injury. 27, 28, 45–48, 52, 56, 57, 60, 98, 99, 139, 159, 160, 172, 181, 183, 184, 186–188, 191–193, 295
NSAID nonsteroidal anti-inflammatory. 87

pmm predictive mean matching. 157, 162, 168–170, 173
PP pin prick. 26, 27, 243, 244
PR phenomenal recovery. 7, 10, 179, 180, 182–184, 186–196, 295–297
PRISMA preferred reporting items for systematic reviews and meta-analysis. 6, 117, 119
PROSPERO international prospective register of systematic reviews. 117

Q1 first quartile. 144, 145, 166, 256
Q3 third quartile. 144, 145, 166, 256

RBF radial basis function. 162, 170, 171
RCT randomized clinical trial. 6, 29–34, 39, 119, 120, 126, 127, 129, 134, 212, 213
RF random forest. 141, 147, 162, 168, 169, 276, 277, 283–294
RHSCIR Rick Hansen spinal cord injury registry. 175
RISCIS riluzole in spinal cord injury study. 32
RMSE root mean squared error. 143, 146, 164, 169, 170, 277, 284, 285, 287–294
RoB risk of bias. 6, 114, 119, 120, 126, 127

ROM range of motion. 27

SCBB spinal cord blood barrier. 91
SCEPs spinal cord evoked potentials. 124
SCI spinal cord injury. 2, 4, 6, 8, 15, 16, 20–22, 24–26, 28, 29, 31–37, 39, 41–46, 49, 50, 54–56, 59–64, 69, 71–74, 80, 81, 83–87, 90–94, 96, 97, 99, 102–106, 108, 114–120, 122–130, 132, 134, 137–140, 149, 150, 153, 155, 157–159, 161, 162, 173–175, 179–182, 195, 196, 199, 212–218
SCIM spinal cord independence measure. 28, 33, 35, 46, 47, 181, 216, 217, 244
SCIMS spinal cord injury model systems. 59, 60
SD standard deviation. 9, 51, 52, 55, 57, 69, 76, 98, 99, 120, 122, 141, 144–147, 149, 166, 167, 169, 187, 188, 190–192, 256, 281–288
SEPs somatosensory evoked potentials. 124
SOC system organ class. 95
STROBE strengthening the reporting of observational studies in epidemiology. 45
SVM support vector machines. 141, 146–148, 162, 170, 171, 276, 277, 283, 285, 286, 288–294

TBI traumatic brain injury. 157, 175
TLT total light touch. 26, 46
TMS total motor score. 26, 46
TPP total pin prick. 26, 46
TRACK-TBI transforming research and clinical knowledge in traumatic brain injury. 175
TSS total sensory score. 46

UEMS upper extremity motor score. 26, 29, 32, 35, 46, 56, 159, 181, 183, 187, 188, 191–193, 216, 217
USA United States of America. 25, 31, 32, 34, 49, 59, 64, 73, 93, 97, 99, 159, 181

VAC voluntary anal contraction. 27

WISCI walking index for spinal cord injury. 33, 46, 47, 244

XGBoost extreme gradient boosting. 141, 148, 276, 277, 283, 285, 286, 288–294

Summary

Spinal cord injury (SCI) is a medical condition resulting from damage to the spinal cord. As the spinal cord represents the primary connection between the brain and peripheral organ systems, a disruption leads to numerous impairments in locomotion, sensation, and organ functions. A SCI therefore undermines the overall quality of life and independence of the individuals affected and their families. This realization is particularly relevant since the field still lacks an intervention, pharmacological or otherwise, to promote the restoration of functions and/or regeneration of the damaged spinal cord. While clinical trials conducted to date did not find any promising intervention, they support the field by thoroughly collecting large amounts of data. The surge of data science, including statistical and machine learning (ML) methods, holds the promise to uncover new insights in better defining and enhancing recovery following SCI by extensively investigating retrospective data.

This thesis aimed to leverage the potential of data science to maximize clinical impact in SCI research. This effort was pursued around three pillars: (i) enlarging the surveillance within clinical studies, (ii) promoting best methodological practices from data science applied to SCI research, and (iii) highlighting the importance of effective research dissemination.

Firstly, the general context in which this thesis fits is outlined in **Part A**. Then, **Part B** sets benchmarks through the secondary analyses of major datasets collected in the field, namely the Sygen clinical trial, the European multicenter study on human spinal cord injury (EMSCI), the Murnau center, and SCIRehab cohorts. **Chapter 1** studies how recovery following SCI evolved over the last two decades and showed that, despite an evolving standard of care, neurological recovery has remained largely unchanged in this period. This observation paves the way to using historical patient data to enrich placebo arms in future clinical trials, therefore maximing the exposure to the intervention of interest. **Chapter 2** describes the natural progression of serological markers following SCI, providing an additional surveillance tool when testing pharmacological interventions that might affect individuals beyond the primary injury targeted. Similarly, studies testing the effect of new pharmacological interventions may be affected by interactions

with medications that are prescribed following injury. We therefore exhaustively report in **Chapter 3** what constitutes the current pharmacological standard of care. We reveal an extensive polypharmacy that individuals with SCI are subject to. To characterize the effects of this polypharmacy on SCI recovery, we systematically review the literature in **Chapter 4** and describe both clinical and pre-clinical evidence supporting beneficial or detrimental effects in neurological recovery following SCI.

Secondly, **Part C** adapts known methods from data science to be translated to SCI research applications. We initially investigate the potential of serological biomarkers as predictors of motor recovery in **Chapter 5**. This analysis shows that accounting for clinical characteristics specific to the condition improved predictions, while still being limited by factors such as missing data leading to small cohorts to be studied. We therefore further characterise missing data in the context of SCI in **Chapter 6**. Here we develop guidelines on how to handle missing information based on simulation studies. We demonstrate that last observation carried forward imputation is a viable approach for imputing missing neurological outcomes after SCI, owing to the distinctive plateau in recovery starting around six months after initial trauma. Finally, **Chapter 7** explores the concept of positive deviance to detect individuals recovering beyond clinical expectations. While data extracted from such individuals may impair the performance of ML prediction models, understanding the mechanisms underlying their phenomenal recovery holds the potential to uncover patterns leading to improved recovery.

Lastly, **Part D** underlines the importance of science communication to effectively link research from bench to bedside. **Chapter 8** particularly promotes the use of new tools such as interactive data visualization to elevate the presentation of research outcomes while leaning towards more transparent and accessible research not only for the scientific and clinical communities but also the individuals affected, their families and society.

Overall, this thesis contributes to the in-depth benchmarking of decisive elements guiding clinical studies in SCI, such as neurological recovery, the evolution of serological biomarkers, and medications commonly prescribed as part of the standard of care. This work leads the path towards improved data analyses and recovery prediction following SCI by integrating known characteristics from the condition. In the context of the SCI research field, this thesis participates in revising the approaches employed to discover interventions to improve recovery following SCI.

Résumé

Une lésion de la moelle épinière (LME) est un état pathologique résultant d'une atteinte de la moelle épinière. Cet organe étant la connexion principale entre le cerveau et les organes périphériques, une perturbation de cette connexion entraîne de nombreux déficits: locomoteurs, sensoriels, régulation des organes périphériques. Par conséquant, une LME nuit à la qualité de vie globale et à l'indépendance des personnes touchées et de leur famille. Cette constatation est d'autant plus critique qu'à ce jour, il n'existe pas d'intervention, pharmacologique ou autre, pour promouvoir la restauration des fonctions et/ou la régénération de la moelle épinière endommagée. Si les essais cliniques menés jusqu'à présent n'ont pas permis de découvrir une intervention prometteuse, ils ont néanmoins permis de collecter de grandes quantités de données. L'essor des sciences des données, incluant les méthodes statistiques et d'apprentissage automatique, promet d'apporter de nouvelles approches afin de mieux définir et d'améliorer le rétablissement après une LME en examinant de manière approfondie les données rétrospectives.

Cette thèse vise à exploiter le potentiel de la science des données pour maximiser l'impact clinique de la recherche liée aux LMEs. Cet effort s'est articulé autour de trois piliers : (i) élargir la surveillance au sein des études cliniques, (ii) promouvoir un usage optimal et adapté des méthodes des sciences des données appliquées à la recherche liée aux LMEs, et (iii) mettre en évidence l'importance d'une diffusion efficace de la recherche.

Premièrement, le contexte général dans lequel s'inscrit cette thèse est décrit dans la **Partie A**. Ensuite, la **Partie B** établit des repères épidémiologiques grâce aux analyses secondaires des principaux ensembles de données collectées, à savoir l'essai clinique Sygen, la cohorte European multicenter study on human spinal cord injury (EMSCI), le centre Murnau et la cohorte SCIRehab. Le **Chapitre 1** étudie l'évolution du rétablissement neurologique après une LME au cours des deux dernières décennies et a montré que, malgré les changements en termes de normes de soins, le rétablissement neurologique est resté globalement stable au cours de cette période. Cette observation ouvre la voie vers l'utilisation de patients historiques pour enrichir les groupes placebo dans les futurs essais cliniques, maximisant ainsi l'exposition à l'intervention testée. Le **Chapitre 2** décrit la progression naturelle des marqueurs sérologiques après une LME, offrant un

outil de surveillance supplémentaire lors des essais cliniques évaluant des interventions pharmacologiques susceptibles d'affecter les individus au-delà de la LME. De même, les études testant l'effet de nouvelles interventions pharmacologiques peuvent être affectées par des interactions avec les médicaments prescrits à la suite du traumatisme. Nous rapportons donc ici de manière exhaustive dans le **Chapitre 3** ce qui constitue actuellement la norme en termes de soins pharmacologiques. Nous mettons en évidence une large polypharmacie à laquelle les individus sont soumis. Pour caractériser les effets de cette polypharmacie sur le rétablissement neurologique, nous avons examiné la littérature de manière systématique dans le **Chapitre 4** et décrivons les preuves cliniques et précliniques soutenant les effets bénéfiques ou préjudiciables de ces traitements sur le rétablissement neurologique après une LME.

Deuxièmement, la **Partie C** expose l'adaptation de méthodes issues des sciences des données pour l'étude des LMEs. Nous étudions d'abord le potentiel des biomarqueurs sérologiques en tant que prédicteurs du rétablissement moteur dans le **Chapitre 5**. Cette analyse a montré que la prise en compte des caractéristiques cliniques spécifiques à la maladie améliore les prédictions, tout en étant limitées par des facteurs tels que les données manquantes qui conduisent à de cohortes restreintes par leur taille. Nous avons donc approfondi la caractérisation des données manquantes dans le contexte des LMEs dans le **Chapitre 6**. Nous y élaborons des recommandations sur la manière de traiter les informations manquantes sur la base d'études de simulation. Nous démontrons également que l'imputation à partir de la dernière observation est une approche viable dans le contexte des tests neurologiques, en raison du plateau distinctif dans le rétablissement qui commence environ six mois après le traumatisme initial. Enfin, le **Chapitre 7** explore le concept de déviance positive pour détecter les personnes qui se rétablissent au-delà des attentes cliniques. Bien que les données issues de ces personnes puissent compremettre la performance des modèles de prédiction par apprentissage automatique, la compréhension des mécanismes qui sous-tendent leur rétablissement phénoménal peut permettre de découvrir des caractéristiques menant à une amélioration du rétablissement.

Enfin, la **Partie D** souligne l'importance de la communication scientifique pour faire le lien entre les résultats obtenus en recherche et la pratique clinique. Le **Chapitre 8** encourage particulièrement l'utilisation de nouveaux outils tels que la visualisation interactive des données pour améliorer la présentation des résultats de recherche. Ces outils permettent également de s'orienter vers une recherche plus transparente et accessible non seulement pour les chercheurs et leurs collaborateurs en clinique, mais aussi pour les patients, leur famille et le reste de la société.

Globalement, cette thèse contribue à l'analyse approfondie des éléments décisifs qui guident les études cliniques liées aux LMEs, tels que le rétablissement neurologique,

l'évolution des biomarqueurs sérologiques et les traitements couramment prescrits en pratique clinique. Ces travaux ouvrent la voie à l'amélioration des analyses de données et à la prédiction du rétablissement après une LME en intégrant des caractéristiques cliniques connues de la maladie. Dans le contexte de la recherche sur les LMEs, cette thèse participe à la révision des approches utilisées pour la découverte d'interventions qui amélioreraient le rétablissement après une LME.

Zusammenfassung

Eine Rückenmarksverletzung (engl. spinal cord injury (SCI)) ist eine medizinische Diagnose auf Grund einer Schädigung des Rückenmarks. Da das Rückenmark die primäre Verbindung zwischen dem Gehirn und den peripheren Organsystemen darstellt, führt eine Unterbrechung der Konnektivität zu zahlreichen Beeinträchtigungen in der Fortbewegung, der sensorischen Wahrnehmung und weiteren Organfunktionen. Eine SCI beeinträchtigt daher die gesamte Lebensqualität und Unabhängigkeit der betroffenen Personen und ihrer Angehörigen. Diese Tatsache ist umso bedeutsamer, da es in diesem Gebiet der Medizin noch keine pharmakologische oder sonstige Intervention gibt, die die Wiederherstellung von Funktionen und/oder die Regeneration des geschädigten Rückenmarks fördert. Die bisher durchgeführten klinischen Studien haben nicht zu einer vielversprechenden Intervention geführt, aber sie unterstützen das Feld mit gründlich gesammelten, großen Datenmengen. Der Fortschritt in den Datenwissenschaften, einschließlich statistischer und maschineller Lernmethoden, verspricht neue Erkenntnisse bei der Definition und Verbesserung der Genesung nach einer SCI, indem diese Daten retrospektiv umfassend untersucht werden.

Ziel dieser Arbeit ist es, das Potenzial der Datenwissenschaft zu nutzen, um die Wirkung in der klinischen Forschung zu maximieren. Diese Bemühungen stützten sich auf drei Säulen: (i) Ausweitung der medizinischen Überwachung innerhalb klinischer Studien, (ii) Einsatz bewährter methodischer Verfahren aus den Datenwissenschaften in der Forschung zu SCI, und (iii) Aufzeigen der Bedeutung einer wirksamen Kommunikation von Forschungsergebnissen.

Zu Beginn wird der allgemeine Kontext, in den sich diese Arbeit einfügt, in **Teil A** umrissen. Folgend werden in **Teil B** Maßstäbe gesetzt, die auf Sekundäranalysen wichtiger Datensätze, die in verschiedenen Bereichen der Forscung zu SCI gesammelt wurden, namentlich die klinische Studie Sygen, die European multicenter study on human spinal cord injury (EMSCI), die Daten des Murnau-Zentrums und die SCIRehab-Kohorten, basieren. **Kapitel 1** untersucht, wie sich die Genesung nach SCI in den letzten zwei Jahrzehnten verändert hat, und zeigt, dass die neurologische Genesung trotz eines sich weiterentwickelnden Behandlungsstandards in diesem Zeitraum weitgehend unverändert geblieben ist.

Diese Beobachtung ebnet den Weg für die Verwendung historischer Patientendaten zur Anreicherung von Placebo-Armen in künftigen klinischen Studien, wodurch die Exposition gegenüber der gewünschten Intervention maximiert werden kann. In **Kapitel 2** wird die natürliche Entwicklung serologischer Marker nach SCI beschrieben. Diese Marker stellen ein zusätzliches Überwachungsinstrument bei der Erprobung pharmakologischer Interventionen dar, die sich über die eigentliche Verletzung hinaus auf den Einzelnen auswirken könnten. Ebenso können Studien, in denen die Wirkung neuer pharmakologischer Maßnahmen getestet werden, durch Wechselwirkungen mit Medikamenten, die nach einer Verletzung verschrieben werden, beeinträchtigt werden. Daher legen wir in **Kapitel 3** ausführlich den derzeitigen pharmakologischen Standard der Versorgung dar. Wir haben eine umfangreiche Polypharmazie aufgedeckt, der die Betroffenen ausgesetzt sind. Um die Auswirkungen dieser Polypharmazie auf die Erholung nach einer SCI zu charakterisieren, haben wir in **Kapitel 4** die Literatur systematisch gesichtet und sowohl klinische als auch präklinische Evidenz beschrieben, die positive oder negative Auswirkungen diverser Medikationen auf die neurologische Erholung nach SCI haben können.

In **Teil C** passen wir bekannte Methoden aus den Datenwissenschaften an, um sie auf Anwendungen in der Forschung zu SCI zu übertragen. Wir untersuchen zunächst das Potenzial von serologischen Biomarkern als Prädiktoren für die Erholung der Motorfunktion in **Kapitel 5**. Diese Analyse zeigt, dass die Berücksichtigung von klinischen Merkmalen, die für diese Verletzung spezifisch sind, die Vorhersage verbessert, aber dennoch durch Faktoren wie fehlende Daten, die zu kleinen zu untersuchenden Kohorten führen, eingeschränkt ist. Daher charakterisieren wir fehlende Daten im Kontext von SCI in **Kapitel 6** ausführlich. Hier entwickeln wir auf der Grundlage von Simulationsstudien Leitlinien für den Umgang mit fehlenden Informationen. Wir zeigen, dass die Imputation mittels der letzten Beobachtung ein praktikabler Ansatz für neurologische Testergebnisse ist, da die Genesung etwa sechs Monate nach dem Trauma ein Plateau erreicht. Schließlich wird in **Kapitel 7** das Konzept der positiven Abweichung untersucht, um Personen zu erkennen, die sich über die klinischen Erwartungen hinaus erholen. Während solche Personen die Leistung von Algorithmen des maschinellen Lernens in Prädiktionsproblemen beeinträchtigen können, birgt das Verständnis der Mechanismen, die diesen phänomenalen Genesungen zugrunde liegen, das Potenzial, Muster aufzudecken, die zu der beobachteten besseren Genesung führen.

Schließlich unterstreicht **Teil D** die Bedeutung der Wissenschaftskommunikation, um die Translation der Forschung vom Labor ans Krankenbett effektiv zu gestalten. **Kapitel 8** beschreibt insbesondere den Einsatz neuer Werkzeuge wie interaktive Datenvisualisierung, um die Präsentation von Forschungsergebnissen zu verbessern und gleichzeitig die

Forschung transparenter und zugänglicher zu machen, nicht nur für die Forschenden und die klinische Gemeinschaft, sondern auch für die betroffenen Personen, ihre Familien und die Gesellschaft.

Insgesamt trägt diese Arbeit zu einem eingehenden Benchmarking entscheidender Elemente bei, die für klinische Studien zu SCI maßgeblich sind, wie z. B. die neurologische Erholung, die Entwicklung serologischer Biomarker und die üblicherweise als Teil der Standardbehandlung verschriebenen Medikamente. Diese Arbeit weist den Weg zu einer verbesserten Datenanalyse und Genesungsvorhersage nach SCI, indem sie bekannte Merkmale der Erkrankung integriert. Im Kontext des Forschungsfeldes trägt diese Arbeit dazu bei, Ansätze weiterzuentwickeln, die zur Entdeckung von Interventionen zur Verbesserung der Genesung nach SCI eingesetzt werden können.

Part A

General introduction

1 Spinal cord injury (SCI)

1.1 Definition

The central nervous system (CNS), responsible for receiving, integrating, and reacting to external stimuli, is composed of two organs: the brain, located in the skull, which extends from the foramen magnum into the spinal cord, located in the vertebral column (**Figure 1**) [2]. The spinal cord regulates the transmission of various neurological signals in the sensory, motor, and autonomous systems. Owing to its crucial role as an intermediate between the peripheral body and the brain, any trauma affecting the integrity of the spinal cord leads to impairments in numerous other systems, including musculoskeletal, cardiovascular, respiratory, urinary, and reproductive systems [3]. The initial loss of homeostasis ultimately leaves the affected individual with impairments spamming from dysregulated functions to complete loss, and may occasionally lead to death when a homeostatic state cannot be maintained.

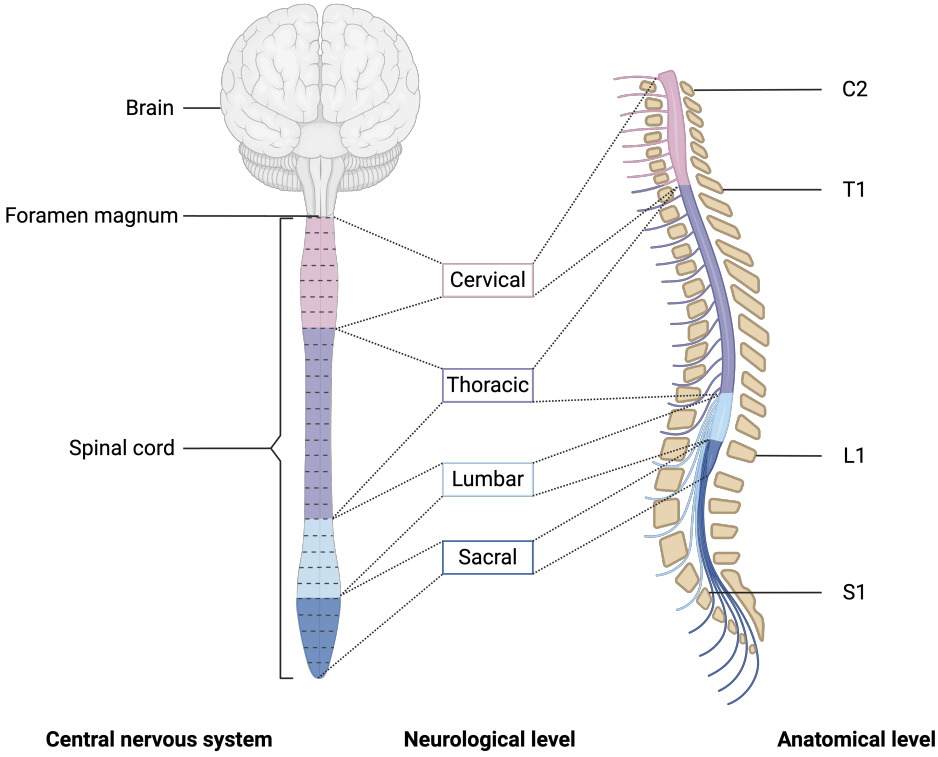


Figure 1: Central nervous system, spinal cord and level of injury. Created with BioRender.com

1.2 Epidemiology

Two main types of spinal cord injuries (SCIs) should be distinguished: the ones from traumatic and non-traumatic origins. Non-traumatic injuries could be the result of tumors, infections, ischemia, or degenerative processes. Traumatic SCIs, on the other hand, are caused by external trauma, such as motor vehicle accidents, sports injuries, gun shots, and falls. The worldwide incidence of traumatic SCIs is estimated around 3.3 to 195.4 per million per year [4], with large variations at the national and subnational level. In total, the prevalence, or total number of individuals living with SCI, is estimated around 250 to 906 per million across Western countries (e.g., Canada, France, United States of America (USA)), with notable differences between geographical areas [5, 6]. Interestingly, the incidence and distribution of the causes for traumatic SCIs vary greatly depending on the geographical area considered, with, for example, an increased incidence of gunshot wounds in the USA compared to Europe, Canada, and Australia [7]. Unlike the primary cause of injury, other epidemiological parameters remain constant across countries. Traumatic SCI is a condition mainly affecting male subjects, with a male:female ratio estimated from 4:1 to 2.3:1 [7, 8]. The two main age groups affected are young adults from 15 to 29 years old and above 50 years of age [3]. Interestingly, the predominant causes of injury differ between the two age groups, with comparatively more falls in the older population [9], leading to further dissimilarities in the injury characteristics (e.g., level of injury, injury severity).

1.3 Characterization

Clinical scores

Numerous assessments exist to specifically characterize a SCI, from neurological assessments to electrophysiological [10] and pain testing [11], imaging (e.g., magnetic resonance imaging [12]), or hand impairment quantification (e.g., graded redefined assessment of strength sensibility and prehension (GRASSP) [13]). Taken together, they contribute to draw a complete clinical description of an individual with SCI. This thesis primarily focuses on scores assessed as part of the international standards for neurological classification of spinal cord injury (ISNCSCI) examination, as presented below.

The ISNCSCI examination, featured in **Figure 2**, contributes to shaping the heterogeneous clinical manifestations. It does not only define sensory and motor functions of each testable dermatome and myotome, but also the level, completeness and severity of the injury.

Dermatomes and myotomes are defined as the projection on a skin area and groups

of muscles, which innervation is provided by a certain spinal root [14]. In sensory testing, each dermatome (n = 56 in total, equally distributed between left and right sides) is assigned a value from 0 (absent) to 2 (normal), when compared to an unaffected area (**Table 1**). The process is repeated for two distinct types of sensations, namely light touch (LT) and sharp-dull discrimination with the pin prick (PP) tests. The sums over all LT (total light touch (TLT)) or PP (total pin prick (TPP)) scores range from 0 to 112, with higher scores representing more preserved sensation. Similarly, motor function is evaluated at each myotome (n = 20 in total), bilaterally, on a scale from 0 (total paralysis) to 5 (normal active movement against full resistance) (see **Table 1** for score level definitions). Combining information from each myotome leads a total motor score (TMS) from 0 to 100, sometimes split between lower extremity motor score (LEMS) and upper extremity motor score (UEMS), both evaluated from 0 to 50.

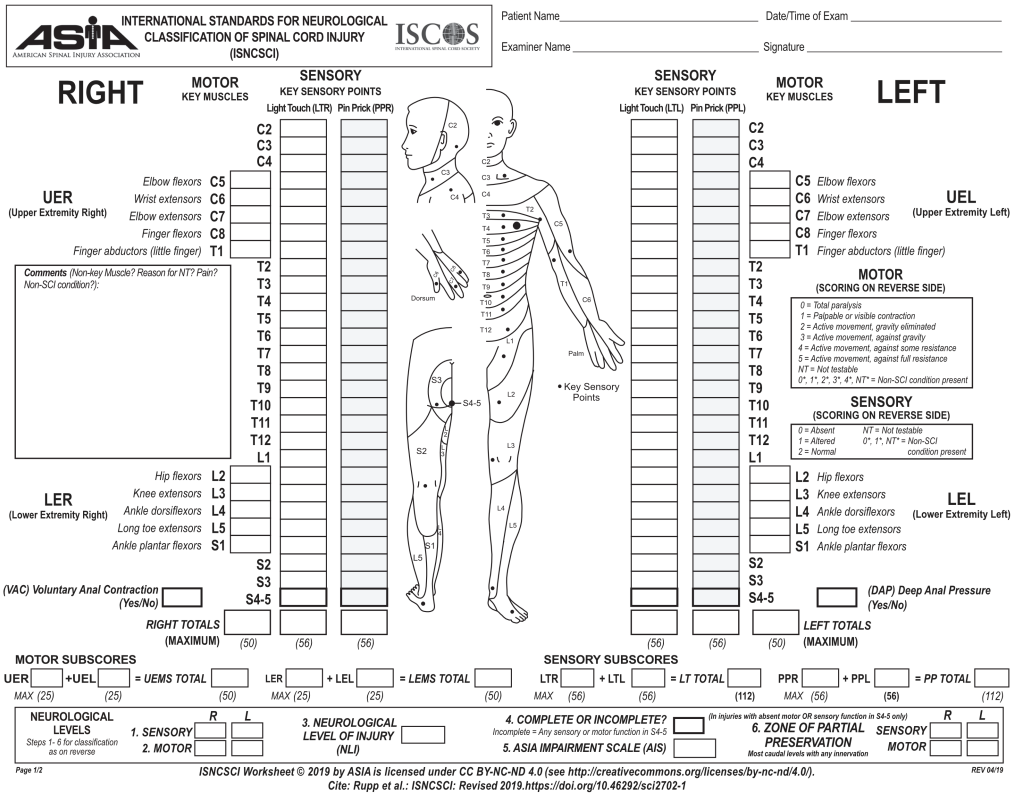


Figure 2: International standards for neurological classification of SCI (ISNCSCI) worksheet. ISNCSCI: Revised 2019 is licensed under CC BY-NC-ND 4.0 and presented in Rupp *et al.* [15].

Table 1: Details on the grading of sensory and motor functions

Score	Sensory function	Motor function
0	absence of sensation	total paralysis
1	altered sensation (either impaired or increased)	palpable or visible contraction
2	normal sensation	active movement, full ROM with gravity eliminated
3		active movement, full ROM against gravity
4		active movement, full ROM against gravity and moderate resistance in a muscle specific position
5		(normal) active movement, full ROM against gravity and full resistance in a functional muscle position expected from an otherwise unimpaired person

Adapted from Rupp *et al.* [15]; range of motion (ROM)

The level of injury, often referred to as neurological level of injury (NLI), is to be distinguished from the anatomical level of injury, as illustrated in **Figure 1**. It refers to the most lowest level of the spinal cord with normal sensory (LT and PP scored as 2) and antigravity motor function (i.e., motor levels scored as 5) on both sides of the body, assuming that there is normal sensory and motor functions in the above levels [16].

Injury completeness is defined according to the sparing of function in the sacral area (see **Figure 1**) and assessed through the deep anal pressure (DAP) and voluntary anal contraction (VAC). For an injury to be classified as neurologically complete, either DAP would be absent, or both VAC and sensory scores at the S4-5 levels would be absent. The injury is otherwise classified as neurologically incomplete.

Combining information from injury completeness, sensory and motor functions allows for the overall grading of the injury severity through the American spinal injury association (ASIA) impairment scale (AIS) grade, ranging from A to E, for the most to least severe injuries. Further details on the definition of each grade are summarised in **Table 2**. The AIS grade is an important assessment as it provides valuable information on the severity of the injury and is associated with potential of recovery [17].

Table 2: Details on the grades constituting the American spinal injury association (ASIA) impairment scale (AIS)

Grade	Type of injury	Description of injury
A	Sensorimotor complete	No sensory or motor function is preserved in the sacral segments S4-5
B	Sensorimotor incomplete	Sensory but no motor function is preserved below the NLI and includes the sacral segments S4-5, AND no motor function is preserved more than three levels below the motor level on either side of the body
C	Motor incomplete	Motor function is preserved below the NLI AND more than half of key muscles functions below the NLI have a muscle grade less than 3
D	Motor incomplete	Motor incomplete status as defined above, with at least half (half or more) of key muscle functions below the NLI having a muscle grade ≥ 3
E	Normal	Normal. If sensation and motor function as tested with the ISNCSCI are graded as normal in all segments, and the patients had prior deficits, then the AIS grade is E. *Someone without an initial SCI does not receive an AIS grade*

Adapted from Rupp *et al.* [15]; American spinal injury association (ASIA) impairment scale (AIS); international standards for neurological classification of spinal cord injury (ISNCSCI); neurological level of injury (NLI)

Beyond neurological assessments, patients can also be evaluated according to functional scores, which reflect their ability to perform daily tasks, such as dressing up, walking, bowel and bladder management, and transfers. Functional scores also tend to relate more to the priority concerns in the spinal cord-injured populations [18], such as regaining or improving, sexual, bladder, and bowel functions. One of the most commonly used functional scores is the spinal cord independence measure (SCIM) score [19]. The latest version of SCIM, SCIM III, is divided in five subscales, assessing self-care abilities, respiration and sphincter management, mobility in room and toilet, and mobility indoors and outdoors [20].

Recovery trajectory and quantification

Assessing clinical scores longitudinally enables the study of recovery trajectories. Specifically after SCI, improvements in aggregate scores such as LEMS or UEMS would occur largely in the first six months following injury, before reaching a plateau between six and 12 months [17]. From this observation, one can define recovery based on the difference between the score obtained shortly after injury and the one obtained around the expected plateau. Similarly, recovery can also be defined based on changes in severity grading, referred to as AIS conversion. Multiple variations of the AIS conversion can be considered, from comparison of actual grades to comparison based on the completeness of the injury.

Taken together, the information characterizing an injury and its recovery, although primarily collected for clinical purposes, are also essential components of clinical studies, in defining outcomes of interest and in studying the natural evolution of a SCI and other affected systems.

2 Clinical studies in SCI

Clinical studies are a type of investigations involving human subjects to assess the safety and performance of an intervention [21] on an outcome of interest. Owing to the lack of intervention, pharmacological or otherwise, to improve recovery following SCI, numerous clinical studies conducted in SCI populations have been and are still conducted to date. These clinical studies can adopt various designs. Here we will focus on the design of a randomized clinical trial (RCT) and an observational study, which are introduced in the following section and summarized in **Figure 3**.

2.1 Randomized controlled trials (RCTs) as gold standard

According to the American national institutes of health (NIH), a clinical trial is *"a research study in which one or more human subjects are prospectively assigned to one or more interventions (which may include placebo or other control) to evaluate the effects of those interventions on health-related biomedical or behavioral outcomes"* [22]. More precisely, RCTs refer to clinical trials in which participants are randomly assigned to the intervention or comparator (typically placebo) group.

Randomization is believed to help reduce bias, which refers to systematic errors that can occur from how the data is collected and/or analyzed [23]. Multiple sources of bias have been identified in the context of RCTs. Firstly, confounding refers to a factor associated with the intervention and/or the outcome of interest (e.g., age, sex, injury

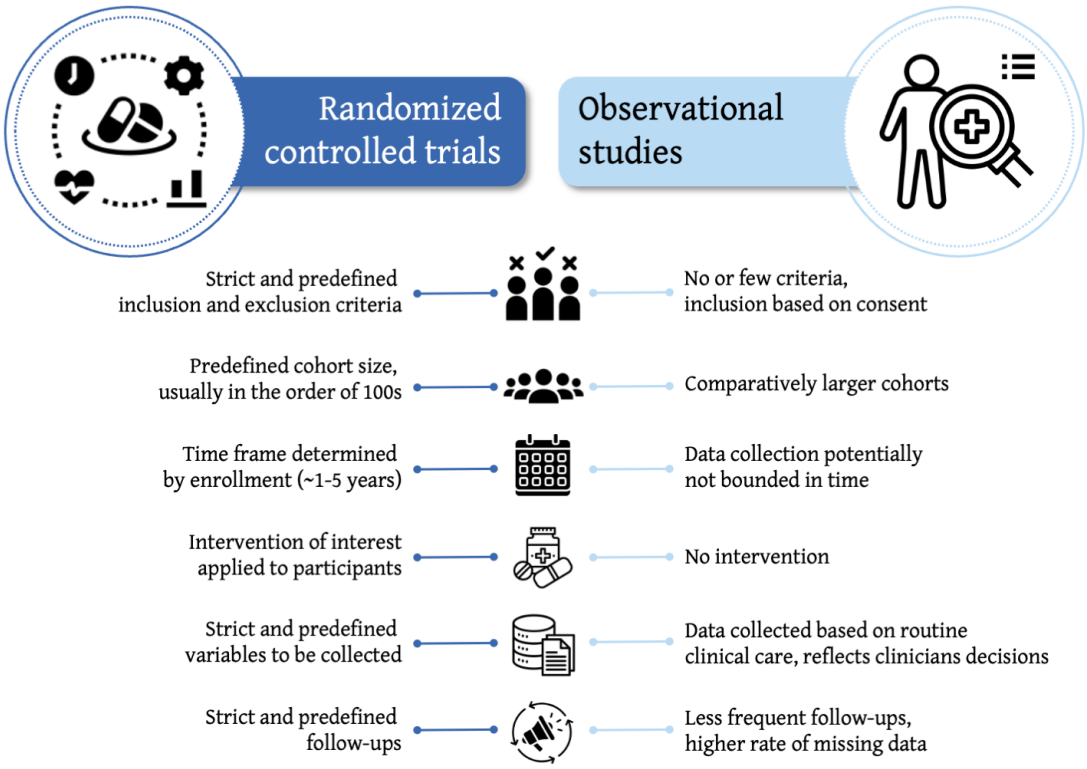


Figure 3: Differences between a randomized clinical trial (RCT) and an observational study.

heterogeneity) [24]. Secondly, selection bias occurs by (sub)consciously favouring the enrollment of participants that are believed to benefit from the treatment [25]. Finally, information bias may result from misdiagnosis or misidentification of the exposure status to the intervention of interest [26].

Since bias may lead to distorted views on the association between the intervention and outcome of interest, study designs such as randomization have been developed to mitigate its effects. Hence, the aim of the random assignment is to obtain groups that would only differ in the intervention tested. The two groups obtained should be otherwise comparable, also referred to as balanced groups (e.g. similar age distributions, proportions in injury severity, sex, missing data). The ultimate goal is to be able to attribute any differences in outcomes observed between the groups to the intervention tested.

A RCT is conceptually the only study design that allows for this direct attribution of effect to the intervention and is therefore, considered the gold standard when testing therapeutical interventions [27]. RCTs are typically conducted in three phases, with increasing cohort sizes: (i) phase I, studying the pharmacokinetics (i.e., the effect of the receiver on the substance after administration [14]), pharmacodynamics (i.e., how the substance affects the receiver [14]) and safety of the intervention; (ii) phase II, aiming to determine the optimal dosage; (iii) phase III, testing drug efficacy. Upon approval of the

intervention on the market, the trial enters phase IV, as known as pharmacosurveillance, monitoring long-term (side) effects in the effective treated population.

RCTs in SCI

Dietz *et al.* [28] reported nearly 1200 entries since 1996 on ClinicalTrials.gov[1] of trials investigating one or more intervention(s) to promote outcomes following SCI. The predominant emphasis in both past and ongoing clinical trials has been on rehabilitation and training, neuromodulation and electrical stimulation, and pharmacological interventions. Dietz *et al.* noted that the latter, while being the most represented intervention in 2007 with about 38% of the trials reported, are proportionally less represented in the early 2020s. This shift in proportion can be explained by two factors: (i) the increased interest in interventions based on neuromodulation and electrical stimulation; and, (ii) the overall increase in clinical trials registered over time.

NASCIS study

The national acute spinal cord injury study (NASCIS) study refers to a series of three trials conducted in the 1980s and 1990s in testing the effect of methylprednisolone sodium succinate (MPSS) on recovery following acute SCI. Results of the second study (NASCIS II) initially presented beneficial effects of high dose of MPSS administered specifically within eight hours after injury [29]. Notably, the results of the NASCIS trials led to MPSS being part of standard of care following SCI, which is the only intervention that ever reached this status for SCI. However, criticisms and new evidence later arose, pointing towards a lack of control group, bias in the results reported and lack of transparency regarding the cohort sizes in the subgroup analyses [30]. At last, MPSS was retracted from treatment guidelines after SCI [31, 32]. The NASCIS trials nonetheless increased the attention of the SCI research community in identifying bias and limitations related to data analyses and interpretation.

Sygen study

Originally set to investigate the effect of gangliosidosis 1 (GM-1), the Sygen trial is a multicenter, randomized, double-blinded clinical trial conducted between 1992 and 1998 in the United States of America (USA). The trial enrolled close to 800 individuals, all subject to the national acute spinal cord injury study (NASCIS) II protocol, and followed for the course of the first year after injury. Standardized time points were set for neurological assessments to be collected: baseline measurement (first 72h), 4, 8, 16, 26, and 52 weeks following injury. Additionally, information spanning from surgical protocols, medications

[1] https://clinicaltrials.gov/

prescribed to serological markers were included in the study protocol. While the trial failed to demonstrate any significant benefit in improving neurological status [33, 34], the data collected have since been reused in numerous secondary analyses.

RISCIS

The riluzole in spinal cord injury study (RISCIS) trial aimed to test the effect of riluzole, a sodium channel-blocking anticonvulsant, in acute SCI across the USA and Canada (ClinicalTrials.gov ID: NCT01597518). This treatment has the peculiarity of being approved and used for the treatment of amyotrophic lateral sclerosis [35]. After a phase I trial providing evidence of safety and suggesting neuroprotective effects in use for SCI (ClinicalTrials.gov ID: NCT00876889) [36, 37], the compound entered a phase II/III trial in 2013, which was discontinued by the sponsor in May 2020 due to the COVID-19 pandemic [38]. The trial originally planned for the enrollment of 351 participants [39], of which 193 were effectively recruited. As a result of the early termination and reduced cohort, the trial led to inconclusive results with the riluzole group showing improved UEMS, an improvement which however failed to reach the predefined statistical criteria for superiority. Further analyses are currently undertaken to investigate the effects in subgroup populations.

NISCI study

The Nogo inhibition in spinal cord injury (NISCI) study is a double-blind, placebo-controlled trial conducted from 2019 to 2023, across Switzerland, Germany, Spain, and Czech Republic (ClinicalTrials.gov ID: NCT03935321). The trial was testing the efficacy of antibodies directed against the growth inhibitory protein Nogo-A in improving movements and quality of life of individuals with acute tetraplegia due to a SCI. The phase II trial is based on preclinical knowledge suggesting enhance axonal sprouting and neuroprotective effects [40], and a phase I trial which demonstrated that the intervention was safe and well tolerated [41]. Notably, the NISCI study included historical controls in its cohort to maximize the number of participants exposed to the intervention tested, allowing for the completion of the trial, despite difficulties in enrollment related to the COVID-19 pandemic []. Owing to its recent completion, definitive results on the effects of anti Nogo antibodies are still unknown. However, the trial allowed for the collection of data from 129 individuals, representing a new data source for future research.

As a result of the failures observed so far, and despite being the gold standard in studying therapeutical interventions, RCT is likely not the only study design that can be used to discover interventions with a therapeutic effect on SCI recovery. The aforementioned limitations are calling for complementary study designs, including the secondary

analyses of data collected through RCTs, such as observational studies.

2.2 Observational studies

Observational studies are studies in which no intervention is introduced by the investigators [42] (see **Figure 3**). By relaxing the need for balanced and comparable groups, observational studies are logistically easier to implement compared to RCTs. They allow for the collection of information from larger cohorts, more representative of the population of interest. Tremendous efforts were deployed in that direction, especially since the early 2000s. We present here examples of observational studies focused on SCI, which contributed to the content of this thesis.

Observational studies in SCI

Secondary analyses of clinical trials data

Secondary analyses of data collected through RCTs are of utmost importance for the field. Indeed, they offer cohorts with a large and thorough collection of data, with sample sizes usually exceeding the ones from in-house data collection, and a more detailed view at the individual level. The Sygen clinical trial is a good example of valuable data source for secondary analyses. Despite the absence of effect found in the intervention tested, the trial nonetheless considerably contributed to the SCI research field. It enabled numerous secondary analyses, unrelated to the initial intervention tested, and contributed to gaining insights in domains as diverse as medications prescribed [43], serological biomarkers [44] or timing of surgical decompression [45].

European multicenter study on human spinal cord injury (EMSCI)

The European multicenter study on human spinal cord injury (EMSCI) [2] is a network of centers specialized in SCI care across Europe and India. Started in 2000, this collaborative effort groups 19 centers in which data from over 6000 individuals have longitudinally collected, as of 2023. Data is collected meticulously throughout the network, relying on specifically-trained staff. This asset makes it one of the highest quality dataset available in the field. The standardized assessments are performed during five distinct time windows: very acute (from 0 to 15 days after injury), acute I (16 to 40 days), acute II (70 to 98 days), acute III (150 to 186 days) and chronic (300 to 546 days). They mainly report neurological scores from the ISNCSCI examination, but also functional scores (e.g., SCIM) and walking ability (e.g., 6-minute walking test (6MWT), walking index for spinal cord injury (WISCI), 10-meter walking test (10MWT)). Notably, specific centers from the network collect

[2] https://www.emsci.org/

additional information, such as the Murnau center in Germany, reporting serological markers.

SCIRehab

The SCIRehab cohort groups data from six partner SCI rehabilitation centers across the USA [46, 47]. The primary aim was to investigate the effectiveness of interventions provided during rehabilitation on recovery following SCI. In pursuing this goal, information on demographics, therapy and medical interventions, including medication prescriptions, patient education and counseling, were collected. Notably, the SCIRehab cohort was followed longitudinally, with outcome data available at six and 12 months after injury. A total of 1500 individuals were enrolled in the cohort, from fall 2007 to end of 2009, making it one of the biggest SCI rehabilitation cohort.

More generally, in the absence of interventions, observational studies are a valuable setting to observe the natural course of a disease or condition. A comprehensive benchmarking is warranted not only to improve clinical practice but also to conduct meaningful and thorough interpretation of data collected as part of RCTs. In particular, having a comprehensive overview of the pharmacological compounds to which individuals are exposed following SCI is essential to consider potential interactions when testing drug-based interventions. Similarly, while many studies report on the epidemiological landscape of SCI [48, 49], relatively little is known about the natural evolution of the neurological landscape over time. This could be an indicator of improved care overall. Such research questions are better suited to observational studies as they require data unaffected by predefined assumptions on effects from interventions of interest. However, observational studies also present specific challenges such as an increased degree of missing data, a potentially high imbalance in the outcomes studied, and a higher heterogeneity of the population observed. Overcoming those challenges is nonetheless made possible by the in-depth study of data through the lens of data-driven approaches.

3 The place of data science in SCI research

3.1 The surge of statistical and machine learning (ML) modeling

Machine learning (ML) is a branch of artifical intelligence (AI) which uses training data and mathematical optimization rules to make predictions on previously unseen testing data [50]. Its development led to numerous successful applications in medical fields, such as infectious disease testing [51] and cancer detection [52, 53]. Applications to the field of SCI naturally followed and prediction models started to arise as early as

1998, with a publication from Roland *et al.* comparing prediction models for ambulation following SCI[54]. Overall, ML models have been used in over 40 publications dealing with SCI, as summarised by Tuci and Håkansson *et al.*, unpublished. They report that the majority of the models published took demographic information (e.g., age, sex) and neurological status, as assessed in the ISNCSCI examination, as input variables to be used for fitting mostly linear models (e.g., linear and logistic regressions). However, it is interesting to note the emergence of more complex models such as neural networks or ensemble architectures in the last ten years [55, 56, 57]. To date, the field has been mainly interested in predicting functional outcomes, especially walking ability [58], or mobility in general taking the example of the functional independence measure (FIM) [59] or SCIM [60]. Other highly represented outcomes of interest are scores derived from the ISNCSCI examination (LEMS, UEMS, motor scores, AIS grade conversion). Overall, we can observe that the applications of data-driven approaches and ML in SCI research remain narrow in their spectrum, partially failing to scope the specific needs expressed by individuals living with a SCI [18].

3.2 Challenges data scientists must face in SCI research

A number of limitations, inherent to the field, contribute to this slow start in integrating data-driven approaches to SCI research. Firstly, while statistical and ML models require large volumes of data to capture meaningful (potentially non-linear) predictive patterns, SCI is a rare condition, automatically limiting the amount of data available. Secondly, SCI is a heterogeneous condition both in its initial clinical presentation (severity, level of injury, demographics) and in the patterns of recovery observed (e.g., from no recovery to an individual being able to walk with an initial severe injury graded with an AIS A). The emergence of more complex models in SCI research will most likely follow the growing number of entries in data registries and contribute to better modeling the heterogeneity of SCI. Despite the increase in data volume, a third limitation remains regarding the quality of the data collected, in particular the quantity of missing data present. Although a vast literature exists on the topic of missing data in statistics, the problem remains largely underreported in the case of SCI research. The example of missing data is in part revealing overlooked best practices as established in the data science and ML community. Taken together, limitations on sample size, heterogeneity and missing data impair not only *prediction performance*, but also *generalizability* of models developed. Further, as reported in Tuci and Håkansson *et al.*, unpublished, code developed for prediction models was shared in only three publications [61, 62, 57], limiting the *reproducilibity* of the research outcomes. Code sharing is an essential part to build on existing models and accurately and critically peer-review models presented. Finally, as raised in other medical applications

[63], ML prediction models, once performant, generalizable, and reproducible, require to be interpretable for use in clinical practice. *Interpretability*, defined as *"the degree to which a human can understand the cause of a decision"* [64], is crucial for data-driven predictions to be accepted in the context of high stakes clinical decisions. Acceptability of models can however be supported by the integration of domain knowledge, making it more relatable by the healthcare providers who are at the interface between the research outputs and affected individuals.

Observing the current challenges faced by researchers in SCI opens avenues for further research, which have been in part pursued in this thesis.

4 Thesis objectives and outline

As described in the previous paragraphs, the field of SCI research currently lacks three key factors to successfully embrace the opportunities provided by the advancements in data science. Firstly, although considerable amount of data is now available, the field is lacking a global overview of the natural course of recovery (neurological and other biomarkers) and standard of care (medications prescribed) of the individuals affected by SCI. This benchmarking is however essential to draw meaningful comparisons with individuals undergoing interventions believed to improve their recovery. Secondly, best practices developed alongside data-driven approaches have not yet been presented in a comprehensive and adapted manner to the SCI research field. Finally, transparency in scientific communication and data presentation is still to be further promoted, especially at the frontier between researchers and clinical partners.

This thesis aims to address these three key factors using data-driven approaches to maximise clinical impact in SCI research.

Further details on the contributions made through this thesis are presented as follows:

Part B	**Surveillance within clinical studies**
Chapter 1	International surveillance study in acute spinal cord injury confirms viability of multinational clinical trials
Chapter 2	Natural progression of routine laboratory markers after spinal trauma: A longitudinal, multi-cohort study
Chapter 3	Pharmacological management of acute spinal cord injury: A longitudinal multi-cohort observational study
Chapter 4	Do commonly administered drugs inadvertently modify the progression of spinal cord injury? A systematic review

Part C **Towards better data analysis for clinical studies**

Chapter 5 Exploring the potential of routine serological markers in predicting neurological outcomes in spinal cord injury

Chapter 6 Studying missingness in spinal cord injury: Challenges and impact of data imputation

Chapter 7 The concept of positive deviance applied to spinal cord injury recovery: An exploratory analysis

Part D **Effectively conveying results through interactive data visualization**

Chapter 8 The interactive manuscript: From tabular to interactive result presentation and data visualization

The thesis is concluded by a general discussion (**Part E**) summarizing the contributions presented in the broader context of SCI research, their limitations and implications for future research endeavours.

Part B

Surveillance within clinical studies

Introduction

Surveillance refers to the process of collecting, managing, analyzing, interpreting, and reporting information relative to the status of a population in terms of a specific disease or condition [65]. Surveillance therefore contributes to comprehensively characterizing and better understanding the natural history of a condition. This represents an essential step towards unravelling potential interventions (pharmacological or otherwise) which aim to improve recovery or disease progression overall. In the case of spinal cord injury (SCI), researchers are mainly working towards either neuroprotective (i.e., dampening secondary injuries caused by ischemia and excitotoxicity leading to tissue inflammation and loss), or neuroregenerative (i.e., promoting axonal growth) interventions [3]. To assess the effectiveness of an intervention, one requires an in-depth understanding of the natural history of the variations that such a trauma causes at all scales: from the neurological recovery, observed at the macroscopic scale, to the chemical variations, observed in parts of the blood, and including interventions taken as part of standard of care, which might interact with the intervention of interest. Surveillance directly enables this in-depth understanding and is therefore an integral part of clinical studies in SCI. Surveillance is itself facilitated by data collection both in randomized clinical trials (RCTs) and observational studies. RCTs allow for systematic and extensive data collection, while observational studies indirectly reflect clinical knowledge (e.g., in testing a specific marker at a given time after injury).

In this **Part B**, we leverage data from the Sygen clinical trial [34, 33], SCIRehab study [46], European multicenter study on human spinal cord injury (EMSCI) [3] and Murnau study center, aiming to analyze, interpret and report on three key components of surveillance in SCI :

(i) *neurological (motor and sensory) recovery* over the last two decades in **Chapter 1**;

(ii) *serological markers variation* over the first year after injury in **Chapter 2**;

(iii) *medications prescribed* in the acute phase following injury (about two months) in **Chapter 3** and their effects on SCI-specific neurological recovery known from and reported in the literature in **Chapter 4**.

[3] http://emsci.org/

Chapter 1

International surveillance study in acute spinal cord injury confirms viability of multinational clinical trials

---

Adapted from:

Lucie Bourguignon, Bobo Tong, Fred Geisler, Martin Schubert, Frank Röhrich, Marion Saur, Norbert Weidner, Rüdiger Rupp, Yorck-Bernhard B. Kalke, Rainer Abel, Doris Maier, Lukas Grassner, Harvinder S. Chhabra, Thomas Liebscher, Jacquelyn J. Cragg, EMSCI study group, John L. K. Kramer, Armin Curt and Catherine R. Jutzeler

BMC Medicine, 2022

Publication: 10.1186/s12916-022-02395-0
Neurosurveillance web site: https://jutzelec.shinyapps.io/Haemosurveillance/
GitHub repository: https://github.com/jutzca/SCI_Neurological_Recovery

Lucie Bourguignon primarily contributed to building the Neurosurveillance web site, on an original idea from Catherine R. Jutzeler; and secondarily contributed to the acquisition, analysis, and interpretation of the data and drafting of the manuscript. Further details can be found in **Section 1.6**.

1.1 Abstract

Background

The epidemiological international landscape of traumatic spinal cord injury (SCI) has evolved over the last decades along with given inherent differences in acute care and rehabilitation across countries and jurisdictions. However, to what extent these differences may influence neurological and functional recovery as well as the integrity of international trials is unclear. The latter also relates to historical clinical data that are exploited to inform clinical trial design and as potential comparative data.

Methods

Epidemiological and clinical data of individuals with traumatic and ischemic SCI enrolled in the European multicenter study on human spinal cord injury (EMSCI) were analyzed. Mixed-effect models were employed to account for the longitudinal nature of the data, efficiently handle missing data, and adjust for covariates. The primary outcomes comprised demographics/injury characteristics and standard scores to quantify neurological (i.e., motor and sensory scores examined according to the international standards for neurological classification of spinal cord injury) and functional recovery (walking function). We externally validated our findings leveraging data from a completed North American landmark clinical trial.

Results

A total of 4601 patients with acute SCI were included. Over the course of 20 years, the ratio of male to female patients remained stable at 3:1, while the distribution of age at injury significantly shifted from unimodal (2001/02) to bimodal distribution (2019). The proportional distribution of injury severities and levels remained stable with the largest percentages of motor complete injuries. Both, the rate and pattern of neurological and functional recovery, remained unchanged throughout the surveillance period despite the increasing age at injury. The findings related to recovery profiles were confirmed by an external validation cohort (n = 791). Lastly, we built an open-access and online surveillance platform (“Neurosurveillance”) to interactively exploit the study results and beyond.

Conclusions

Despite some epidemiological changes and considerable advances in clinical management and rehabilitation, the neurological and functional recovery following SCI has remained stable over the last two decades. Our study, including a newly created open-access and online surveillance tool, constitutes an unparalleled resource to inform clinical practice and implementation of forthcoming clinical trials targeting neural repair and plasticity in acute SCI.

Keywords

Spinal cord injury, Surveillance study, Neurological recovery, Functional recovery, Aging, Epidemiological shift, Benchmark

1.2 Introduction

Traumatic spinal cord injury (SCI) is a devastating neurological disorder that is associated with life-long neurological condition with motor, sensory, and autonomic deficits [66]. Damage to the spinal cord occurs via both mechanical perturbation (so-called primary injury) and a cascade of damaging pathophysiological events (so-called secondary injury) [67, 68]. There are no pharmacological or non-pharmacological interventions available that mitigate the extent of damage in the acutely injured spinal cord. Despite the lack of effective treatment options, considerable progress has been made toward reducing the mortality rate and morbidity among patients with SCI [69, 70]. This progress is chiefly attributable to advances in the acute and long-term care of SCI, including early spine surgery (i.e., decompression and stabilization) [71], blood pressure augmentation within the first week post injury [72], introduction of antibiotics [73], availability of specialized rehabilitation centers [74], rehabilitation practices (e.g., gait training), and the prevention and treatment of secondary complications (e.g., infections and neuropathic pain) [75, 76].

Little is known about the impact of these advances on the rate and pattern of functional and neurological recovery following traumatic SCI. This knowledge gap is partially attributable to the data sources available, which are often limited in consistency and sample size, lack follow-up measures, and/or non-standardized data collection [77]. Various recent studies have reported changes in demographics and injury characteristics over the past decades. Most of these, however, have focused on regional epidemiology for a limited number of outcome measures, spanning only a relatively short time period [78, 79]. There is a paucity of validated long-term and comprehensive longitudinal studies.

Our study addressed this knowledge gap by leveraging data from the European multicenter study on human spinal cord injury (EMSCI) — the largest and most comprehensive longitudinal international data source in the field of SCI [1]. The first aim was to investigate changes in the epidemiological landscape of traumatic SCI over the last 20 years with a focus on changes in demographics and geographical and injury characteristics. Based on previous evidence [78, 79], we hypothesized a shift to older and less severe injuries along with an invariable ratio of female to male patients. The second aim was to establish a benchmark for the rate and pattern of neurological and functional recovery after a SCI. To this end, we investigated the extent that functional and neurological recovery following traumatic SCI has changed over the last two decades. We hypothesized that changes in acute and rehabilitation practices have led to improved outcomes during the transition from acute to chronic SCI. External validation was conducted using data from a landmark clinical trial.

[1] http://emsci.org/

Lastly, we developed the *Neurosurveillance* web platform for the SCI community, researchers, authorities, and policymakers that offers an open-access resource for benchmarking recovery and inform the design and implementation of clinical trials.

1.3 Methods

1.3.1 Study design and data source

We performed a prospective and longitudinal observational cohort study of individuals enrolled in the EMSCI [2] (ClinicalTrials.gov Identifier: NCT01571531). The design and reporting of this study adhere to the strengthening the reporting of observational studies in epidemiology (STROBE) guidelines for observational studies [80]. Founded in 2001, the EMSCI comprises 30 participating trauma and rehabilitation centers from across Europe and India that have collected data from more than 5000 individuals with SCI. Detailed neurological and functional outcomes are comprehensively tracked in individuals with traumatic or ischemic SCI at fixed time points over the first year of injury (i.e., very acute [within 2 weeks], acute I [4 weeks], acute II [3 months], acute III [6 months], and chronic [12 months]). Further details on the EMSCI study (e.g., inclusion and exclusion criteria, active centers per year) can be found in **Additional file 1: Table S1**.

1.3.2 Cohort definition: inclusion and exclusion criteria

To be included in our study, patients enrolled in the EMSCI had to meet the following inclusion criteria: (i) available baseline information on sex, age at injury, and year of injury; (ii) defined cause of SCI (e.g., disc herniation, traumatic, ischemic, hemorrhagic); (iii) neurological level of injury (NLI) either "cervical," "thoracic," or "lumbar" (i.e., L1 and L2); and (iv) neurological assessment of injury severity according to the American spinal injury association (ASIA) impairment scale (AIS) [1] (for details see **Table 1.1**) at exam stage "very acute" (i.e., <2 weeks post injury) and/or "acute I" (i.e., 2–4 weeks post injury). The NLI refers to the most caudal segment of the cord with intact sensation and antigravity muscle function strength, provided that there is normal (intact) sensory and motor function rostrally [16]. We excluded patients who had sustained a non-traumatic SCI (with the exception of ischemic injuries), in whom damage was below the level L2 of the spinal cord, and missing information on injury completeness at the very acute or acute I stage. Ischemic injuries with a determinable disease onset were included owing to the fact that this type of injury is characterized by a sudden disease onset and the rate and pattern of recovery is comparable to traumatic SCI [81]. The workflow for the individuals included/excluded from our analysis is highlighted in **Figure 1.1A**.

[2] http://emsci.org/

Table 1.1: American spinal injury association (ASIA) impairment scale (AIS) describes the functional impairment as a results of spinal cord injury (SCI) [1].

Grade	Type of injury	Description of injury
A	Sensorimotor complete	No sensory or motor function is preserved in the sacral segments S4-5.
B	Sensorimotor incomplete	Sensory but no motor function is preserved below the NLI and includes the sacral segments S4-5, AND no motor function is preserved more than three levels below the motor level on either side of the body.
C	Motor incomplete	Motor function is preserved below the NLI AND more than half of key muscles functions below the NLI have a muscle grade less than 3.
D	Motor incomplete	Motor incomplete status as defined above, with at least half (half or more) of key muscle functions below the NLI having a muscle grade ≥ 3.
E	Normal	Normal. If sensation and motor function as tested with the ISNCSCI are graded as normal in all segments, and the patients had prior deficits, then the AIS grade is E. *Someone without an initial SCI does not receive an AIS grade.*

international standards for neurological classification of spinal cord injury (ISNCSCI); neurological level of injury (NLI)

1.3.3 Primary outcome (dependent) variable

The primary outcomes were common neurological (total motor score (TMS), lower extremity motor score (LEMS), upper extremity motor score (UEMS), total pin prick (TPP), total light touch (TLT), total sensory score (TSS)) and functional outcome scores (spinal cord independence measure (SCIM), walking index for spinal cord injury (WISCI), 10-meter walking test (10MWT), and the 6-minute walking test (6MWT)). For motor scores, key muscles in the upper and lower extremities were examined according to the international standards for neurological classification of spinal cord injury (ISNCSCI) [16], with a maximum score of 50 points for each, the upper and lower extremities (for a

maximum total motor score of 100). Light touch and pin prick (sharp-dull discrimination) scores were also assessed according to ISNCSCI, with a maximum score of 112 each (for a maximum total sensory score of 224) [16]. It is important to note that between 2001 and 2019, different ISNCSCI versions were used to assess the sensorimotor scores. For our analysis, we standardized and recalculated the ISNCSCI data by using the EMSCI ISNCSCI calculator [82] to comply with the 2015 ISNCSCI revision [16]. The SCIM is a scale for the assessment of activities of daily function. Throughout the duration of this study (2001–2019), two different versions of the SCIM were used: between 2001 and 2007 SCIM II [83] and since 2008 SCIM III [19]. The major difference between the versions is that SCIM II does not consider intercultural differences. Both versions contain 19 tasks related to activities of daily living organized in four areas of function (subscales): self-care (scored 0–20); respiration and sphincter management (0–40); mobility in room and toilet (0–10); and mobility indoors and outdoors (0–30). For the longitudinal analysis, we pooled the SCIMII and SCIMIII variables. WISCI has an original scale that quantifies a patient's dependency on walking aids to travel a distance of 10 m; a score of 0 indicates that a patient cannot stand and walk 10 m and the highest score of 20 is assigned if a patient can walk 10 m without walking aids of assistance [84]. Lastly, 10MWT measures the time (in seconds) it takes a patient to walk 10 meters at a self-selected walking speed, and the 6MWT quantifies the distance (in meters) covered by the patient within 6 minutes [85]. The 10MWT and 6MWT were only analyzed for ambulatory patients.

1.3.4 Input (independent) variables

Year of injury and exam stage (i.e., time since injury) were selected as the independent variables. Exam stage comprises four levels: very acute (≤2 weeks post injury), acute I (1-month post injury), acute II (3 months post injury), acute III (6 months post injury), and chronic (12 months post injury). The exam stage variable was coded as continuous variable for the estimation of temporal recovery trajectories. As with all observational studies, there is potential for confounding effects and bias. Potential confounders included age, sex, injury completeness (at time of injury) according to the AIS grade [86], and NLI (cervical, thoracic, or lumbar).

1.3.5 Data preprocessing and statistical analyses

As part of the preprocessing, we assessed the type and pattern of missing data. Briefly, we tested the hypothesis that the missing data are missing completely at random (MCAR) using the `LittleMCAR` function of the R package `BaylorEdPsych`. To visually explore the pattern of missing data as well as combinations of missingness across cases, we used

the R package `naniar`.

In the first step of analysis, descriptive statistics (mean, standard deviations, median, min, max, percentage, and proportions) were used to provide summary information on the demographics, baseline injury characteristics, and baseline functional and neurological outcomes. Independent 2-group Mann-Whitney-U and χ-squared tests were used to assess whether there was a difference in demographics and injury characteristics between included and excluded cohorts. Prior to the regression analyses, we normalized and standardized our data (i.e., ExamStage, YEARDOI, AgeAtDOI). Specifically, normalization refers to scaling a variable to have a value between 0 and 1, while standardization transforms data to have a mean of zero and a standard deviation of 1. These two steps are important as they improve the interpretability and computational performance of the described statistical models. Employing linear and logistic regression analysis, we assessed if demographics (i.e., age at injury, ratio of male and female patients) and injury characteristics (i.e., injury severity and NLI) differed between 2001 and 2019. Variability in injury characteristics were assessed separately for male and female patients. Specifically, the proportions (in percent) of the different injury severities (AIS A to AIS D), injury level (cervical, thoracic, and lumbar), and plegia (paraplegia, tetraplegia) were calculated for each year of the surveillance period. Subsequently, we fit a linear regression model with the proportion of AIS A as the response, and time since injury as the predictor to assess if the confidence interval of the beta coefficient includes zero or not. This was repeated for each AIS grade and all injury levels (i.e., cervical, thoracic, and lumbar). The second step of the analysis entailed the employment of non-linear mixed effect models to address the question if and to what extent the functional and/or neurological recovery were subject to change over the course of the last two decades. We assumed a random intercept and random effect for time since injury [87]. Moreover, we assumed a continuous time autoregressive process of order 1 for within-patient correlation structure and assumed a power function of the mean value for within-patient heteroscedasticity structure [88]. The model was fitted using restricted maximum likelihood for unbiased estimates of variance components. Dependent variables were all primary outcome variables described above, independent variables were year of injury (YEARDOI) and exam stages (ExamStage). To assess time-dependent changes in the recovery trajectories, the independent variables were included as interaction effect (YEARDOI*ExamStage). These analyses were performed for the overall cohort and stratified by sex, plegia, and AIS grades. Confounders of not interest included age and sex. If applicable, we also adjusted for AIS grades. The significance threshold was set at $\alpha = 0.05$. Post-hoc pairwise comparisons were Bonferroni corrected to account for multiple comparisons [89]. Lastly, as we expected a covariate-shift in terms of age, we performed a sensitivity analysis to determ-

ine if the recovery trajectories of sensorimotor and functional recovery changed in an age-dependent manner throughout the surveillance period. A second sensitivity analysis aimed at testing for sex-specific differences in recovery profiles. The third sensitivity analysis was performed to test the assumption that patients with ischemic and traumatic SCI recover in a comparable fashion. For all analyses and figures, R Statistical Software Version 3.5.2 for Mac Os Mojave was used. All analyses were run locally (MacBook Pro, Memory 16GB, Processor 2.3GHz Intel Core i5).

1.3.6 External validation cohort

In order to externally validate our findings related to the epidemiology as well as neurological recovery trajectories, we analyzed an independent clinical trial dataset [33]. Specially, the Sygen trial was a randomized, prospective, phase III, placebo-controlled, multi-center study testing the efficacy of gangliosidosis 1 (GM-1) therapy in acute, traumatic SCI. Clinically active from 1992 to 1998, the Sygen trial failed to demonstrate a superior treatment effect of GM-1 over placebo treatment. The Sygen clinical trial enrolled patients with traumatic SCI who were admitted to trauma centers across the United States of America (USA) and followed them over a year. Detailed information regarding the trial can be found in the **Additional file 3**. It is noteworthy to mention that the Sygen clinical trial is particularly well-suited to serve as an external validation data set for EMSCI owing to similar granularity in data, timepoints of assessment, duration of follow-up period, and standardized assessments across participating trauma and rehabilitation centers. There is no contemporary dataset that offers comparable data granularity, quality, and depth as the Sygen trial. The workflow for the individuals included/excluded from our analysis is highlighted in **Additional file 3: Figure S1**. To maximize the interpretability of cross-data sources comparisons, the same inclusion/exclusion criteria to be included in our analysis as for EMSCI were applied. Similar to the EMSCI data, we standardized and recalculated the ISNCSCI data by using the EMSCI ISNCSCI calculator [82] to comply with the 2015 ISNCSCI revision [16]. The validation was focused on the sensorimotor recovery owing to the comparable assessment methods (i.e., ISNCSCI). In the Sygen trial, functional recovery was assessed with different outcome measures (i.e., Modified Benzel Score, functional independence measure (FIM)) compared to the EMSCI study making a proper validation of the functional recovery profiles impossible. Lastly, we performed a sensitivity analysis to assess if the recovery trajectories are different for patients who had early surgery (<24h) vs. those with late surgery (>24 h). In light of that, we added the timing of surgery as an independent variable to the models described above.

1.3.7 Interactive web platform *Neurosurveillance*

In order to enable the SCI community, researchers, authorities, and policymakers to fully explore the data and results of this study (and beyond), we developed the freely available and open source Neurosurveillance web platform. Neurosurveillance was implemented with the Shiny framework [90], which combines the computational power of the free statistical software R with friendly and interactive web interfaces. Both, the front- and back-end of Neurosurveillance have been built using the shiny dashboard package [91]. Neurosurveillance is available as an online application and is hosted at https://jutzelec.shinyapps.io/neurosurveillance/ and can be accessed via any web browser on any device (e.g., desktop computers, laptops, tablets, smartphones). Neurosurveillance is published under the BSD 3-Clause License. The source code of Neurosurveillance is available through Github at https://github.com/jutzca/Neurosurveillance/. Further details on the technical implementation can be found in **Additional file 4**.

1.3.8 Data sharing and code availability

The data used for this study, including de-identified individual participant data and a data dictionary defining each field or variable within the dataset, can be made available upon reasonable request to the corresponding author (CRJ). Data will be made available following publication of this work. Written proposals will be evaluated by the authors, who will render a decision regarding suitability and appropriateness of the use of data. Approval of all authors will be required and a data sharing agreement must be signed before any data are shared. The code to run the analysis as well as create the figures and tables can be found on our Github repository [3].

1.3.9 Role of funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

[3] https://github.com/jutzca/SCI_Neurological_Recovery

1.4 Results

1.4.1 Cohort summary

Between 2001 and 2019, a total of 5220 individuals were enrolled in the EMSCI (**Figure 1.1A**). Based on our initial inclusion criteria, 4601 patients were eligible for our analysis (mean age at injury, 47.2 ± 19.0 years; 77.0% male); 53.9% were injured at the cervical level, and 51.5% had a motor complete injury at the initial ISNCSCI examination (i.e., AIS A and AIS B). Detailed cohort characteristics are provided in **Table 1.2**. The average number of patients enrolled per year was 242.2 ± 101.6 (**Additional file 2: Figure S1**). As shown in **Figure 1.1B** and summarized in **Additional file 2: Table S1**, the majority of the patients were admitted to EMSCI centers located in Germany (n = 2949, 64.1%), followed by Switzerland (n = 451, 9.8%), and the Czech Republic (n = 297, 6.5%). **Additional file 2: Table S2** provides the demographics and injury characteristics stratified by age groups.

Table 1.2: Demographics and injury characteristics of included European multicenter study on human spinal cord injury (EMSCI) cohort stratified by sex.

	Female (n = 1059)	Male (n = 3542)	Overall (n = 4601)
Sex			
Female	1059 (100%)	0 (0%)	1059 (23.0%)
Male	0 (0%)	3542 (100%)	3542 (77.0%)
Age (years)			
Mean (SD)	51.1 (20.2)	46.0 (18.4)	47.2 (19.0)
Median [Min, Max]	52.0 [9.0, 94.0]	46.0 [9.0, 92.0]	47.0 [9.0, 94.0]
Cause			
Disc herniation	3 (0.3%)	10 (0.3%)	13 (0.3%)
Hemorrhagic	12 (1.1%)	3 (0.1%)	15 (0.3%)
Ischemic	129 (12.2%)	202 (5.7%)	331 (7.2%)
Traumatic	915 (86.4%)	3327 (93.9%)	4242 (92.2%)
AIS grade			
A	360 (34.0%)	1459 (41.2%)	1819 (39.5%)
B	136 (12.8%)	418 (11.8%)	554 (12.0%)

Continued on next page

Table 1.2: Demographics and injury characteristics of included European multicenter study on human spinal cord injury (EMSCI) cohort stratified by sex. (Continued)

	Female ($n = 1059$)	Male ($n = 3542$)	Overall ($n = 4601$)
C	227 (21.4%)	644 (18.2%)	871 (18.9%)
D	336 (31.7%)	1021 (28.8%)	1357 (29.5%)
NLI			
Cervical	539 (50.9%)	1899 (53.6%)	2438 (53.0%)
Thoracic	387 (36.5%)	1256 (35.5%)	1643 (35.7%)
Lumbar	133 (12.6%)	387 (10.9%)	520 (11.3%)

American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description.
European multicenter study on human spinal cord injury (EMSCI), standard deviation (SD), neurological level of injury (NLI)

A total of 619 EMSCI patients (mean age at injury, 49.7 ± 20.5 years; 77.1% male) were excluded from our analysis (**Additional file 2: Table S3**). The ratio of male and female patients was comparable between included and excluded cohorts ($\chi^2 = 0.006$, df = 1, p-value = 0.939). However, the cohorts were different in terms of age ($t = 2.779$, df = 697.900, p-value = 0.006) and injury characteristics ($\chi^2 = 14.106$, df = 3, p-value = 0.003), with the excluded cohort being older and represented by a larger proportion of AIS D injuries. For detailed information on the missing data, see **Additional file 1: Figures S1** and **S2**.

1.4.2 Epidemiological landscape between 2001 and 2009

The overall ratio between female and male patients remained constant over the last 20 years ($\beta = 0.102$, standard error = 0.665, p-value = 0.880, **Figure 1.1C**). Along these lines, the ratio between female and male patients remained unchanged stratified according to cervical and thoracic/lumbar spine levels (i.e., tetraplegia [ratio 1:3] and paraplegia [ratio 1:3], **Additional file 2: Figure S2A**) as well as injury severity (AIS A [ratio 1:4], B [ratio 1:4], C [ratio 1:3], and D [ratio 1:3], **Additional file 2: Figure S2B**). In contrast, the overall distribution of age at injury changed significantly over the years ($\beta = 8.603$, standard error = 1.045, p-value < 0.001). Between 2001 and 2019, there was a shift towards older age at injury (**Figure 1.1D**, **Additional file 2: Table S4**), which was more

prominent in male compared to female patients (interaction effect YEARDOI*Sexmale: $\beta = 5.306$, standard error = 2.433, p = 0.029, **Additional file 2: Figure S3**). This shift in age remained evident after stratifying patients according to their plegia (**Additional file 2: Figure S4A**) and injury severity (**Additional file 2: Figure S4B**). In terms of the baseline injury severity, the overall proportion (in percentage) of AIS A, AIS B, AIS C, and AIS D remained constant throughout the study duration (**Figure 1.1E**). The proportions of cervical, thoracic, and lumbar injuries were also unchanged (**Figure 1.1F**). These findings remained constant in post hoc sensitivity analyses of subgroups according to AIS grades (**Additional file 2: Figure S5A**) and plegia (**Additional file 2: Figure S5B**). When stratified by age groups, linear regression models revealed significant changes in the proportion of injury severities as a function of time (**Additional file 2: Figure S6**), with more motor-complete injuries (AIS A, AIS B) among female and male patients older than 50 years of age. Summary statistics of all models can be found in the **Additional file 2: Table S5**.

1.4.3 Temporal progression of neurological and functional outcomes

The mixed-effect models revealed that recovery trajectories (i.e., fitted regression lines) of all neurological and functional outcomes remained comparable between 2001 and 2019 (**Figure 1.2**). Dependent on the injury severity, the recovery trajectories within a year were characterized by an improvement in function between baseline (i.e., very acute and acute I) and 6 months followed by a plateau phase up to 12 months post injury (**Additional file 2: Figures S7-S10**). In addition to the pattern, the rate of sensorimotor recovery remained comparable between the years of the surveillance period (**Figure 1.3A, B**, and **Additional file 2: Table S6**). This was also true when stratifying patients based on sex, plegia, and AIS grades. Summary statistics of all models are provided in **Additional file 2: Tables S7-S15**. Our sensitivity analyses revealed that the neurological and functional recovery profiles were comparable throughout the surveillance period between different age groups (**Additional file 2: Figure S11** and **Table S17**), male and female patients (**Additional file 2: Figure S12** and **Table S17**), and cause of injury (traumatic vs. ischemic, **Additional file 2: Figure S13** and **Table S18**). The results can be further interactively explored on our open access and online Neurosurveillance platform.

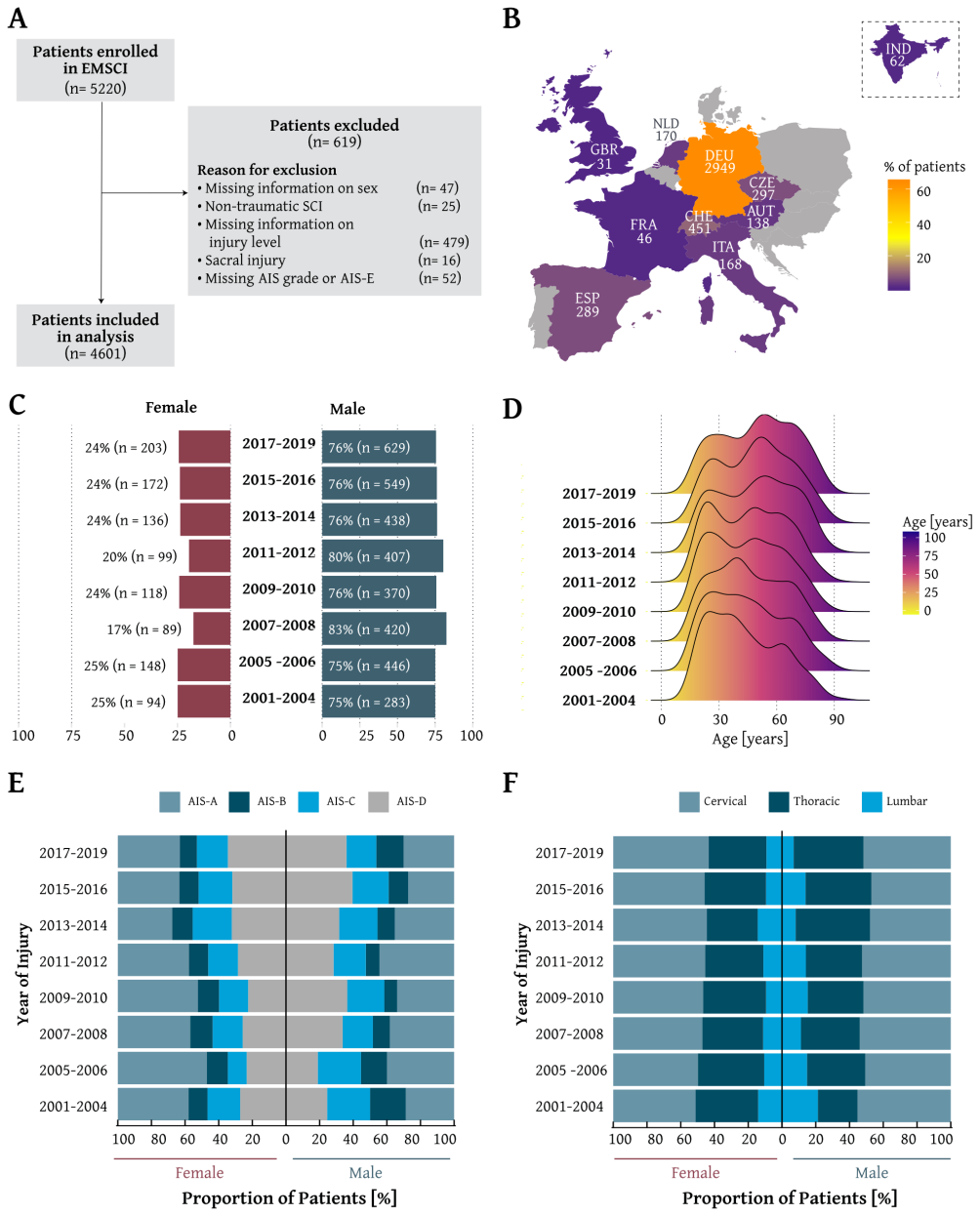


Figure 1.1: Study overview and result from the main cohort. A. Flowchart of the included and excluded patients that were originally enrolled in the European multicenter study on human spinal cord injury (EMSCI) study. Almost 90% of the EMSCI patients met our inclusion criteria; **B.** Number of patients recruited between 2001 and 2019 per country. The majority of patients were admitted to centers in Germany, Switzerland, and Czech Republic. Note: The Indian center joined the EMSCI network only in 2011; **C.** Annual ratio between female and male individuals with spinal cord injury (SCI) enrolled in the EMSCI. Between 2001 and 2019, the ratio between men and women sustaining a traumatic or ischemic SCI remained comparable at 3:1; **D.** Change in distribution of age at injury. Over the last two decades, a shift in age at injury was observed for individuals with SCI. In comparison to early 2000s, which were characterized by a unimodal distribution, the proportion of elderly people sustaining a traumatic SCI increased significantly; **E.** Baseline injury severity. While there are some fluctuations, the proportions of injury severities, as measured by American spinal injury association (ASIA) impairment scale (AIS) scores, remained constant across the study period; **F.** Baseline level of injury. The proportion of cervical, thoracic, and lumbar injuries did not significantly change as a function of time.

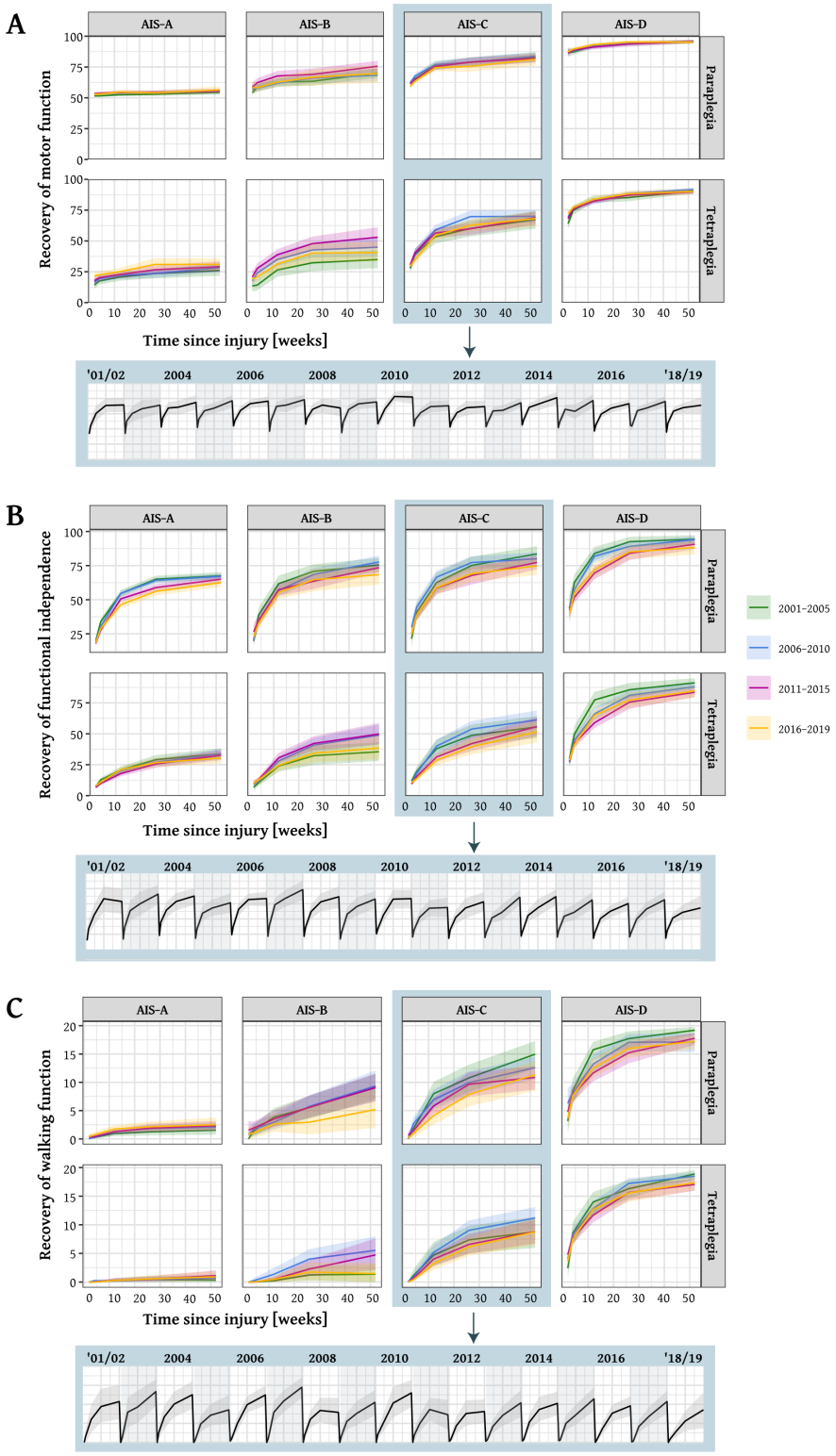


Figure 1.2: Neurological and functional recovery throughout the surveillance period. The recovery trajectory profiles of **A.** the motor function; **B.** functional independence; and **C.** walking function remained comparable across the surveillance period. In other words, the degree a person with spinal cord injury spontaneously recovers motor and walking function as well as functional independence within 1-year post-injury is the same now as it was two decades ago. The solid lines represent the fitted models and the shaded areas, the standard deviation. The inserted boxes illustrate the robustness of the recovery profiles across all years for patients with AIS C injuries. For all other injury severities, please refer to the **supplementary material section 11.1**.

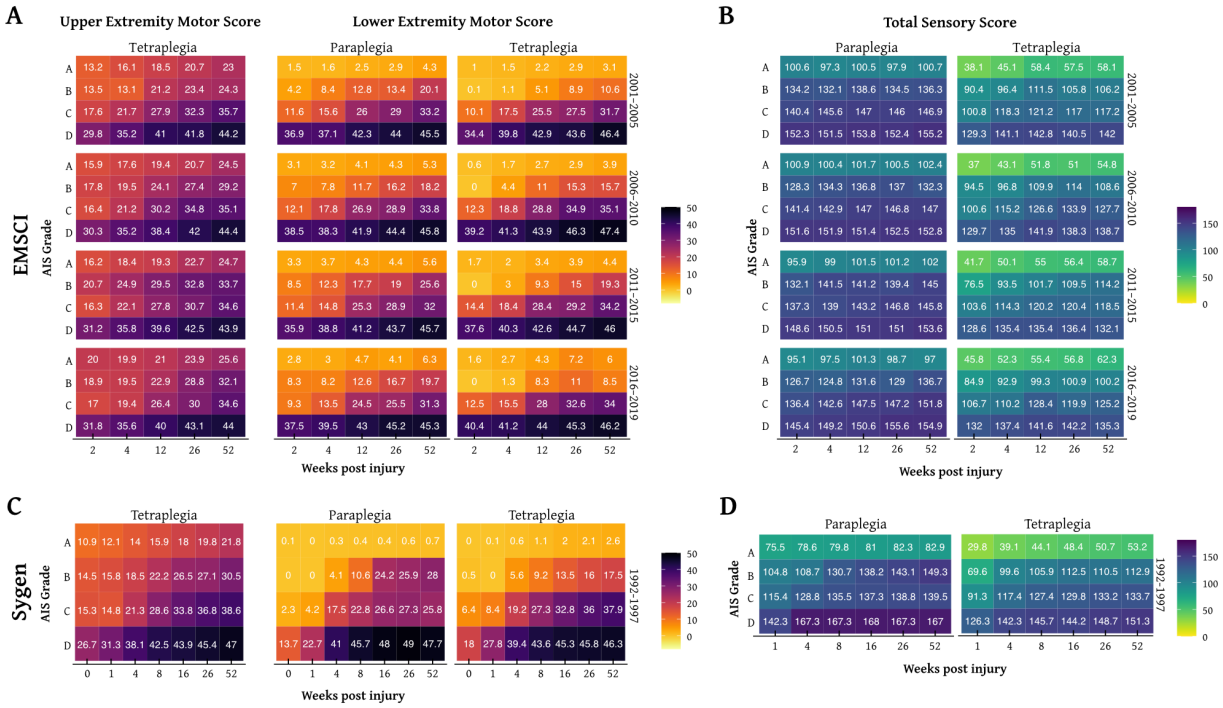


Figure 1.3: Comparison of sensorimotor recovery between data sources. **A.** The pattern and degree recovery of motor; and **B.** sensory function of patients enrolled in the European multicenter study on human spinal cord injury (EMSCI) were comparable to those of patients from the Sygen study (**C.** and **D.**). The heat plots and the number in the tiles represent the mean of motor and sensory scores, respectively. The progression of upper extremity motor score (UEMS) is only shown for individuals with a tetraplegic spinal cord injury. Note: Individuals with paraplegic spinal cord injury have, by definition, full function in the upper extremities (i.e., UEMS of 50)

1.4.4 Validation study

As summarized in **Table 1.3**, the validation cohort comprised 703 patients (mean age at injury, 32.9 ± 13.5 years; 79.7% male, 74.4% motor complete injury). In comparison to EMSCI cohort, the Sygen cohort exhibited a comparable ratio of male and female patients ($\chi^2 = 3.176$, df = 1, p-value = 0.074). However, the cohorts were different in terms of age ($t = 2.779$, df = 697.900, p-value = 0.006) and injury characteristics (AIS grades: $\chi^2 = 301.44$, df = 3, p-value < 0.001, NLI: $\chi^2 = 219.12$, df = 2, p-value < 0.001), with the Sygen cohort being younger and represented by a larger proportion of AIS A and cervical injuries.

Table 1.3: Demographics and injury characteristics of Sygen cohort per year and overall.

	1992	1993	1994	1995	1996	1997	Overall
	$n = 104$	$n = 161$	$n = 128$	$n = 139$	$n = 159$	$n = 12$	$n = 703$
Sex, n (%)							
Female	23 (22.1)	32 (19.9)	30 (23.4)	24 (17.3)	32 (20.1)	2 (16.7)	143 (20.3)
Male	81 (77.9)	129 (80.1)	98 (76.6)	115 (82.7)	127 (79.9)	10 (83.3)	560 (79.7)
Age (years)							
Mean (SD)	33.6 (13.8)	32.0 (13.4)	32.7 (12.9)	32.6 (13.3)	34.2 (14.0)	26.3 (13.2)	32.9 (13.5)
Median[Min, Max]	31.0 [15.0, 69.0]	30.0 [15.0, 66.0]	30.0 [15.0, 69.0]	30.0 [15.0, 67.0]	33.0 [13.0, 69.0]	23.5 [13.0, 60.0]	30.0 [11.0, 69.0]
AIS grade							
A	69 (66.3)	102 (63.4)	75 (58.6)	83 (59.7)	106 (66.7)	11 (91.7)	446 (63.4)
B	9 (8.7)	14 (8.7)	16 (12.5)	19 (13.7)	19 (11.9)	0 (0)	77 (11.0)
C	22 (21.2)	34 (21.1)	27 (21.1)	34 (24.5)	31 (19.5)	1 (8.3)	149 (21.2)
D	4 (3.8)	11 (6.8)	10 (7.8)	3 (2.2)	3 (1.9)	0 (0)	31 (4.4)
NLI, n(%)							
Cervical	81 (77.9)	115 (71.4)	103 (80.5)	112 (80.6)	119 (74.8)	10 (83.3)	540 (76.8)
Thoracic	23 (22.1)	46 (28.6)	25 (19.5)	27 (19.4)	40 (25.2)	2 (16.7)	163 (23.2)

American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description.

European multicenter study on human spinal cord injury (EMSCI), standard deviation (SD), neurological level of injury (NLI)

The ratio of male to female patients remained constant at 3 : 1 ($\beta = 1.247$, standard error = 0.668, p-value = 0.135; **Figure 1.4A**) and there was no significant change in the distribution of age at injury between 1992 and 1997 ($\beta = 0.392$, standard error = 1.782, p-value = 0.826, **Figure 1.4B**). The proportion (%) of AIS grades remained comparable during the trial period (**Figure 1.4C**) (AIS A: $\beta = 9.833$, standard error = 20.484, p-value = 0.634; AIS B: $\beta = 2.891$, standard error = 3.955, p-value = 0.486; AIS C: $\beta = -1.156$, standard error = 6.622, p-value = 0.865; AIS D: $\beta = -2.148$, standard error = 2.707, p-value = 0.454). The ratio between patients sustaining cervical and thoracic injuries ($\beta = 2.375$, standard error = 2.471, p-value = 0.454; **Figure 1.4D**) was comparable across the 6 years of study duration. An overview of the excluded cohort (**Additional file 2: Table S1**) as well as detailed information on the missing data (**Additional file 2: Table S2** and **Figures S2** and **S3**) and model summaries of demographics (**Additional file 2: Table S3**) can be found in the **Additional files 2** and **3**. As shown in **Figure 1.4E and F**, the motor and sensory recovery, respectively, were dependent on the injury severity and level (**Additional file 2: Figure 4**). The direct comparison with the EMSCI revealed similar pattern and rates of motor (**Figure 1.3C**, **Additional file 2: Tables S4-S6**) and sensory recovery (**Figure 1.3D**, **Additional file 2: Tables S7** and **S8**). Age and sex had no effect on the rate and pattern of sensorimotor recovery. Owing to a significant degree of missingness in the functional scores (i.e., Benzel score, > 30% data was missing), we refrained from computing functional recovery profiles for the patients enrolled in the Sygen clinical trial. In terms of the surgical timing, there was no statistical difference in the sensorimotor recovery trajectory between the early and late surgery group (**Additional file 2: Table S9**).

1.4.5 Interactive web platform *Neurosurveillance*

The Neurosurveillance web platform is hosted online and contains three main data visualization parts: (i) epidemiological features, including demographics and injury characteristics (**Additional file 4: Figure S1**); (ii) functional and neurological profiles (**Additional file 4: Figure S2**); and (iii) recovery monitoring of single patients or a group thereof. All data from the EMSCI study and the Sygen clinical trial, which was used in this study, can be explored in a customized fashion (e.g., customized selection of patient groups, one time point vs. multiple time-points).

1.5 Discussion

The primary aim of this study was to outline the epidemiological landscape of acute SCI over the last 20 years, as well as to provide a benchmark for the expected changes in standardized neurological and functional SCI outcomes. In line with our first hypothesis, the ratio between female and male patients remained fairly stable at approximately 1 : 3 throughout the surveillance period. The mean age at injury, however, has been steadily increasing over the last 20 years, which is consistent with an aging general population at risk. This increase was accompanied by a shift from a unimodal (i.e., young patients) to a bimodal distribution of age at injury (i.e., young and elderly patients). In terms of injury characteristics, the proportional distribution of injury severities and levels remained stable with the largest percentages of motor complete injuries. Our second hypothesis was not confirmed as neither the rate nor the pattern of neurological and functional recovery has changed since 2001—even after adjusting for injury characteristics and demographics. In essence, the degree a patient with SCI recovers sensory, motor, and walking function within 1-year post-injury remained stable over the last two decades. With the exception of the change in age at injury, all findings derived from the EMSCI study were confirmed through the external validation analysis of a secondary source of data (i.e., Sygen clinical trial performed in the USA). The similarity of results from these different data sources affirm that our findings are not markedly influenced by temporal or geographical biases or confounding factors related to the study design, timing of data collection, or population structure.

Confirming previous findings, the age at injury progressively increased throughout the surveillance period in both, male and female patients [78, 92, 93]. A shift from an unimodal (i.e., young patients) to a bimodal distribution of age at injury (i.e., young and elderly patients) was observed between 2001 and 2019. A cursory glance at the one of the largest US data sources, spinal cord injury model systems (SCIMS) [93], suggests that this upward trend in age at injury is evident since the early 1970s. Possible explanations for this observation are the increasing longevity in the general population along with an increase in propensity for risk taking among the elderly population [94]. Furthermore, the elevated susceptibility for SCI among elderly is also attributable to the increasing risk of falling with ageing [95]. In fact, the majority of SCI among elderly are sustained traumatically through falls [9]. Comparable to trends in the general population, the changed age structure of the SCI population has major implications on the medical and nursing services required in prevention and treatment of SCI and associated complications [96]. The latter is of particular concern, as the frequency of secondary health complications in older patients with SCI is markedly higher compared to younger patients [97]. Older age

at injury is not only associated with greater number of infections, cardiovascular and metabolic complications, but also more fatigue and a greater risk for cognitive impairments [98]. Moreover, the shift towards bimodal distribution of age at injury also has implications on the design of clinical trials and the stratification of patients as it is imperative that data collected from clinical trials are applicable to the patient population to be treated. Thus, forthcoming clinical trials must ensure an appropriate representation of elderly to provide meaningful and generalizable evidence and knowledge regarding the trialed treatment strategy. A proportionate participation of the elderly individuals in clinical trials is further desirable to allow for statistically meaningful subgroup analyses to account for age-related differences in treatment response (e.g., altered affect pharmacokinetics and pharmacodynamics, adverse drug events due to comorbidities or concomitant drugs). While the epidemiological landscape has been changing in terms of age, traumatic SCI remains much more common in men, with incidence rates that are three to four times higher compared to women. Along with reports from the SCIMS [93], the data from the Sygen clinical trial study further corroborate the robustness of the sex ratio. Our findings partially contrast previous reports suggesting an increase in the proportion of female patients since the early 2000s [99]. These divergent observations can likely be explained by the differences in study size (smaller studies are more prone to outliers), study population (e.g., focus on subgroups vs entire cohort), and duration of observation period (longer time windows allow to account for seasonal fluctuations). Independent of age and sex, the incidence of cervical injures remained higher than that of thoracic/lumbar injuries, as has been reported in other studies [99]. Although not reaching statistical significance, the annual proportion of lower thoracic spine injuries steadily decreased, while a greater number of cervical injuries was consistently recorded over the last two decades. In contrast, no such trend was detected for the injury severities, as their distribution remained fairly stable for both male and female patients and independent of the NLI. Results from the Sygen clinical trial study further suggest that the proportion of sensorimotor complete injuries are following a declining trend since the early 1990s.

Both the rate and pattern of neurological and functional recovery have been extensively studied over the last couple of decades [100, 101]. Generally speaking, recovery after acute SCI is characterized by an initial period of rapid improvement, with a plateau in sensory and motor function by one year, leaving most patients with some permanent neurological and functional deficits [100, 102]. Outcome-modifying factors include injury characteristics (level and severity), age, acute care concepts (early surgical decompression, blood pressure regulation), comorbidities, and medication administered to treat secondary complications (e.g., gabapentionoids) [103, 104]. Our international surveil-

lance study revealed that rate, pattern, and variability of neurological and functional recovery remained stable between 2001 and 2019. As a matter of fact, our validation analysis further suggests that this pattern has been unchanged since the early 1990s and is independent of geographical region, study design (observational vs. controlled clinical trial), and changes in population structure. Independent of the data source and year of injury, changes in neurological score were the greatest for tetra- and paraplegic AIS C patients. A markedly smaller increase was observed for patients with AIS D injuries owing to ceiling effects of the neurological scores. In contrast, the AIS C and AIS D showed the greatest increase in the functional scores, which are less prone to ceiling effects. Our findings are remarkable considering the ongoing changes in the acute care [75, 76, 105] and neurorehabilitation practices aiming at maximizing the functional recovery following a SCI [106]. However, the mainly applied concept in SCI rehabilitation still relies on fostering mechanisms of compensation and adaptation, while interventions of true neural repair and induced regeneration have not yet reached clinical practice. It is noteworthy that our study does not allow to make any assumption of the effects of potential changes to the very early acute care (e.g., surgical decompression, specialized transportation from scene of accident to hospital) on recovery. While our study indicates consistent patterns and robust trends for injury characteristics-dependent neurological and functional recovery during early rehabilitation in the sub-acute time period and long-term follow-up of one year, it does not capture the immediate effects of very early interventions on the recovery or outcome-modifying factors. Nevertheless, it is noteworthy that our sensitivity analysis revealed that there is yet no significant effect of early surgery on the longitudinal recovery trajectory. This is in line with a recent study by Jaja and colleagues [107] employing group-based longitudinal trajectory modeling. Additionally, the effect of outcome-modifying factors, such as medication, comorbidities, and readmissions, has not been assessed owing to the lack of this data in the EMSCI study. However, given the observed robustness of the recovery patterns, rate, and variability over the years and a fairly large cohort, we carefully conclude that these effects are marginal and might be specific to subpopulations. Future studies, powered to detect effects of outcome-modifying factors, are warranted to investigate the validity of this conclusion.

With recovery rates remaining rather consistent over recent decades, the data from the EMSCI can be pooled across the years making it the largest longitudinal observational study world-wide. EMSCI constitutes an unparalleled resource to inform real-time clinical practice as well as guide the design and implementation forthcoming clinical trials targeting neural repair and neural plasticity [108]. Gauging a patient's recovery trajectory is challenging owing to the high variability in neurological and functional recovery after

injury. Heterogeneous recovery makes accuracy in prognosis at early time-points after injury very difficult and creates a dilemma for clinicians asked to provide a prediction of long-term outcomes to patients and their families. Undoubtedly, there is a great need for accurate and reliable early injury exams or surrogates (e.g., blood biomarkers) thereof. With data from EMSCI, patients that share similar demographics and injury characteristics, physicians can provide a reference context with greater confidence to newly injured patients (i.e., concept of digital twins/siblings) [109]. Having a "digital twin" also allows tracking a patient's progress, detecting deviations from the projected trajectory, and initiating timely interventions (e.g., treatment of infections) if required. In the context of clinical trials, heterogeneity also adds variability to recovery trajectories, limiting the effectiveness of patient stratification methods, and potentially masking subtle treatment effects. Thus, the provided surveillance data will be instrumental to refine the patient selection and stratification for future clinical trials clinical trials targeting neural repair and neural plasticity.

Beyond this, our study suggests that observational data, such as the EMSCI, could be implemented as historical control data in clinical trials to, at least partially, replace a concurrent control. For rare conditions like acute SCI, there are a number of distinct advantages to the incorporation of historical control data into clinical trials. Chief among them is increasing the number of participants exposed to treatment and thereby, avoiding early termination of trials owing to difficulties with patient enrollment [110]. Moreover, the incorporation of quality external historical control data (e.g., EMSCI data) allows for reduced mean square error, increased power, and reduced type I error within the current trial [111]. In contrast, should the historical data be inconsistent with current trial control arm data, there is a potential for bias and inflated type I error. Residual confounding cannot just reliably be adjusted away, and misleading (causal and non-causal) associations may not be ruled out. Owing to the standardized data collection and curation by highly trained staff, the EMSCI constitutes a unique source for real-world evidence, particularly for clinical trials that are conducted at EMSCI centers. This is highlighted by the ongoing Nogo inhibition in spinal cord injury (NISCI) trial (clinicaltrial.gov identifier: NCT03935321). Accumulating evidence suggests that the appropriate usage of real-world evidence can increase the probability of successfully completing a clinical trial and even support regulatory decisions [112].

Our study has limitations. Firstly, the EMSCI database lacks information on mortality, which is an important factor when investigating how modifications to the standard of care change the epidemiological landscape. This limitation is mainly driven by the fact that the majority of the participating centers of EMSCI dedicated comprehensive SCI care centers to which patients are transferred from trauma centers, where they received

acute medical and surgical care. Trauma-related deaths would be recorded in the trauma centers and thus not collected within the EMSCI. Secondly, the standard of care after SCI (e.g., surgery and timing of surgery, rehabilitation training) was not standardized across the EMSCI centers. Non-uniform standard of care can potentially confound the data and results. In contrast, the Sygen study was completed in a rigorous manner, using a randomized clinical trial protocol designed to limit confounding variables. Despite these differences in study design, the findings related to neurological outcomes were comparable. Thirdly, neither the EMSCI nor the Sygen trial included non-traumatic SCI, with the exception of ischemic injuries. Longitudinal studies are warranted to shed light on potential changes in epidemiology and recovery profiles of non-traumatic spinal cord injuries. Lastly, EMSCI data have not undergone a thorough monitoring process as typically applied in controlled trials, which is a concern as it might impact the results of the study. Data missingness is inherent to any clinical study and particular observational studies. We addressed this concern by performing a comprehensive examination of the variables and patterns of missing data, which revealed that, in comparison to other observational studies, the degree of missingness is remarkably low.

1.6 Conclusion

In conclusion, the goal of this surveillance study was to provide an unparalleled overview of how the epidemiological landscape of SCI evolved between 2001 and 2019. Additionally, we addressed the questions whether and to what extent the rate and pattern of neurological and functional recovery changed over the last two decades. Leveraging the largest longitudinal observational SCI study, we observed a continuation in the previously reported trend toward increasing mean age at injury of new cases, while the ratio between male and female patients as well as the acute injury characteristics remained stable. Most interestingly, the rate and the pattern of neurological and functional recovery did not change throughout the surveillance period. External validation using the data from a landmark clinical trial conducted in the USA corroborated our findings regarding forecastable neurological recovery. It further suggests that our findings are not significantly confounded by geography, study design, and population structure and change thereof. In addition to the longitudinal quantification of the change in the population structure, our study provides a benchmark for expected changes in standardized outcomes after traumatic SCI. These seminal findings will inform and guide the development and implementation of future clinical trials assessing the safety and effectiveness of novel therapies — with the potential applicability in a multinational setting.

Acknowledgements

The authors would like to acknowledge the participating centers in the EMSCI network and Sygen clinical trial that were involved in the patient care and collection of data necessary for this study. Special thanks are given to Prof HJ Gerner (Heidelberg, Germany) as well as Prof V Dietz and Prof M Schwab (Zurich, Switzerland) for envisioning the value of the EMSCI network and laying the foundation for this research. Furthermore, the authors like to thank Prof. Andrew Maas for his insightful comments and suggestions.
EMSCI active members:

- Uniklinik Balgrist, Zentrum für Paraplegie, Zurich, Switzerland
- Klinik für Paraplegiologie, Universitätsklinikum Heidelberg, Heidelberg, Germany
- Klinikum Bayreuth, Bayreuth, Germany
- UMC St. Radboud, Nijmegen, Netherlands
- RKU - Universitäts- und Rehabilitationskliniken Ulm, Ulm, Germany
- BG Kliniken Bergmannstrost Halle, Halle, Germany
- Institut Guttmann, Barcelona, Spain
- SRH Klinikum Karlsbad-Langensteinbach, Karlsbad, Germany
- BG Unfallklinik Murnau, Murnau, Germany
- Werner-Wicker-Klinik, Bad Wildungen, Germany
- Motol Hospital, Prague, Czech Republic
- Orthopädische Klinik Hessisch Lichtenau, Hessisch-Lichtenau, Germany
- Hospital Nacional de Parapléjicos, Toledo, Spain
- Schweizer Paraplegiker-Zentrum, Nottwil, Switzerland
- Ospedale Fondazione Santa Lucia, Rome, Italy
- Indian Spinal Injury Center, New Delhi, India
- BG Unfallklinik Tübingen, Tübingen, Germany
- Unfallkrankenhaus Berlin, Berlin, Germany
- Klinik Bavaria, Kreischa, Germany
- REHAB Basel, Basel, Switzerland
- The Robert Jones and Agnes Hunt Orthopaedic Hospital, Oswestry, Great Britain
- AUVA Wien, Vienna Austria
- Fondazione Salvatore Maugeri, Pavia, Italy

Authors' contributions

LB: substantial contributions to acquisition, analysis, and interpretation of data; drafting the manuscript; and data visualization
BT: acquisition, curation, and interpretation of data and revising the manuscript critically for important intellectual content
FG: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
MS: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
FR: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
MS2: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
NW: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
RR: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
KY: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
RA: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
DM: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
LG: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
CH: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
TL: acquisition and interpretation of data and revising the manuscript critically for important intellectual content
JC: substantial contributions to the interpretation of data and revising the manuscript critically for important intellectual content
JK: substantial contributions to the interpretation of data and revising the manuscript critically for important intellectual content
AC: substantial contributions to the conception and design of the study; acquisition, analysis, and interpretation of data; and drafting the manuscript
CJ: substantial contributions to the conception and design of the study; acquisition,

analysis, and interpretation of data; data visualization; and drafting the manuscript. All authors read and approved the final manuscript.

Funding information

This study was supported by the Swiss National Science Foundation (Ambizione Grant, #PZ00P3_186101), Wings for Life Research Foundation (#2017_044, #ID 2020_118), the International Foundation for Research in Paraplegia (IRP, Curt), and the Craig H. Neilsen Foundation. This study has further received funding from the European Union's Horizon 2020 research and innovation program under the grant agreement No 681094 and is supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 15.0255. Dr. Cragg holds a Tier 2 Canada Research Chair. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Open access funding provided by Swiss Federal Institute of Technology Zurich.

Availability of data and materials

The data used for this study, including de-identified individual participant data and a data dictionary defining each field or variable within the dataset, can be made available upon reasonable request to the corresponding author (CRJ). Data will be made available following publication of this work. Written proposals will be evaluated by the authors, who will render a decision regarding suitability and appropriateness of the use of data. Approval of all authors will be required and a data sharing agreement must be signed before any data are shared. The code to run the analysis as well as create the figures and tables can be found on our Github repository[4].

Declarations

Ethics approval and consent to participate

The study was performed in accordance with the Declaration of Helsinki and was approved by all responsible institutional review boards. Throughout the study duration, EMSCI followed the ethics guidelines of the participating countries and implemented changes and new policies as required. Patients gave their written informed consent before being included in the EMSCI database.

[4]https://github.com/jutzca/SCI_Neurological_Recovery

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

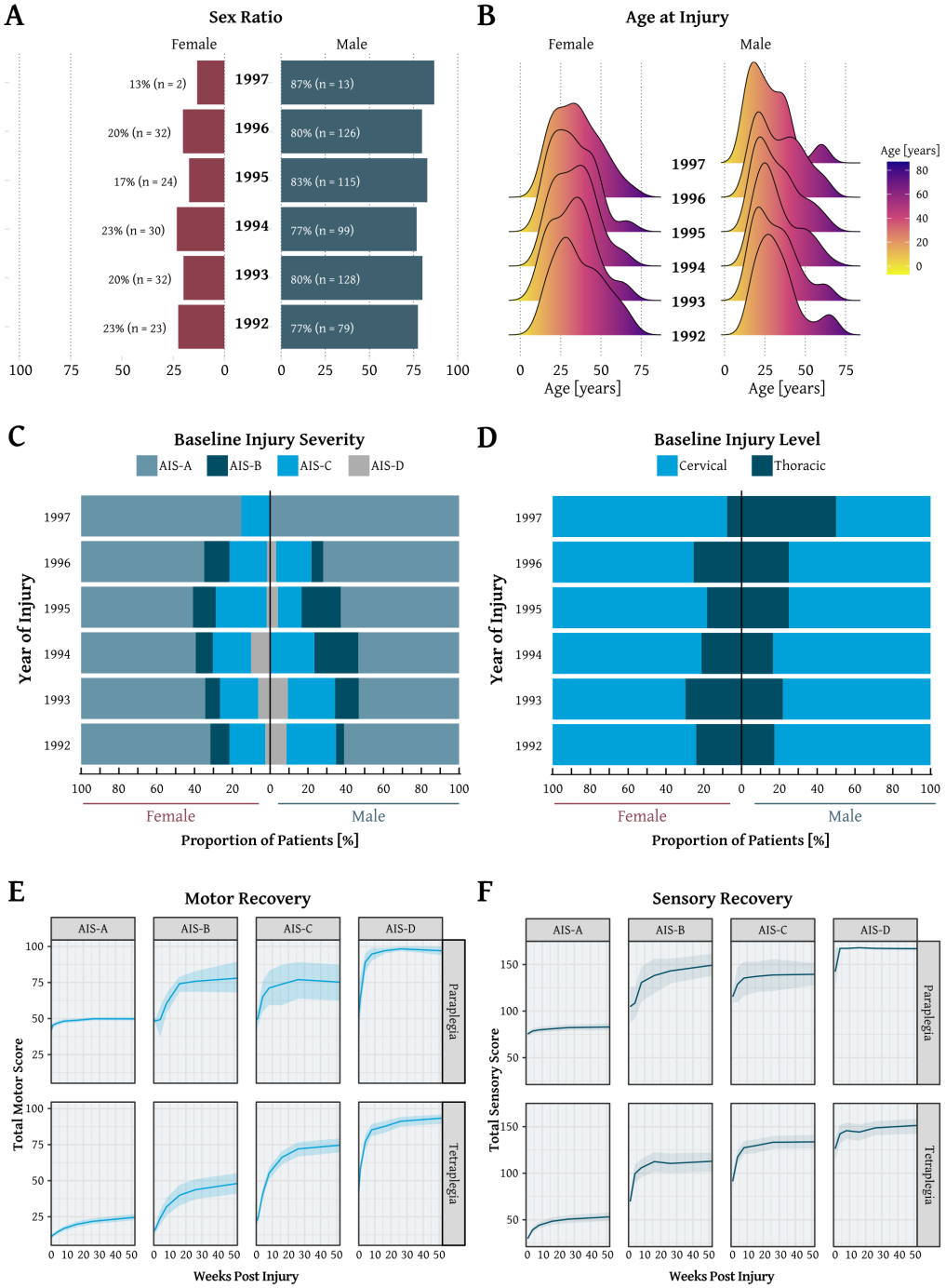


Figure 1.4: Overview of the validation study. A. The ratio between male and female individuals with a spinal cord injury. Depending on the year, the ratio of male and female spinal cord injury individuals changed between 3:1 and 4:1; **B.** Distribution of age at injury. Throughout the clinical trial period, there was no change in distribution of age at injury. Important to note, the average age at injury of the Sygen clinical trial cohort, independent of sex, was significantly lower compared to the European multicenter study on human spinal cord injury (EMSCI) cohort; **C.** Baseline injury severity; and **D.** injury level: The proportions of injury characteristics remained constant between 1992 and 1997; **E.** Motor; and **F.** sensory recovery stratified by American spinal injury association (ASIA) impairment scale (AIS) grade and plegia (i.e., paraplegia or tetraplegia). The solid lines represent the fitted models and the shaded areas the standard deviation.

Chapter 2

Natural progression of routine laboratory markers after spinal trauma: A longitudinal, multi-cohort study

---

Adapted from:

Lucie Bourguignon, Anh Khoa Vo, Bobo Tong, Fred Geisler, Orpheus Mach, Doris Maier, John L. K. Kramer, Lukas Grassner, and Catherine R. Jutzeler

Journal of Neurotrauma, 2021

Publication: 10.1089/neu.2021.0012
Haemosurveillance web site: https://jutzelec.shinyapps.io/Haemosurveillance/
GitHub repository: https://github.com/jutzca/Systemic-effects-of-Spinal-Cord-Injury

Lucie Bourguignon primarily contributed to the data cleaning and analyses, interpretation, and visualization of the results including building the Haemosurveillance web site, and drafting the manuscript, on an original idea and with contributions from Catherine R. Jutzeler. Further details can be found in **Section 2.6**.

2.1 Abstract

Our objective was to track and quantify the natural course of serological markers over the first year following spinal cord injury. For that purpose, data on serological markers, demographics, and injury characteristics were extracted from medical records of a clinical trial (Sygen) and an ongoing observational cohort study (Murnau study). The primary outcomes were concentration/levels/amount of commonly collected serological markers at multiple time points. Two-way analysis of variance (ANOVA) and mixed-effects regression techniques were used to account for the longitudinal data and adjust for potential confounders. Trajectories of serological markers contained in both data sources were compared using the slope of progression.

Our results show that, at baseline (≤ 2 weeks post-injury), most serological markers were at pathological levels, but returned to normal values over the course of 6–12 months post-injury. The baseline levels and longitudinal trajectories were dependent on injury severity. More complete injuries were associated with more pathological values (e.g., hematocrit, ANOVA test; $\chi^2 = 68.93$, degree of freedom (df) = 3, adjusted p value < 0.001, and $\chi^2 = 73.80$, df = 3, adjusted p value < 0.001, in the Sygen and Murnau studies, respectively). Comparing the two databases revealed some differences in the serological markers, which are likely attributable to differences in study design, sample size, and standard of care. We conclude that because of trauma-induced physiological perturbations, serological markers undergo marked changes over the course of recovery, from initial pathological levels that normalize within a year. The findings from this study are important, as they provide a benchmark for clinical decision making and prospective clinical trials. All results can be interactively explored on the Haemosurveillance website [1] and GitHub repository [2].

[1] https://jutzelec.shinyapps.io/Haemosurveillance/
[2] https://github.com/jutzca/Systemic-effects-of-Spinal-Cord-Injury

2.2 Introduction

Because of its crucial role in the coordination of bodily functions, damage to the spinal cord can lead to severe dysfunction or failure in single or multiple organs, including the heart, kidney, and liver [113]. As a consequence of altered functions, levels or concentration of biomarkers derived from conventional serological tests are modified [44, 114]. Their readiness and straightforward collection make these serological markers, which encompass both hematological (complete blood count (CBC)) and biochemical indices, ideally suited for evaluating the trauma-induced systemic perturbations. Laboratory tests are routinely conducted in the acute phase of injury to assess the initial magnitude of systemic damage and to monitor the bodily functions. However, little is known about how the systemic effects and their respective serological markers progress as a function of time. This paucity of knowledge is even more surprising, considering that these serological markers have the potential to guide the design (patient stratification) and implementation of clinical trials (safety assessment of trialed drug) [114, 115, 116]. To address this knowledge gap, the aim of this study was to determine the natural progression of serological markers following a spinal cord injury (SCI). We hypothesized that, by disruption of normal innervation of vital organs after a traumatic SCI, there will be time-dependent and injury-specific alterations in serological markers characterized by an initial pathological change that normalizes over time (i.e., reaches norm values of healthy able-bodied people). Lastly, we provide the scientific and medical community with a first-of-its-kind surveillance tool "Haemosurveillance" which aims to generate novel research questions as well as to inform clinical decision making and clinical trial design.

2.3 Methods

2.3.1 Study design and data source

To determine the natural progression of serological markers following SCI, we performed an observational study of prospectively collected data. Therefore, we analyzed two different data sources, one each from the United States of America (USA) and Germany. The first data source was a prospective phase III, placebo-controlled, multi-center study assessing the efficacy of gangliosidosis 1 (GM-1) therapy in acute traumatic SCI [33, 34]. Running from 1992 to 1998, the Sygen trial failed to demonstrate a superior treatment effect of GM-1 over placebo treatment. Full design, recruitment, and enrollment details of the Sygen trial have been described previously [117]. A total of 797 patients across the United States were included in the randomization. Within the framework of this USA food and drug administration (FDA) regulated trial, detailed information concerning neurological scores and blood chemistry were meticulously collected. The second data source was an observational cohort study conducted at the over-regional level-I trauma center in Murnau, Germany (hereafter referred to as the Murnau study). Between 2004 and 2017, 363 patients were enrolled and followed up for one-year post-injury. All patients enrolled in the Murnau study received standard rehabilitation care.

2.3.2 Ethics approval

The study was performed in accordance with the Declaration of Helsinki. Approval for the secondary analysis of the Sygen trial was received by an institutional ethical standards committee on human experimentation at the University of British Columbia. The original Sygen clinical trial (results published elsewhere) also received ethical approval, but was conducted before clinical trials were required to be registered [34, 117, 118]. The data received from the original clinical trial were de-identified. The Murnau study was approved by the Bavarian Medical Chamber (#2018-077).

2.3.3 Cohort definition: Inclusion and exclusion criteria

To be included in our study, patients needed to have blood values at three different time points as well as information on sex, age, and injury characteristics (i.e., injury severity, injury level, and baseline motor and sensory scores). Baseline was defined as the first 72 hours after injury for the Sygen trial and the first two weeks post-injury for the Murnau study. Patients were excluded if any of these data were missing or if they had sustained a non-traumatic injury (e.g., a tumor), or had decided to withdraw their data over the

course of the study.

2.3.4 Outcome, predictor, and confounding variables

The primary outcomes were serological markers with data available for at least 50 patients at each time point. This threshold was chosen to ensure that the model output was interpretable, statistically powerful enough to make inferences, and clinically relevant. Independent variables were time points post-injury at which serological markers were collected. As an FDA requirement for the Sygen trial, detailed information regarding routine blood chemistry was collected at admission to the trauma center (hereinafter referred to as week 0), and at 1, 2, 4, 8, and 52 weeks post-injury. The laboratory analyses were all performed by SmithKline Beecham between February 1997 and April 1993 using the available clinical machines in this time period (**Table S1**). In the Murnau study, information on serological markers was collected upon the request of the attending physicians (i.e., not at standardized time points). As a consequence, different numbers of blood draws were collected for each patient on different days post-injury. All laboratory analyses were performed in-house at the berufsgenossenschaftliche unfallklinik (BGU) Murnau. Normal ranges for the serological markers were provided by the manufacturer of the analytic devices (**Table S1**). Normal ranges derived from the Murnau study were also applied to the analysis of the Sygen study. The rationale for that stems from the fact that the original upper and lower bound values in Sygen are not available anymore. Potential confounders included age, sex, injury completeness (at time of injury) according to the American spinal injury association (ASIA) Impairment scale (AIS) [119], level of injury (at/above T6 vs below T6), and presence or absence of polytrauma. Polytrauma was defined as significant injuries of three or more points in two or more different anatomic regions in addition to the SCI [120]. In the Sygen trial, the injury severity was assessed using the Frankel Scale, whereas in the Murnau study the AIS grading scale was employed. In order to facilitate a comparison between the two data sources, we recalculated the AIS grades for all patients enrolled in the Sygen trial using the European multicenter study on human spinal cord injury (EMSCI) international standards for neurological classification of spinal cord injury (ISNCSCI) calculator [3].

2.3.5 Statistical analyses

Two-way analysis of variance (ANOVA) and mixed-effects regression models were chosen for the primary analyses. These models were naturally suited to account for the longitudinal nature of the data as well as to adjust for potential confounders. Dependent variables

[3] https://ais.emsci.org/

were all serological markers that met our inclusion criteria. In the Murnau study, blood values were averaged per week, from week 0 to week 7 post-injury. In both studies, if, for a certain marker, patient, and time point, no data were available, the time point for this patient's marker was excluded. For analyses comparing both studies, we examined the percentage of deviation from the mean of the normal range, collected from the Murnau study. The rationale for this normalizing procedure was to make the data of the two cohorts comparable despite having different units. Independent variables were time post injury, AIS grade, or level of injury, when examining data from the individual studies. When comparing the serological markers from both studies, we added the data sources as an independent variable. For mixed-effects regression models, pairwise comparisons of the different levels of the independent variable of interest were performed. Hence, significance levels were adjusted for multiple comparisons using Tukey's test, and $p < 0.05$ after adjustment, was regarded as statistical significance. For one-study two-way ANOVA tests, we applied Bonferroni correction for testing for six independent variables together. Thus, we adjusted p values, and $p < 0.05$ was regarded as statistical significance. In the same way, when comparing the two studies, no correction was applied, as only the data source was considered to be an independent variable. Thus, $p < 0.05$ was regarded as statistical significance. For all analyses, R Statistical Software, version 3.6.3 (running under: macOS Mojave 10.13.6), was used.

2.3.6 Data visualization

Using the R package `Shiny` and `ShinyDashboard`, we created an online interface to visualize the results of the current study and to interactively explore the data used for this study.

2.3.7 Data and code availability statement

Anonymized data used in this study will be made available upon request to the corresponding author and in compliance with the European general data protection regulation (EU GDPR). The code describing the analysis can be accessed on our GitHub repository [4].

[4] https://github.com/jutzca/Systemic-effects-of-Spinal-Cord-Injury

Table 2.1: Subject and injury characteristics of patients included in our analysis and enrolled in the Sygen trial and Murnau study, respectively

	Sygen trial	**Murnau study**	**p value**
Subject characteristics			
Total, *n*	703	239	
Sex, *n* (%)			0.786
Male	560 (79.7)	193 (80.8)	
Female	143 (20.3)	46 (19.2)	
Age in years at injury			**< 0.001**
Mean±SD	33±14	51±19	
Neurological/functional outcomes			
Baseline ASIA impairment scale[a], *n* (%)			**< 0.001**
A	446 (63.4)	81 (33.9)	
B	77 (11.0)	22 (9.2)	
C	149 (21.1)	26 (10.9)	
D	31 (4.4)	110 (46.0)	
Lower extremity motor score, mean±SD			
Baseline	2.82±7.3	19.5±19.9	**< 0.001**
After one year	12.8±19.3	28.1±21.9	**< 0.001**
NA, *n*	140	105	
Serological markers, *n*	47	39	

[a] American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description;

Significant p values are highlighted in **bold**.
standard deviation (SD)

2.4 Results

2.4.1 Cohort summary: Included patients

Subject and injury characteristics of both cohorts (Sygen: 679; Murnau: 239) are summarized in **Table 2.1**. A comparison revealed a comparable ratio of male and female patients (Pearson's χ^2 test, $\chi^2 = 0.07$, df = 1, $p = 0.786$). However, significant differences were found in terms of age distribution (two-sided t-test, $t = 13.63$, df = 322.55, $p < 0.001$, **Figure S1**) and injury severity distribution (Pearson's χ^2 test, $\chi^2 = 244.9$, df = 3, $p < 0.001$).

2.4.2 Cohort summary: Excluded patients

A total of 94 and 124 patients in the Sygen trial and Murnau study, respectively, did not meet the inclusion criteria and were excluded. Reasons for exclusion comprised normal AIS grade (AIS E, $n = 5$) and missing information on baseline AIS grade ($n = 192$). **Table S2** provides a detailed overview of the excluded cohorts. Excluded and included cohorts were significantly different in terms of age distribution (two-sided t-test; $t = 2.03$, df $= 124.56$, $p = 0.04$, with excluded cohort younger than included cohort; and, $t = -1.8852$, df $= 123.91$, $p = 0.06$, with excluded cohort older than included cohort), in the Sygen trial and Murnau study, respectively. Excluded and included cohorts were comparable in terms of ratio of male and female patients (Pearson's χ^2 test; $\chi^2 = 3.43$, df $= 1$, $p = 0.06$), in the Sygen trial, but significantly different in the Murnau study (Pearson's χ^2 test; $\chi^2 = 8.73$, df $= 1$, $p = 0.003$).

2.4.3 Serological markers

A total of 32 and 28 routinely assessed blood markers were available in the Sygen trial and Murnau study, respectively. Among these, 14 and 8 blood markers, respectively, were part of the CBC, which is a test that evaluates the cells that circulate in blood. Notably, it includes counts of platelets, red and white blood cells, hemoglobin, and hematocrit. The remaining blood markers reflect renal function (5 and 4 markers in the Sygen trial and Murnau study, respectively), hepatic function (5 and 6 markers), pancreatic function (1 and 2 markers), and muscle damages (2 and 3 markers). Overall, 20 blood markers were shared among the two data sources. **Table 2.2** provides an overview of all collected markers.

Table 2.2: Serological markers collected in the Sygen trial and Murnau study

	Sygen trial	**Murnau study**
Complete blood count		
	Erythrocytes	**Erythrocytes**
	Hemoglobin	**Hemoglobin**
	Hematocrit	**Hematocrit**
	MCHC	**MCHC**
	MCV	**MCV**
	Thrombocytes	**Thrombocytes**

Continued on next page

Table 2.2: Serological markers collected in the Sygen trial and Murnau study (Continued)

	Sygen trial	**Murnau study**
	Leucocytes	**Leucocytes**
	Lymphocytes	Hemoglobin per erythrocyte
	Monocytes	
	Neutrophils	
	Eosinophils	
	Basophils	
	MCH	
	Total serum	
Liver		
	Alkaline phosphatase	**Alkaline phosphatase**
	ASAT	**ASAT**
	ALAT	**ALAT**
	Total bilirubin	**Total bilirubin**
	Chloride	Gamma-GT
		Lactate dehydrogenase
Kidney		
	Calcium	**Calcium**
	Creatinine	**Creatinine**
	Albumin	**Total proteins**
	Blood urea nitrogen	**Blood urea nitrogen**
	Uric acid	
Muscle		
	Potassium	**Potassium**
	Sodium	**Sodium**
		Cholinesterase
Pancreas		
	Amylase	**Amylase**

Continued on next page

Table 2.2: Serological markers collected in the Sygen trial and Murnau study (Continued)

	Sygen trial	**Murnau study**
		Lipase
Others		
	Glucose	**Glucose**
	Prothrombin time	**INR**
	Cholesterol	Partial thromboplasmin time
	Triglycerides	CRP
	Carbon dioxide	Quick test
Serological markers, n	32	28

A total of 32 and 28 serological markers were available in the Sygen trial and Murnau study, respectively. Overall, 20 serological markers were collected in both studies (highlighted in **bold**). mean corpuscular hemoglobin concentration (MCHC); mean corpuscular volume (MCV); mean corpuscular hemoglobin (MCH); aspartate aminotransferase (ASAT); alanine aminotransferase (ALAT); γ-glutamyl transferase (Gamma-GT); international normalized ratio (INR); C-reactive protein (CRP)

2.4.4 Natural progression of serological markers post-injury

With the exception of amylase, c-glutamyl transferase (GGT), glucose, lipase, and alanine aminotransferase (ALAT) in the Murnau study ($p = 0.624$, $p = 1$, $p = 0.081$, $p = 1$, $p = 0.242$, respectively) and alkaline phosphatase, potassium, and thrombocyte levels in the Sygen trial ($p = 0.685$, $p = 1$, $p = 1$, respectively), the concentrations of serological markers significantly changed as a function of time since injury (**Tables S3** and **S4**). For 28 serological markers, these changes occurred within the normal range. The remaining 24 serological markers had baseline values outside the normal range, which normalized over the course of recovery (**Figures 2.1** and **2.2**). One serological marker (i.e., hematocrit) remained outside the normal range at one-year post-injury.

2.4.5 Relationship between serological levels and injury characteristics

In line with our hypothesis, ANOVA revealed a global effect of injury severity (i.e., AIS grade). Our post-hoc analysis revealed that the serological values were dependent on

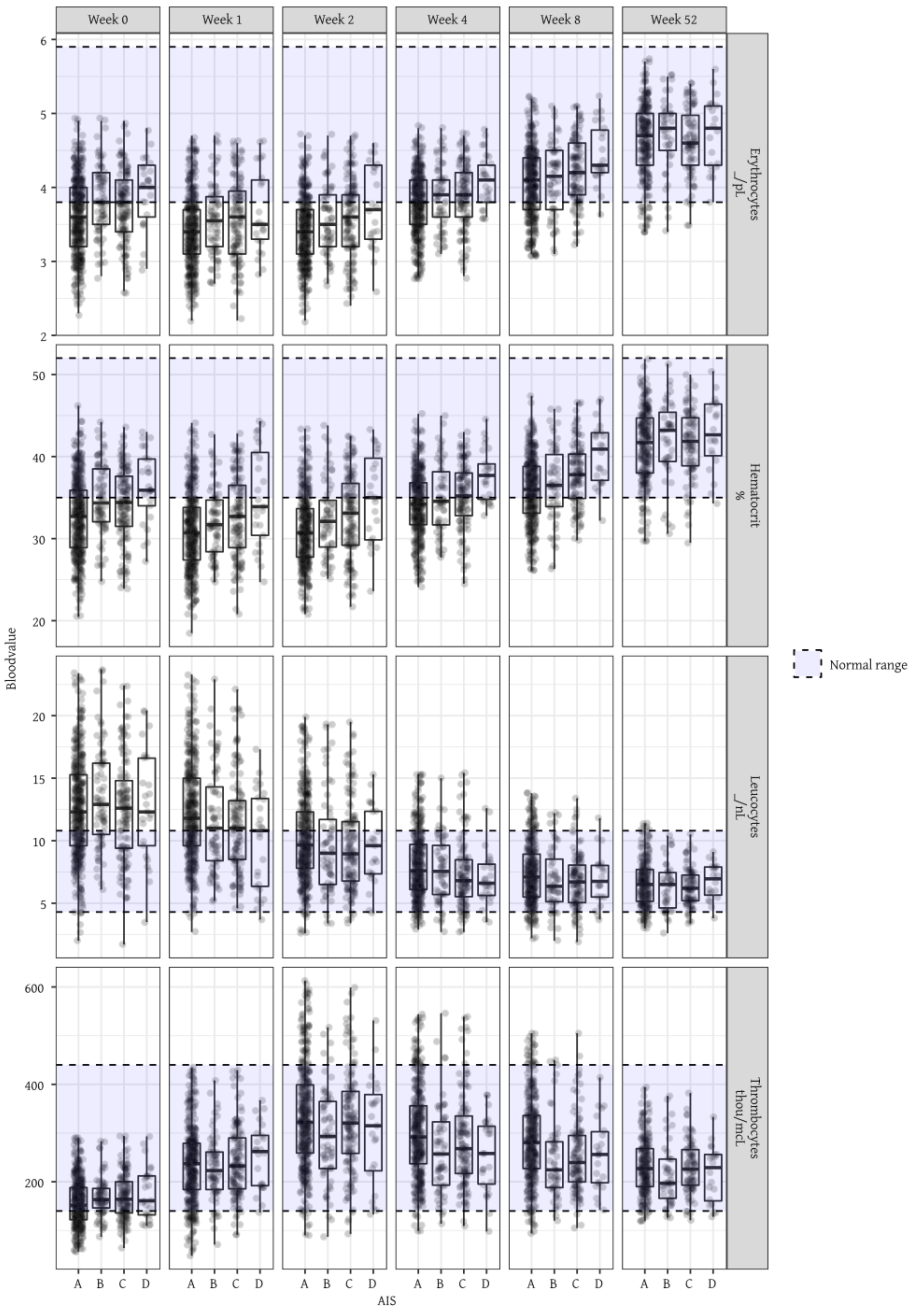


Figure 2.1: Natural progression of the complete blood count in patients with spinal cord injury (SCI) who were enrolled in the Sygen trial. Three different patterns of progression were observed. First, the blood markers, such as thrombocytes, remained constant and within the range of able-bodied people. Second, blood markers were pathological immediately after the trauma, but recovered over the course of a year and reached the normal range. Erythrocytes, hemoglobin, and leucocytes are characterized by such a course. Third, values were initially within the normal range, but as a function of time they became pathological when compared with those of able-bodied people. Hematocrit is one such example (not shown here). For clinical decision making as well as the design and implementation of clinical trials, it is of utmost importance to know the temporal progression of these blood markers. For further exploration of the data, please refer to the web application Haemosurveillance (https://jutzelec.shinyapps.io/Haemosurveillance/)

.

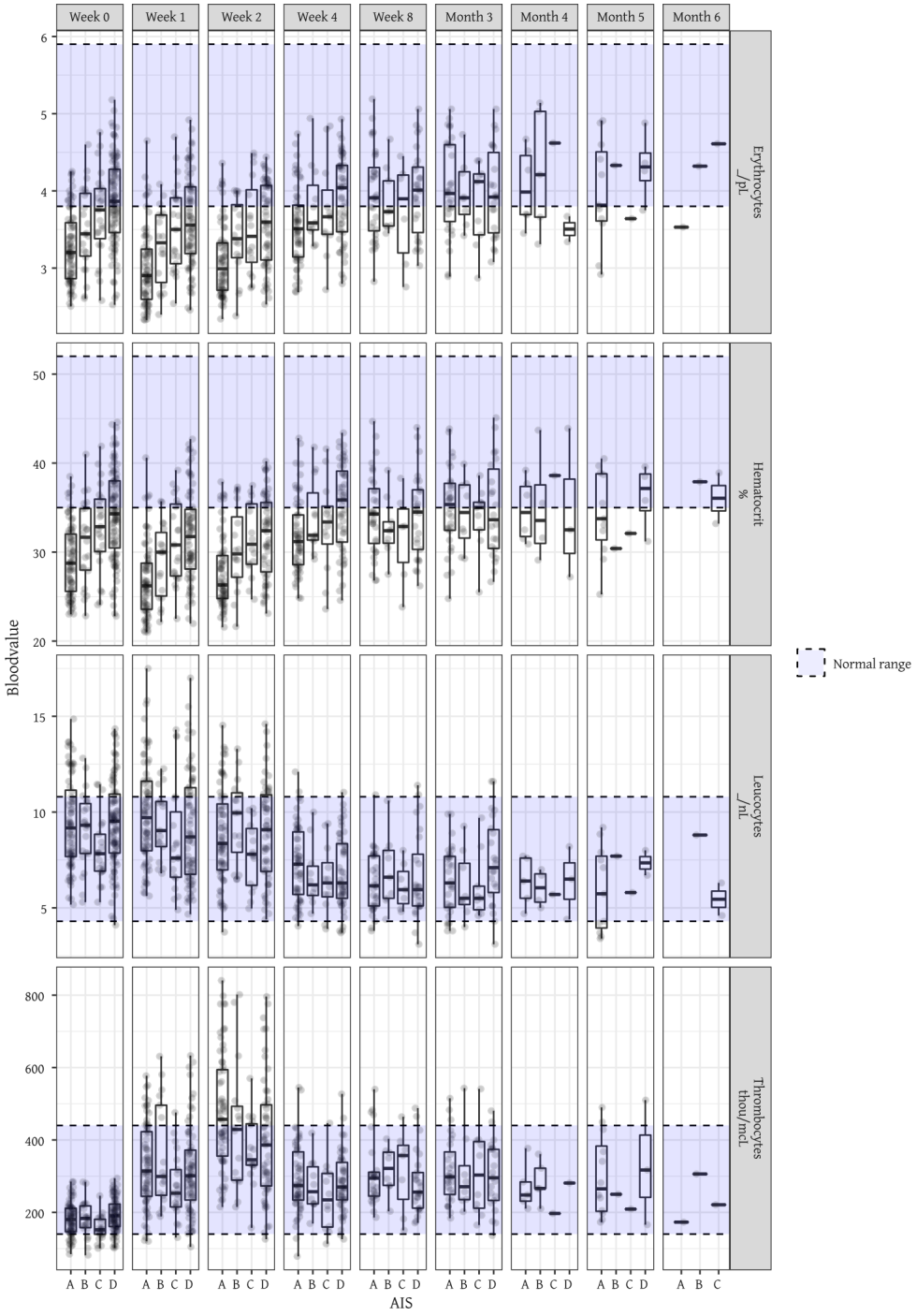


Figure 2.2: Natural progression of the complete blood count in patients with spinal cord injury (SCI) who were enrolled in the Murnau study.

the AIS grades, calcium ($p < 0.001$ and $p = 0.007$), hematocrit ($p < 0.001$ and $p < 0.001$), hemoglobin ($p < 0.001$ and $p < 0.001$), erythrocytes count ($p < 0.001$ and $p < 0.001$), and total protein/albumin levels ($p < 0.001$ and $p < 0.001$), in both the Murnau study and the Sygen trial, respectively (**Tables S3** and **S4**). The pairwise comparisons between the AIS grades yielded that calcium, hematocrit, hemoglobin, erythrocyte count, and total

protein/albumin levels were significantly different between patients classified as AIS A and those classified as AIS D. In all cases, higher values for these markers, closer to the normal range, were associated with less severe injury (AIS D), as illustrated in **Figure 2.1**. Additionally, hematocrit, hemoglobin, erythrocyte count, and total protein/albumin were significantly different between patients classified as AIS A and those classified as AIS B, C, and D. All results are reported in **Tables S5** and **S6** and illustrated in **Figures S2–S7**. In terms of injury level, we found no significant differences in serological values between patients with injuries at/above T6 and those with injuries below T6 in both the Murnau study and the Sygen trial (**Tables S3** and **S4**). Lastly, the presence or absence of a polytrauma had a significant impact on some of the serological values (**Tables S3** and **S4**).

2.4.6 Comparison between historical and contemporary cohort

As described, the Murnau study and Sygen trial have a number of major differences in their design. As illustrated in **Figure 2.3**, there were significant differences in the serological markers and their progression (**Table S7**), with the exception of amylase ($p = 0.114$), alkaline phosphatase ($p = 0.409$), MCHC ($p = 0.053$), sodium ($p = 0.476$), and ALAT levels ($p = 0.746$).

2.4.7 Data visualization

All results can be explored interactively on the Haemosurveillance website [5]. Information is presented in separate tabs for patients enrolled in the Sygen and Murnau studies, respectively. The interactive interface also allows visualization of the data stratified by demographics (sex and age group) and injury characteristics (i.e., injury severity and type of plegia). Additionally, the interface facilitates a direct comparison of the two data sources.

[5] https://jutzelec.shinyapps.io/Haemosurveillance/

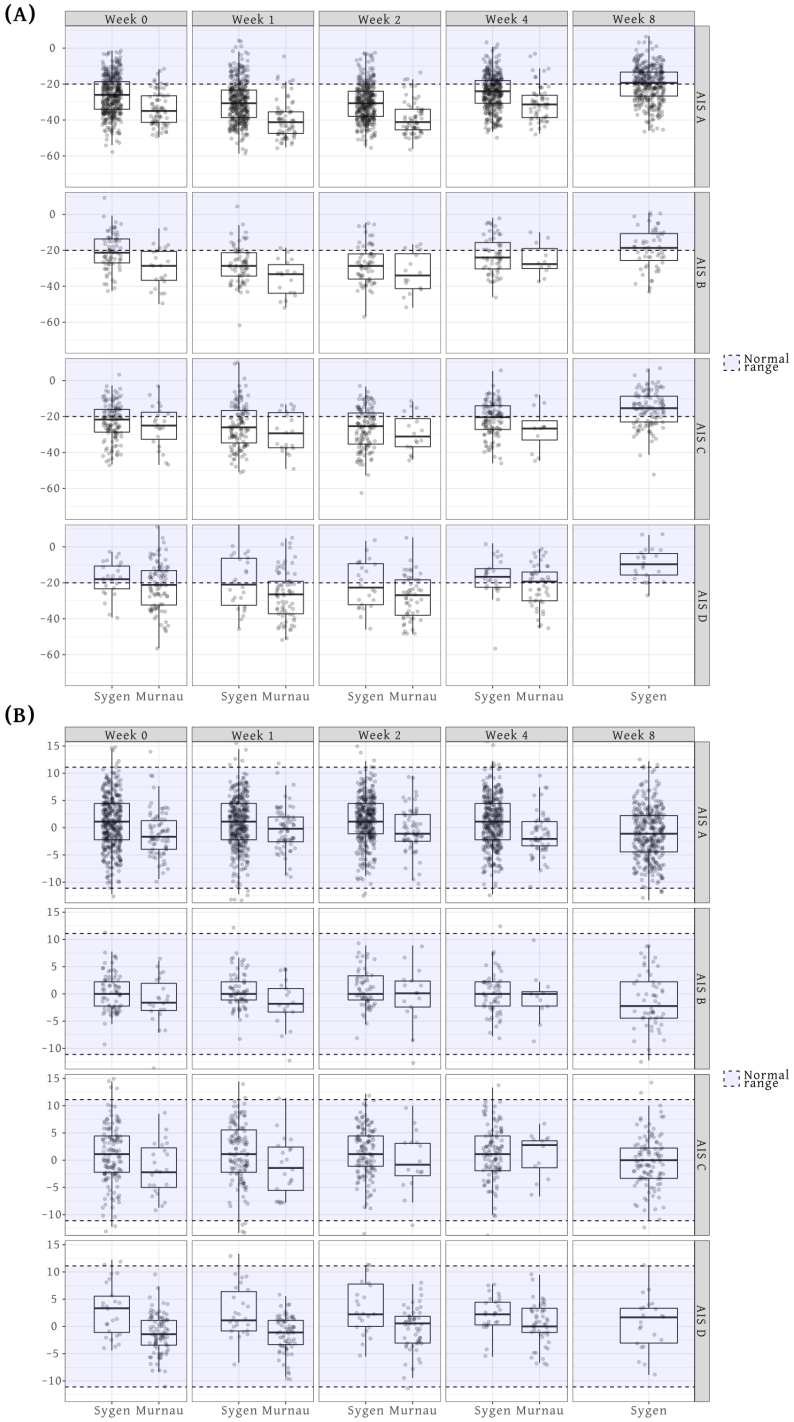


Figure 2.3: Comparison of the natural progression of hemoglobin (A) and mean corpuscular volume (MCV) (B) in patients with spinal cord injury (SCI) enrolled in the Sygen trial and the Murnau study, respectively.

2.5 Discussion

The present study describes the natural progression of serological parameters that are routinely assessed on admission and in the days to weeks following acute SCI. Consistent

with our first hypothesis, we found trauma-induced changes in routinely collected serological markers (e.g., hemoglobin, glucose). By and large, most of the markers normalized at one-year post-injury (i.e., reached the normal values of healthy able-body people). Our second hypothesis was also confirmed, insofar as the observed changes in markers were dependent on age at injury, sex, and injury severity, but not injury level. This suggests that these changes, in addition to reflecting the polytrauma and the consequent recovery process, are also capturing the severity of the SCI. Additionally, age at injury can be considered as a potential confounder for both the serological levels and the injury severity, which, itself impacts significantly the observed changes in serological markers. Collectively, this study provides new insights that will aid the design and implementation of clinical trials.

2.5.1 Natural progression and the relationship between serological levels and injury severity

In the present study, the majority of the serological markers reach pathological level shortly after the traumatic event and then normalize within a year post-injury. At baseline (within two weeks post-injury), the degree of alterations in the serological markers was associated with the injury severity, in such a way that patients with complete injuries exhibited more pronounced abnormalities in serological markers than those with incomplete injuries. This relationship between serological markers and degree of injury severity underpins the notion that serological markers may be utilized as measurable indicators of the severity. As such, they bear the potential to aid the diagnosis of SCI severity, particularly in cases in which standard neurological examination is not possible (e.g., intoxicated or unresponsive patients) [114]. Moreover, abnormalities in certain serological markers (e.g., albumin) [44, 121] may also induce further damage or delay the recovery process and, therefore, need to be addressed. In a recent study, Tong and colleagues detected that patients with prolonged hypoalbuminemia recovered to a lesser degree than those patients with normal albumin levels [44, 121]. Timely substitution of albumin might have beneficial effects on the functional and neurological recovery of the patient, as suggested by findings from animal studies [122]. Although the return to normal serological levels occurs along the same timeline as the neurological and functional recovery, for many serological markers there is no longer an association between serological levels and injury severity. This lack of association in the chronic phase of injury suggests that the serological markers are more representative of the initial polytrauma and the recovery from it as opposed to being specific indices of the SCI.

2.5.2 Serological markers in the design and implementation of clinical trials

Our study provides an important framework for the implementation of serological markers in the design and conduct of clinical trials. Conventionally, the safety and tolerability of trialed treatments are assessed by means of specific abnormalities of routinely collected serological and cerebrospinal fluid (CSF) markers [123]. As the majority of drugs, including the currently trialed riluzole [124, 125] and minocycline [126, 127], are metabolized and cleared by the liver and kidney, respectively, regulatory agencies released guidelines for the assessment of risk surrounding drug-induced liver injuries (DILI) [128, 129] and nephrotoxicity [130] in clinical trials. Multiple scheduled blood draws facilitate the early detection, tracking, and management of drug-induced organ damage. Typically, any deviation from the norm values of healthy able-bodied people would alert the investigators. In SCI, however, baseline values of numerous serological markers are pathological (**Figures 2.1** and **2.2**), which, when ignored or unknown, can substantially bias assumptions on drug safety. Our Haemosurveillance tool offers a first-of-its-kind platform to accurately disentangle drug-induced from trauma-driven perturbations in routinely collected serological markers. This tool is particularly useful for (i) clinical trials without a control group (i.e., placebo) and (ii) clinical trials with a control group that is not being managed by a standard of care. In the former situation, historical data can aid the evaluation of the safety of the trialed drug, whereas in the latter situation, the effect of the deviation from the standard of care can be measured. For example, in the ongoing Nogo inhibition in spinal cord injury (NISCI) trial [6], all enrolled patients are subject to repeated lumbar puncture regardless of their allocation. As repeated lumbar puncture is not a standard of care, historical data can be leveraged to assess their impact on health (e.g., rate of infections).

In addition to providing guidance on drug safety and tolerability, serological markers bear the potential to refine the stratification of patients and increase the likelihood of detecting a significant treatment effect [131, 132]. A major barrier to detecting small treatment effects in clinical trials is the extensive heterogeneity of the neurological recovery and the scarcity of reliable predictors, such as the initial damage to the spinal cord (i.e., AIS grades), that can fully capture the extent of the injury. Therefore, utilizing a biological correlate (e.g., blood or central nervous system (CNS) marker) is potentially advantageous and informative because of its representation of the trauma and indirect involvement in the CNS.

[6] https://nisci-2020.eu/index.php?id=1449

2.5.3 Differences between data sources

In the current study, we analyzed data from two different data sources to validate our findings regarding temporal trajectories of the serological markers. Overall, these trajectories show comparable trends. However, some differences were uncovered that are likely attributable to differences in the study design, study period, standard of care, population structure, and sample size. The Sygen trial, our first data source, was conducted in the 1990s and had five pre-defined time points of blood collection. Moreover, as part of the standard of care at the time, all patients sustaining a SCI received methylprednisolone, a corticosteroid, to reduce inflammation and secondary damage [133, 134]. Corticosteroids have been reported to alter the concentration of certain serological markers, including bilirubin, albumin, and leukocytes [135, 136, 137]. Patients enrolled in the Sygen trial exhibited reduced bilirubin levels and leukocytosis (i.e., an increase in the number of white cells in the blood) compared with the patients in the Murnau study, who did not receive acute treatment with methylprednisolone. Moreover, the time points of blood draw could have contributed to the differences observed. Whereas the Sygen trial collected blood samples at pre-defined time points, the patients in the observational Murnau study were subject to blood draws when indicated by the treating physician. Lastly, it is well known that organ function declines with age and is correlated with changes in laboratory values. A larger proportion of elderly patients was enrolled in the Murnau study (Figure S1), which could have contributed to the divergent findings [138, 139].

2.5.4 Limitations

The primary limitation of the current study is that we utilized nearly 20-year-old retrospective data, collected in clinical trial conditions, which might compromise the translation of our results to the current clinical context. We partially address this limitation by prospectively collecting contemporary data in the framework of the Murnau study. Potential bias introduced by changes in standards of care over the last decades can be, at least in part, mitigated. However, time points of data collection were not standardized in the Murnau study. As a consequence, the time-varying sample size complicated the analyses. For example, the chosen cutoff of 50 patients for the analyses was largely driven by the sample size. Future studies with larger and more consistent sample sizes at each time point of data collection are warranted to validate our findings and provide the optimal cutoff values in a data-driven fashion. The small sample size further prevented a meaningful subgroup analysis stratified by sex and age, considering that many serological markers have different normal ranges for women and men as well as being subject to age-related changes. It should also be noted that excluding patients because

of missing AIS grade (e.g., because the patient was unconscious at baseline) represents a loss of information and introduces a potential bias toward patients with slightly less severe injuries. Studies with large sample sizes at baseline and follow-up time points are warranted to address this in further detail. Additionally, our study is focused on correlations at the population level, which does not guarantee the translation of our findings at the individual level. Further investigations are needed to assess the potential of serological markers in individual recovery prediction. Moreover, we did not account for any of the medications that were administered to the patients to treat secondary complications associated with SCI [140, 141]. Some medications (e.g., corticosteroids and nonsteroidal anti-inflammatory (NSAID) medication) can affect the concentration of the serological markers. Future studies should also address the impact of medication on the serological markers, particularly in the acute phase of injury.

2.6 Conclusion

To our best knowledge, this is the first study to comprehensively investigate the natural progression of serological markers in patients with a traumatic SCI. As a consequence of the sustained trauma, numerous routinely collected serological markers are altered in their concentration. The majority of these markers return to a normal range after 6–12 months post-injury. The current study provides a first step toward establishing a benchmark for serological markers and their natural course, which can inform clinical decision making and prospective clinical trials. Our online surveillance platform (Haemosurveillance) provides a tool for the SCI community, researchers, authorities, and policy-makers to interactively exploit the natural progression of serological markers and compare different data sets with each other. The platform is configured such that existing or newly generated data sets can be added if they comply with EU GDPR.

Acknowledgments

The authors acknowledge the participating centers in the EMSCI network [7] that were involved in the patient care and collection of data necessary for this study.

Authors' Contributions

Lucie Bourguignon: data cleaning, data analyses, interpretation of data, and drafting the manuscript

[7] http://emsci.org/members

Anh Kho Vo: data analyses, interpretation of data, and revising the manuscript for intellectual content
Bobo Tong: data cleaning, interpretation of data, and revising the manuscript for intellectual content
Fred Geisler: primary data collection, interpretation of data, and revising the manuscript for intellectual content
Orpheus Mach: primary data collection, and revising the manuscript for intellectual content
Doris Maier: primary data collection, and revising the manuscript for intellectual content
John L. K. Kramer: study concept/design, interpretation of data, and revising the manuscript for intellectual content
Lukas Grassner: primary data collection, interpretation of data, and revising the manuscript for intellectual content
Catherine R. Jutzeler: data entry, data cleaning, data analyses, interpretation of data, and drafting the manuscript
Statistical analyses were completed by Lucie Bourguignon and Catherine R. Jutzeler (Swiss Federal Institute of Technology, ETH Zurich).

Funding Information

This study was funded by research grants from the Swiss National Science Foundation (Ambizione Grant #PZ00P3_186101, Dr. Jutzeler) and Wings for Life Research Foundation (#2017_044, Drs. Jutzeler and Kramer).

Chapter 3

Pharmacological management of acute spinal cord injury: A longitudinal multi-cohort observational study

---

Adapted from:

Catherine R. Jutzeler†[1], **Lucie Bourguignon†**, Bobo Tong, Elias Ronca, Eric Bailey, Noam Y. Harel, Fred Geisler, Adam R. Ferguson, Brian K. Kwon, Jacquelyn J. Cragg‡, Lukas Grassner‡and John L. K. Kramer‡[2]

Scientific Reports, 2023

Publication: 10.1038/s41598-023-31773-8
RxSCI web site: https://jutzelec.shinyapps.io/RxSCI
GitHub repository: https://github.com/jutzca/Acute-Pharmacological-Treatment-in-SCI

Lucie Bourguignon primarily contributed to building the RxSCI web site, on an original idea from Catherine R. Jutzeler; and secondarily contributed to interpretation of the data and drafting of the manuscript. Further details can be found in **Section 3.6**.

---

[1] †indicates shared first authorship
[2] ‡indicates shared last authorship

3.1 Abstract

Multiple types and classes of medications are administered in the acute management of traumatic spinal cord injury (SCI). Prior clinical studies and evidence from animal models suggest that several of these medications could modify (i.e., enhance or impede) neurological recovery. We aimed to systematically determine the types of medications commonly administered, alone or in combination, in the transition from acute to subacute SCI. For that purpose, type, class, dosage, timing, and reason for administration were extracted from two large SCI datasets. Descriptive statistics were used to describe the medications administered within the first 60 days after SCI. Across 2040 individuals with SCI, 775 unique medications were administered within the two months after injury. On average, patients enrolled in a clinical trial were administered 9.9 ± 4.9 (range 0–34), 14.3 ± 6.3 (range 1–40), 18.6 ± 8.2 (range 0–58), and 21.5 ± 9.7 (range 0–59) medications within the first 7, 14, 30, and 60 days post-injury, respectively. Those enrolled in an observational study were administered on average 1.7 ± 1.7 (range 0–11), 3.7 ± 3.7 (range 0–24), 8.5 ± 6.3 (range 0–42), and 13.5 ± 8.3 (range 0–52) medications within the first 7, 14, 30, and 60 days post-injury, respectively. Polypharmacy was commonplace (up to 43 medications per day per patient). Approximately 10% of medications were administered acutely as prophylaxis (e.g., against the development of pain or infections). To our knowledge, this was the first time acute pharmacological practices have been comprehensively examined after SCI. Our study revealed a high degree of polypharmacy in the acute stages of SCI, raising the potential to impact neurological recovery. All results can be interactively explored on the R_XSCI web site [3] and GitHub repository [4].

[3] https://jutzelec.shinyapps.io/RxSCI/
[4] https://github.com/jutzca/Acute-Pharmacological-Treatment-in-SCI/

3.2 Introduction

Traumatic spinal cord injury (SCI) is a neurological condition associated with varying degrees of motor, sensory and autonomic deficits. At present, there are no pharmacological interventions available to enhance the extent a person neurologically or functionally recovers from acute SCI [117, 142]. In the absence of interventions that enhance neurological recovery, acute care of SCI chiefly focuses on managing neurological sequela (e.g., neuropathic pain) and secondary complications (e.g., infections). As SCI ultimately affects every organ system of the human body, a multidisciplinary treatment strategy is necessary. In accordance with existing treatment guidelines, these necessitate the administration of various drugs, including narcotics, analgesics, sympathomimetics, antibiotics, muscarinic antagonists, antithrombotics, anticonvulsants, and antidepressants to manage pain, infections, urinary tract dysfunction, deep venous thrombosis, and psychological disorders. To date, little is known to what degree common drugs used in the management of acute SCI have downstream and potentially unintended effects, which modify neurological recovery. This is surprising in light of the fact that numerous drugs: (i) are spinal cord blood barrier (SCBB) permeable and/or gain access to the central nervous system via a leaky SCBB after injury, (ii) act on targets in the central nervous system, and (iii) are administered during the window of opportunity to promote neural repair and plasticity (i.e., in the initial hours to weeks post injury).

Recent observational studies have reported a potential beneficial effect of acutely administered gabapentionoid medications (but not other anticonvulsants) on long-term neurological outcomes after SCI [103, 143, 104]. Subsequent preclinical studies demonstrated a potential gabapentionoids-meditated mechanism for enhanced recovery, as well as confirmed behavioral benefits in animal models [144, 145]. While efficacy awaits confirmation in prospective clinical trials, these collective observations point to the promise of a reverse translational approach (bedside-to-bench) to restore neurological function after SCI. Identifying other opportunities for drug repurposing depends, in part, on knowledge regarding specific medications commonly administered in the acute phase. Additionally, if promising pharmacologic agents are to be proposed for human evaluation in clinical trials of acute SCI, it is important to consider the spectrum of other concomitant medications that are routinely administered in the care of these patients, as they may have known interactions with the promising agent in question.

The aim of this study was to characterize what constitutes the "acute pharmacological management of SCI" leveraging available clinical trial and observational study data. Specifically, we determined the types of timing, and reason of administration for drugs commonly administered, alone or in combination, in the acute to subacute phase (i.e.,

first 2 months) of SCI.

3.3 Methods

3.3.1 Study design

The design and reporting of this analysis adhered to the relevant guidelines for observational studies [146].

3.3.2 Data source and cohort definition

To quantify medications commonly administered in the acute management of SCI, we analyzed two sources of data. Both sources represent collections of data from the United States of America (USA); the first (i.e., trial) between 1992 and 1998 and the second (i.e., observational) from 2007 to 2009.

The first source comprised details of concomitant medications administered in a clinical trial—the Sygen trial—delivering gangliosidosis 1 (GM-1) in acute SCI [117, 34]. The Sygen trial was a randomized, prospective, phase III, placebo controlled, multi-center study testing the efficacy of GM-1 therapy in acute, traumatic SCI [117, 34]. Full design, recruitment, and enrollment details have been published previously [117]. Briefly, to be included in the Sygen trial patients were required to have at least one lower extremity with a substantial motor deficit. Patients with spinal cord transection or penetration, head trauma, major chest trauma, or intubation were excluded, as were patients with a cauda equina, brachial or lumbosacral plexus, or peripheral nerve injury. Multiple trauma cases were included as long as they were not so severe as to preclude neurologic evaluation. Patients were also excluded when they suffered from significant systemic disease such as lung, liver, gastrointestinal, or kidney disease; or active malignancy or any other condition as determined by history or laboratory investigation that could alter the distribution, accumulation, metabolism, or excretion of the study medication, cause a neurologic deficit, or result in the patient's life expectancy being less than 2 years. The full list of inclusion and exclusion criteria can be found elsewhere [117]. All patients were to receive the national acute spinal cord injury study (NASCIS) II dose regimen of methylprednisolone sodium succinate (MPSS) starting within eight hours after the SCI. To avoid any possible untoward interaction between MPSS and Sygen® [134], the study medication was not started until after completion of MPSS administration. With 797 enrolled patients followed over the first year following injury, the Sygen trial was the largest clinical trial ever conducted in the field of SCI. The Sygen trial, which followed patients over the first year following injury, was clinically active from 1992 to 1998, and showed no differences between treatment and placebo groups in terms of neurological recovery [33]. The negative finding of the Sygen study is considered Class

I Medical Evidence by the SCI Committee of the american association of neurological surgeons (AANS) and the congress of neurological surgeons (CNS) [147, 148]. Subsequent analyses of the Sygen data have been performed to characterize the trajectory and extent of spontaneous recovery from acute SCI [149, 150].

The second source of data was from a large, observational study (i.e., SCIRehab), which abstracted information pertaining to medication use in the acute phase of SCI from patient medical records [46]. The SCIRehab study enrolled, upon consent, individuals aged ≥ 12 years with traumatic SCI who were rehabilitated at six participating rehabilitation centers from 2007 through 2009 [47]. Participating centers included Rocky Mountain Regional Spinal Injury System at Craig Hospital, Shepherd Center, Atlanta GA; Rehabilitation Institute of Chicago, Chicago, IL; Carolinas Rehabilitation, Charlotte, NC; the Mount Sinai Medical Center, New York, NY; and National Rehabilitation Hospital, Washington, DC. Patients were followed for the first-year post-injury and were excluded if they spent two or more weeks at a non-participating rehabilitation center. Details of more than 460,000 interventions provided to 1500 patients were documented by over 1000 clinicians at the six participating centers. Patient demographics and injury characteristics were extracted from the patient medical record (part of the National Institute on Disability and Rehabilitation Research Spinal Cord Injury Model Systems Form I). Design, recruitment, inclusion criteria, and enrollment details have been previously described in detail [47].

To be included in our study, information on medications administered needed to be available for the patients.

3.3.3 Commonly administered medications

In the Sygen trial, alongside serious adverse events, concomitant medication information was routinely tracked following standardized case report forms by trained examiners in clinical trials as a measure of safety. For each concomitant medication administered during the trial, the reason for administration, dosage, dosing (i.e., start and end date, frequency), and reason for conclusion were recorded. It was also documented in case medications were administered for prophylactic reasons (e.g., to prevent deep vein thrombosis). Note that, although patients were randomized to GM-1 therapy, individuals were not randomized to any concomitant medication administered and were managed according to the conventional care protocols of the enrolling center. The SCIRehab study documented the use of all commonly administered medications. For each medication administered, route, dosage, and dosing (i.e., start and end date, frequency) were abstracted directly from medical records. However, medication indication was not recorded.

3.3.4 Medication data cleaning and organizing

Medication data from the Sygen trial and SCIRehab study were separately cleaned and organized. From the medication files, which exist for each patient in the Sygen trial and SCIRehab, we extracted generic medication name and information on dosing (i.e., start and end date, frequency). As information on medication indication (i.e., reasons for administering a medication) was not entered in a standardized fashion during data collection, we classified the medication indication according to the common terminology criteria for adverse events (CTCAE) [151]. Briefly, each indication was assigned to a system organ class (SOC), the highest level of the medical dictionary for regulatory activities (MedDRA) hierarchy [5]. The SOC is identified by anatomical or physiological system, etiology, or purpose (e.g., SOC Investigations for laboratory test results) and comprises 26 different categories. We added a separate class for trauma-related pain (i.e., nociceptive and neuropathic). The rationale for this amendment stems from the fact that the CTCAE does not sufficiently cover this category. After carefully reviewing the medication list, we have also consulted study clinicians of both data sources to identify any discrepancies, including missing or duplicate medications, changes in dosages, and drug interactions (i.e., medication reconciliation).

3.3.5 Assessment of blood brain barrier (BBB) permeability

Leveraging the information from the DrugBank database [6], the permeability of medications to cross the blood brain barrier (BBB) was determined. In case corresponding information was missing in the DrugBank, a PubMed search was performed to consider studies that have evaluated BBB permeability.

3.3.6 Statistical analysis and data visualization

R Statistical Software version 3.6.3 (Running under: macOS Mojave 10.14.4) was used for all analyses and to visualize the results. Descriptive statistics (mean, standard deviation, ranges, and proportions) were used to describe the patients' demographics, injury characteristics, and medication information. For the latter, this included the number and type of medications administered, reason for administration, and how many medications each patient received per day (i.e., point prevalence). Type and frequency of medications that were administered prophylactically were also computed.

[5] medical dictionary for regulatory activities (MedDRA)® terminology is the international medical terminology developed under the auspices of the international council for harmonisation of technical requirements for pharmaceuticals for human use (ICH)

[6] www.drugbank.ca

3.3.7 Interactive web platform R_XSCI

In order to enable the SCI community, researchers, authorities, and policymakers to fully explore the data and results of this study (and beyond), we developed the freely available and open source R_XSCI web platform. R_XSCI was implemented with the Shiny framework [90], which combines the computational power of the free statistical software R [152] with friendly and interactive web interfaces. Both, the front- and back-end of R_XSCI have been built using the shiny dashboard package [91]. R_XSCI is available as an online application and is hosted at https://jutzelec.shinyapps.io/RxSCI/ and can be accessed via any web browser on any device (e.g., desktop computers, laptops, tablets, smartphones). R_XSCI is published under the BSD 3-Clause License. The source code of R_XSCI is available through Github at https://github.com/jutzca/Acute-Pharmacological-Treatment-in-SCI/tree/master/shinyapp.

3.3.8 Data sharing and code availability

Full anonymized data of both data sources will be shared at the request from any qualified investigator (please contact CRJ). The code for the data analysis and visualization is available in our GitHub repository [7].

3.3.9 Standard protocol approvals, registrations, and patient consents

Approval for this study (secondary analysis) was received by an institutional ethical standards committee on human experimentation at the University of British Columbia. The original Sygen clinical trial (results published elsewhere) also received ethical approval, but was conducted before clinical trials were required to be registered (i.e., no clinicaltrial.gov identifier available) [33]. Each participating center of the SCIRehab study received institutional review board approval for this study and obtained informed consent from each patient (or their parent/guardian).

[7] https://github.com/jutzca/Acute-Pharmacological-Treatment-in-SCI/

3.4 Results

3.4.1 Patient characteristics and summary statistics

797 and 1243 patients from the Sygen clinical trial and SCIRehab observational study, respectively, were included in our analysis. While all patients from the Sygen study were included in our analysis, we had to exclude 257 patients from the SCIRehab study due to missing data on medications ($n = 242$) or SCI with no sensory or motor impairments (i.e., American spinal injury association (ASIA) impairment scale (AIS) E, cauda equine or peripheral nervous system injuries, $n = 15$). In both cohorts, the ratio between male and female patients was approximately 4 : 1, the majority of the patients were injured at the cervical levels (Sygen: 75.2%; SCIRehab: 60.4%), and motor complete (Sygen: 65.7%; SCIRehab: 65.6%). The most frequent cause of injury was car accidents (Sygen: 47.9%; SCIRehab: 35.5%) followed by falls (Sygen: 16.2%; SCIRehab: 24.1%). Detailed description of both cohorts is provided in **Table 3.1**.

Table 3.1: Demographics and injury characteristics of the included cohorts

	Sygen clinical trial (n = 797)	**SCIRehab study (n = 1243)**
Study details		
Study type	Prospective, double-blind, randomized, stratified, multicenter trial	Prospective observational study
Study outcome	No differences between treatment and placebo groups in terms of neurological recovery	Not applicable
Running time	1992-1998	2007-2010
Country	USA	USA
Time of enrollment	<72h	Admission to rehabilitation center (30 ± 27 days post-injury)
Follow-up	1-year post-injury	Discharge from rehabilitation center

Continued on next page

Table 3.1: Demographics and injury characteristics of the included cohorts (Continued)

	Sygen clinical trial (*n* = 797)	**SCIRehab study (*n* = 1243)**
References	[117, 34]	[46]
Sex, n(%)		
Female	153 (19.2)	231 (18.6)
Male	642 (80.6)	1012 (81.4)
Missing	2 (0.3)	Not applicable
Age or age groups (years)		
Mean (SD)	32.5 (13.4)	Not applicable
Median [Min, Max]	30.0 [11.0, 69.0]	Not applicable
Missing	2 (0.3%)	Not applicable
12-19	150 (18.8%)	183 (14.7%)
20-29	236 (29.6%)	340 (27.4%)
30-39	194 (24.3%)	190 (15.3%)
40-49	118 (14.8%)	201 (16.2%)
50-59	55 (6.9%)	165 (13.3%)
60-69	44 (5.5%)	106 (8.5%)
70-79	Not applicable	45 (3.6%)
80+	Not applicable	13 (1.0%)
AIS grade, n(%)		
A	446 (56.0%)	624 (50.2%)
B	77 (9.7%)	192 (15.4%)
C	149 (18.7%)	230 (18.5%)
D	31 (3.9%)	197 (15.8%)
Missing	94 (11.8%)	Not applicable
NLI, n(%)		
Cervical	599 (75.2)	751 (60.4)
Thoracic	196 (24.6)	46 (3.7)
Lumbar	Not applicable	446 (35.9)

Continued on next page

Table 3.1: Demographics and injury characteristics of the included cohorts (Continued)

	Sygen clinical trial (*n* = 797)	**SCIRehab study (*n* = 1243)**
Missing	2 (0.3)	Not applicable
Paraplegia/tetraplegia, n(%)		
Paraplegia	189 (50.9)	461 (37.1)
Tetraplegia	602 (36.5)	782 (62.9)
Unknown	2 (0.3)	Not applicable
Cause, n(%)		
Automobile	382 (47.9)	441 (35.5)
Blunt trauma	9 (1.1)	Not applicable
Fall	129 (16.2)	300 (24.1)
Gunshhot wound	36 (4.5)	125 (10.1)
Motorcycle	48 (6.0)	110 (8.8)
Sports	35 (4.4)	125 (10.1)
Others	61 (7.7)	51 (4.1)
Pedestrian	10 (1.3)	20 (1.6)
Person-to-person contact	Not applicable	10 (0.8)
Water related	85 (10.7)	61 (4.9)
Missing	2 (0.3)	Not applicable

American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description;
United States of America (USA); standard deviation (SD); neurological level of injury (NLI)

3.4.2 Acute pharmacological management after SCI

In total, 489 (trial) and 575 (observational study) unique medications were administered over the course of 60 days after SCI. More than a third ($n = 289$[~ 37.3%]) of the medications administered were common to both data sources (for details see **Supplementary Table 1**). Medications were administered to manage secondary complications arising from 21 different system organ classes or to facilitate surgical and medical procedures (**Figure 3.1A** and **Supplementary Table 2**). No medications were administered for the

following five organ systems: (i) congenital, familial and genetic disorders; (ii) injury; (iii) hepatobiliary disorders, poisoning and procedural complications; (iv) pregnancy, puerperium and perinatal conditions; and (v) social circumstances. On average, patients enrolled in the Sygen trial received 9.9 ± 4.9 (range 0–34), 14.3 ± 6.3 (range 1–40), 18.6 ± 8.2 (range 0–58), and 21.5 ± 9.7 (range 0–59) medications within the first 7, 14, 30, and 60 days post-injury, respectively (**Figure 3.1B**). Patients enrolled in the SCIRehab cohort study received on average 1.7 ± 1.7 (range 0–11), 3.7 ± 3.7 (range 0–24), 8.5 ± 6.3 (range 0–42), and 13.5 ± 8.3 (range 0–52) medications within the first 7, 14, 30, and 60 days post-injury, respectively (**Figure 3.1C**). **Supplementary Figure 1** shows the absolute and cumulative number of unique drugs per day for the Sygen (Supplementary Fig. 1A) and the SCIRehab (**Supplementary Figure 1B**). The disparity between Sygen and SCIRehab in the first month post injury can be attributed to different time-points of patient enrollment, with the Sygen trial enrolling patients within 72 h, compared to SCIRehab, which enrolled patients within days or weeks of injury (**Table 3.1**). As a result, medications for first-line trauma management (e.g., nitroglycerin, dopamine) as well as surgical and medical procedures (e.g., isoflurane, vecuronium bromide) are only captured by the Sygen trial. Acetaminophen (analgesic, $n = 674$ patients), morphine (analgesic, $n = 664$ patients), and heparin (anticoagulant, $n = 505$ patients) were the three most commonly administered medications in the Sygen trial (**Figure 3.1D**). Similarly, in the SCIRehab study, the analgesic acetaminophen ($n = 924$ patients) was the most commonly administered medication, followed by the laxative docusate ($n = 620$ patients) and the analgesic combination medicine acetaminophen and oxycodone ($n = 603$ patients) (**Figure 3.1E**).

The majority of patients enrolled in the Sygen trial required medications to treat secondary complications arising from the gastrointestinal system ($n = 752$, 95.1%), pain ($n = 742$, 93.8%), infections ($n = 737$, 93.2%), and psychiatric issues ($n = 650$, 82.2%) (**Figure 3.2A**, **Supplementary Table 3**). A total of 150, 99, and 93 unique medications were administered to treat a variety of secondary complications arising from infections, respiratory system, and gastrointestinal system, respectively. Moreover, pain (e.g., musculoskeletal), gastrointestinal complications (e.g., heartburn, ulcers), and infections (i.e., bacteria, viral, and fungal) were the most frequently managed problems (**Figure 3.2B**, **Supplementary Table 4**). This was also true when stratifying for injury severity (AIS grades, **Supplementary Table 5**). While infections were mainly treated with antibiotics, antifungal, and antiviral medications depending on their nature, complications arising from gastrointestinal tract were targeted with analgesics, antibiotics, antacids, antiulcer, anti-anemics, anticholinergics, and antispasmodics (see detailed overview in **Supplementary Table 6**).

3.4.3 Polypharmacy

As illustrated in **Figure 3.3**, polypharmacy was commonplace. Almost every patient enrolled the Sygen trial or the SCIRehab study received multiple medications per day (**Figure 3.3A**). Patients with more severe injuries (AIS A and B) received more medications per day than those with less severe injuries (AIS D). The number of medications administered per day per patient ranged between 1 and 30 for patients enrolled in Sygen trial (**Figure 3.3B**) and between 1 and 43 for patients enrolled in the SCIRehab study (**Figure 3.3B**). Individual patient examples of the extend of polypharmacy is shown in **Figure 3.3C**. The complexity of the combination of medications administered is illustrated in **Figure 3.3D**. In the Sygen trial, the three most common combinations of medications were acetaminophen and morphine (n = 164 patients), morphine and ranitidine (n = 128 patients), as well as acetaminophen and heparin (n = 123 patients). In the SCIRehab study, acetaminophen and acetaminophen oxycodone was the most common combination of medications (n = 480 patients), followed by acetaminophen and acetaminophen hydrocodone (n = 407 patients), as well as acetaminophen and ibuprofen (n = 346 patients). The complexity of the combination of medications administered to patients in the SCIRehab study is illustrated in **Figure 3.3E**.

3.4.4 BBB permeability

Out of the 775 unique medications, 59.4% (n = 460) have the ability to cross the BBB while 20.6% (n = 160) are not permeable for the BBB. No information regarding the BBB permeability was identified for the remaining 20.0% (n = 155). Detailed information on the permeability can be found in **Supplementary Table 7**.

3.4.5 Prophylactic administration of medications

Approximately 10% (n = 2838) of all recorded indications in the Sygen trial (**Figure 3.4A**) were labelled 'prophylactic' or 'preventative'. A total of 137 unique medications were administered for prophylactic treatment to prevent a wide range of secondary complications (**Figure 3.4B**). The major medication groups included antihistamines (ranitidine, famotidine), anticoagulants (heparin, warfarin), and antibiotics (cefazolin, gentamicin) for the prevention of secondary complications arising from the gastrointestinal system (e.g., heart burn, gastric ulcers), blood and vasculature system (e.g., deep vein thrombosis), and infections, respectively (**Figure 3.4C**). The majority of patient enrolled in the Sygen trial (n = 666 [83.6%]) received prophylactic treatments ($\text{mean}_{medications/patient}$ = 3 [range 1–21]; $\text{mean}_{indications/patient}$ = 4.3 [range 1–33]) (**Figure 3.4D**). **Supplementary**

Table 8 provides a comprehensive overview of all medications (and their respective indications) that were administered prophylactically.

3.4.6 Interactive web platform R_XSCI

The R_XSCI web platform is hosted online [8] and contains three main data visualization parts: (1) epidemiological features, including demographics and injury characteristic; (2) information on the pharmacological treatment of SCI patients on daily basis, including medication administration patterns; and (3) visualization of the polypharmacy. All data from the Sygen clinical trial and the SCIRehab study, which was used in this study, can be explored in a customized fashion (e.g., customized selection of patient groups). The platform is configured such that existing or newly generated data sets can be added if they comply with European general data protection regulation (EU GDPR).

[8] https://jutzelec.shinyapps.io/RxSCI/

3.5 Discussion

The aim of the current study was to comprehensively evaluate pharmacological management practices in acute SCI. To this end, two large data sources were examined, one from a clinical trial and the other from an observational study. Our analysis revealed an incredibly high rate of polypharmacy spread over the course of the first 60 days' post injury, which was administered to manage various health conditions arising directly or indirectly from acute SCI. Various medications were administered, including those that readily cross the BBB (e.g., pregabalin [153], morphine [154]) to manage the sequela of SCI (e.g., neuropathic pain), as well as other complex medical complications. Drugs that cross the BBB may be more likely to have effects (positive or negative) on neural recovery pathways after injury.

To our knowledge, this was the first time acute pharmacological practices have been comprehensively examined after SCI. Even considering its extreme and traumatic nature, the sheer number of medications administered in a short window of time after SCI, over the course of the 2 months, was remarkably high. This led to a very high degree of polypharmacy. For comparison, polypharmacy in other complex health conditions is generally considered more than five medications [155, 156] — the average for acute SCI patients was approximately double that threshold. While perhaps startling, the complexity of managing SCI requires aggressive pharmacological management. Nevertheless, the lack of attention paid to the question of "neurological safety" (i.e., whether use of a medication or its interaction with other medications in the acute phase of injury will have long-term and detrimental neurological consequences) is surprising, as is the fact that few attempts have been made to discern potential beneficial (or detrimental) effects of medications that readily cross the BBB. Furthermore, one must consider potential interactions between the high number of clinically used concomitant medications with novel medications and biologics being trialed for improving recovery from SCI.

The limited knowledge about the potential effects of acutely administered medications on recovery in humans becomes all the more curious considering that a number of these medications alter outcomes in animal studies. As an example, pregabalin, a potent calcium channel blocker and anticonvulsant administered for neuropathic pain, has been repeatedly shown to benefit recovery after SCI in animal and human SCI [103, 143, 104, 144]. Detrimental effects were also observed for some medications, including opioids, which attenuated the recovery of locomotor function and exacerbated pathophysiological processes in rodent models of SCI [157, 158, 159]. A detrimental opioid effect is in line with beneficial effects of naloxone (i.e., opioid antagonist) [157, 160], and is highly concerning in light of the fact that opioids are ubiquitously administered for pain

management in the early stages of injury (to > 80% of the patients). While completely removing or restricting opioids would be highly problematic and present with serious ethical concerns (i.e., weighing the management of acute pain with long-term neurological effects), opioids were among medications commonly administered to prevent the onset of pain. This suggests that opioids, at least in a proportion of patients, were prescribed with the intention to prevent the onset of pain, despite a lack of evidence [161]. Among these individuals, neurological recovery could perhaps be facilitated by minimizing the administration of opioids. Many other common medications (up to 10%) are prophylactically administered, including acetaminophen, cefazolin, and famotidine for pain/fever, infection, and ulcer prophylaxis, respectively.

Despite years of use in clinical routine, safety information with respect to neurological outcomes of many concomitant medications is currently not available. This is highly concerning because fundamental assumptions of pharmacokinetics and -dynamics may not apply as in other (healthy) individuals [162]. Alterations in physiology lead to prolonged absorption as a consequence of slowed gastric emptying and gastrointestinal motility [162], altered distribution due to leaky blood spinal cord barrier [163], hampered metabolism [164, 165], and slowed excretion are hallmarks of this altered physiology [162, 164, 165]. Examples of medications with changed pharmacokinetics are amikacin, baclofen, carbamazepine, cefotiam, ciprofloxacin, diazepam, diclofenac, doxycycline, ketamine, lorazepam, naproxen, and vancomycin. A major issue with these injury-induced modifications in pharmacokinetics is that some medications do not reach desired therapeutic effects, whereas others may reach potentially toxic levels. In addition to potential toxicity, also common side effects of medications (e.g., gastric emptying and gastrointestinal motility caused by opioids) may worsen the natural pathophysiology of injury. Post-marketing surveillance and risk assessment programs aim at detecting previously unrecognized positive or negative effects that may be associated with a medication—within real-world populations. To our knowledge, few of these studies have examined effects after SCI. An exception is a recent study that established neurological safety profile of baclofen, an antispasmodic to treat debilitating muscle spasms [43]. Cragg et al. performed a secondary analysis of clinical trial data to provide data reaffirming that baclofen is neurologically, hepatically, and renally safe to use in patients sustaining a SCI [43]. Complementing the existing safety profile, neurological safety medication profiles in the context of concomitant medications in real-world settings will enable health care providers to provide an informed, evidence-based response regarding the use of medications such as baclofen in the acute phase of SCI.

3.5.1 Limitations

There are multiple limitations that are noteworthy. Firstly, in this study, we compared two cohorts which were collected a decade apart. It cannot be excluded that changes in the management, in particular pharmacological management, of SCI occurred over this period. However, it has been shown that the recovery rate did not change [166]. Thus, we can hypothesize that the potential changes in the standard of care did not significantly improve or deteriorate the recovery of the SCI itself. Secondly, all medications administered after SCI were meticulously tracked in the Sygen trial. However, there is no information on medications prescribed prior to the injury. Thirdly, the two studies involve dissimilar populations of people with acute SCI and data from two drastically different periods (1992 versus 2007), both of which are dated. Another limitation was the differences between the two study cohorts in reporting of demographics (i.e., age, time since injury, etc.) at the time of enrollment. Thus, more contemporary studies are warranted to establish the extent to which polypharmacy during acute SCI management may have changed within the last 30 years. Lastly, there might be potential confounding factors that may undermine the legitimacy of the data used in this study, including comorbidities, patient characteristics (age, sex, race, or genetics), concomitant diseases or conditions, non-adherence of patients, variance in physician prescribing practices, timing and duration of concomitant medication use, and dosage and potency of concomitant medications. These confounding factors must be considered when analyzing the concomitant drug data of clinical trials and observational studies.

3.6 Conclusion and implications for other neurological disorders

Our study revealed a dramatic degree of polypharmacy after acute SCI that potentially impacts recovery and the potency of novel treatments of SCI. It should be noted that in the testing of novel drug agents in preclinical models of SCI, the experiments are typically designed to minimize (and of course standardize) the concomitant medications administered to the animals. How starkly different this is from clinical reality is revealed in our analysis. SCI is a complex condition and as such, the pharmacologic needs are understandably high. While we are not arguing for an arbitrary "reduction" in the use of various medications in the management of these individuals, evaluating current standards of acute care and understanding what pharmacologic agents patients are typically exposed to does represent an intriguing alternative strategy to improve the lives of individuals with SCI. Knowledge gained from our study has major implications

for other diseases hallmarked by polypharmacy, including Parkinson's disease [167], Alzheimer's disease [168], Multiple Sclerosis [169], traumatic brain injury [170, 171], cancer [172], and sepsis [173]. Similar to SCI, these diseases are complex conditions associated with a wide range of symptoms (e.g., functional impairment) and secondary complications (e.g., gastrointestinal and cardiovascular complications, pain) necessitating pharmacological treatment — at times simultaneously. Many of these diseases are not yet curable, but effective disease modifying treatments that relieve symptoms, slow down disease progression, and improve quality of life are available [174, 175, 176, 177]. A cursory glance at the literature corroborates that the knowledge gap regarding the effect of commonly used medications on disease progression and their potential to alter the effectiveness of disease modifying treatments is not unique to SCI.

Data availability

Fully anonymized data of both data sources will be shared at the request from any qualified investigator (please contact CRJ). The code for the data analysis and visualization is available in our GitHub repository [9].

Acknowledgements

The authors would like to acknowledge the participating centers in the Sygen trial and SCIRehab network that were involved in the patient care and collection of data necessary for this study.

Author contributions

C.R.J.: Study concept/design, data entry, data cleaning, data analyses, interpretation of data, and drafting the manuscript
L.B.: Data visualisation, interpretation of data and revising the manuscript for intellectual content
B.T.: Data cleaning, interpretation of data and revising the manuscript for intellectual content
E.R.: Interpretation of data and revising the manuscript for intellectual content
E.B.: Primary data collection, interpretation of data, and revising the manuscript for intellectual content
N.Y.H.: Data cleaning, interpretation of data, and revising the manuscript for intellectual

[9] https://github.com/jutzca/Acute-Pharmacological-Treatment-in-SCI/

content
F.G.: Data cleaning, interpretation of data, and revising the manuscript for intellectual content
K.B.: Interpretation of data, revising the manuscript for intellectual content
A.R.F.: Interpretation of data, revising the manuscript for intellectual content
B.K.K.: Interpretation of data, revising the manuscript for intellectual content
J.J.C.: Interpretation of data, revising the manuscript for intellectual content
L.G.: Study concept/design, interpretation of data, revising the manuscript for intellectual content
J.L.K.K.: study concept/design, interpretation of data, and revising the manuscript for intellectual content

Funding

This study was funded by research grants from the Swiss National Science Foundation (Ambizione Grant #PZ00P3_186101, Jutzeler), Wings for Life Research Foundation (#2017_044, Jutzeler and Kramer), and Craig H. Nielsen Foundation (Cragg and Jutzeler).

Competing interests

CR. Jutzeler reports no disclosures relevant to the manuscript. L. Bourguignon reports no disclosures relevant to the manuscript. B. Tong reports no disclosures. E. Ronca reports no disclosures relevant to the manuscript. E. Bailey reports no disclosures relevant to the manuscript. NY. Harel reports no disclosures relevant to the manuscript. F. Geisler reports no disclosures relevant to the manuscript. AR. Ferguson reports no disclosures relevant to the manuscript. BK. Kwon reports no disclosures relevant to the manuscript. JJ. Cragg reports no disclosures relevant to the manuscript. L. Grassner reports no disclosures relevant to the manuscript. JLK. Kramer reports no disclosures relevant to the manuscript.

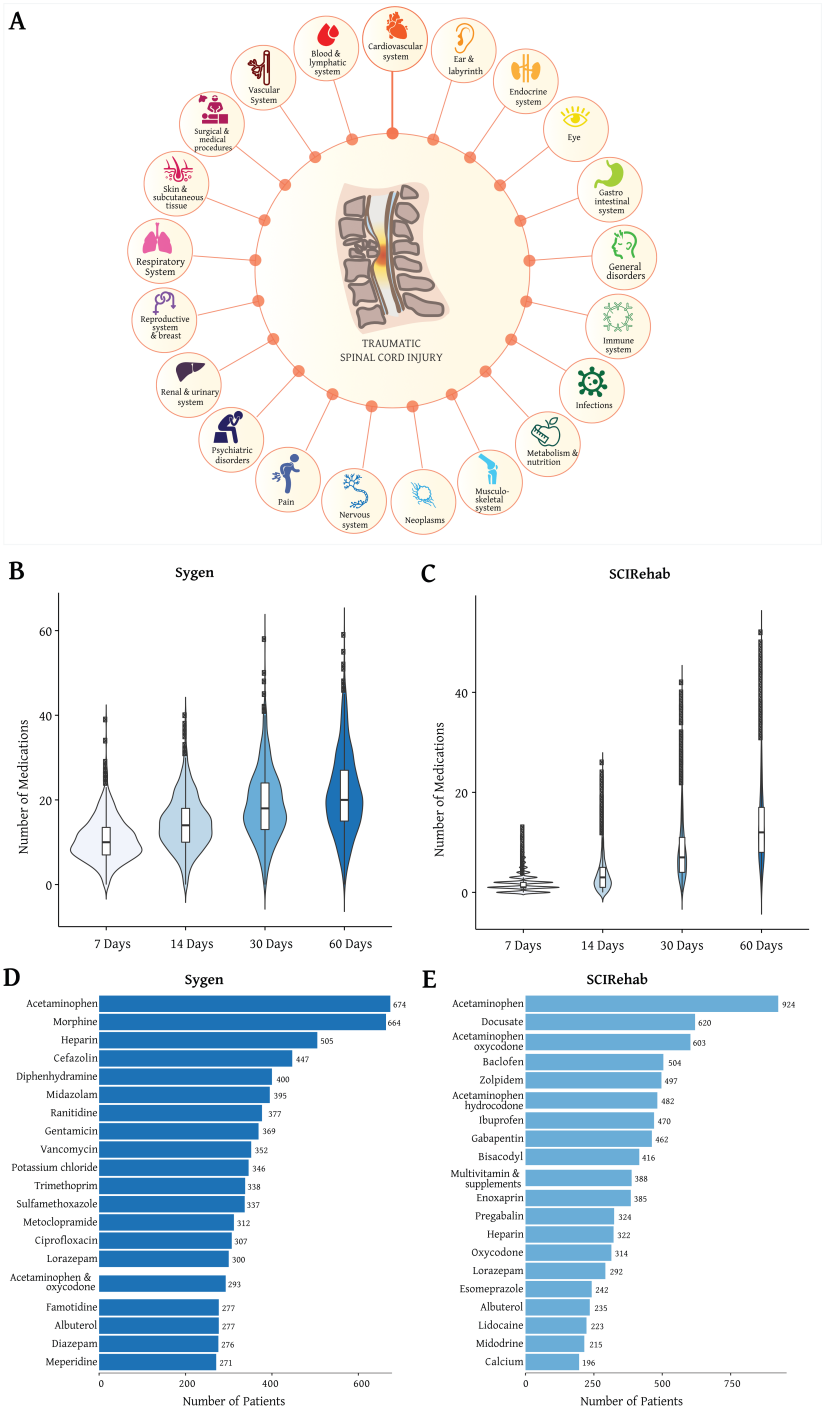


Figure 3.1: Pharmacological management of acute spinal cord injury (SCI). A. Secondary complications. Spinal cord injury is associated with a large number of secondary complications that arise from 20 organ systems as defined by common terminology criteria for adverse events (CTCAE) published by the U.S. Department of Health and Human Services [151]. Many medications were also administered to facilitate medical and surgical procedures, such as decompression surgeries, laminectomy, and computer tomography scans. **B.** Number of medications administered to patients enrolled in the Sygen trial within the first 7, 14, 30, and 60 days post-injury. **C.** Number of medications administered to patients enrolled in the SCIRehab study within the first 7, 14, 30, and 60 days post-injury. **D.** Frequency of medications administered. The majority of patients enrolled in the Sygen trial received acetaminophen, morphine, and heparin to treat secondary complications, such as pain and deep venous thrombosis. **E.** Frequency of medications administered. Pain killers (acetaminophen and acetaminophen oxycodone) as well as the laxative docusate were among the most frequently administered medications in the SCIRehab study.

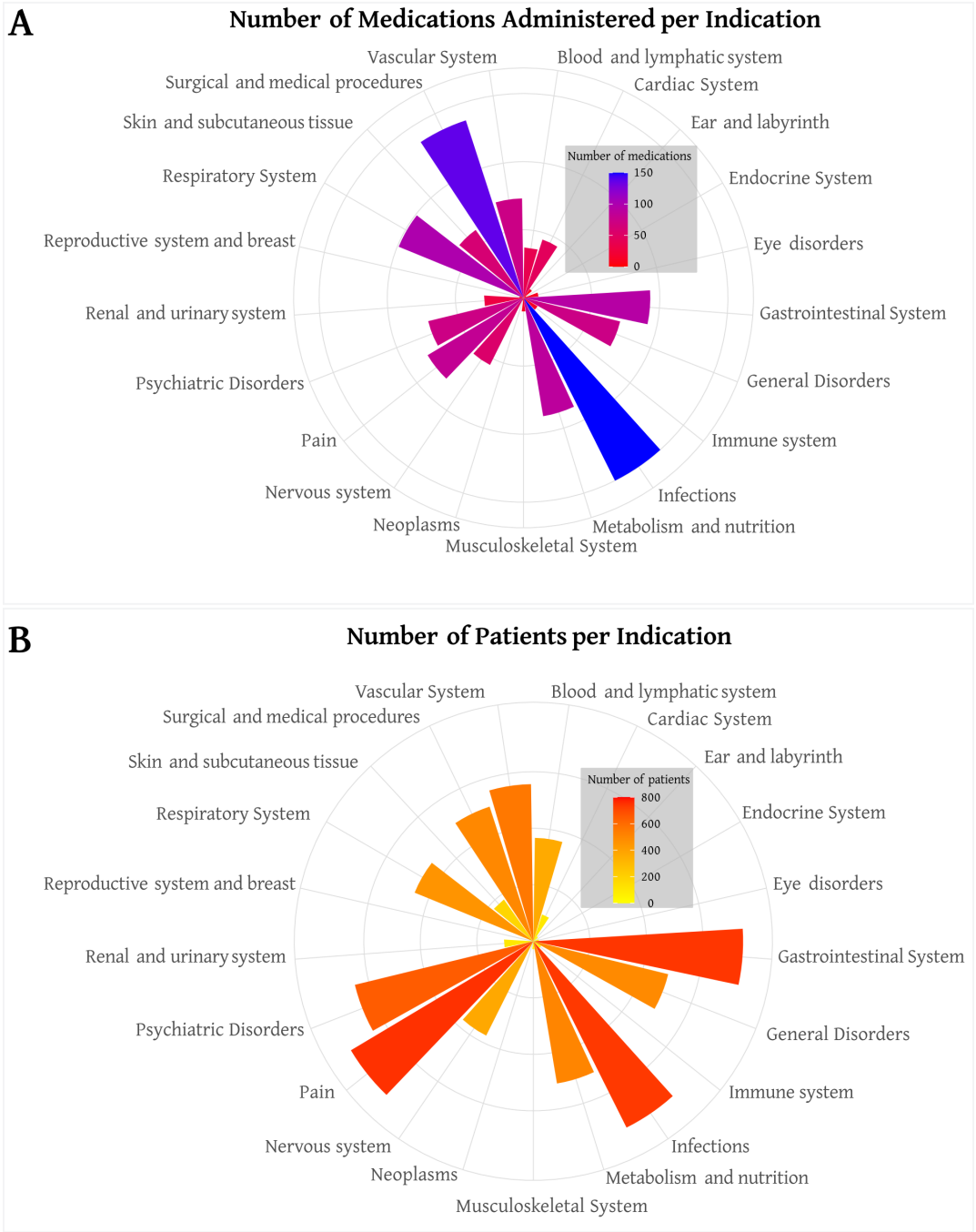


Figure 3.2: Indication of medications administered. A. Number of unique medications administered per organ system for patients enrolled in the Sygen clinical trial. Note the diversity of medications administered within each category of complications. For instance, over 100 different medications were administered to treat infections and infestations as well as for surgical and medical procedures. **B.** Number of patients of the Sygen clinical trial that required treatment per organ system. The three most frequently treated secondary complications were pain, gastro-intestinal system disorders, as well as infections. The SCIRehab database did not track the indications for which medications were prescribed.

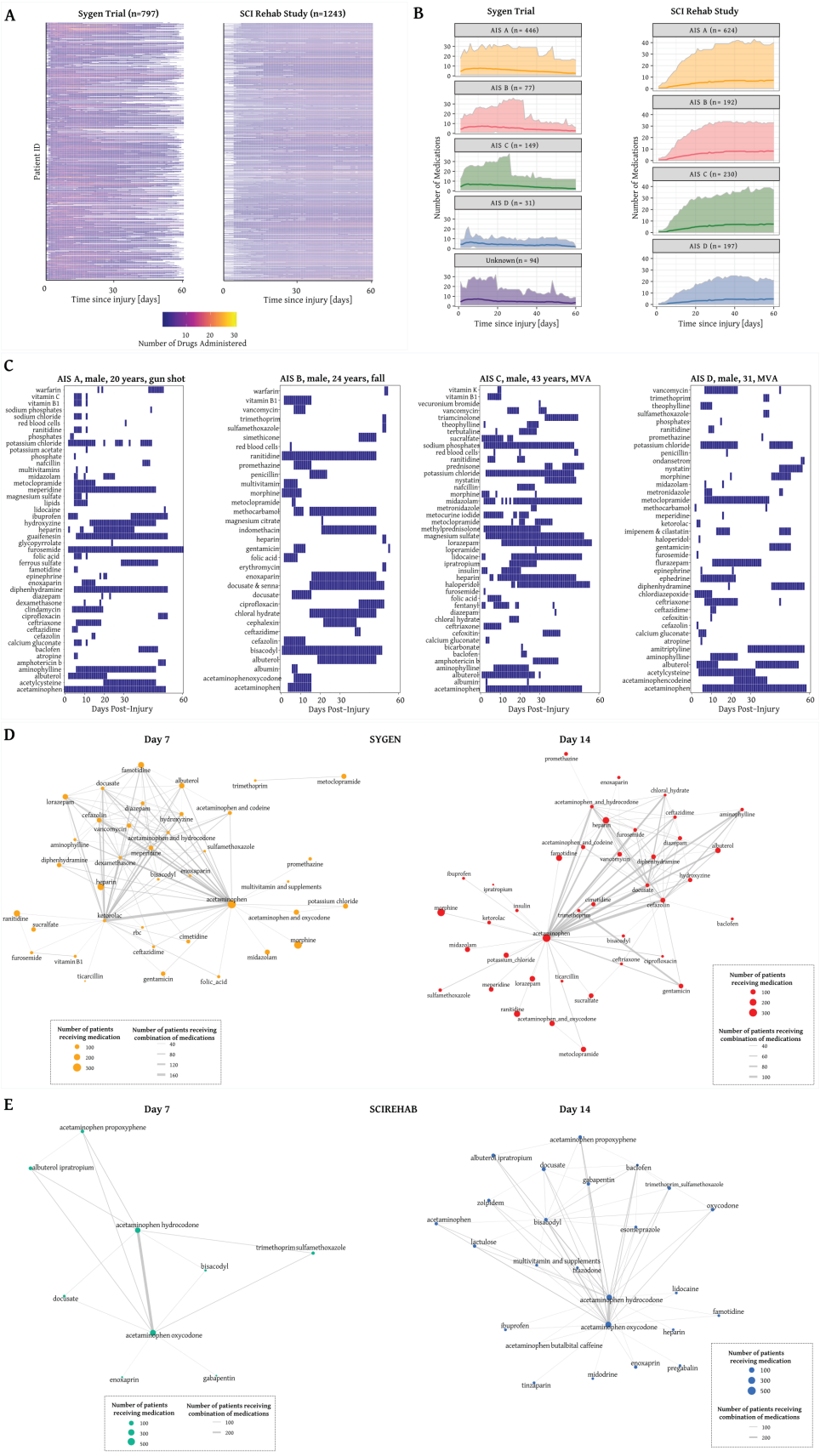


Figure 3.3: Polypharmacy. Caption continues on the next page.

Figure 3.3: (Previous page.) **A.** Point prevalence of commonly administered medications. The number of medications administered per day per patient in the first 60 days post injury varied between 1 and 30 for the clinical trial and between 1 and 43 in the observational study. Each line represents one patient and the color white indicates that no medication was administered or no data was available for that time period. **B.** Daily average number of medications administered. Patients with motor complete injuries (AIS A and B) received on average more medications per day compared to patients with motor incomplete injuries. The range medications administered varies quite drastically. The dashed line denotes the average number of medications and the solid lines the minimum and maximum number of medications, respectively. Patients with no information on AIS grades at baseline were grouped together in the category 'unknown'. **C.** Examples longitudinal medication profiles for four patients in the first 60 days post injury. Polypharmacy was commonplace across different injury severities and aetiologies. The pattern of medication administration varied between continuous, intermittent, and single-use indications. Medications were often co-administered bearing a high risk of pharmacological interactions between medications. While some are well-understood, the majority of these interactions (particularly combinations of three and more medications) have not yet been explored. **D.** Network of medications administered in combination to patients enrolled in the Sygen trial. The nodes of the network represent the medications. The size of the nodes represents the number of patients that have received this particular medication on day 7 or 14, respectively. Medications that were administered together on a specific day, either 7 or 14, are connected via an edge. The width of the edge represents the number of patients that have received the two medications (acetaminophen and ketorolac) in combination on the day of interest. **E.** Network of medications administered in combination to patients enrolled in the SCIRehab study. The nodes of the network represent the medications. The size of the nodes represents the number of patients that have received this particular medication on day 7 or 14, respectively.

.

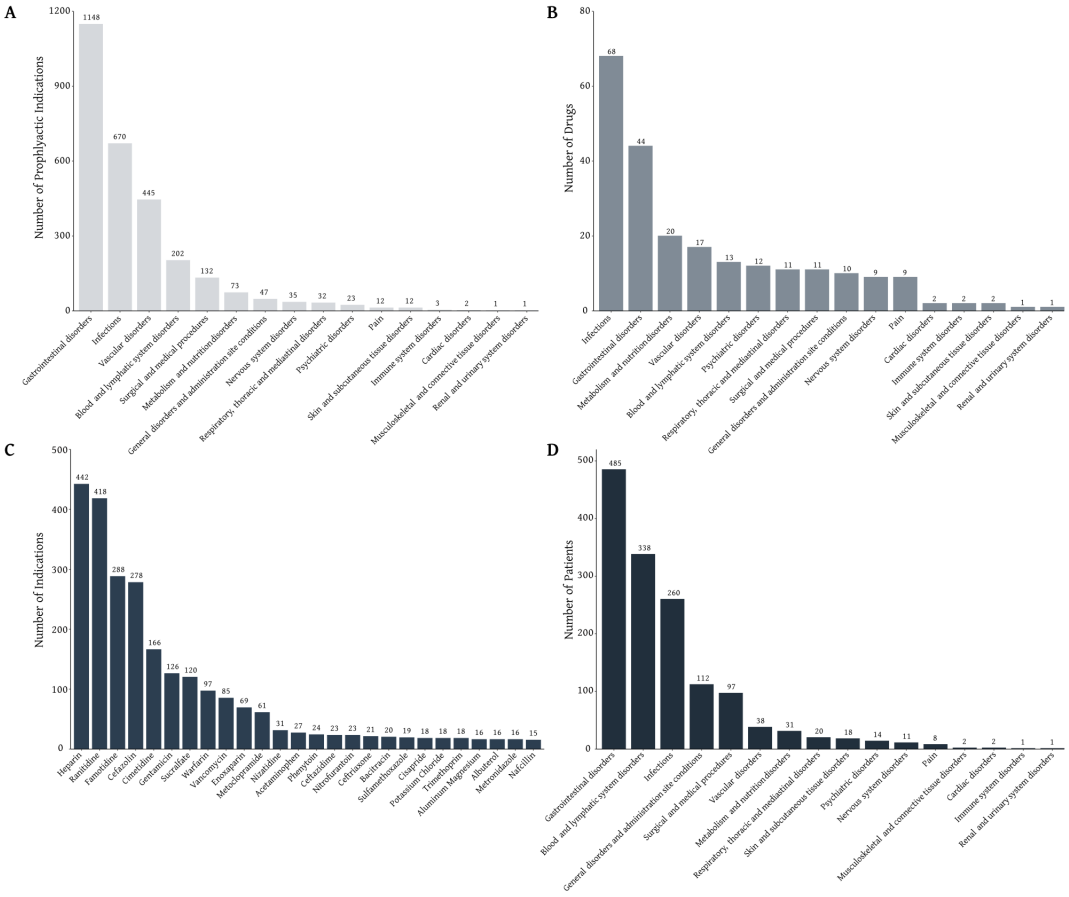


Figure 3.4: Prophylactic pharmacological treatment to prevent secondary complications from occurring. A. Number of indications per organ system. The majority of prophylactic indications were related to the gastrointestinal and vascular system as well as infections of all sorts. **B.** Number of unique medications administered for disease prophylaxis. **C.** Number of indications per medications. Anticoagulants, antihistamines, and antibiotics were amongst the most frequently administered medication classes. **D.** Number of patients that received prophylactic treatment per organ system. The majority of the patients enrolled in the Sygen trial ($n = 666$ [83.6%]) received at least one medication for disease prophylaxis. The average number of medications per patient was 3 (range 1–21) and average number of indications per patient was 4.3 (1–33).

.

Chapter 4

Do commonly administered drugs inadvertently modify the progression of spinal cord injury? A systematic review

---

Adapted from:

Lucie Bourguignon†, Louis P. Lukas†, Bethany R. Kondiles, Bobo Tong, Jaimie J. Lee, Tomás Gomes, Wolfram Tetzlaff, John L. K. Kramer, Matthias Walter, Catherine R. Jutzeler

Submitted to *Communications Medicine*, 2023

Publication:
GitHub repository: https://gitlab.ethz.ch/BMDSlab/publications/SCI-drug-review-publication

Lucie Bourguignon & Louis P. Lukas jointly surveyed the literature, extracted, analyzed, interpreted, and visualized the data, and wrote the manuscript, on an original idea and with contributions from Catherine R. Jutzeler. Further details can be found in **Section 4.5**.

4.1 Abstract

Background and objectives

Complications arising from acute traumatic spinal cord injury (SCI) are routinely managed by various pharmacological interventions. Despite decades of clinical application, the potential impact on neurological recovery has been largely overlooked. This study aims to highlight commonly administered drugs with potential disease modifying effects.

Methods

This systematic literature review included studies referenced in PubMed, Scopus and Web of Science from inception to March 31st, 2021, and assessing disease-modifying properties on neurological and/or functional recovery of drugs routinely administered following SCI. Drug effects were classified as positive, negative, mixed, no effect, not (statistically) reported. Due to the vast heterogeneity in study protocols and reporting standards, a meta analysis could not be conducted. Risk of bias (RoB) was assessed separately for animal, randomized clinical trials and observational human studies.

Results

Our literature review revealed 394 studies conducting 486 experiments that evaluated 144 unique or combinations of drugs. 195 of the 464 experiments conducted on animals (42%) and one study in humans demonstrate positive disease-modifying properties on neurological andor functional outcomes. Methylprednisolone, melatonin, estradiol and atorvastatin were the most common drugs associated with positive effects. Two studies on morphine and ethanol reported negative effects on recovery compared to control.

Discussion

Despite a large heterogeneity observed in study protocols, research from bed to bench and back to bedside provides an alternative approach to identify new candidate drugs in the context of SCI. Future research in human populations is warranted to determine if introducing drugs like melatonin, estradiol or atorvastatin would contribute to enhancing neurological outcomes after acute SCI.

Trial registration information

The study protocol was registered on PROSPERO (CRD42021231851).

4.2 Introduction

Spinal cord injury (SCI) is a devastating condition that often leads to severe and permanent neurological and functional impairments. Despite recent advancements, effective treatments promoting neurological and functional recovery are urgently needed [178, 179]. Over the last decades, interest in exploring the disease-modifying effects of commonly administered drugs in this context has grown [104, 43, 143, 103]. Nearly every individual sustaining a traumatic SCI receives multiple types and classes of drugs to manage a wide range of secondary complications associated with the neurotrauma [180, 181, 182]. These range from drugs to manage blood pressure, to analgesics for concomitant traumatic injuries, to anticholinergics for spasms. A recent study showed that patients receive up to 60 unique drugs within the first two months, often in combinatorial fashion [180]. Despite extensive polypharmacy, little is known to what degree common drugs used in the management of acute SCI have downstream, unintended, beneficial or detrimental, effects on neurological and functional outcomes.

The acute phase of SCI represents a crucial window of opportunity for therapeutic intervention. Consequently, understanding the potential therapeutic benefits of routinely administered drugs on neurological and functional recovery is paramount in the development of effective treatment strategies for SCI. The detrimental effects of SCI extend beyond the initial damage, as a cascade of secondary injury processes like inflammation, oxidative stress, excitotoxicity, and apoptosis is triggered further compromising neural tissue and impeding recovery. Identifying drugs that can modify these secondary injury mechanisms while promoting neural repair and regeneration presents a promising avenue of research. Commonly administered drugs, already approved for various medical conditions, offer the advantage of established safety profiles and known pharmacokinetics. These drugs have been extensively studied in their primary therapeutic indications, but emerging evidence suggests that some possess additional neuroprotective, neuroregenerative, or anti-inflammatory properties potentially promoting recovery after SCI [103, 183]. Disease-modifying effects of these drugs can be multifaceted. Some drugs may act directly on the injured spinal cord by reducing inflammation [184], inhibiting cell death pathways [185], or promoting axonal regeneration [186]. Others may exert their effects indirectly by modulating the surrounding environment, such as promoting angiogenesis or altering the immune response [187, 188] to create a more conducive environment for neural repair.

To bridge this knowledge gap, we conducted a comprehensive systematic review of preclinical and clinical studies examining the effects of commonly administered drugs on functional and neurological recovery following SCI. Our study aimed to provide a

thorough synthesis of the existing literature and identify potential therapeutic agents that could improve outcomes in individuals with SCI.

4.3 Methods

The study protocol was registered with and approved by the international prospective register of systematic reviews (PROSPERO) (registration number: CRD42021231851). This review conforms to the preferred reporting items for systematic reviews and meta-analysis (PRISMA) guidelines [189].

4.3.1 Selection of drugs

The list of commonly administered drugs to treat secondary complications after SCI was extracted from our recent publication [180]. The subset of drugs for which studies could be retrieved and were included in this review is provided in **Supplementary Table 8.1**.

4.3.2 Search methods for identification of studies

Using "Publish or Perish" (version 7.23.2852.7498 [1]), PubMed, Scopus, and Web of Science were searched using the time range from their individual inception dates (1977, 1960, and 1945 respectively) to March 31st, 2021. Search terms were "spinal cord injury", "recovery", and name of a drug of interest (**Section 4.3.1**), joined with AND. A manual search was also performed to include matching references of relevant trials.

4.3.3 Selection of studies

Articles were independently screened in two stages: initial screening of titles and abstracts (MW, CRJ), and full-text assessments (LB, LPL, MW, CRJ) using criteria described in **Section 4.3.4**. In case multiple articles reported on a single cohort, the article providing the most data or detail was selected for further synthesis [190]. Disagreements were discussed and resolved at multiple consensus meetings.

4.3.4 Inclusion and exclusion criteria

All full-text, peer-reviewed studies investigating the disease-modifying effect of a drug of interest (**Section 4.3.1**) on relevant neurological or functional outcomes (**Section 4.3.5**) after acute SCI were included. Where original articles were not published in English, screening and data extraction were performed by native speakers. We excluded duplicates, non peer-reviewed articles, reviews, meta-analyses, abstracts, editorials, commentaries, perspectives, patents, letters with insufficient data reporting, studies

[1] https://harzing.com/resources/publish-or-perish

exclusively on children/neonates, or out of scope studies (see **Figure 4.1** for full definition). We only included studies comparing the treatment group to a placebo control group, and excluded experiments using active compounds as the only control as it is impossible to compare drug effects between studies using different comparators (i.e., different active controls in studies A and B instead of placebo). Authors of articles that were indexed but not accessible either through institutional library access (ETH Zurich) or open source publishing, were contacted to obtain a copy of the full article. In case no copy was provided, the article was excluded (see "not accessible" in **Figure 4.1**). Subsequent data extraction was performed by six investigators (LB, LPL, BT, JL, TG, and CRJ).

4.3.5 Assessments and outcomes

The review focused on studies reporting drug effects on recovery as assessed by locomotor function, skilled fore- or upper limb function, sensory function as well as electrophysiology. Details about the assessments included in the analysis are reported in **Supplementary Tables 8.2** and **8.3**. Assessments used to track recovery outcomes in animals with SCI were grouped into categories based on the deficits measured. Tasks that assess spontaneous and voluntary motor function were differentiated between quadrupedal locomotion or skilled reaching or forelimb usage. Sensory assessments were grouped, including sensory reflex arcs, regardless of the type of sensory input eliciting the reflex. Assessments of electrical activity of muscle fibers or circuits were grouped under electrophysiology assessments to mirror comparable assessments in humans and reflect neural excitability. Too few papers assessed reflexes or utilized electrophysiology to warrant distinguishing between proprioceptive or pain/withdrawal reflexes, or between assessments of single units vs. monosynaptic or polysynaptic potentials or motor vs. sensory circuits. Assessments spanning multiple categories (e.g., Gale scale) or used in only a few studies were grouped together. In cases of ambiguity, the methods and results of the paper were closely reviewed to ascertain the feature of the deficit being assessed (e.g., toe spread as a measure of reflexes vs. weight bearing during locomotion).

4.3.6 Data extraction and synthesis

The following information was extracted from all studies: (i) study characteristics (first author's last name, publication year, language), (ii) study population (species, group sizes [total/control/treatment], sex, age, weight), (iii) injury characteristics (level, severity, mechanism, duration), (iv) drug administration (drug name, dose, route of administration, timing of start of treatment relative to injury, duration of treatment), and (v) neurological and functional assessment outcomes (name, time point(s), investigators blinded to

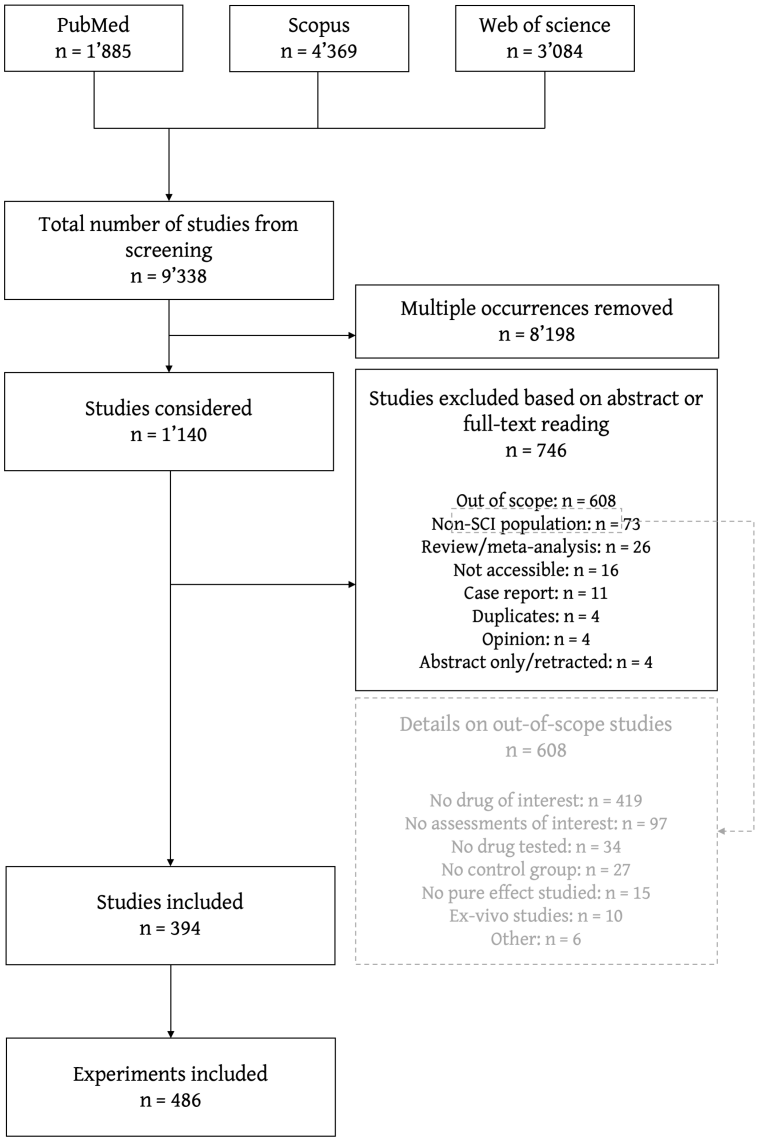


Figure 4.1: Preferred reporting items for systematic reviews and meta-analysis (PRISMA) flowchart. Protocols ($n = 4$), non-standardised spinal cord injury model ($n = 1$) and capsaicin-based transient receptor potential vanilloid 1 study ($n = 1$) are grouped under "other" out of scope excluded studies.

.

treatment, drug effect). A full list of extracted variables is provided in **Supplementary Table 8.4**. Studies analyzing multiple drugs of interest (e.g., drug A, drug B, and control, with drugs A and B of interest) were separated into multiple experiments (e.g., experiment 1: drug A vs. control, experiment 2: drug B vs. control) and extracted individually. risk of bias (RoB) was assessed for each experiment, considering animal, randomized clinical trial (RCT) [191], and intervention (observational) human studies [192] separately. Animal experiments were assigned a score from 0 (no bias) to 20 (highest RoB) according

to criteria listed in **Supplementary Table 8.5**. Visualizations for RoB assessments of RCTs and intervention studies were created using `robvis` [193].

4.3.7 Statistical analysis

Drug effects were classified for each experiment in one of six categories (**Table 4.1**). Descriptive statistics (mean, standard deviation (SD), median, min, max, percentage, and proportions) were used to provide summary information on the study characteristics, the studied drugs, and their effect on recovery after SCI.

Table 4.1: Classification of drug effect

Drug effect	Description
Positive	Treatment with the drug of interest resulted in improved/increased functional/neurological outcomes compared to control.
Negative	Treatment with the drug of interest resulted in worse/decreased functional/neurological outcomes compared to control.
No effect	Treatment with the drug of interest did not impact the functional/neurological outcomes compared to control in a statistically significant manner.
No statistics	Qualitative comparison between treatment and control groups were performed, but no statistical test results were reported.
Not reported	Functional/neurological outcomes were defined in Methods but results of comparison between treatment and control groups were not reported.
Mixed	Combination of positive, negative, no effects and/or no statistics was reported, depending on the assessments, dosage, timing, regimen or a combination of those situations.

4.3.8 Data and code availability statement

The data used in this study and source code of the analysis performed (including visualizations) can be accessed on our GitLab repository [2]. R Statistical Software version 4.3.1 and Python version 3.10.10 were used.

[2] https://gitlab.ethz.ch/BMDSlab/publications/SCI-drug-review-publication

4.3.9 Role of funding source

Funding sources of the study had no role in study design, data collection, analysis, interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

4.4 Results

Initially, 9338 studies were screened and 1140 qualified for full-text reading. 394 unique studies, reporting 486 experiments, met our inclusion criteria (see **4.3**, Figure 4.1). Sixty-four studies (16%) reported more than one experiment (see **Section 4.3.6**). Studies were published between June 1975 and March 2021, with the majority after 2010 (238 studies, 60%, **Supplementary Figure 8.3A**). While most studies were published in English (n = 381, 96.7%), some were also written in Mandarin (n = 7, 1.8%), Turkish (n = 2, 0.5%), Portuguese (n = 2, 0.5%), Persian (n = 1, 0.3%), and Korean (n = 1, 0.3%).

Most studies addressed the effect of medications in animal models (n = 377, 96%). Seventeen (4%) studies, reporting 22 experiments (5%), reported results in humans. 774 drugs are known to be administered in the acute phase of SCI. 7116 (15%) of those drugs were included in experiments identified in our review. 110 drugs were examined individually and 33 in combination (**Supplementary Table 8.1**). Six drugs were only tested as part of combinatorial treatments - aminocaproic acid [194], rosuvastatin [195], magnesium chloride [196, 197], ketamine [198], isoflurane [198] and nitroprusside [199].

Rat models were most extensively investigated (n = 382/464 experiments, 82%). Larger mammals (i.e., cats, dogs) were mainly used before 2001 (n = 19/22 experiments conducted on cats and dogs, 86%, **Supplementary Figure 8.3B**). By contrast, all experiments performed on mice (n = 38) were published after 2000. Sample size, age, and sex were partly or fully missing in 77 (17%), 341 (73%) and 61 (13%) experiments, respectively. Partly missing entries included sample size bounded or expressed as ranges, age described as “adult” or “young”, and samples comprising both male and female in unknown proportions. Likewise, exclusion or death of animals was only reported for 51 (11%) experiments. Among experiments reporting sample size, cohorts included a mean of 63 animals (standard deviation (SD): 52, median: 48, Q1-Q3: 32-80). Studies using larger mammals exhibited smaller cohorts (**Supplementary Table 8.6**). When reported, age was commonly expressed in weeks (n = 87, 19%). Rats had a mean age of 10 weeks (10.69 weeks when mean age is reported [n = 31/77], 8.92 weeks for lower bound and 10.76 weeks for upper bound when ranges are reported [n = 57/77]). Mice were also 10 weeks of age (mean of 10.00 [n = 2/10], 8.25 [n = 8/10] and 10.13 [n = 8/10] weeks when mean, minimum and maximum are reported, respectively). A majority of studies included exclusively male or female animals (n = 387, 83%), with more experiments being performed on exclusively male populations (n = 206, 44%). Details on the use of male, female and mixed populations over time are reported in **Supplementary Figure 8.3C**.

SCI models have been previously categorized into contusion, compression, distraction, dislocation, transection and chemical models [200]. 278 (60%), 132 (28%), 27 (6%), 16 (3%),

7 (2%), 5 (1%) experiments reported a contusion, compression, transection, ischemia, multiple or other injury mechanisms (photochemical lesion [201, 202, 203], irradiation [204], electrolytic lesion [205]), respectively. Although protocols used to induce injuries were often described in detail, information about the corresponding severity of the injury was missing for most experiments (n = 257, 55%). The level of injury was typically reported either precisely (n = 262, 56%) or in ranges (n = 172, 37%). Most experiments studied injuries at the thoracic level, predominantly at or below T5 (n = 222, 85% and n = 151, 88%, of experiments reporting unique and range levels respectively, **Figure 4.2A**).

109 individual drugs and 32 combinations were tested in SCI animal models. Methylprednisolone (MP) and methylprednisolone sodium succinate (MPSS) were most prevalent among experiments reported with 71 (15%) and 23 (5%) experiments, respectively (**Figure 4.2B**). A total of 60 (43%) unique drugs or combinations were tested in more than one experiment.

Drug effects were evaluated by a wide range of neurological and locomotor assessments. The most common choice was the Basso Beattie Bresnahan (BBB) [206] scale, developed and employed for rats. Its original or modified versions (e.g. Basso mouse scale (BMS) [207], canine BBB locomotor scale [208]) were used in 275 (59%) of the experiments (**Figure 4.2C**). Overall, most tests performed (n = 620/848, 73%) evaluated locomotor function. One experiment or study could include more than one assessment and 174 (46%) unique studies tested more than one category among locomotion, skilled forelimb function, sensory function, electrophysiology and other functional assessments. While assessment protocols were mostly well described, timing, number of repeats and follow-up period varied widely between experiments.

Figure 4.2B illustrates the drug effect reported for the most prevalent drugs in our review. One can notice that diverging findings were reported when testing the same drug in different experiments. Using MP as an example, 31 experiments reported positive effects, while 28 experiments found no effect for MP. Similarly, metformin, atorvastatin, lithium, valproic acid, melatonin and estradiol were investigated in more than five independent experiments and the majority (> 50%) of those experiments reported a positive effect of the treatment (80%, 78%, 63%, 60%, 57%, 56%, respectively). Interestingly, we identified two drugs with negative effects reported (morphine [158], ethanol [198]). However, most of the experiments published and reviewed here found their respective drugs of interest to have a positive (n = 195, 42%) or no effect (n = 115, 25%) on neurological or functional recovery following SCI. Details of mixed effects reported are presented in **Supplementary Figure 8.4**.

We extracted information from 17 studies reporting 22 experiments conducted on hu-

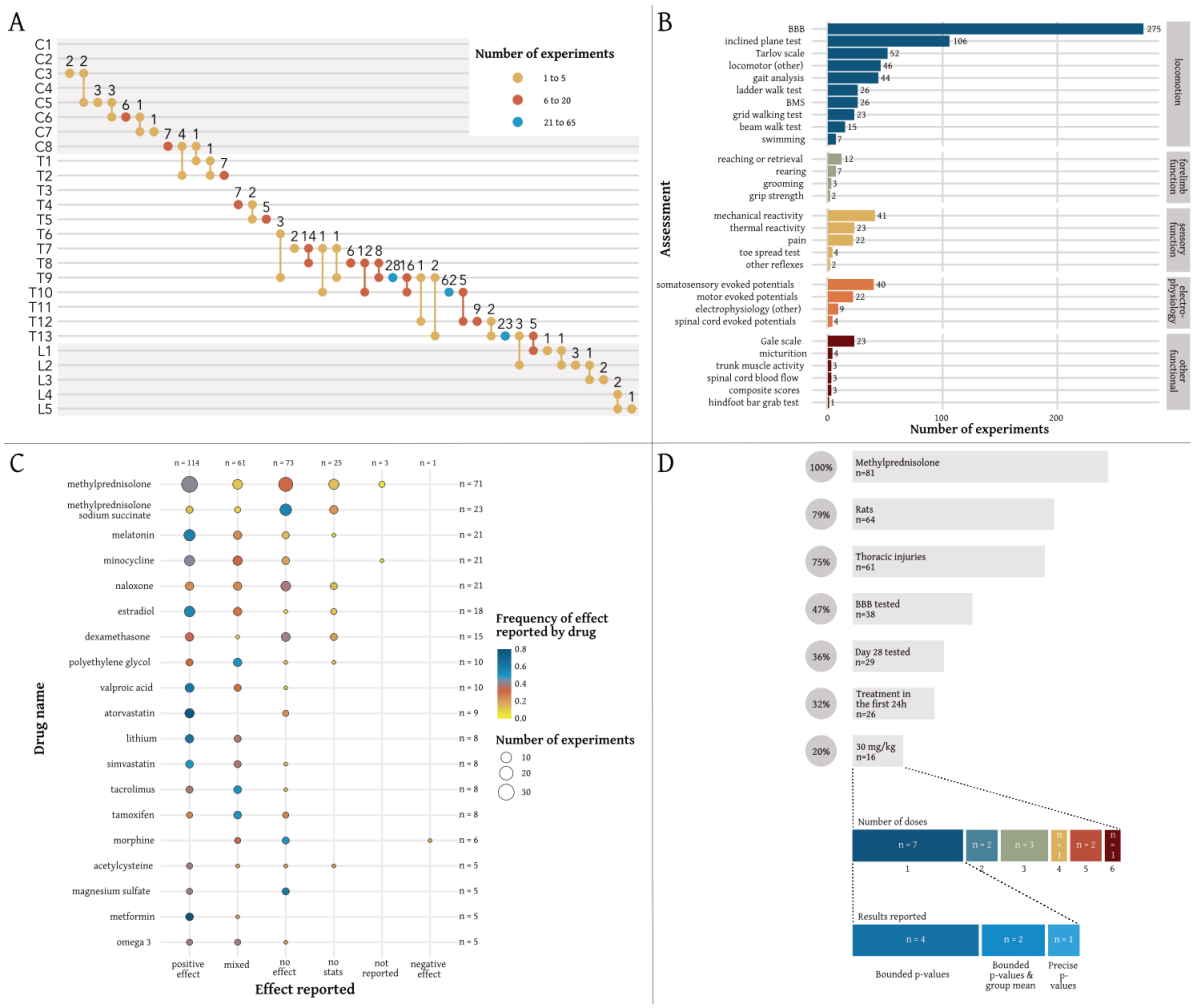


Figure 4.2: Analysis of animal studies reporting drug effects on neurological and functional recovery following spinal cord injury (SCI). A. Number of experiments per level of injury studied. Shaded areas distinguish between cervical, thoracic and lumbar injuries. Notably, thoracic injuries were the most prevalent in animal experiments. **B.** Drug effects reported for drugs studied in at least five experiments. Circle size is proportional to the number of experiments reporting the effect of interest. Circles are colored proportionally to the frequency that the effect of interest represents among all experiments studying the drug of interest. **C.** Number of experiments per assessment reported, classified in locomotion, skilled forelimb function, sensory function, electrophysiology (EP) and other functional assessments. Basso Beattie Bresnahan (BBB) locomotor scale; Basso mouse scale (BMS); somatosensory evoked potentials (SEPs); motor evoked potentials (MEPs); spinal cord evoked potentials (SCEPs). **D.** Illustration of the heterogeneity observed among experiments reporting effects of methylprednisolone on neurological and functional recovery after SCI.

man cohorts with SCIs (**Figure 4.3**). Cohort sizes varied greatly ($n = 10$ [209] to $n > 2000$ [210]). Sex distributions were consistently skewed towards male population (from 53.4% to 100% male), in line with the sex distribution observed in the general SCI population [211, 49, 77, 166]. While one study (two experiments) explicitly included pediatric participants [29], most experiments considered only adult participants with mean age between

32.5 [43] and 57.6 [212] years, matching the age distribution reported in the literature [77, 166, 213, 214].

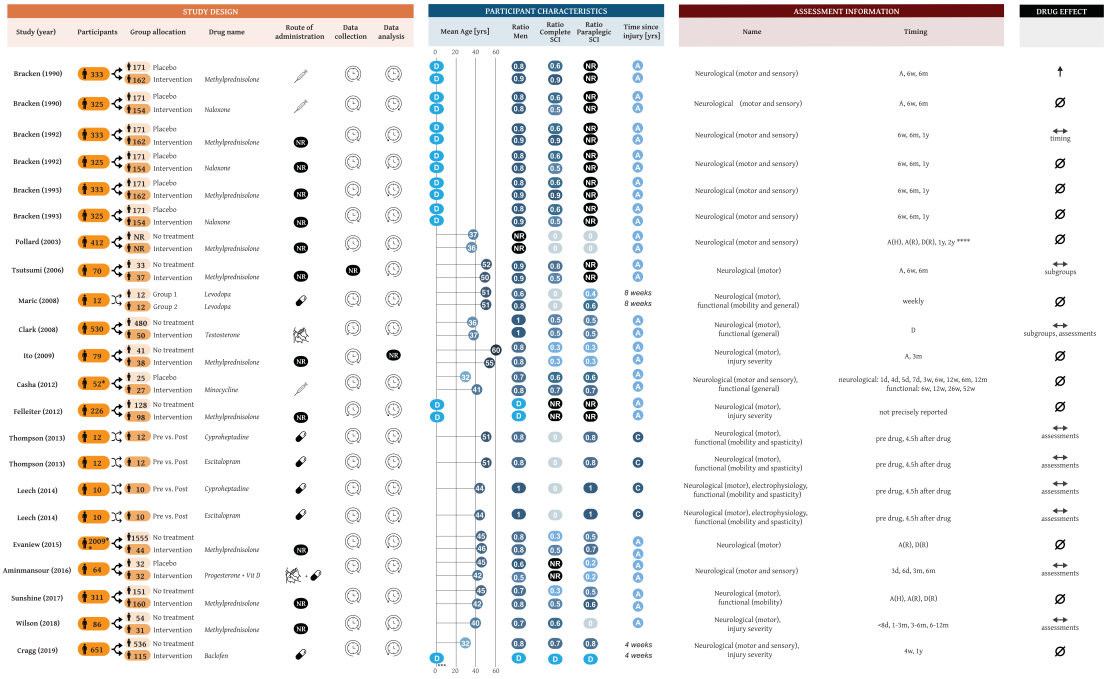
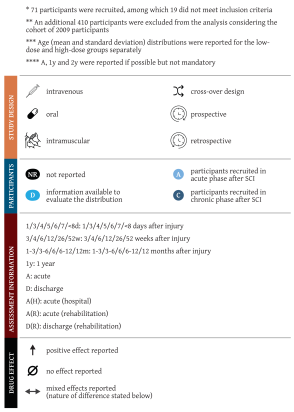


Figure 4.3: Overview of the human experiments included and its corresponding legend.



As expected and in contrast to animal studies, most human experiments were performed on heterogeneous groups with regards to their injury characteristics (neurological level of injury, severity, mechanism of injury). The majority of the studies (n = 15, 18 experiments) investigated patients with acute SCI. Only two studies (four experiments) specifically enrolled participants with chronic incomplete injuries [209, 215] comparing test performances pre- and post-exposure to the drugs of interest.

Drugs tested included naloxone [29, 160, 216] (n = 3, 14%), cyproheptadine [209, 215] (n = 2, 9%), escitalopram [209, 215] (n = 2, 9%), baclofen [43], minocycline [217], levodopa

[218], testosterone [219] and a combination of progesterone and vitamin D [220] ($n = 1$, each, 5%). Methylprednisolone was the most studied drug ($n = 10$, 45%) with publications between 1990 [29] and 2018 [108].

All studies evaluated drug effects through neurological assessments. Additionally, functional outcomes such as mobility [209, 215, 221] or spasticity [209, 215] were tested in eight experiments, and one study (two experiments) [209] reported electrophysiological outcomes. Lastly, recovery was assessed based on changes in injury severity in four experiments [43, 212, 108, 222].

Results reported for the effects of MP diverged from the animal studies with only one experiment recording positive results [29], which was part of the oldest study of MP in humans. Most of the experiments on MP reported no effect ($n = 6$, 60%) and three observed mixed effects depending on subgroup [223], assessment [108] or timing of treatment [160]. A similar trend was observed when considering all drugs tested in human populations: a total of 12 experiments reported no effect (55%) and 9 described mixed results (41%), mainly due to differences between assessments ($n = 7$, 32%). Notably, most of the data from human populations were collected prospectively ($n = 18/22$, 82%), i.e., individuals were followed and data was collected over time, while they were most often analyzed retrospectively ($n = 12/22$, 55%), i.e., data were analyzed after the final outcome was known. This hints towards few clinical trials testing pharmacological treatments for SCI.

RoB was assessed for animal, RCTs and observational human studies separately. Overall, animal studies exhibited scores ranging from zero to 12, with 36 experiments (7.8%) having a score greater or equal to six (**Supplementary Table 8.7**). Variables most affected by a potential bias were age and blinding of recovery assessments (**Figure 4.4A** and **B**). Among observational human studies, only one showed critical RoB (**Figure 4.4C**), while most RCTs showed high RoB in the selection of the reported results (**Figure 4.4D**).

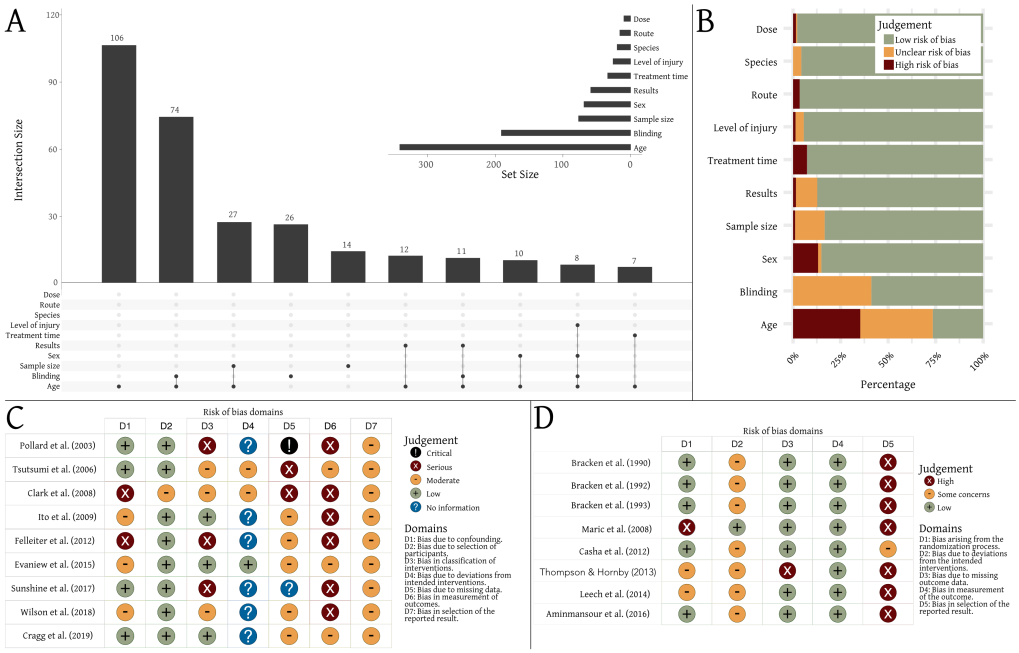


Figure 4.4: Assessment of the risk of bias (RoB) for included experiments. A. Co-occurrence of potential bias (grading as low or high risk) within animal experiments. RoB was most prevalent in reported age, followed by a combination of age and blinding status. Conversely, information on species, route of drug administration and dose showed lower RoB. **B.** Proportion of each RoB (low, unclear or high) by domain of bias studied. Age represents the domain with the most prevalent high RoB. **C.** RoB for human intervention studies (observational). Only one study showed a critical RoB (domain 5: bias due to missing data), whilst most studies ($n = 6$, 67%) did not provide sufficient information to assess the RoB due to deviations from intended interventions (domain 4). Additionally, the majority of the studies ($n = 8$, 89%) had a low RoB due to selection of participants (domain 2). **D.** RoB for randomized clinical trial (RCT). High RoB was detected in 7 studies (88%) for bias in selection of the reported results.

4.5 Discussion

The current study aimed to systematically review existing literature assessing the effects of drugs commonly administered in the acute phase of SCI. Encouragingly, several drugs have been investigated across multiple animal models and have consistently demonstrated positive effects [183, 224, 225, 226, 227]. This convergence of evidence prompted the formulation of drug repositioning as a novel translational approach in the field of acute SCI care. Repositioning has emerged as a successful strategy in other fields (e.g., amantadine in Parkinson's Disease [228] and Lintuzumab in Alzheimer's disease [229]) to improve neurological outcomes in the absence of novel therapies. Drug repositioning aims at identifying new uses for approved or investigational drugs that are outside the scope of the original drug indication [230]. A clear advantage of this approach is the use of de-risked compounds with established safety and biological activity profiles, thereby

reducing overall development costs and shortening timelines [231, 232]. While drug repositioning utilizes existing evidence to accelerate the development of new treatments, it is still affected by challenges of translational research. We identified 377 studies considering the effects of drugs previously identified as administered to human patients with acute SCI. Evidence exists for 112 (77.78%) unique compounds or combinations to exert beneficial and/or detrimental effects. For example, metformin is routinely used in humans to manage high blood sugar levels caused by type 2 diabetes [233]. Preclinical studies have identified enhanced regeneration in the spinal cord related to metformin-induced autophagy via the mTOR signaling pathway [183, 234, 235]. These observations suggest that administering metformin early after injury could potentially improve long-term neurological outcomes.

Detrimental effects were also observed for some drugs, including opioids, which attenuated the recovery of locomotor function and exacerbated pathophysiological processes in rodent models of SCI [236, 159, 237, 228]. A detrimental opioid effect is in line with beneficial effects of naloxone, an opioid antagonist [229, 230], and highly concerning in light of the ubiquitous administration of opioids for pain management in the early stages of SCI. Completely removing or restricting opioids presents serious ethical concerns (i.e., weighing the management of acute pain with long-term neurological effects). However, minimizing the administration of opioids could potentially facilitate neurological recovery [238, 239]. To allow for a comprehensive characterisation of potential effects of commonly administered drugs on neurological recovery, we deliberately decide to include preclinical studies involving animal models and clinical studies in humans. Nonetheless, the high degree of heterogeneity across studies, even in a single species, was surprising. A large variability in population characteristics, exact administration parameters and timing of assessment is observed. In combination with a wide range of spinal levels subjected to injury and different species being studied, comparisons between experiments are challenging or impossible. One exception is the study by Popovich et al. [240] aiming to replicate findings, which noted a strong connection between initial injury characteristics and detectable drug effects. This highlights the need for varying as few parameters as possible to allow for meaningful comparisons. Currently, meta-analyses are not feasible, even for the most commonly studied drugs (**Figure 4.2D**), constituting a notable limitation as the large fraction of positive effects reported might hint towards a publication bias. In human studies we suspect that the majority reporting mixed or no effects reflects the heterogeneity in injury patterns included. This likely results in effects which vary widely between individuals and cannot be detected in a group-level analysis. The lack of an effective pharmacological treatment for SCI highlights the discrepancy between largely positive pre-clinical results and unsuccessful translation to human subjects. The present

review allows to formulate a number of hypotheses that could explain this divergence. One noticeable difference concerns basic study parameters such as the age of the cohort studied or the level of injury. While the age distribution in humans affected by SCI is moving towards a bimodal shape [166], studies in animal models are typically performed on more homogeneous groups of younger individuals [241, 242, 243, 186, 244]. While the use of young animals might be a result of ethical guidelines, it may affect transferability. Further, SCI in humans occurs predominantly in the cervical segment of the spinal cord [166], while animals are mostly injured in the thoracic region (**Figure 4.2A**). Similarly, injury severity, frequently not reported, has been named as a critical parameter to control for in animal studies to ensure transferability of findings to the human population [245]. Noticeable differences also exist in the administration of drugs. Animal studies typically follow a weight-based dosing regime while humans receive a standardized dose. Similarly, many animal studies initiate treatment immediately after injury [227, 246, 247], which appears infeasible in the human population. These issues in the transfer from animal to human studies might contribute to the majority of human studies reporting mixed effects. While beneficial effects might still exist in humans, they could go undetected due to the scarcity of RCTs. While RCTs require substantial resources, and can be challenging to conduct in a rare and heterogeneous condition like SCI, advancements in the treatment of SCI will only be possible if efforts extend from preclinical studies to systematic prospective data collection and analysis in humans. Finally, only a small subset of studies in humans considers the effect of drugs in the chronic phase. As chronic injuries cannot be investigated in animal studies due to ethical restrictions, studies of chronic human SCI populations should be expanded to address debilitating secondary complications [245]. A noteworthy limitation of the current review was that literature search was limited to articles listed in PubMed/Medline, Scopus, and Web of Science, or identified by hand searches. Considering the pace at which research in this area advances, it is likely that the findings of the publications described in this paper will be quickly complemented by further research. The literature search also excluded gray literature (e.g., preprints, reports, conference proceedings), the importance of which to this topic is unknown, and thus might have introduced another source of search bias. Publication bias is likely to result in studies with positive results being preferentially submitted and accepted for publication. The present review provides an extensive summary of existing evidence on effects of drugs administered to individuals affected by SCI. In particular, results highlight melatonin, estradiol, and valproic acid as commonly investigated drugs with largely positive effects, indicating the inherent potential to advance treatment through drug repurposing. Simultaneously, we observed and extensively characterized sources of heterogeneity among the valuable resources provided by existing studies. In light of

the current lack of an effective pharmacological treatment for SCI and failed attempts to develop new treatments, the field would benefit from further standardization in studying and reporting drug effects investigated in animal models.

Authors' contributions

All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. LB: selection of studies, extraction of data, statistical analysis, visualization, interpretation of the data, drafting of the manuscript
LPL: selection of studies, extraction of data, statistical analysis, visualization, interpretation of the data, drafting of the manuscript
BK: interpretation of data, revising the manuscript for intellectual content
BT: extraction of data, interpretation of data, revising the manuscript for intellectual content
JL: extraction of data, interpretation of data, revising the manuscript for intellectual content
TG: extraction of data, interpretation of data, revising the manuscript for intellectual content
WT: interpretation of data, revising the manuscript for intellectual content
JLKK: interpretation of data, revising the manuscript for intellectual content
MW: study design, selection of studies, interpretation of data, revising the manuscript for intellectual content
CRJ: study design, selection of studies, extraction of data, visualization, interpretation of the data, drafting of the manuscript

Declaration of interests

The authors declare no competing interests.

Data sharing

The GitLab repository [3] including all code and data is publicly available.

[3] https://gitlab.ethz.ch/BMDSlab/publications/SCI-drug-review-publication

Acknowledgements

Funding

This study was supported by the Swiss National Science Foundation (Ambizione Grant, #PZ00P3_186101 and #IZLIZ3_200275, Jutzeler), Wings for Life Research Foundation (#2017_044, #ID 2020_118, Jutzeler and Kramer), the International Foundation for Research in Paraplegia (#P192, Jutzeler). Lucie Bourguignon was supported by an ICORD student exchange award provided by the Praxis Spinal Cord Injury Foundation. Bethany R. Kondiles was supported by the Paralyzed Veterans of America Research Foundation (#3195).

Icons

The following icons from thenounproject.com were used:

- “Average” icon by James Bickerton CC BY 3.0
- “Less or equal” icon by Julia Holmberg CC BY 3.0
- “Pin” icon by Nice Design CC BY 3.0
- “Clock” icon by Made x Made Icons CC BY 3.0
- “Past” icon by Made x Made Icons CC BY 3.0
- “Pill” icon by Three Six Five CC BY 3.0
- “Shuffle” icon by Gregor Cresnar CC BY 3.0
- “Syringe” icon by arman maulana CC BY 3.0
- “Vaccination” icon by WiStudio CC BY 3.0
- “Empty set” icon by Gregor Cresnar CC BY 3.0

Part B in a nutshell

As presented in **Part B**, data collected as part of clinical trials and observational studies are valuable assets to better characterize SCI. However, handling real-world data also comes with challenges. Some of these challenges are inherent to the SCI field, such as the heterogeneous presentation of the condition and subsequent recovery. Other aspects, however, are known from other fields, such as the influence of missing data. Importantly, observational studies are, by design, prone to missing information (e.g., different markers needed for making clinical decisions, less controlled environment). The SCI research field currently lacks methods adapted to the peculiarities of data collected from individuals with SCI. Hence, **Part C** presents the effort pursued towards tailoring existing methods to their applications in SCI research.

Part C

Towards better data analysis for clinical studies

Introduction

At the heart of the analysis of a randomized clinical trial (RCT) lies the definition of a treatment effect. An average treatment effect is determined by comparing the outcomes observed, or recovery achieved, in the treatment versus placebo arm [248]. This comparison assumes that the two groups were comparable at baseline, and one could therefore expect equally comparable recovery in the absence of an intervention. Any difference between the two groups can in turn be attributed to the intervention tested. This assumption of groups being comparable is, however, often questioned in cases like spinal cord injury (SCI), where the injured population is largely heterogeneous [116]. Obtaining truly comparable groups would require the recruitment of large samples, which is challenging in light of the low prevalence of traumatic SCI. A complementary measure of a treatment effect can therefore include individual-level estimation of the expected recovery, which is then compared to the one observed in presence of an intervention. The individual-level comparison would be made possible thanks to accurate, explainable and transferable predictive models using machine learning (ML) algorithms. Predicting the expected recovery of an individual would first require to train a model, i.e., to use existing data from which the outcome is known, to infer patterns that are associated with the outcome. The setting where the outcome is known in the population used to train the algorithms is referred to as supervised ML[249]. Various models, differing in their architecture and complexity, will be trained, optimized and later compared using performance metrics to determine the combination of parameters that allows for the best performing algorithm, i.e., predicting outcomes closest to the ones observed.

However, building predictive models for healthcare poses specific challenges [250]. Firstly, they should be capable of capturing the large heterogeneity of the population on which the model would later be applied. Secondly, the expected prediction error should fall under what is considered a clinically significant change [251], such that any potential error made by the model would not result in clinically observable or significant differences. Lastly, transferability and interpretability, namely the ability for a model to achieve a good prediction when given previously unseen cases [252] and the possibility for a human to understand the decision made by the algorithm [253], are crucial for a successful deployment in a clinical setting.

Part C exposes our contribution towards improved data analyses for better characterization and prediction of recovery at the individual level following SCI. It is organized in three chapters, where we:

(i) present our attempt to build a prediction model for lower extremity motor score (LEMS) observed 52 weeks after injury in **Chapter 5**. In this chapter, we expand the

insights drawn from the study of serological markers (**Chapter 2**) and question the predictive power of these biomarkers;

(ii) following from the limitations identified in **Chapter 5**, examine the impact of missing data on results reported through a simulation study detailed in **Chapter 6**;

(iii) explore the uniqueness of individuals recovering beyond clinical expectations, with the intent to formulate new hypotheses for future research, in **Chapter 7**.

Chapter 5

Exploring the potential of routine serological markers in predicting neurological outcomes in spinal cord injury

---

Adapted from:

Jan Matthias, Louis P. Lukas, Sarah C. Brüningk, Doris Maier, Orpheus Mach, Lukas Grassner, John L. K. Kramer, **Lucie Bourguignon‡**, Catherine R. Jutzeler‡

Submitted to *Experimental Neurology*, 2024

Publication:
GitHub repository: https://gitlab.ethz.ch/jmatthias/Serological_markers_LEMS

Lucie Bourguignon was responsible for the design of the project and contributed to the data analyses, interpretation, and visualization of the results and the writing of the manuscript. Further details can be found in **Section 5.5.2**.

5.1 Abstract

Spinal cord injury (SCI) is an orphan condition with a heterogeneous presentation, making the search for a pharmacological cure challenging. With the increased amount of clinical routine data available our investigation aimed to assess the feasibility of predicting lower extremity motor score (LEMS) at chronic stage (52 weeks after initial injury) in patients with SCI using routine serological markers.

Serological markers, assessed within the initial seven days post-injury in the observational cohort study from the Trauma Hospital Murnau underwent diverse feature engineering approaches. These involved arithmetic measurements such as mean, median, minimum, maximum, and range, as well as considerations of the frequency of marker testing and whether values fell within the normal range. To predict LEMS scores at the chronic stage, eight different regression models (including linear, tree-based, and ensemble models) were used to quantify the predictive value of serological markers relative to a baseline model that relied on the very acute LEMS score and patient age alone.

The inclusion of serological markers did not improve the performance of the prediction model. The best-performing approach including serological markers achieved a mean absolute error (MAE) of 6.59 (2.14) which was equivalent to the performance of the baseline model. Stratifying the models based on the acute-phase LEMS exceeding zero, led to a mean improvement in MAE across all cohorts and models, of 1.2 (2.13).

We conclude that routine serological markers hold limited predictive power in our study. However, the implementation of model stratification by the very acute LEMS markedly enhanced prediction performance. This observation supports the inclusion of clinical knowledge in the modeling of prediction tasks for SCI recovery. Additionally, it lays the path for future research to consider stratified analyses when investigating the predictive power of potential biomarkers.

5.2 Introduction

Spinal cord injury (SCI) is a rare, yet severe medical condition with far-reaching implications for affected individuals, caregivers, and society. The consequences of SCI are often profound, impacting motor function, sensation, and autonomic functions [3]. Despite recent advances in acute care and rehabilitation, the extent of recovery during the first year post-injury has remained unchanged over the past two decades [166]. During the acute phases of the SCI, a cascade of molecular mechanisms are activated including demyelination, apoptosis, inflammatory pathways, lipid peroxidation and reactive oxygen species creation [254]. These changes can be measured both in the cerebrospinal fluid (CSF) and in the serum [255]. Unraveling the dynamics of fluid serological markers offers a promising avenue for predicting SCI related outcomes in a more objective and cost-effective manner compared to conventional imaging techniques like magnetic resonance imaging or invasive methods like CSF markers [256].

CSF biomarkers, such as neurofilament light chain (NF-L) and glial fibrillary acidic protein (GFAP) have been demonstrated to be associated with injury severity as measured by the American spinal injury association (ASIA) impairment scale (AIS) [256]. The concentration of these biomarkers in CSF and serum have been used by Stukas *et al.* [256] and Leister *et al.* [257] to predict AIS grade conversion (change in severity category). Nevertheless, CSF biomarkers pose challenges in routine and straightforward collection compared to peripheral blood. Serological markers, derived from routine blood draws, have also been demonstrated to be indicative of injury severity. For instance, Tong *et al.* [44] revealed a significant association between elevated serum albumin concentrations at one, two, and four weeks post-injury and higher lower extremity motor score (LEMS) at the 52-week post-injury mark. Further analyses [258, 259] examined 28 routine serological markers and identified some correlations with injury severity. Specifically, higher values closer to the normal range of calcium, hematocrit, hemoglobin, erythrocyte count, and total proteins were associated with less severe injury. This suggests a potential for utilizing serological markers to define SCI severity, which is in itself an indicator of the expected recovery, and therefore enhance the prediction of possible outcomes.

We hypothesized that incorporation of information from serological markers would enhance the performance of machine learning models in predicting LEMS at the chronic stage, surpassing prediction accuracies based solely on baseline patient characteristics (LEMS and age). Additionally, we explored the idea that the frequency of serological draws conducted during the initial seven days acts as a proxy of the overall patient health status, potentially leading to improved prediction performance.

5.3 Methods

5.3.1 Study design and data source

The Murnau Study is an observational cohort study, conducted at the level 1 trauma center in Murnau, Germany. Between 2004 and 2017, a total of 363 SCI patients were enrolled and followed for the course of one year after initial injury. During that period, standard rehabilitation care was given to all patients.

5.3.2 Ethics approval

The study was performed in accordance with the Declaration of Helsinki. The Murnau study was approved by the Bavarian Medical Chamber (#2018-077).

5.3.3 Cohort definition: Inclusion and exclusion criteria

To be eligible for inclusion in the analysis presented, patients were required to have LEMS assessed at both the very acute (within two weeks post-injury) and chronic (26 to 52 weeks post-injury) stages. The utilization of the LEMS score facilitated the inclusion of both paraplegic and tetraplegic individuals. All patients who satisfied these criteria were included, independent of neurological level of injury (NLI) or age.

5.3.4 Outcome, features, and confounding variables

The primary outcome was LEMS at 52 weeks post injury (i.e., chronic stage of injury). LEMS is evaluated as part of the international standards for neurological classification of international standards for neurological classification of spinal cord injury (ISNCSCI), where five key muscles of each lower limb are tested. Each muscle group has a maximum score of five (active movement against full resistance) and minimum of zero (full paralysis), for a total score per limb ranging from zero to 25 and total LEMS ranging from zero to 50 [16]. Following the data collection protocol of the European multicenter study on human spinal cord injury (EMSCI) [1], the Murnau study assesses LEMS at five distinct stages following SCI: very acute (0 to 15 days post injury), acute I (16 to 40 days), acute II (70 to 98 days), acute III (150 to 186 days) and chronic (300 to 546 days). Representing only the muscle activity of the lower extremities, LEMS therefore offers better walking capability prediction [260] and allows for the inclusion of both para- and tetraplegic patients, thereby making it a suitable outcome of interest.

[1] http://emsci.org/

A total of 28 routine serological markers (**Supplementary Methods 8.8**) were used as features, based on previous work by Bourguignon *et al.* [258]. All serological marker analyses were conducted by the berufsgenossenschaftliche unfallklinik (BGU) Murnau following an in-house protocol. Serological marker samples were collected upon request from the attending physician, resulting in a heterogeneous amount of blood draws and serological markers data sets across all patients, where some patients had multiple draws per day, while others had none. In order to mitigate the heterogeneous sampling frequency, three different feature engineering strategies were used.

5.3.5 Data preprocessing, feature engineering, and feature selection

Data imputation

To maximize our sample size, data imputation was employed. Imputation of missing LEMS at chronic stage was obtained using the last observation carried forward (LOCF) method from LEMS evaluated at the acute III stage, as it significantly outperforms other imputation methods for chronic LEMS imputation [261]. This approach is also consistent with observed SCI recovery trajectories showing that most recovery happens within the first six months, followed by a plateau [262].

For patients lacking LEMS scores at the initial acute stage, but presenting acute I and acute II scores of 0, subsequently leading to a final acute III or chronic LEMS score of 0, the missing LEMS score at the initial acute stage was backwards filled as 0.

Feature engineering 1: arithmetic transform of serological markers

The first feature engineering strategy was to calculate the average, median, minimum, maximum, and range of the serological markers values across the first seven days. Opting for a seven-day time frame was a compromise between data collected in temporal proximity to the injury, sample size, and the number of missing values. In this time window, patients are likely to have blood tests at least twice, therefore providing the opportunity to extract a mean, median, minimum, maximum, and range. It should be noted that minimizing the amount of missing values is crucial as the downstream prediction tasks demand complete data (see **Section 5.3.6**). This first feature engineering step resulted in five different data sets for the serological markers; one for each calculation method of the serological marker values: mean, median, minimum, maximum and range cohorts, respectively.

Feature engineering 2: sampling frequency features

The second feature engineering approach was to calculate the sampling frequency, where the number of times a specific serological marker was tested over the first seven days after injury, was counted. We will refer to this cohort hereinafter as the sampling frequency cohort.

Feature engineering 3: values inside/outside normal range

A seventh set of input features was created where each value of the serological markers were encoded as a 1, 0 or -1, for abnormal, normal or missing values, respectively (**Supplementary Methods 8.8**). This approach accounts for both sampling frequency and for a serological marker to be out of the norm. Owing to the categorical nature of these features, they were one-hot encoded and no feature selection was performed.

Feature selection

The features were filtered based on a moderate Pearson correlation threshold of 0.7 [263] (**Figure 5.1** and **Supplementary Table 8.10**). Additionally, an eighth cohort was created where features from all feature engineering strategies were combined, followed by a forward feature selection with a linear regressor (p-value threshold = 0.05).

5.3.6 Statistical analysis

Considering LEMS as a continuous variable, a range of commonly-used regression models were employed for the prediction task. These included linear, least absolute shrinkage and selection operator (LASSO) and ridge regressions, random forest (RF), support vector machines (SVM) with a linear kernel, gradient boosting regressor (GBR), extreme gradient boosting (XGBoost) regressors, and light gradient boosting machine (LightGBM) as implemented in `scikit-learn` (version 1.0.2). All hyperparameters were optimized using a five fold cross-validation scheme (**Supplementary Methods 8.9**). All model scores are reported as mean (standard deviation (SD)) through 50 random seed iterations. As the LEMS is lower (0) and upper bounded (50), all predicted scores below 0 or above 50 were capped to 0 or 50, respectively. As a confounding variable, age was added as a feature, as older age has been associated with negative impact on recovery [17].

The prediction task was first approached considering **Equation 5.1** applied to the different cohorts independently. The features of the regression models are the very acute LEMS, age, and the serological markers (after feature selection).

$$LEMS_{chronic} \sim LEMS_{very acute} + \text{age} \pm \text{serological marker features} \quad (5.1)$$

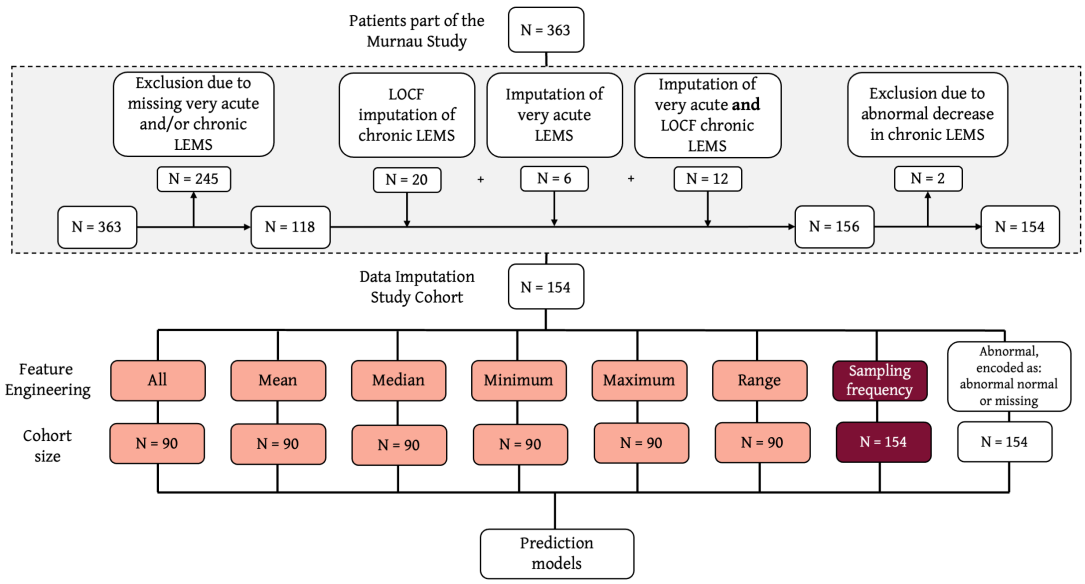


Figure 5.1: Overview of the inclusion criteria, data imputation, and feature engineering strategies. Number of patients in the specific cohort is indicated with N. In orange are the cohorts obtained from arithmetic transformations of the serological values (all, mean, median, minimum, maximum, and range) cohorts, followed by the sampling frequency cohort in dark red. Subsequently in white is the cohort that encodes the serological marker, based on the serological value being abnormal, normal, or missing. lower extremity motor score (LEMS); last observation carried forward (LOCF)

In a second approach, a stratified analysis based on values of LEMS at the very acute stage was employed since we observed two distinct groups in our population (see **Supplementary Figure 8.5**). Here, in order to account for this skewed distribution, two separate models were trained for all eight regression models based on **Equation 5.1**: the first model only included patients with a LEMS of 0 at the very acute stage, while the second model included all patients with a score above 0. After prediction, the respective test predictions were merged for evaluation.

As the aim of the study was to evaluate the predictive power of serological markers, two baseline models were also created, as the distinct feature engineering strategies resulted in two different patient cohort sizes due to patient exclusion in case of missing features (see **Figure 5.1**). The baseline models relied exclusively on the very acute LEMS and age (**Equation 5.1**, in the case where serological marker features are not included). This allows a direct comparison to quantify the predictive power of serological markers in predicting chronic LEMS scores.

5.3.7 Noise

Irrelevant or meaningless data, also known as noise [264], can significantly affect various machine learning (ML) tasks by rendering them less efficient and more computationally demanding [265]. In order to test whether the serological marker information holds predictive power, a random binary noise variable was introduced as a feature in separate models. This parameter allows one to determine how robust the models are to random noise [266, 267] and to quantify the feature importance of the serological markers in this context. A feature importance for the random noise variable similar to one for the serological marker information would indicate that serological markers lack predictive value.

5.3.8 Evaluation of models

All models were scored and compared using the root mean squared error (RMSE) and mean absolute error (MAE) between the true and predicted LEMS. Both scoring methods are on the same scale of the measurement itself and negatively-oriented for ease of interpretation. A lower score, approaching zero, is indicative that on average the predicted value is closer to the true value. However, the RMSE, which squares the difference between the root and predicted values, penalizes larger errors more.

5.3.9 Data and code availability statement

Anonymized data used in this study will be made available upon request to the corresponding author and in compliance with the European general data protection regulation (EU GDPR). The code describing the analysis and library versions can be accessed on our GitHub repository [2].

5.3.10 Role of funding source

The funding sources of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

[2] https://gitlab.ethz.ch/jmatthias/lems_prediction_serological_markers

5.4 Results

5.4.1 Cohort summary

With the fulfillment of the inclusion criteria (i.e., LEMS score at very acute and at chronic stage), the original data set was initially reduced from 363 patients to 118 patients. After imputation of missing LEMS at chronic stage using LOCF, two patients had an abnormal decrease of 40 points or above between very acute and chronic LEMS, without a clear justification, which led to their exclusion from the final cohort. With imputation for the acute and chronic LEMS (**Supplementary Table 8.11**), the final cohort is composed of 154 patients. A summary of the steps taken to define the final cohort is provided in **Figure 5.1**.

Demographics and injury characteristics of the original Murnau and final study cohort are summarized in **Table 5.1**. Comparing the two cohorts revealed a similar ratio of female to male patients (Pearson's χ^2 test, $\chi^2 = 1.947$, degree of freedom (df) = 1, $p = 0.1629$). Furthermore, the AIS grade distribution at the very acute stage does not present significant differences (Pearson's χ^2 test, $\chi^2 = 3.538$, df = 4, $p = 0.4721$). The very acute LEMS and chronic LEMS scores in the two cohorts also did not show any significant differences. However, the mean age revealed a significantly younger population in the study cohort (Wilcoxon rank-sums test, stat = 3.429, $p = 0.0006$).

Table 5.1: Detailed description of the final study cohorts

	Murnau	**Study cohort**	**p value**
Subject characteristics			
Total, *n*	363	154	
Age in years at injury			**0.0006**
Mean±SD	54.01±19.99	47.64±18.52	
Sex, n (%)			0.1629
Male	275 (76)	126 (82)	
Female	88 (24)	28 (18)	
LEMS			
Very acute			
Time in days (median)	5.00	4.00	0.0915
Time in days [Q1 - Q3]	[2.00 - 8.00]	[2.00 - 7.00]	

Continued on next page

Table 5.1: Detailed description of the final study cohorts (Continued)

	Murnau	**Study cohort**	**p value**
Score (mean±SD)	22.65 ± 20.60	17.37 ± 20.25	0.0159
NA, n	140	0	
Chronic			
Time in days (median)	363.00	346.50	0.0151
Time in days [Q1 - Q3]	[326.25 - 396.00]	[305.25 - 394.75]	
Score (mean±SD)	28.60 ± 21.87	25.69 ± 22.40	0.2605
NA, n	201	0	
AIS grade at very acute stage, *n* (%)			0.4721
A	67 (18.5)	48 (31.2)	
B	22 (6.1)	17 (11.0)	
C	27 (7.4)	14 (9.1)	
D	103 (28.4)	54 (35.1)	
E	2 (0.6)	0 (0)	
NA	142 (39.1)	21 (13.6)	

American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description. P values significant after Bonferroni adjustment are indicated in **bold**; standard deviation (SD), lower extremity motor score (LEMS), first quartile (Q1), third quartile (Q3), not available (NA)

The LEMS score of the included patients was measured at very acute stage within approximately 2 weeks post injury (median [Q1, Q3]: $4.00[2.00, 7.00]$ days), acute III (approximately 6 months, $159.50[154.25, 171.00]$ days) and chronic (approximately 12 months, $346.50[305.25, 394.75]$ days).

Arithmetic transformation of the serological markers cohorts

Stemming from exclusion of patients who presented with missing values in their serological markers, the serological marker cohorts for which arithmetic calculations were performed, present a smaller cohort size compared to the study cohort. The mean, median, min, and max, range and the cohort which includes all features cohorts comprise 90 patients each (**Supplementary Table 8.12**).

Sampling frequency cohort and encoded cohort

Due to the categorical nature of its encoding, the cohort based on normal ranges did not undergo feature selection and includes 415 serological marker features. The sampling frequency cohort (with sampling frequency ranging from 0 to 15) and the cohort that encodes the serological marker based on the normal range, each consist of 154 patients (**Supplementary Table 8.12**).

5.4.2 Model performance (non stratified approach)

Figure 5.2A-C summarizes the model performance. Detailed results can be found in **Supplementary Table 8.13**. The baseline model for the arithmetic calculation cohort ($n = 90$) achieves its best MAE performance with the SVM regressor (mean (SD): 9.69 (2.88)), and best RMSE with the LightGBM (13.98 (3.49)). Similarly, the baseline model for sampling frequency and normal range encoded cohorts ($n = 154$) performs best in terms of MAE with the SVM regressor (8.15 (2.23)) and with LightGBM for the RMSE (12.43 (2.50)).

We failed to demonstrate statistically significant improvements in prediction performance when comparing models including serological information to their respective baseline (t-test, $p > 0.05$, Bonferroni corrected).

5.4.3 Stratification models

Figure 5.2D-F summarizes the stratified model performance. All results are presented in **Supplementary Table 8.14**. In the cohorts including 90 and 154 patients respectively, baseline models after stratification showed a best average MAE when fitting a SVM regressor (7.41 (3.04) and 6.59 (2.14)) and best average RMSE with the LASSO (13.55 (3.62) and 12.09 (2.60)). In the stratified models, all best performing combinations between cohorts and models fail to significantly outperform the respective baseline (t-test, $p > 0.05$, Bonferroni corrected).

5.4.4 Comparison between stratified and non-stratified model fitting

When comparing the average performance of the cohorts across all models, the stratified models, on average, perform significantly better (comparisons with t-test resulted in 67 and 42 out of 80 models with $p < 0.05$ after Bonferroni correction, for MAE and RMSE, respectively; **Supplementary Table 8.15**). This is especially evident with the cohorts based on arithmetic transformation of the serological markers (i.e., mean, median, minimum, maximum and range) with the LASSO regression and Ridge regression models.

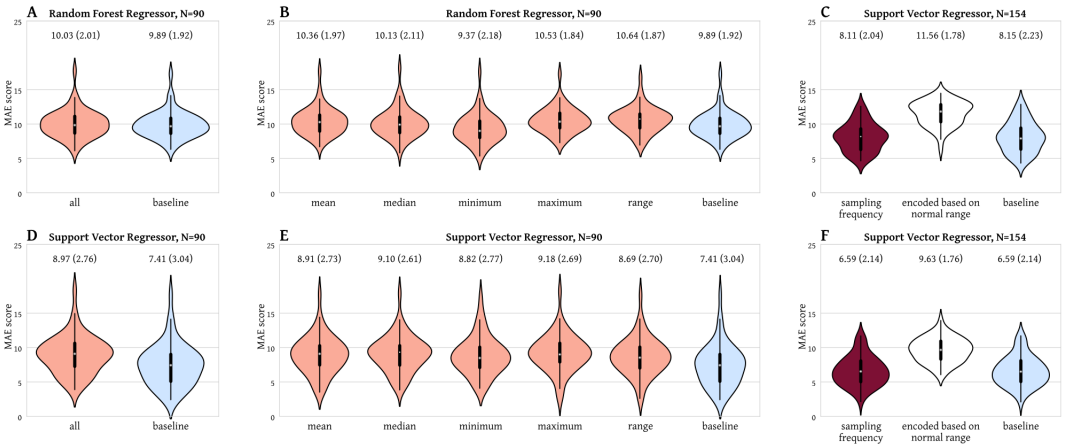


Figure 5.2: Best performing models, scored with MAE, in the different cohorts in the non stratified and stratified models, compared to the baseline model of the specific cohort (in light blue). In orange are the mean, median, minimum, maximum, range of the serological markers and the cohorts which includes all features. Dark red marks the sampling frequency cohort, while the cohort encoded based on normal range is displayed in white. **A.** RF regressor, non stratified, comparing the cohort which includes all features and the baseline; **B.** RF regressor, non stratified, comparing the mean, median, minimum, maximum and range cohort to its respective baseline model; **C.** SVM regressor, non stratified, with the sampling frequency cohort and the cohort which encodes the serological marker based on its normal range; **D.** SVM regressor, stratified, with the cohort which includes all the features compared to its respective baseline; **E.** SVM regressor, stratified, with the serological marker mean, median, minimum, maximum and range cohort; **F.** SVM regressor, stratified, sampling frequency cohort and the cohort which encodes the serological marker based on normal range; mean (SD) is the mean and standard deviation of the MAE across 50 random seed iterations; mean absolute error (MAE), random forest (RF), support vector machines (SVM)

In the mean cohort the stratified LASSO average MAE score is 9.39 (1.92), compared to the MAE score of 11.38 (1.70) with the non stratified LASSO model. The respective baseline has a stratified MAE score of 9.33 (1.91) compared to the non stratified MAE value of 11.06 (1.73). As seen in **Figure 5.3**, most stratified models (5 out of 8 regression models) significantly outperform the non-stratified models on the same cohort when evaluated with the MAE metric. All results are presented in the **Supplementary Table 8.15**.

5.4.5 Introduction of random noise

When introducing a noise variable to the best performing combination between cohort (i.e., cohort based on sampling frequency), and regression model (i.e., stratified SVM regressor), the mean MAE score increases from 6.59 (2.14) to 6.62 (2.13). This increase is statistically significant (dependent t-test, $t = 3.58$, $p = 0.0007$). Additionally the mean feature importance of noise in the model stratified based on a very acute LEMS score above zero is -0.18 (0.27), putting it in a similar range as the feature importance from

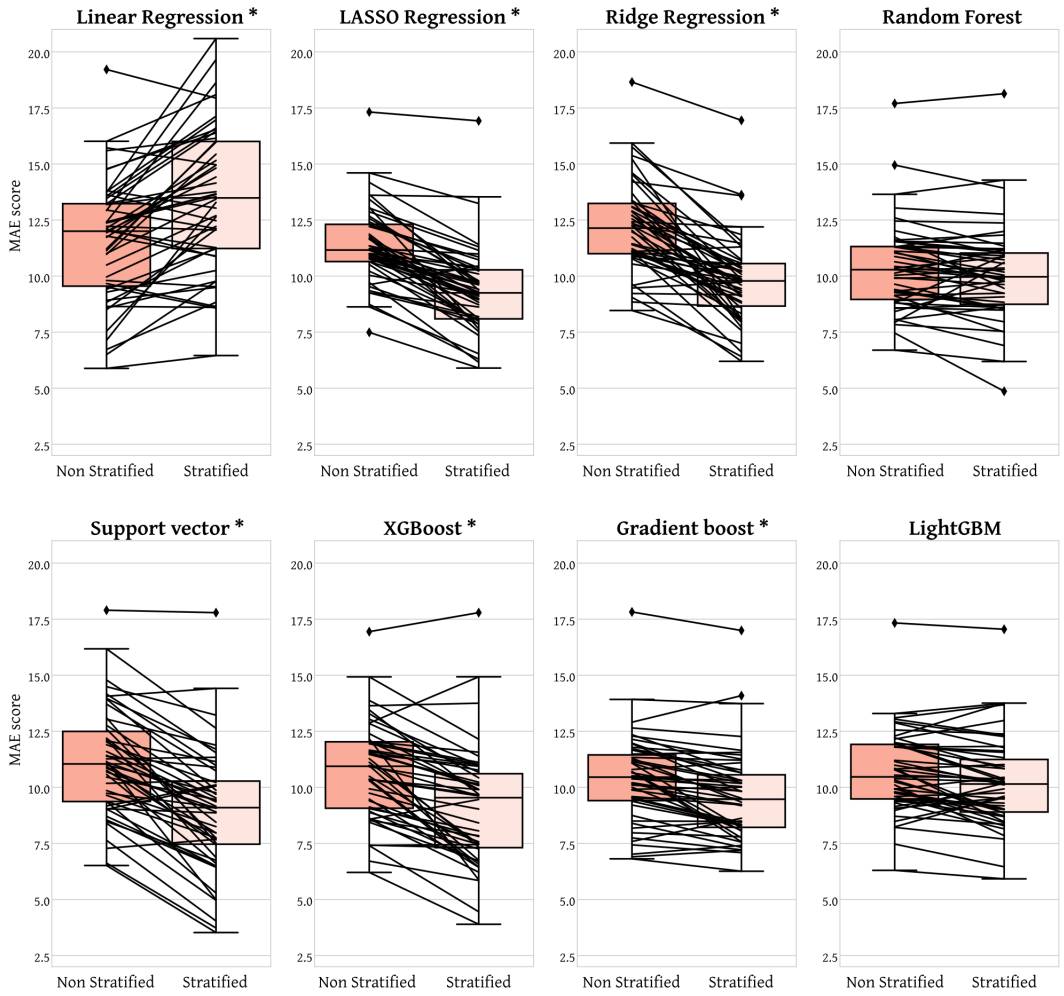


Figure 5.3: Comparison of the mean cohort across non stratified and stratified regression models. In darker orange is the performance of non stratified models, in lighter orange the performance of stratified models, the lines connect the various seed iterations. Significant differences between MAE scores (linear regression: $t = -6.00$, $p < 0.001$, LASSO regression: $t = 11.80$, $p < 0.001$, ridge regression: $t = 9.67$, $p < 0.001$, SVM regressor: $t = 9.36$, $p < 0.001$, XGBoost: $t = 7.87$, $p < 0.001$, GBR: $t = 5.82$, $p < 0.001$) between the two distributions, calculated with a dependent t test, are indicated with an $*$; mean absolute error (MAE), least absolute shrinkage and selection operator (LASSO), support vector machines (SVM), extreme gradient boosting (XGBoost), gradient boosting regressor (GBR), light gradient boosting machine (LightGBM).

serological information (0.12 (0.19), for the sampling frequency of erythrocytes). In comparison, this model showed a feature importance of 4.42 (0.78) for the very acute LEMS score.

5.5 Discussion

The objective of our study was to determine whether serological marker information can predict LEMS at the chronic stage after SCI.

In the used data set we failed to demonstrate a benefit by including serological markers for the prediction of chronic stage LEMS indicated by no statistically significant improvement in performance in models including serological markers relative to the corresponding baseline models.

Interestingly, stratifying the regression models based on the bimodal LEMS distribution observed at the very acute stage improved the regression models' performance, independent of the set of input features. One possible explanation is that the serological markers are not representative of the severity of the SCI itself, but rather of the severity of the overall trauma. Patients with a very acute LEMS score of 0, had more severe injuries, which is associated with serological marker values deviating more from the normal range compared to the less severe injuries [258]. By stratifying the cohorts before fitting, the variability in serological marker values driven by the initial severity of the injury is reduced, which may be reflected in the prediction. A second explanation for the improved scores with the stratified models, is that we are reducing the variability in the outcome score, which would explain the more accurate predictions. For example, in the sampling cohort (n = 154), patients who have a very acute score of 0 have a mean (SD) chronic LEMS score of 7.88 (15.61), whereas the patients who have a very acute score above 0, have a mean (SD) chronic LEMS of 43.97 (10.25). This can be compared to the entire cohort's mean chronic LEMS of 25.69 (22.40).

We initially hypothesized that the frequency of blood draws could be indicative of injury severity. Thus this feature would have offered predictive value for the chronic LEMS score, which is enabled by the observational nature of the data. For example, higher numbers of white blood cells entries could be indicative of infectious episodes and correlated to injury severity, as more severe injuries tend to lead to more infections [268]. However, this hypothesis could not be validated with the sampling frequency cohort performing in a comparable manner to its respective baseline. This result was further verified by the feature importance of sampling frequencies being similar to an added random noise feature. It should also be reiterated that this model only has two more features compared to its respective baseline regression.

5.5.1 Limitations

Our study was primarily limited by the small cohort size, which is a typical problem for statistical and predictive studies done with rare and heterogenous diseases [269]. In order

to mitigate this, we used data imputation at the very acute stage and at the chronic stage, which resulted in 36 additional patients meeting the inclusion criteria. Additionally, the different feature engineering strategies for the serological markers cohort enabled us to to include all 154 patients in two cohorts: the sampling frequency cohort and the cohort which encodes the serological markers based on their normal range. Furthermore, by using LEMS as the outcome score, we were able to include both paraplegic and tetraplegic patients within our study. However, although we used three different approaches to mitigate the small cohort size, the largest cohorts only included 154 patients.

Due to the different feature engineering strategies the cohorts are composed of different numbers of patients, which in turn has an influence on the interpretation of our results. As shown by the differences in prediction performance between baseline cohorts, results obtained on different cohorts are not readily comparable given the small cohort sizes. This lack of comparability is justified to maximize the sample size in each cohort. Additionally, it should be noted that in the cohort that encodes the serological value based on the normal range, 415 serological features for 154 patients were included ($p > n$). This introduces a challenge when modeling, where we risk overfitting on our training set leading to increased error rates [270].

In this study we considered LEMS as the outcome to be predicted. More precisely, we examined LEMS as a continuous score, implicitly assuming a linear scale in LEMS improvement, which is not clinically accurate (e.g., recovering from a LEMS score of zero to five is less likely than recovering from 45 to 50). Moreover, by its composite nature, the LEMS score is masking heterogeneity encountered in the recovery (e.g., two patients with an improvement from zero to eight might clinically present differently).

While the initial findings are promising, they do not yet meet the required level of accuracy for clinical application. With models leading to average errors of seven to 11 points in predicting a LEMS score ranging from zero to 50, the current error range is equal to or larger than the clinically accepted threshold of five, which signifies a clinically-relevant improvement in LEMS score [271].

5.5.2 Conclusions

Despite considering a large panel of prediction models and feature selection, this work failed to show significant improvement from serological markers collected in the first seven days following SCI in predicting LEMS 52 weeks after injury. However, considering that SCI is a highly heterogeneous disease, stratifying prediction algorithms by the very acute LEMS is a promising strategy to increase prediction performance. Here we demonstrated that stratification on the bimodal very acute LEMS score enabled more accurate prediction. This approach could be used for other diseases, especially outcomes

that show non-normally or skewed distributions. In the future it would be interesting to include other biomarkers (e.g., GFAP and NF-L). Although not routinely collected, they hold the potential of a more precise representation of the spinal injury itself since they are collected closer to the injury site compared to biomarkers present in the serum. Furthermore, future studies might consider constructing new outcomes based on the LEMS which would take into account its sequential nature and hence potentially contribute to improving prediction performance.

Acknowledgement

The authors acknowledge the Murnau center and patients involved collection of data necessary for this study.

Funding

This study was supported by the Swiss National Science Foundation (Ambizione Grant, #PZ00P3_186101 and #IZLIZ3_200275, Jutzeler), Wings for Life Research Foundation (#2017_044, #ID 2020_118, Jutzeler and Kramer), the International Foundation for Research in Paraplegia (#P192, Brüningk and Jutzeler).

Declaration of competing interest

The authors do not report any (financial or otherwise) conflict of interest.

Authors' contribution

Jan Matthias: data preprocessing, data analyses, visualization and interpretation of the data, and drafting of the manuscript
Louis P. Lukas: interpretation of the data, and revising of the manuscript for intellectual content
Sarah C. Brüningk: interpretation of the data, and revising of the manuscript for intellectual content
Lukas Grassner: primary data collection, and revising of the manuscript for intellectual content; Lucie Bourguignon: study concept/design, data analyses, interpretation of the data, and drafting of the manuscript
Catherine R. Jutzeler: study concept/design, interpretation of the data, and revising of the manuscript for intellectual content

Chapter 6

Studying missingness in spinal cord injury data: Challenges and impact of data imputation

---

Adapted from:

Lucie Bourguignon, Louis P. Lukas, James D. Guest, Fred H. Geisler, Vanessa Noonan, Armin Curt, Sarah C. Brüningk‡, Catherine R. Jutzeler‡

BMC Medical Research Methodology, 2024

Publication: 10.1186/s12874-023-02125-x
GitHub repository: https://github.com/lbourguignon/missingness-in-SCI-data

Lucie Bourguignon was responsible for the design and execution of the project, the interpretation and visualization of the results and writing of the manuscript, with continuous guidance from Louis P. Lukas, Sarah C. Brüningk and Catherine R. Jutzeler. Further details can be found in **Section 6.5.2**.

6.1 Abstract

Background

In the last decades, medical research fields studying rare conditions such as spinal cord injury (SCI) have made extensive efforts to collect large-scale data. However, most analysis methods rely on complete data. This is particularly troublesome when studying clinical data as they are prone to missingness. Often, researchers mitigate this problem by removing patients with missing data from the analyses. Less commonly, imputation methods to infer likely values are applied.

Objective

Our objective was to study how handling missing data influences the results reported, taking the example of SCI registries. We aimed to raise awareness on the effects of missing data and provide guidelines to be applied for future research projects, in SCI research and beyond.

Methods

Using the Sygen clinical trial data (n = 797), we analyzed the impact of the type of variable in which data is missing, the pattern according to which data is missing, and the imputation strategy (e.g., mean imputation, last observation carried forward, multiple imputation).

Results

Our simulations show that mean imputation may lead to results strongly deviating from the underlying expected results. For repeated measures missing at late stages (≥ 6 months after injury in this simulation study), carrying the last observation forward seems the preferable option for the imputation. This simulation study could show that a one-size-fit-all imputation strategy falls short in SCI datasets.

Conclusions

Data-tailored imputation strategies are required (e.g., characterisation of the missingness pattern, last observation carried forward for repeated measures evolving to a plateau over time). Therefore, systematically reporting the extent, kind and decisions made

regarding missing data will be essential to improve the interpretation, transparency, and reproducibility of the research presented.

6.2 Introduction

In the era of big data, medical research fields are facing a data challenge. The surge of new mathematical and statistical methods promises to help understand the progression of patients' recovery following a medical event, improve diagnosis and prognosis, thereby enhancing patients' care. However, such models require sufficient data, preferable in the magnitude of thousands of entries, to identify recurring patterns and infer prediction rules. In a number of medical fields, such as the ones studying rare conditions (e.g., spinal cord injury (SCI)) or rehabilitation, the sample size available is typically smaller and further limited by the presence of missing data, with only a fraction of the overall data being available. With its low prevalence and particular recovery pattern (i.e., time of onset precisely defined followed by recovery which plateaus between six to 12 months after the initial event), traumatic SCI constitutes an ideal study case for missing data, which can be transferred to other medical fields. The last few decades saw the emergence of SCI datasets, such as the European multicenter study on human spinal cord injury (EMSCI) [1] or National Spinal Cord Injury Model Systems [2], including over 5,000 and 50,000 patients, respectively, partially filling the gap of data availability. However, these registries, like most medical data, are prone to missing entries (e.g., patients lost to follow-up, incomplete data entry, injury conditions making it impossible to perform certain tests, different medication schemes etc.).

According to Rubin [272], missing data is categorized into three patterns, missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) (see **Table 6.1**). More precisely, MCAR refers to values, which are missing not only independently of their true unknown value, but also of the value of the other variables present in the data. In other words, data MCAR are equivalent to sampling a representative subset of the complete population. When data is MAR, a missing entry is not directly related to the underlying value, but related to other variables collected along with the variable in which missing data is observed, i.e., the proportion of missing entries differs between identifiable subgroups in the data. Finally, data are MNAR when the underlying missing value is directly related to the entry being missing. Previous studies have shown that MNAR could lead to biased interpretation of the results of statistical analysis [273, 274, 275, 276]. Bias is defined as a deviation from the truth (e.g., either over- or underestimating an effect) which can lead to erroneous conclusions [277]. This phenomenon is important when dealing with medical data, as they are prone to data MAR and MNAR [278, 279].

[1] http://emsci.org/
[2] https://msktc.org/about-model-systems/sci

Table 6.1: Missing data patterns. lower extremity motor score (LEMS)

Pattern	Definition	Example
Missingness completely at random (MCAR)	Values are missing independently of their true unknown value and independently of other variables	A LEMS value is missing for participant A with no underlying reason
Missingness at random (MAR)	Values are missing independently of their true unknown value but the pattern depends on other variables	A LEMS value is missing for participant A because they had a cast at the time of assessment, i.e., knowing the cast status gives information on whether LEMS value will be missing or not
Missingness not at random (MNAR)	Having a value missing depends on the true unknown value	A LEMS value is missing for participant A because their injury was so severe that they could not come for the assessment, i.e., the underlying true LEMS gives information on whether LEMS value will be missing or not

Independent of the missing data pattern, incomplete reports often lead to the exclusion of patients as most mathematical models require so called "complete data", effectively performing complete case analysis (CCA). This does not only represent a missed opportunity to benefit from the entire sample available, but can also lead to conclusions that are not representative of the entire population, and/or transferable to other populations. Despite those limitations, CCA is the most frequent strategy applied when handling missing data in SCI registries, although the resulting limitations are not always explicitly acknowledged [280, 55, 115]. It has been shown that this strategy, when applied to other medical research questions, could introduce bias in the results reported [281, 282]. Beyond performing a complete case analysis, there exist multiple ways of handling missing data. Imputation, in particular, refers to the procedure of inferring likely values of the missing entries [283]. These strategies can be categorized into single or multiple imputation, which would infer one or multiple likely value(s), respectively. Likewise, imputation methods can consider only one variable (e.g., mean imputation) or multiple variables at a time (e.g., model-based imputation such as predictive mean matching (pmm)). Previous studies have reported better performances of multiple imputation compared to single imputation strategies when data was missing in a human immunodeficiency virus (HIV) cohort [284] or in oncogene expression profiles [285]. Those results are in line with the underlying motivation for multiple imputation. Having multiple plausible imputed values allows to take into account the uncertainty when estimating missing values. On the other hand, single imputation might impute falsely precise values [283].

A particularity of traumatic SCI disease progression is that patients do recover to some extent over time. Most of the recovery takes place in the first six months after injury followed by a plateau between six and 12 months after injury [262]. The recovery is characterized by non-linear and highly heterogeneous recovery patterns. Owing to a scarcity of studies, the effect of missing data and imputation is not well understood for SCI datasets. Importantly, other medical scenarios involving repeated measures may show a similar plateau in the evolution of variables over time (e.g., observational studies characterising recovery in rehabilitation centers following stroke [286] or traumatic brain injury (TBI) [287], partial recovery following relapses in multiple sclerosis [288]).

To address this knowledge gap, we designed a simulation study characterizing the impact of three key parameters on the results reported, namely the variable in which data is missing, the pattern of missingness, and finally the imputation strategy applied. Firstly, considering the recovery pattern following SCI, we hypothesized that performing an imputation by last observation carried forward (LOCF) for the outcome variable evaluated at week 52 would not significantly affect the models' outcomes. However, we expected

that carrying an observation from earlier time points (e.g., 16 weeks post injury) would introduce bias in the interpretation of a model owing to the non-linearity of the recovery trajectory. Secondly, we suspected that, while CCA is an efficient and unbiased way of handling missing data when it is MCAR, it would introduce bias when data is MAR or MNAR in the field of SCI as well. When data is MAR or MNAR, we hypothesized that multiple imputation strategies, which consider the uncertainty in the imputation process, would outperform ad-hoc and single imputation strategies. Finally, we hypothesized that mean imputation is not a suitable strategy to handle missing SCI data, regardless of the missingness pattern, since the assumption of normally distributed data is not met for many SCI-related outcomes, such as the lower extremity motor score (LEMS).

Overall, our study evaluates extensively the impact of missingness on the analysis of medical data, taking the example of SCI. Using data from the Sygen clinical trial, a well established SCI data source, provides an opportunity to reconsider the importance of missing data when studying SCI data and beyond.

6.3 Methods

6.3.1 Data source

Sygen cohort

The Sygen project was a multicenter, randomized, double-blinded clinical trial conducted between 1992 and 1998 in the United States of America (USA), to evaluate the effect of gangliosidosis 1 (GM-1) on recovery following acute SCI [33, 117, 118]. Failing to demonstrate superiority over placebo in terms of recovery following SCI, the Sygen study has emerged as a valuable data source for research projects owing to the diligent data collection and the size of the cohort, which is considerably larger than many contemporary cohorts [166]. All enrolled patients were treated with methylprednisolone sodium succinate (MPSS) according to the national acute spinal cord injury study (NASCIS) II protocol as part of the standard of care [29]. The design of this clinical trial included the assessment of neurological status at predefined time points. A baseline measurement (before 72 hours from injury and after the competition of the NASCIS II [34]), 4, 8, 16, 26, and 52 weeks following injury. The delayed baseline exam was centered around 48 hours after injury. This time delay in baseline exam allowed a complete neurological examination, also considering any recovery from hemodynamic normalization occurring between the emergency room and 48 hours after injury. Among other variables, neurological level of injury (NLI), motor scores (LEMS and upper extremity motor score (UEMS)), sensory scores (pin prick and light touch) [15] and the American spinal injury association (ASIA) impairment scale (AIS) [289] were reported. Overall, the cohort includes 797 participants, with a majority of severe injuries (AIS A, 56%).

6.3.2 Simulation study

We conducted a simulation study where missing values were artificially introduced in data otherwise complete. We assessed three key characteristics of the missing data: the type of variable in which data is missing (i.e., outcome versus explanatory variable), the patterns of missingness and the imputation strategy. We summarized the simulation study in **Figure 6.1**.

Definition of the bootstrap subsets

We first selected all patients, who had data for LEMS at delayed baseline exam (referred to as “baseline”) stage and chronic/52 weeks stage (referred to as “chronic”), as well as

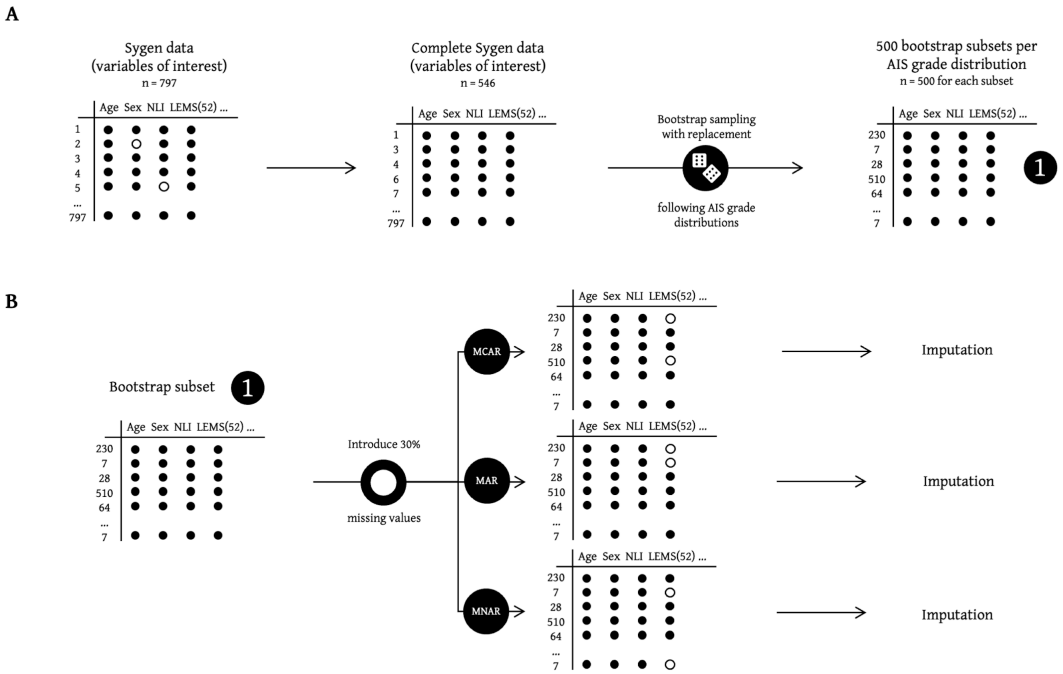


Figure 6.1: Simulation study overview. A. The first step leads to the definition of 500 bootstrap subsets, with $n = 500$ in each subset; **B.** In each bootstrap subset, 30% missing data is introduced in the variable for lower extremity motor score (LEMS) (either at baseline or at recovery) according to one of the three patterns of missingness (missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAR)), independently, before being imputed. Empty circles represent missing entries, while plain circles represent known values. neurological level of injury (NLI)

AIS grade, NLI, sex and age at baseline. To emulate a plausible research hypothesis, we considered the following model:

$$LEMS_{chronic} \sim LEMS_{baseline} + AISgrade_{baseline} + NLI + age + sex \qquad (6.1)$$

where we intended to study the association between LEMS at the chronic stage (outcome variable) and LEMS at baseline (explanatory variable), taking into account potential confounders, such as the AIS grade, NLI, age, and sex at baseline. Note, for simplicity LEMS scores were considered to be continuous scores and NLI a binary variable, taking either the value "cervical" or "thoracic". Patients with lower injuries (i.e., at and below L1) were excluded in the original study [33]. In order to assess how variable the effects of missing data and data imputation are, we performed a bootstrap sampling with replacement to create 500 bootstrap subsets ($n = 500$ entries for each) with fixed AIS grade distributions (**Figure 6.1A**). The distributions followed either the original AIS grade distribution from the complete Sygen data for the variables of interest, or with balanced AIS grade groups (25% of grade A, B, C and D in the final cohorts, AIS grade E was not

included as this category is not present in the original study [33]).

Introduction of missing data

In each bootstrap subset, we introduced 30% missing values in two of the variables, namely LEMS at chronic stage and LEMS at baseline (**Figure 6.1B**). This percentage was chosen based on the percentage of missing data observed in the Sygen cohort (24.1%) and was set to a higher, more conservative percentage. In this study, we focused on simulations in which missing data would be introduced in one variable at a time, whilst the rest of the variables would be complete, as a way to simplify the task at hand. The choice of these variables was motivated by their different status in the example model (outcome and explanatory, respectively).

Missing values were introduced according to three patterns: MCAR, MAR, and MNAR as described in Rubin in 1976 [272]. For modeling MCAR, if LEMS at baseline is missing for a specific patient, the missing value would be unrelated to all other variables, including the outcome variable LEMS at chronic stage (i.e., 52 weeks after injury). As such, disregarding those entries should not introduce bias, provided a sufficient sample size [290]. In the case of values MAR, information about the missing value can be retrieved by studying the other variables. To simulate this behavior values MAR were introduced depending on the variable sex, where being male made it twice more likely to have a value missing, compared to being female. Finally, MNAR is a pattern, in which the unknown true value influences whether the value is missing or not. In this study, we simulated that patients with less severe injuries would be more likely to be missing. Specifically, high LEMS (i.e., above the LEMS 30th percentile) were four times more likely to be missing compared to low LEMS (i.e., below the LEMS 30th percentile). The four times difference reflects the four AIS grade categories (from A to D), closely related to the LEMS [291]. The 30th percentile threshold was chosen to match the 30% missing data introduced, easily allowing for a change in percentage of missing data introduced in future studies.

Imputation strategies

The introduced missing values were imputed with three types of procedures: ad-hoc methods, single imputations and multiple imputations.

Ad-hoc methods included mean imputation [292] and LOCF [293]. The latter was used for imputation of the outcome variable only (LEMS at chronic stage), where missing data were replaced by LEMS assessed 26 weeks after injury. Intending to test the time sensitivity of the LOCF, we repeated the analysis using LEMS at 26 weeks as the primary outcome variable, and imputed it using LEMS available at week 16. We hypothesized, based on the recovery profile following SCI [262], that LOCF from week 26 to chronic stage

would be more relevant than LOCF from week 16 to week 26, where a substantial amount of recovery is still likely to occur. We focused our analyses on outcomes measured at week 26 and week 52 after initial injury as they are the reference timepoints used in clinical trials to assess recovery following SCI [33] [41].

Single imputation consisted of three main steps: (i) taking the set of patients, which are not missing for a defined variable; (ii) fitting a model to describe this variable according to all others; and (iii) predict likely values for the missing ones, based on the fitted model. For example, if one imputes missing LEMS score at baseline, the fitted model would be:

$$LEMS_{baseline} \sim AISgrade_{baseline} + NLI + age + sex \tag{6.2}$$

Note that we excluded the outcome variable as it represents information that is not available at baseline. Different models can be used to fit the data available for imputation. In this simulation study, we focused on linear regression (LR) [294], k-nearest neighbors (k-NN) [295], and support vector machines (SVM) [296] using two types of kernel (linear and radial basis function (RBF)) and random forest (RF) [297], as they represent a commonly used set of machine learning models for prediction tasks [298]. All models included a 5-fold cross-validation scheme for hyperparameter optimization. The corresponding parameter grids can be found in **Additional File 1**.

Single imputation is inherently limited as it does not provide uncertainty related to the imputed value. Multiple imputation addresses this challenge: the imputation is performed multiple times (25 times here, as a compromise between increased power and reasonable run time [299]) before being pooled. Similarly as for single imputation, the outcome variable was excluded from the imputation of the explanatory variable. Models including pmm, LR imputation (norm.predict) and tree-based method (RF) were chosen, as implemented in the R `mice` package [300]. Models for multiple imputation were chosen to match the models used in the single imputation with the aim to increase the comparability between the two approaches.

Based on our study design, we chose to pool the data before fitting the example model featured in **Equation 6.1**, with the final imputed value being the mean of all imputed values for the LEMS continuous variables [301]. This approach was taken in order to obtain a single imputed value for each missing entry to allow for the computation of metrics (see 6.3.3). However, it does not match the flow advocated in the implementation for multiple imputation as presented in the R `mice` package. In order to ensure that this change in procedure does not impair the outcome of the multiple imputation, we compared both approaches in the **Additional File 2**.

Finally, we performed a CCA, where any patient (case) with at least one missing value

among the variables described in **Section 6.3.2** would be disregarded and the analysis performed solely using patients for whom the entire set of variables was observed.

All imputation strategies were compared to their corresponding bootstrap subset when complete, designated as baseline subset.

6.3.3 Evaluation of data imputation

Following imputation, we sought to evaluate and compare the different imputation strategies tested. We employed various methods to both examine population- (i.e., statistical tests, β coefficient comparisons) and individual-level (i.e., metrics) performance of imputation methods in restoring the missing entries.

Statistical tests

Two-sample Kolmogorov-Smirnov test

We tested the null hypothesis considering that the two sets of observations were drawn from the same unknown probability distribution, using a two-sample Kolmogorov-Smirnov test [302], as implemented by the `ks.test` function in the `stats` R package. The two sets of observations considered were either the variable before and after introducing the missing values, or the variable before introducing the missing values compared with the variable after imputation, or the variable after introducing missing values and the variable after imputation. Note that in a typical imputation situation, true values are not available, and thus only comparison between the set of non-missing values and the set of values after imputation would be possible.

chi-squared (χ^2) goodness of fit test

The χ^2 goodness of fit test, `chisq.test` in R, was employed to compare the proportion of categorical variables between two cohorts. Its null hypothesis states that the sample to be tested follows the hypothesized distribution from the other cohort.

Little's test

Little's test was first described in 1988 [303]. It tests the null hypothesis that data is missing completely at random in a given cohort. In our framework, the `mcar_test` function, implemented in the `naniar` R package [304], first allowed us to ensure that the missing data was introduced as intended, i.e., MCAR or not (**Section 6.4.1**), and was further used to describe the missingness in the original Sygen cohort (**Section 6.4.3**).

Metrics

We used the following metrics for a quantitative comparison of variables, continuous and categorical, in their complete version versus after imputation. All imputation methods

were subsequently ranked to determine, for each metric, which imputation method would consistently lead to imputations closer to the true values across repeated runs.

mean absolute error (MAE)

The MAE computes the average absolute difference between a known true value y_i and its corresponding imputed version $\hat{y}_i$ for all n entries i for which missingness was introduced:

$$MAE = \frac{1}{n}\sum_{i}^{n}|y_i - \hat{y}_i| \tag{6.3}$$

MAE is a negatively-oriented score, which means lower values indicate better imputation performance. This metric has the advantage of being intuitively interpretable as it is expressed in the units of the variables, i.e., a MAE of 3.5 for LEMS at baseline would mean that, on average, the imputed values for LEMS missing at baseline are 3.5 points away from their true values.

root mean squared error (RMSE)

The RMSE differs from the MAE as it squares the difference between true and imputed values, thus penalizing large errors more. By taking the square root of the overall average of differences, it allows one to interpret the RMSE on the scale of the initial values, similarly to the MAE. Likewise, a RMSE of 0 corresponds to the best possible imputation.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(y_i - \hat{y}_i)^2} \tag{6.4}$$

Comparison of β coefficients after linear regression (LR) using imputed data

The last method we employed to assess the quality and impact of imputation was to fit a linear regression (LR) based on the simulated research question stated in **Equation 6.1** and compare the β coefficients for the explanatory variables estimated from a LR based on the complete set of data and the imputed data. This method allowed us to highlight the difference in the conclusion drawn from a research question according to its study design regarding the way to handle missing data. We considered the 95% confidence interval (CI) and mean difference in β coefficients for each explanatory variable (i.e., LEMS at baseline). For an imputation method to be considered unbiased, the CI should include the value 0 (i.e., it is likely that the true difference between the β coefficients is negligible) and be as small as possible.

For all tests, the threshold of $p < 0.05$ was considered significant and led to rejecting the corresponding null hypothesis. Analyses were performed with R Statistical Software (version 3.6.0) and Python (version 3.7.4).

6.4 Results

6.4.1 Description of the data

Full cohort and selected complete case cohort from the Sygen trial

Summary statistics of the variables of interest for our simulation study are presented in **Table 6.2**. After including only complete cases for the variables of interest, the cohort was reduced from 797 to 546 patients. Comparing the two cohorts did not yield significant differences in terms of the proportion of sex (χ^2 test, $\chi^2 = 0.66$, df = 1, p-value = 0.42), age (two-sample Kolmogorov-Smirnov test, $D = 0.02$, p-value = 0.99), level of injury (χ^2 test, $\chi^2 = 0.25$, df = 1, p-value = 0.62) or LEMS at baseline and at recovery (two-sample Kolomogorov-Smirnov test, $D = 0.01$, p-value = 1, for both variables). When comparing the proportions of AIS grades and considering missing data as a category in itself, which would not be present by design in the cohort with only complete data, a significant difference is reported between the two cohorts (χ^2 test, $\chi^2 = 73.96$, df = 4, p-value < 0.001). Since this difference is likely to be driven by the additional missing category, we performed the same test using only the actual grades available. It revealed no significant difference in the proportions of each grade between the two cohorts ($\chi^2 = 0.84$, df = 3, p-value = 0.84).

Subsets from the cohort of complete cases

Variables of interest are summarized for every AIS grade distribution in **Table 6.3**. Each value is reported as the mean of the variable's values across the 500 subsets drawn according to the same AIS grade distribution as in the cohort with complete cases from the Sygen data, or with balanced AIS grade groups.

In order to test whether the missing data were introduced as intended (i.e., following MCAR, MAR and MNAR patterns, respectively), we performed a Little's test for each subset and for each variable in which missing data was introduced, separately. As expected, the null hypothesis, stating that the data is MCAR, is mostly rejected when missingness is introduced at random or not at random (range: 491-500 subsets out of 500, **Additional File 3**). When missingness is introduced completely at random, it is expected that the null hypothesis would be rejected in 5% of the 500 subsets since we defined our significance threshold to be less than 0.05. That represents a 5% probability that the null hypothesis, whilst being correct, is rejected. This expectation matches the observation across subsets in which missingness was introduced completely at random, with the null being rejected in 26 (5.2%) and 32 (6.4%) bootstrap subsets, depending on the AIS distribution

Table 6.2: Characteristics of the Sygen cohort for the variables of interest, before and after selecting for complete cases.

		Entire cohort	Complete cases only	p value
Number of patients	*n*	797	546	
Sex				0.42
	n (% male)	643 (80.7)	433 (79.3)	
	NA, *n* (%)	0 (0.0)	0 (0.0)	
Age				1
	mean (SD)	32.5 (13.4)	32.0 (13.3)	
	NA, *n* (%)	0 (0.0)	0 (0.0)	
LEMS at week 01				1
	mean (SD)	2.7 (7.2)	2.7 (7.1)	
	median [Q1 - Q3]	0 [0-0]	0 [0-0]	
	NA, *n* (%)	74 (9.3)	0 (0.0)	
LEMS at week 26				1
	mean (SD)	12.1 (18.7)	11.9 (18.9)	
	median [Q1 - Q3]	0 [0-29]	0 [0-28]	
	NA, *n* (%)	168 (21.1)	27 (4.9)	
LEMS at week 52				1
	mean (SD)	12.8 (19.3)	12.6 (19.3)	
	median [Q1 - Q3]	0 [0-32]	0 [0-31]	
	NA, *n* (%)	192 (24.1)	0 (0.0)	
Level of injury				0.62
	Cervical, *n* (%)	600 (75.3)	406 (74.4)	
	Thoracic, *n* (%)	197 (24.7)	140 (25.6)	
AIS grade				0.84
	A, *n* (%)	446 (56.0)	356 (65.2)	
	B, *n* (%)	77 (9.7)	59 (10.8)	
	C, *n* (%)	149 (18.7)	108 (19.8)	
	D, *n* (%)	31 (3.9)	23 (4.2)	
	NA, *n* (%)	94 (11.8)	0 (0.0)	**< 0.001**

Significant p values are highlighted in **bold**.
Continuous variables (LEMS and age) were compared using a two-sample Kolmogorov-Smirnov test, categorical variables (sex, level of injury and AIS grade) were compared using a χ^2 goodness of fit test.
lower extremity motor score (LEMS); American spinal injury association (ASIA) impairment scale (AIS); not available (NA); standard deviation (SD); first quartile (Q1); third quartile (Q3)

Table 6.3: Characteristics of the 500 bootstrap subsets (500 entries each) created according to the American spinal injury association (ASIA) impairment scale (AIS) grade distributions present in the Sygen cohort including only complete cases for the variables of interest, and a balanced cohort, where all four grades are present in equal proportions.

Outcome at week 52		**Sygen subsets**	**Balanced subsets**
Number of patients	*n*	500	500
Number of male	mean (SD)	369.8 (8.6)	385.5 (9.2)
Age [years]	mean (SD)	32.0 (13.3)	34.2 (14.0)
LEMS at week 01			
	mean (SD)	2.7 (7.0)	8.5 (12.1)
	median [95% CI]	0 [0-0]	0 [0-0]
LEMS at week 26			
	mean (SD)	11.9 (18.8)	11.9 (18.9)
	median [95% CI]	0 [0-0]	34 [33-34]
	NA, *n* (%)	24.9 (5.2)	22.8 (4.6)
LEMS at week 52			
	mean (SD)	12.6 (19.3)	26.2 (21.7)
	median [95% CI]	0 [0-0]	35 [35-36]
Level of injury			
	cervical, mean (SD)	372.3 (9.8)	412.7 (8.4)
	thoracic, mean (SD)	128.7 (9.8)	87.3 (8.4)
AIS grade			
	A, *n* (%)	325 (65.0)	125 (25.0)
	B, *n* (%)	55 (11.0)	125 (25.0)
	C, *n* (%)	100 (20.0)	125 (25.0)
	D, *n* (%)	20 (4.0)	125 (25.0)

lower extremity motor score (LEMS); American spinal injury association (ASIA) impairment scale (AIS); not available (NA); standard deviation (SD); confidence interval (CI)

(**Additional File 3**). Overall, this step allows us to assume that the missingness patterns were introduced appropriately.

Following the introduction of the missing data, we evaluated the impact of the missing data on the distribution of the variable in which it was introduced. When tested with the two-sided Kolmogorov-Smirnov test, introducing MCAR and MAR did mostly not significantly change the distributions of the two variables (LEMS at baseline and recovery) (**Additional File 4**). By contrast, introducing MNAR introduced a shift in the distribution of the variables for the majority (500 and 305/500 when AIS grade distribution follows the complete Sygen data's distribution and a balanced AIS grade distribution, respectively) of the bootstrap subsets. Introducing MNAR in LEMS at recovery in a population where the proportions of AIS grades are balanced (25% for each group), was an exemption to that observation. In this particular case, the null hypothesis of the two-sided Kolmogorov-Smirnov test, stating that the values of LEMS at recovery before and after introducing MNAR were drawn from the same underlying population, was rejected for 305 subsets out of 500. In comparison, it was rejected for all subsets in a similar population AIS grade distribution, when missingness was introduced at random.

6.4.2 Performance of imputation methods

Statistical tests

The results comparing the distributions of the true and imputed values after introducing missing data are summarized in **Additional Files 5** and **6**. While introducing data MCAR or MAR did not lead to significant shifts in distributions (see **Section 6.4.1**), we observed that the imputation methods introduced shifts irrespective of the underlying AIS grade distribution in the population or the variable with missing entries. Similarly, we noted that across variables, underlying AIS grade distributions in the samples and missingness patterns, the majority or mean imputation systematically shifted the distribution of the imputed variable.

When data was MNAR, the distributions of the resulting population were often significantly different from the initial population (from 305 to 500 out 500 subsets, **Section 6.4.1** and **Additional File 4**). Following imputation, this shift was more likely to be reversed as the underlying population structure approached balanced proportions in AIS grades (e.g., 150 versus 295 subsets out of 500 had a significantly different population distribution after imputation with multiple RF when data is MNAR in LEMS at baseline, **Additional File 5**). The imputation method that led to the least number of subsets in which a shift was still observed was imputation using a RF (simple imputation, four subsets when data MNAR, **Additional File 5**), followed by pmm (multiple imputation, 14 subsets when

data MNAR, **Additional File 5**) for the LEMS at baseline. One exception arose when missingness was introduced in the outcome variables, LEMS at the chronic stage, where imputation with LOCF led to sample distributions that were never significantly different from the true population (**Additional File 6**). This observation also held true when the outcome variable was measured 26 weeks after injury and imputation was based on data collected 16 weeks after injury (**Additional File 7**). However, when both LEMS at chronic stage and week 26 were missing, LOCF could not be performed and led to the exclusion of a mean of 6.8 (standard deviation (SD): 2.5), 6.8 (SD: 2.5) and 6.5 (SD: 2.6) entries per bootstrap subset, when LEMS at chronic stage was MCAR, MAR and MNAR, respectively.

Metrics

Testing for difference in distributions is equivalent to looking at the performance of the imputation at a population level. It is, however, also interesting to see at the scale of the individual imputed values how the imputation performs. For that purpose, we computed various metrics to quantify the agreement between individual imputed values and their true counterpart, across bootstrap subsets.

Two main observations were similar to the ones obtained when comparing imputation methods at the population level by means of statistical tests. Firstly, LOCF was the imputation method leading to the lowest MAE and RMSE, when imputing the outcome variable evaluated at week 52 (**Figure 6.2A**). When the outcome was measured at week 26 after injury, LOCF was still consistently among the top four imputation methods but was outperformed by pmm (**Figure 6.2B**). Secondly, mean imputation led to the lowest ranked metric values in most of the scenarios, regardless of the other three parameters to be studied in this simulation (i.e., AIS grade distribution, missingness patterns, variables to be imputed, **Additional Files 8** and **9**, **Figure 6.2**). Multiple imputation, on the contrary, was always ranked the highest (following LOCF if present), across all metrics, with a slight advantage to pmm and norm.predict (ranked in the top two, after LOCF, in all the simulations) over multiple RF (ranked third, or fourth when LOCF is present, in over 90% of the simulations), when imputing LEMS variables (**Figure 6.2**). We also observed that the distribution of the metrics values were less variable with multiple imputation when repeating the process in 500 bootstrap subsets compared to the single imputation methods (SD of distribution of MAE when LEMS at chronic stage MAR: 0.97, 0.71, 0.24 and 0.21 when imputed using k-NN, LR, pmm and norm.predict, respectively, **Additional File 9**).

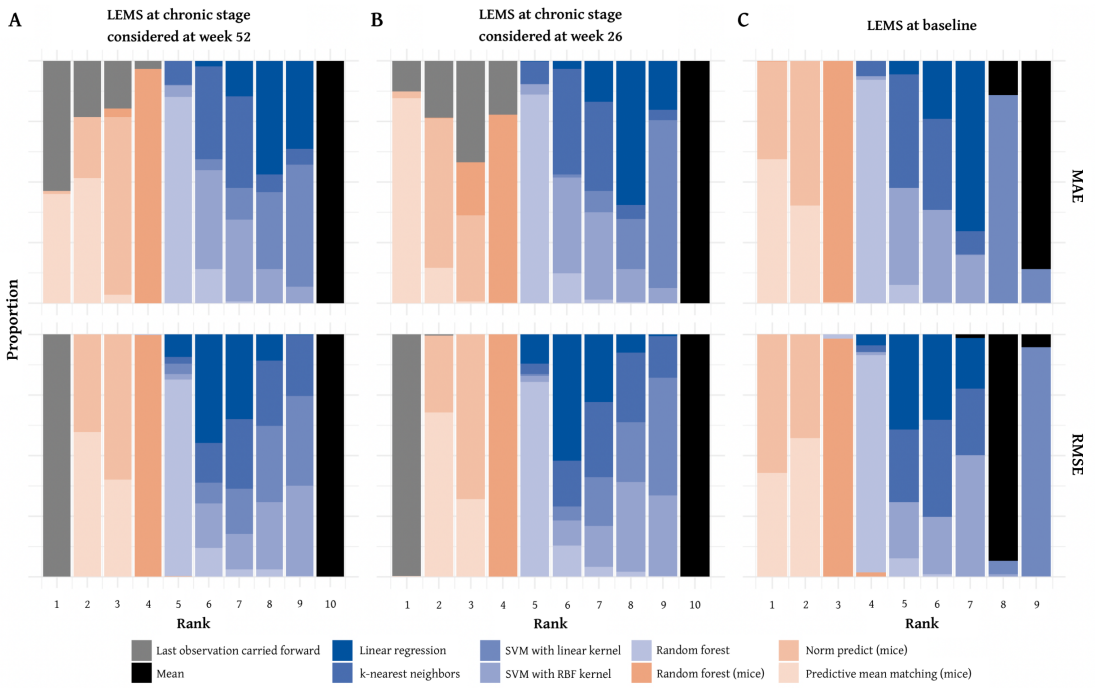


Figure 6.2: Imputation methods ranked from lowest (1) to highest (9 or 10) metrics' values when introducing missing data not at random in A. LEMS at outcome considered at week 52. **B.** LEMS at outcome considered at week 26. **C.** LEMS at baseline. For each subset ($n = 500$), missing data is introduced and imputed using all methods. Within each subset, imputation performance is compared between imputation methods and ranked from best performance (i.e., closest to 0 and ranked 1) to lowest performance (i.e., highest metric value and ranked 9 or 10). We display the proportion of subsets (out of 500) per rank and imputation method. lower extremity motor score (LEMS); mean absolute error (MAE); root mean squared error (RMSE); support vector machines (SVM); radial basis function (RBF).

Comparison of β coefficients after linear regression (LR) using imputed data

As shown in **Figure 6.3**, mean imputation for LEMS missing at baseline consistently introduced a bias in coefficients estimated via LR, with the magnitude of the bias increasing from data MCAR to MAR to MNAR (mean difference between estimates of beta for LEMS at baseline of −0.33, −0.35, and −0.50 when data MCAR, MAR and MNAR, respectively). In contrast, bias would not be introduced when performing a CCA, i.e., zero would also be present in the CI. This imputation method, however, led to wide CIs in the difference between coefficients estimated on the entire data versus on the imputed data (e.g., when estimating the effect of AIS grade D in comparison with AIS grade A, 95% CI of [−6.5; 5.3], [−5.8; 6.6] and [−30.1; 13.1] for data MCAR, MAR and MNAR, respectively). Taken together, **Figure 6.3** supports the use of multiple imputation methods such as pmm and norm.predict in imputing missing LEMS at baseline, as those methods did not introduce bias and resulted in smaller CI, especially with data MNAR ([−17.9; 13.5] for estimates of the effect of AIS grade D in comparison with AIS grade A). When imputing

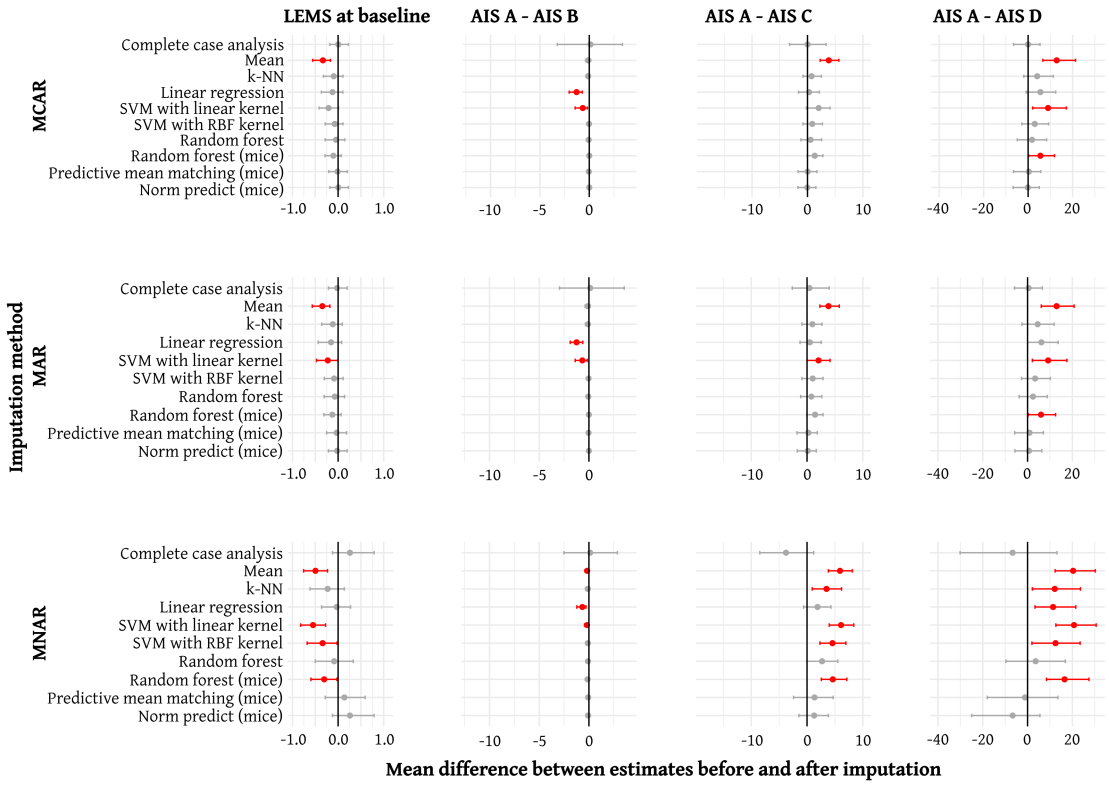


Figure 6.3: Mean and CI of the difference between estimates from the data before introducing missingness in LEMS at baseline and after imputation. Each row corresponds to missing data being introduced using a different missingness pattern (MCAR, MAR and MNAR from top to bottom). Each column corresponds to the estimate of one explanatory variable (LEMS at baseline, AIS B compared to AIS A, AIS C compared to AIS A and AIS D compared to AIS A, from left to right). Intervals displayed in red do not contain the value 0. missing completely at random (MCAR); missing at random (MAR); missing not at random (MNAR); lower extremity motor score (LEMS); American spinal injury association (ASIA) impairment scale (AIS); k-nearest neighbors (k-NN); support vector machines (SVM); radial basis function (RBF); confidence interval (CI).

missing LEMS at week 52, LOCF produced estimates of β that were both unbiased and close to the estimates derived from the entire data (**Additional File 10**). If the outcome is evaluated at week 26, imputation of missing data using LOCF uses information available at week 16. Despite using information from an earlier time in the recovery process, it appears to still be the most reliable imputation method with no bias introduced, except when data is MNAR. In that case, although the estimates repeatedly deviated from the expected ones, the bootstrap CI is tight compared to CI obtained with other imputation methods ($[-2.5; -0.3]$ and $[-3.1; -0.5]$ for the estimates for AIS B and C versus AIS grade A, respectively, **Additional File 11**).

6.4.3 Application of studying missing values to real-world data

As presented in **Table 6.2**, the full Sygen cohort (n = 797) presents missing entries for LEMS at both time points and AIS grade at week 1, when taking into account the variables studied in our example model. Sex, age and NLI, however, had no missing data. **Additional File 12** illustrates the co-occurrence of missing data across the variables considered. The hypothesis of the data being MCAR was rejected when taking all the variables of the model together (Little's test, statistic = 76.4, df = 44, number of missing patterns = 8, p-value = 0.002). The variable with most missing entries, LEMS evaluated at week 52, presents with 24.1% missing data, making our simulation with 30% missing data more conservative. Notably, both LEMS at week 52 and 26 were missing for 136 (17.1%) participants. It is important to highlight that this subset could not benefit from a reliable imputation based on the LOCF strategy. However, 56 participants (7.0%) could be included in such an analysis by imputing the missing outcome variable using the LOCF strategy. For the participants in which either LEMS or AIS grade at baseline was missing, imputation could be envisaged, preferably through multiple imputation. General consideration on how to apprehend missing data, both based on knowledge from the literature and results from the simulation study described here, are presented in **Figure 6.4**.

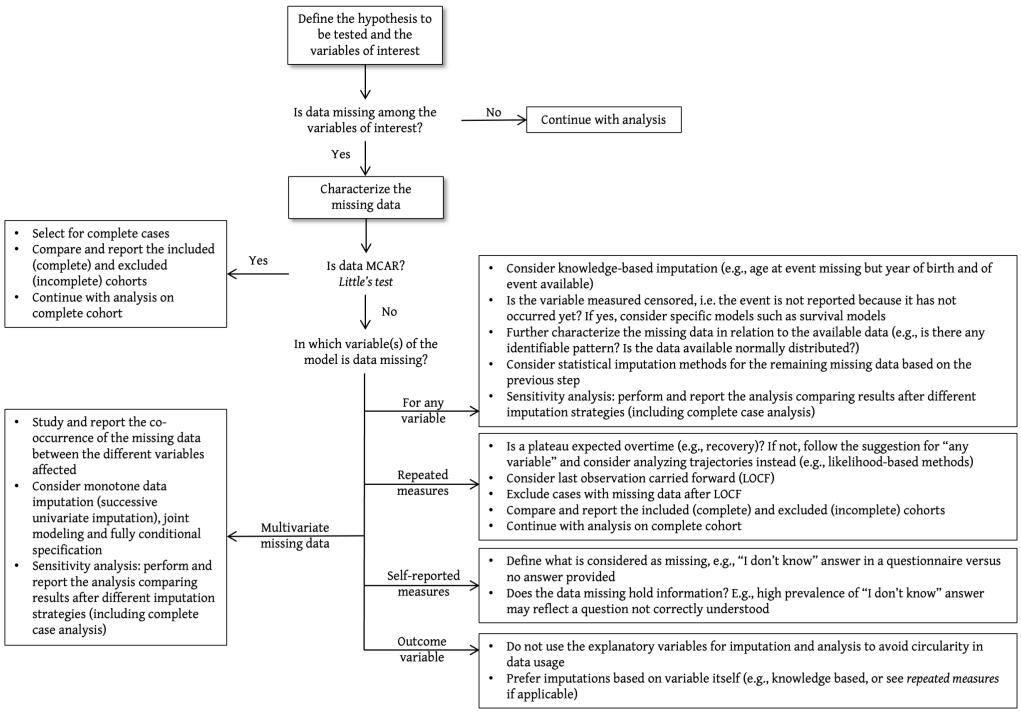


Figure 6.4: General consideration when facing missingness in medical data.

6.5 Discussion

In this simulation study, we aimed to address the impact of missing data in SCI data sources on the results reported. We specifically focused on three key components that could affect the analysis itself and the interpretation of the results: the type of variable in which data is missing, the pattern in which the data is missing (i.e., MCAR, MAR, MNAR), and the imputation strategy applied.

In agreement with reports from other medical fields [273, 305, 306], we showed that data MNAR is more likely to lead to biased subsequent analysis as it might change the distribution of the data available for analysis (**Additional File 4**). Likewise, disregarding the presence of data MNAR by performing an analysis based on complete cases can also lead to erroneous conclusions compared to an analysis that would have been performed on the entire sample with no missing data (**Figure 6.3**) considering the large CI of the difference between the true estimate and the estimate obtained from the imputed data. This point is particularly crucial as most studies currently perform complete case analysis [258, 44, 307, 308], and we reported absolute effect sizes greater than 5 (when MNAR and considering AIS D or C compared to A), surpassing the threshold of 5 points considered as clinically significant for LEMS [271]. It could also not be excluded that data was MNAR in the Sygen data (**Section 6.4.3**). The latter is likely to hold true in most SCI datasets (EMSCI database [166]) owing to the nature of the data itself (i.e., observational medical data). When dealing with MNAR, CCA did not consistently lead to the introduction of bias in the estimates of β coefficients, contradicting our initial hypothesis. We observed, however, that multiple imputation strategies, in particular pmm and norm.predict equally led to unbiased estimates but with narrower CIs, suggesting a lower variance in the estimates. Similarly, multiple imputation methods were more likely to generate distributions closer to the initial true distribution. Taken together, it seems that, for this cohort, handling MNAR with multiple imputation would be more appropriate than to use CCA.

While LOCF is only possible in the case of variables being observed at multiple time points and may not be appropriate for other medical outcomes [309, 310], our study supports the use of this imputation for SCI-related outcomes such as the LEMS. We were able to show that performing LOCF from week 26 to week 52 leads to a population similar to the true underlying population in terms of distribution (**Additional File 7**), individual values imputed and estimated β coefficients from the LR model (**Additional File 9**). This is likely attributable to the very characteristic recovery trajectory of SCI, including a plateau starting six to 12 months after the initial injury. Contrary to our initial expectation, this observation still held true when performing LOCF from week 16 to week 26 (**Additional File 10**). However, it is important to note that this conclusion

might be specific to outcomes with this particular recovery trajectory, and might not be transferable to outcomes where no plateau can be observed (both SCI-related or unrelated outcomes). LOCF is also a well suited imputation method for outcome variables as it only relies on data that will be further used at a later stage of the data analysis or modeling process. This effectively prevents introducing circularity, which in turn improves the potential transferability of the reported results to a clinical setting. However, it should be noted that in longitudinal studies, one may take advantage of the repeated measures and analyze the entire recovery trajectory rather than the mere association between baseline and chronic measures. In such cases, likelihood-based methods (e.g., mixed-effects models) would be advantageous. Indeed, they inherently allow for MCAR and MAR, or specification of the joint distribution between the data present and missing data when data is MNAR, thus not requiring imputation [311].

6.5.1 Limitations

It is important to note that the interpretation of this study might be limited by a few factors. Firstly, we studied imputation for repeated measures in the context of SCI using LOCF considering carrying forward information from week 26 to week 52, and from week 16 to week 26. However, we have not explored whether carrying forward values from earlier timepoints (e.g., week 4 or 8) would lead to equally reliable imputed values. Additionally, the exact time points to consider for a valid LOCF will depend on the spacing between repeated measures available and the expected trajectory and timeframe of the variable of interest. Secondly, we restricted our analysis to a fixed amount of missing data (i.e., 30%). This percentage was chosen based on the actual percentages of missing data observed in the variables studied in the original data used and was fixed to a higher percentage to be more conservative while being able to compare our results across variables. Thirdly, we only investigated continuous variables. Dealing with missing data in categorical variables (e.g., AIS grade, assessing SCI severity) would require the use of other models (e.g., proportional odds logistic regression for multiple imputation) and give rise to specific challenges (e.g., how to impute a category that is not present in the data but theoretically possible). Additionally, we did not study missingness in self-reported variables, which can carry information and should therefore be studied beyond imputation [312, 313]. These points have not been explored as a means to limit the complexity of our primary analysis, but constitute the starting point of future work. Finally, we focused on missing data being present in one variable at a time, i.e., univariate imputation. Investigating the multivariate missing data problem poses additional challenges, including but not restricted to combining different missingness patterns, introducing circularity when imputing outcomes based on explanatory variables, or

potentially masking meaningful information from the co-occurrence of missing entries. In such cases, imputation strategies can range from combining multiple univariate imputation (i.e., monotone data imputation), conditional univariate models or modeling the joint distributions within the entire dataset [314]. Similarly, exploring different research questions or at the scale of larger databases was beyond the scope of this initial analysis but would benefit from their own study. Accordingly, it would be interesting to extend this simulation study and further analysis of missing data using additional SCI datasets such as the EMSCI or the Rick Hansen spinal cord injury registry (RHSCIR) [315], and similar observational datasets beyond SCI such as the transforming research and clinical knowledge in traumatic brain injury (TRACK-TBI) initiative [316] focusing on TBI.

6.5.2 Conclusions

Our study raises awareness regarding the presence and impact of missing data in medical data sources (e.g., clinical trials, registries), taking the example of SCI. We demonstrated that disregarding missing data could not only result in a significant loss of information, but also lead to erroneous conclusions. Hence, we see this work as a first step towards systematically considering and reporting the presence of missing data as part of good practices in SCI data analysis and beyond.

Declarations

Ethics approval and consent to participate

Approval for this study (secondary analysis) was received by an institutional ethical standards committee on human experimentation at the University of British Columbia. The original Sygen clinical trial [117, 118, 33] also received ethical approval, but was conducted before clinical trials were required to be registered (i.e., no clinicaltrial.gov identifier available). All experiments were performed in accordance with relevant guidelines and regulations. Informed consent to participate was obtained from all of the participants as part of the original study.

Availability of data and materials

Anonymized data used in this study is available upon request to the corresponding author and in compliance with the European general data protection regulation (EU GDPR). The code describing the analysis can be accessed on our GitHub repository [3].

Competing interests

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Funding

This study was supported by the Swiss National Science Foundation (Ambizione Grant [PZ00P3_186101, Jutzeler]). SB was supported by the Botnar Research Centre for Child Health Postdoctoral Excellence Programme (#PEP-2021-1008). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The funding sources of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Authors' contribution

Lucie Bourguignon: design, execution and interpretation of the simulation study; data visualization; writing of the manuscript

[3] https://github.com/lbourguignon/missingness-in-SCI-data

Louis Lukas: substantial contributions to the design and interpretation of the simulation study; revising the manuscript critically for important intellectual content
James Guest: revising the manuscript critically for important intellectual content
Fred Geisler: revising the manuscript critically for important intellectual content
Vanessa Noonan: revising the manuscript critically for important intellectual content
Armin Curt: revising the manuscript critically for important intellectual content
Sarah C. Brüningk: substantial contributions to the design and interpretation of the simulation study; revising the manuscript critically for important intellectual content
Catherine R. Jutzeler: substantial contributions to the design and interpretation of the simulation study; revising the manuscript critically for important intellectual content

Acknowledgements

The authors would like to acknowledge and thank Ainesh Sewak and Prof. Dr. Torsten Hothorn for their statistical inputs.

Chapter 7

The concept of positive deviance applied to spinal cord injury recovery: An exploratory analysis

---

Adapted from:

Lucie Bourguignon, Louis P. Lukas, Sarah C. Brüningk, Fred H. Geisler, EMSCI study group, Catherine R. Jutzeler‡, John L.K. Kramer‡

Manuscript in preparation, 2024

Publication:
GitHub repository: https://github.com/lbourguignon/outstanding-recovery-SCI

Lucie Bourguignon was responsible for the data analyses, interpretation and visualization of the results and the writing of the manuscript, on an original idea from John L.K. Kramer. Further details can be found in **Section 7.5.2**.

7.1 Abstract

Background

The concept of positive deviance studies individuals achieving better outcomes than expected. Applied to spinal cord injury (SCI), it allows for an exploration and characterization of individuals that recover beyond what is clinically explainable.

Methods

In this study, we defined positive deviance as phenomenal recovery (PR) following SCI in two cohorts, namely the Sygen clinical trial and the European multicenter study on human spinal cord injury (EMSCI) cohort. The definitions of PR followed two strategies: based on clinical knowledge, and on statistical characterization. After defining PR and comparator groups, we reported on demographics, patterns in motor scores recovery - with a specific focus on the comparison between scores in the upper and lower extremities -, and prescriptions of antibiotics.

Results

We observe that phenomenal recoveries tend to occur more in individuals with cervical injuries (from 61.7% to 100% of the PR groups defined). They exhibit a higher prevalence of central cord patterns in motor scores at recovery, especially when using more refined definition of central cord syndrome, based on the level of injury. However, we could not identify consistent patterns in antibiotics prescription across the different PR groups or in comparing PR groups to their respective comparator group.

Conclusions

Further explorations in additional cohorts are warranted to confirm or infirm the trends observe and thus better characterize patients with a potential for PR. This characterization would be crucial in the context of clinical trials, where such PRs should not be mistaken for a treatment effect.

7.2 Introduction

The concept of positive deviance refers to the observation that a few individuals will achieve better outcomes thanks to unique behaviors or characteristics [317]. While this concept was originally developed in behavioral and social sciences, particularly applied to nutrition education [318], its pertinence extends to other fields, including recovery following trauma such as spinal cord injury (SCI). SCI is a devastating condition where the impairment of the spinal cord leads to loss of functions in all major systems of the body (neurological, musculoskeletal, cardiovascular, pulmonary, urinary, among others). Following the initial trauma, the severity of a SCI is summarized on a five-grade scale, from A most severe to E least severe, according to the American spinal injury association (ASIA) impairment scale (AIS) grading system. The AIS grade has been shown to be a good indicator of the expected recovery [17], with more injured patients being expected to recover less than a patient with a less severe injury, all other characteristics being otherwise similar. Likewise, infections such as pneumonia and postoperative wounds in the acute stage after initial trauma seem to be associated with worse recovery in the most severely injured patients (AIS grades A and B) [105]. Despite the knowledge acquired on factors influencing recovery following SCI, there still exists exceptions where a few individuals with SCI recover beyond clinical expectations and explanations.

In this project, we intend to explore the characteristics of such individuals, referred to as phenomenal recovery (PR) group, focusing on neurological recovery profiles and medications received. We define the PR group according to two distinct methods: (i) a clinical definition, matching the clinical experience of positive deviance; and (ii) a statistical definition, based on the identification of outliers in the amount of motor recovery observed. Our exploration includes a comparison between the methods used to define the PR group and between groups defined in two distinct data sources, namely the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts.

We hypothesize that individuals from the PR group do not only exhibit increased motor recovery as defined by aggregate scores such as the lower extremity motor score (LEMS) but also unique patterns in motor recovery at the myotome level. Furthermore, taking advantage of information about medication received at the acute stage after injury in the Sygen cohort, we hypothesize that the PR group differs from the rest of the cohort in the medications prescribed, particularly antibiotics used as proxy for infections.

Defining and studying individuals recovering to a greater extent than expected holds the potential to formulate new hypotheses on how to enhance recovery following SCI and further informs the detailed analyses of patient subgroups within clinical trials.

7.3 Methods

7.3.1 Data sources

Sygen cohort

The Sygen clinical trial was conducted in the 1990s, in the United States of America (USA), to assess the effect of the gangliosidosis 1 (GM-1) on the motor recovery of individuals with SCI [33, 34, 117, 118]. While it failed to prove significant improvement over placebo, since then, it has been a valuable data source for secondary analyses in the field. In this study, we particularly rely on information on the longitudinal neurological assessments, and prescribed medications in the first 30 days following injury. Neurological assessments were collected following the international standards for neurological classification of spinal cord injury (ISNCSCI) examination standards, making available both aggregate scores such as the lower extremity motor score (LEMS) and upper extremity motor score (UEMS), but also gradings at the myotome and dermatome levels. Assessments were performed according to a predefined schedule which involved measurements 2, 4, 8, 16, 26, and 52 weeks after injury. Here, we particularly focused on week 2 and 4, considered together as baseline, and week 26 and 52, considered together as recovery time. Additionally, the neurological level of injury (NLI) was recalculated to match the definition given by Rupp *et al.* in [15].

EMSCI cohort

The EMSCI is a observational cohort collecting data across centers for SCI in Europe and India [1]. Started in the early 2000s, this initiative now represents data from over 6000 individuals. Information such as demographics, neurological assessments following the ISNCSCI examination and functional scores (e.g., spinal cord independence measure (SCIM)) are collected at predefined time points: very acute (~2 weeks, 0 to 15 days post injury), acute I (~4 weeks, 16 to 40 days), acute II (70 to 98 days), acute III (~26 weeks, 150 to 186 days) and chronic (~52 weeks, 300 to 546 days). Mimicking the criteria applied to the Sygen data, we consider here very acute and acute I stages as baseline, and acute III and chronic stages as recovery.

[1] http://emsci.org/

7.3.2 Definitions of positive deviance

Clinical approach

The clinical definition of positive deviance aimed to identify individuals who had experienced a sustained very severe injury and manifested a PR. Sustained very severe injury was defined as an injury which was consistently graded as AIS grade A 2 and 4 weeks after initial injury, with no remaining motor or sensory functions (all scores evaluated as 0) below L1 in the same timeframe. A PR following a sustained very severe injury was defined based on two observations. Firstly, spinal cord injury (SCI) initially graded as severe (AIS grade A) tend to remain severe after one year following initial injury [17]. Secondly, myotomes graded with a 0 after initial injury tend to remain 0, while initial scores above 0 have higher likelihood to improve (Lukas *et al.*, unpublished). Therefore, PR was fixed with an improvement of at least five points in LEMS from baseline to recovery [271].

$$\Delta\text{LEMS} = \text{LEMS}_{\text{LOCF}} - max(\text{LEMS}_2, \text{LEMS}_4) \tag{7.1}$$

The comparator group corresponded to individuals a sustained very severe injury at early stages after injury with improvement in LEMS less than five points. The full decision tree used to define the PR and comparator groups is shown in **Figure 7.1**.

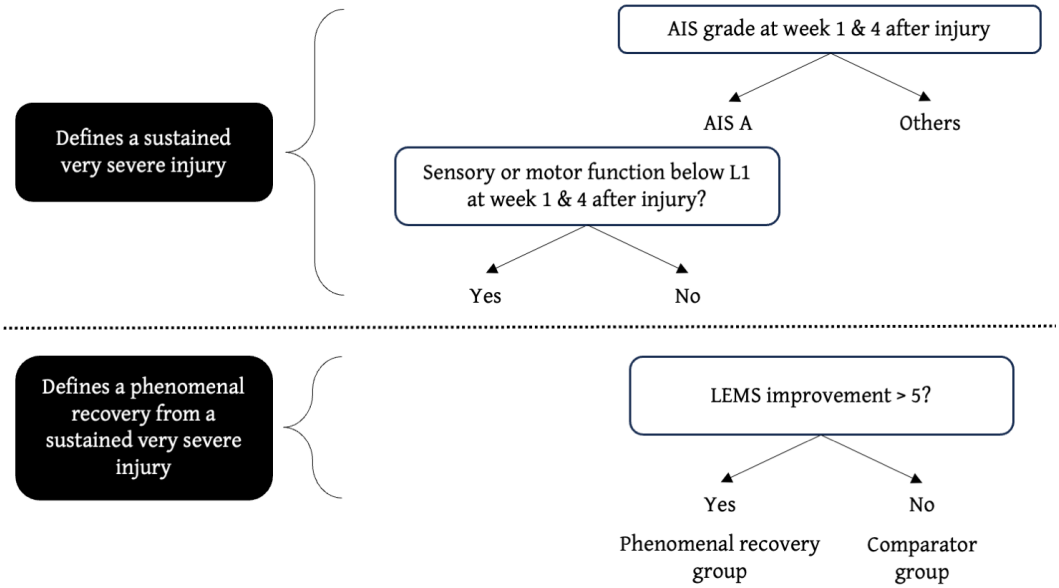


Figure 7.1: Decision tree illustrating the clinical definition of a PR. phenomenal recovery (PR); American spinal injury association (ASIA) impairment scale (AIS); lower extremity motor score (LEMS).

Statistical approach

With the intention to define PR beyond sustained very severe injuries, we introduce a definition of PR based on the distribution of mean improvement of motor scores (MS), across all myotomes M, below the NLI, between scores evaluated at week 4 versus recovery:

$$\Delta \mathrm{MS}_{\mathrm{belowNLI}} = \frac{1}{|M|} \sum_{m \in M} (\mathrm{MS}_4^m - \mathrm{MS}_{\mathrm{LOCF}}^m) \tag{7.2}$$

PR is then defined as individuals who fall beyond the 95th percentile of the distribution observed, to match an approximate 5% of the population studies which would be defined as experiencing a PR. Notably, this approach accounts for (i) the ceiling effect observed in scores such as the lower extremity motor score (LEMS), and (ii) the neurological level of injury (NLI) by considering on the myotomes below this level.

The comparator group is drawn from the rest of the distribution, using propensity score matching as implemented in the package `MatchIt` in R. Matching was performed on sex, age, NLI, LEMS, UEMS and AIS grade evaluated 4 weeks after injury. We chose a 4:1 allocation with no replacement. The ratio was limited by the AIS grade distribution from the entire population and the absence of replacement was chosen to avoid potential over representation of prescribed medications from a single individual.

7.3.3 Data processing

Inclusion and exclusion criteria

To be included in our study, individuals needed to have information about LEMS and AIS grade evaluated at baseline (2 and/or 4 weeks after injury), and LEMS at recovery. The statistical approach required information about the NLI to be available. Finally, individuals with lumbar or sacral injuries were excluded from the EMSCI cohort, for better comparability with the Sygen cohort, including only cervical and thoracic injuries by design.

Last observation carried forward

To maximize the size of the cohort to study, we performed a last observation carried forward (LOCF) imputation from week 26 to week 52 when outcomes of interest, such as the LEMS, evaluated 52 weeks after injury was missing [261].

Exploring neurological recovery

Neurological recovery was studied by comparing myotome and dermatome gradings at week 4 versus recovery after LOCF imputation. When comparing the incidence of central cord syndrome (CCS), five definitions of CCS were used as defined by Engel-Haber *et al.* [319]:

(i) disproportionate weakness with LEMS – UEMS ≥ 1, later referred to as *CCS-1*

(ii) disproportionate weakness with LEMS – UEMS ≥ 5, later referred to as *CCS-5*

(iii) disproportionate weakness with LEMS – UEMS ≥ 10, later referred to as *CCS-10*

(iv) disproportionate weakness with LEMS – UEMS ≥ 19, later referred to as *CCS-19*

(v) NLI-based definition with $(1 - \frac{\text{meanUEMS}_{\text{blwNLI}}}{\text{meanLEMS}}) > 10\%$, later referred to as *NLI-based CCS*

Exploring antibiotics prescribed

The exploration of antibiotics prescribed was limited to the first 30 days after injury, period during which the PR and its comparator groups are designed to be similar in their injury characteristics. Various strategies were employed to compare the two groups in terms of the antibiotics prescribed, both from a quantitative and qualitative point-of-view. This analysis included the comparison in the incidence of each antibiotic prescribed, in the number of unique antibiotics prescribed, and in number of days with no antibiotics prescribed. Additionally, we examined the number of individuals with antibiotics prescribed on day 0 and day 1 after injury, in an effort to identify prophylaxis prescriptions. Finally, we studied the number of cumulative antibiotics-days, defined as the sum of unique antibiotics given across the first 30 days after injury.

Statistical analyses

The clinical definition of positive deviance did not allow for statistical testing due to low sample size. Kolmogorov-Smirnov (`ks.test` and Fisher's exact tests (`fisher.test`) were performed, for continuous and categorical variables, respectively, for comparison between PR and comparator groups as defined according to the statistical approach. Bonferroni multiple testing correction was used, and the adjusted p-values < 0.05 were considered significant. All analyses were conducted with R Statistical Software, version 4.3.2 (running under: macOS Sonoma 14.2.1).

7.3.4 Data and code availability statement

Anonymized data used in this study will be made available upon request to the corresponding author and in compliance with the European general data protection regulation (EU GDPR). The code describing the analysis can be accessed on our GitHub repository [2].

[2] https://github.com/lbourguignon/SCI-phenomenal-recovery

7.4 Results

7.4.1 Clinical approach

The clinical definition of PR, based on LEMS improvement after a sustained very severe injury, yields an incidence of about 2% both in the Sygen ($n = 6$) and EMSCI ($n = 8$) cohorts. Although the small sample sizes prevent the use of statistical tests, we observe that the PR groups tend to include more men and older individuals. One noticeable difference was in the level of injury, with PR occurring predominantly after a cervical injury ($n = 6$, 100%, and $n = 7$, 87%, in the Sygen and EMSCI cohorts, respectively). Further details are reported in **Table 7.1**.

Table 7.1: Demographic and injury characteristics of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the clinical definition

	Sygen cohort		**EMSCI cohort**	
	PR group	**Comparator group**	**PR group**	**Comparator group**
Sample size				
	6 (2)	293 (98)	8 (2)	393 (98)
Sex, N (%)				
	0 (0)	53 (18.1)	0 (0)	81 (20.6)
Male	6 (100)	240 (81.9)	8 (100)	312 (79.4)
Age (years)				
	37.0 (11.7)	30.6 (12.6)	51.0 (16.9)	41.2 (17.3)
Median [Min, Max]	22.0 [18.0, 49.0]	28.0 [13.0, 69.0]	56.0 [18.0, 70.0]	40.0 [13.0, 93.0]
AIS grade 4 weeks after injury, N (%)				
	6 (100)	293 (100)	8 (100)	495 (100)
B	0 (0)	0 (0)	0 (0)	0 (0)
C	0 (0)	0 (0)	0 (0)	0 (0)
D	0 (0)	0 (0)	0 (0)	0 (0)
NLI, N (%)				

Continued on next page

Table 7.1: Demographic and injury characteristics of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the clinical definition (Continued)

	Sygen cohort		**EMSCI cohort**	
	PR group	**Comparator group**	**PR group**	**Comparator group**
	6 (100)	210 (71.7)	7 (87)	166 (42.2)
Thoracic	0 (0)	69 (23.5)	1 (13)	222 (56.5)
Missing	0 (0)	14 (4.8)	0 (0)	5 (1.3)
LEMS 4 weeks after injury				
	0 (0)	0 (0.12)	0 (0)	0 (0)
Median [Min, Max]	0 [0, 0]	0 [0, 2]	0 [0, 0]	0 [0, 0]
Missing, N (%)	0 (0)	3 (1.0)	0 (0)	0 (0)
UEMS 4 weeks after injury				
	9.2 (14.5)	25.0 (19.5)	24.8 (16.7)	35.5 (19.6)
Median [Min, Max]	4 [0, 38]	19 [0, 50]	23.5 [6, 50]	50.0 [0, 50.0]
Missing, N (%)	0 (0)	20 (6.8)	0 (0)	6 (1.5)
LEMS at recovery (with LOCF)				
	16.5 (8.7)	0.1 (0.5)	20.9 (13.6)	0.1 (0.4)
Median [Min, Max]	14.5 [6, 32]	0 [0, 5]	20 [6, 45]	0 [0, 4]
UEMS at recovery (with LOCF)				
	22.0 (14.5)	29.6 (17.8)	35.1 (13.4)	37.6 (17.7)
Median [Min, Max]	21.5 [5, 46]	26 [0, 50]	37 [17, 50]	50 [0, 50]
Missing, N (%)	0 (0)	1 (0.3)	0 (0)	1 (0.3)

American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description.
European multicenter study on human spinal cord injury (EMSCI), phenomenal recovery (PR), standard deviation (SD), neurological level of injury (NLI), lower extremity motor score (LEMS), upper extremity motor score (UEMS), last observation carried forward (LOCF)

When investigating motor recovery at the myotome level, one can note that some

individuals classified in the PR group recover motor abilities further in the distal lower limb myotomes, as illustrated in **Figure 7.2**. This observation can be quantified according to the different CCS definitions, as summarized in **Supplementary Table 8.16**. Overall, we observe higher proportions of CCS patterns at recovery in the PR groups, both in the Sygen and EMSCI cohorts.

A total of 19 unique antibiotics were prescribed in both the Sygen PR and comparator groups, with 17 being prescribed most often in the PR group. Among them, vancomycin was the antibiotic with the largest difference between the two groups, with 100% ($n = 6$) versus 43.7% ($n = 128$) of the PR versus comparator which received vancomycin at least once in the first 30 days after injury. Details of the antibiotics and their prevalence are summarised in **Supplementary Table 8.17**. We observe that more unique antibiotics were prescribed in the PR group (mean: 6.7, standard deviation (SD): 3.3, median: 6.0, minimum: 2, maximum: 11; and mean: 5.0, SD: 2.4, median: 5.0, minimum: 1, maximum: 17). The proportion of individuals who received antibiotics on day 0 or 1 after injury was also higher in the PR group ($n = 3$, 50% and $n = 83$, 28.3%, for PR and comparator groups, respectively). Interestingly, the number of cumulative antibiotics-days tends to be lower in the PR group compared to the comparator group (mean: 45.2, SD: 28.9, median: 44.5, minimum: 12, maximum: 82; and mean: 56.8, SD: 37.0, median: 50.0, minimum: 0, maximum: 194; respectively).

7.4.2 Statistical approach

Individuals located beyond the 95th percentile in mean improvement of motor scores (MS) below the NLI were mainly less severely injured individuals, with a majority of AIS grade C both in the Sygen and EMSCI cohort ($n = 21$, 75.0%, and $n = 86$, 67.2%, respectively). Similarly as in the clinical-PR groups, cervical injuries were the most represented ($n = 24$, 85.7%, and $n = 79$, 61.7% in Sygen and EMSCI, respectively). As expected based on the definition used, PR and comparator groups showed statistically significant differences in LEMS and UEMS at recovery (Kolmogorov-Swirnov tests, both $p < 0.001$). All comparisons, demographic and injury characteristics are summarised in **Table 7.2**.

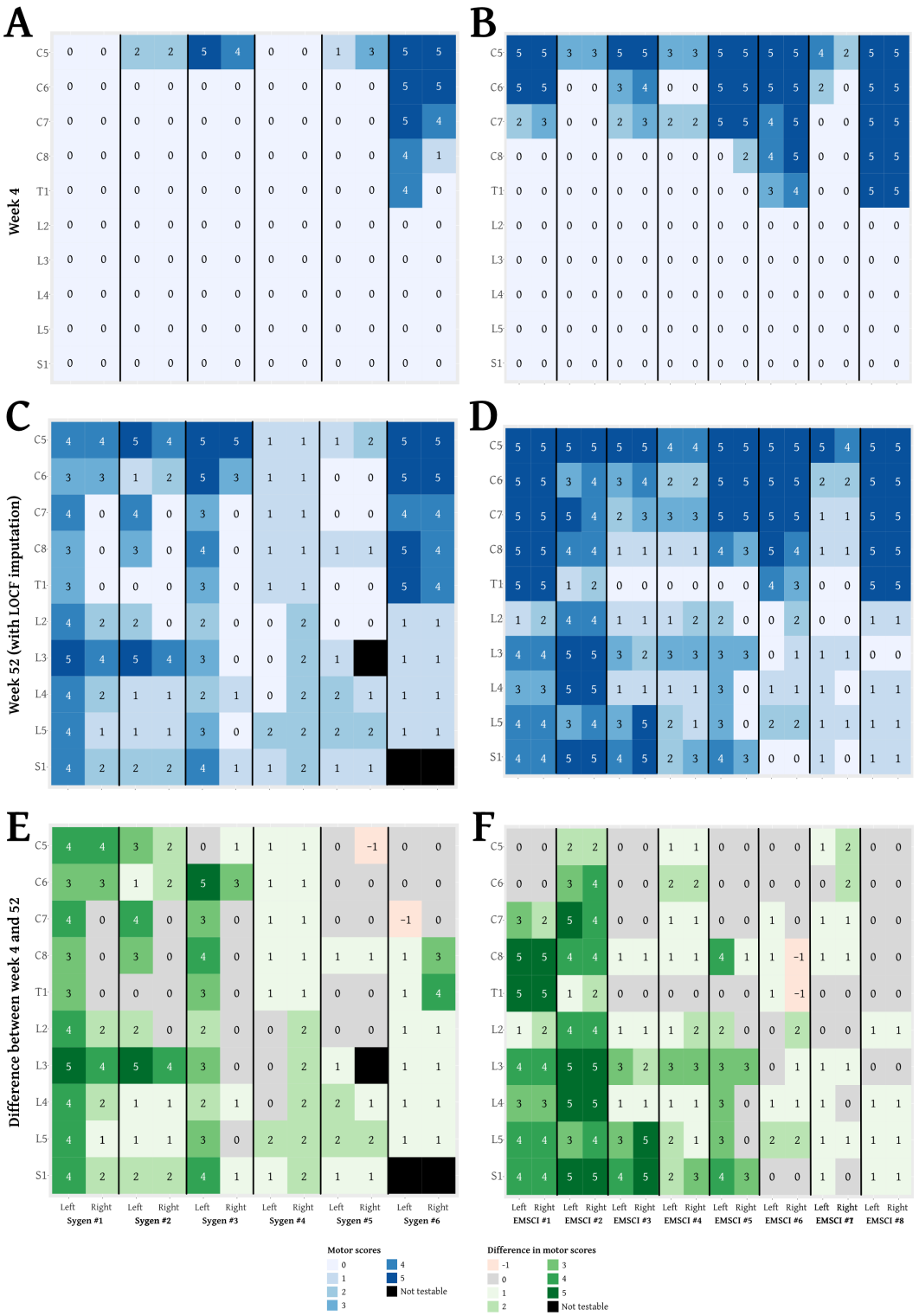


Figure 7.2: Motor scores at the myotome level for PR groups based on clinical definition. **A.** Evaluated at week 4 in the Sygen PR group; **B.** Evaluated at week 4 in the EMSCI PR group; **C.** Evaluated at recovery after LOCF imputation in the Sygen PR group; **D.** Evaluated at recovery after LOCF imputation in the EMSCI PR group; **E.** Difference between scores at week 4 and at recovery in the Sygen PR group; **F.** Difference between scores at week 4 and at recovery in the EMSCI PR group; phenomenal recovery (PR); European multicenter study on human spinal cord injury (EMSCI); last observation carried forward (LOCF).

Table 7.2: Demographic and injury characteristics of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the statistical definition

	Sygen cohort			EMSCI cohort		
	PR group	Comparator group	p-value	PR group	Comparator group	p-value
Sample size						
N (%)	28 (20)	112 (80)		128 (20)	512 (80)	
Sex, N (%)						
Female	2 (7.1)	18 (16.1)	0.37	21 (16.4)	99 (19.3)	0.53
Male	26 (92.9)	94 (83.9)		107 (83.6)	413 (80.7)	
Age (years)						
Mean (SD)	32.1 (13.6)	33.4 (14.5)	0.97	42.5 (18.9)	45.0 (18.7)	0.53
Median [Min, Max]	30.5 [13.0, 62.0]	30.5 [14.0, 69.0]		42.0 [9.0, 78.0]	44.5 [14.0, 89.0]	
AIS grade 4 weeks after injury, N (%)						
A	2 (7.1)	8 (7.1)	0.34	15 (11.7)	56 (10.9)	0.72
B	5 (17.9)	37 (33.0)		24 (18.8)	121 (23.6)	
C	21 (75.0)	67 (59.8)		86 (67.2)	320 (62.5)	
D	0 (0)	0 (0)		2 (1.6)	12 (2.3)	
NT	0 (0)	0 (0)		1 (0.8)	3 (0.6)	

Continued on next page

Table 7.2: Demographic and injury characteristics of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the statistical definition (Continued)

	Sygen cohort			EMSCI cohort		
	PR group	Comparator group	p-value	PR group	Comparator group	p-value
NLI, N (%)						
Cervical	24 (85.7)	101 (90.2)	0.50	79 (61.7)	323 (63.1)	0.84
Thoracic	4 (14.3)	11 (9.8)		49 (38.3)	189 (36.9)	
LEMS 4 weeks after injury						
Mean (SD)	6.89 (6.33)	7.62 (9.93)	0.01	6.57 (6.91)	7.43 (8.87)	0.17
Median [Min, Max]	6 [0, 22]	0 [0, 36]		4.0 [0, 33]	4.0 [0, 37]	
UEMS 4 weeks after injury						
Mean (SD)	16.9 (15.9)	18.8 (13.3)	0.19	29.7 (18.9)	29.8 (18.4)	0.90
Median [Min, Max]	12 [0, 50]	17 [0, 50]		31.5 [0, 50]	27.0 [0, 50.0]	
LEMS at recovery (with LOCF)						
Mean (SD)	42.0 (7.0)	19.2 (18.8)	**<0.001**	41.9 (6.53)	17.5 (15.4)	**<0.001**
Median [Min, Max]	42.5 [24, 50]	17 [0, 50]		43 [6, 50]	17 [0, 50]	
UEMS at recovery (with LOCF)						

Continued on next page

Table 7.2: Demographic and injury characteristics of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the statistical definition (Continued)

	Sygen cohort			EMSCI cohort		
	PR group	Comparator group	p-value	PR group	Comparator group	p-value
Mean (SD)	40.4 (7.9)	30.2 (13.8)	**0.001**	44.7 (6.26)	35.6 (14.9)	**<0.001**
Median [Min, Max]	40.5 [24, 50]	32 [0, 50]		47 [27, 50]	40 [1, 50]	

Note that p-values reported are not Bonferroni-corrected, in bold are the p-values significant after multiple testing correction; American spinal injury association (ASIA) impairment scale (AIS): see **Table 1.1** for full description, European multicenter study on human spinal cord injury (EMSCI), phenomenal recovery (PR), standard deviation (SD), neurological level of injury (NLI), lower extremity motor score (LEMS), upper extremity motor score (UEMS), last observation carried forward (LOCF)

Table 7.3 quantifies the differences in incidence of CCS between PR and comparator groups. We note a trend towards higher percentage of individuals with CCS patterns in their motor recovery in the PR groups, especially when considered more complex definitions taking into account patterns rather than aggregate scores alone as when comparing UEMS and LEMS only. Hence, in the Sygen cohort, NLI-based CCS was found in 15 individuals (53.6%) in the PR group, and only in 32 individuals (28.6%) in the comparator group. Similarly, in the EMSCI cohort, 49 (38.3%) and 102 (19.9%) were classified as CCS when considering the definition based on NLI in the PR and comparator groups, respectively. It is however important to note that only the differences in proportions in CCS-1 and NLI-based CCSs in the EMSCI are statistically significant after multiple testing correction (Fisher's exact test, both $p < 0.001$ uncorrected and $p < 0.001$ considering adjustment for five tests).

Table 7.3: Central cord syndrome (CCS) of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the clinical definition

	Sygen cohort			EMSCI cohort		
	PR group	Comparatc group	p-value	PR group	Comparatc group	p-value
CCS-1	15 (53.6)	32 (28.6)	0.02	49 (38.3)	102 (19.9)	**<0.001**
CCS-5	10 (35.7)	18 (16.1)	0.04	24 (18.2)	72 (14.1)	0.21
CCS-10	4 (14.3)	10 (8.9)	0.45	8 (6.3)	36 (7.0)	0.85
CCS-19	0 (0)	2 (1.8)	1	1 (0.8)	3 (0.6)	1
NLI-based CCS	15 (53.6)	35 (31.3)	0.06	51 (39.8)	106 (20.7)	**<0.001**

Note that p-values reported are not Bonferroni-corrected, in **bold** are the p-values significant after multiple testing correction; central cord syndrome (CCS), neurological level of injury (NLI)

Among the antibiotics prescribed in both PR and comparator groups, only ceftriaxone showed a significant difference (Fisher's exact test, p< 0.001) with it being more prescribed in the PR group (n = 9, 32.1%) compared to the comparator group (n = 7, 6.3%). Interestingly, vancomycin did not differ, opposite to what was observed in the comparison of groups defined according to the clinical definition of PR. All antibiotics prescribed in both groups are summarized in **Supplementary Table 8.18**. The number of unique antibiotics prescribed across the first 30 days after injury and number of cumulative antibiotic-days did not significantly differ between PR and comparator group

($p = 0.71$ and $p = 0.27$, respectively). The number of individuals who received antibiotics on day 0 or 1 after injury was equally comparable ($n = 7$, 33% and $n = 13$, 12% for PR and comparator groups, respectively, $p = 0.11$).

7.5 Discussion

In this project, we aimed to explore characteristics from individuals who recover beyond clinical expectations and explanations following a SCI. The first challenge in this task was to define what we would consider a so-called PR. To address this challenge, we followed two distinct approaches, based on clinical knowledge, and using a statistical approach. We further applied those definitions in two distinct cohorts, from the Sygen clinical trial and the EMSCI cohort. In multiplying the definitions and cohorts studied, we could compare and draw meaningful trends in the characteristics observed.

Firstly, we observed that PR was more likely to occur following a cervical injury, but was otherwise as frequent in both sex and all age groups. Secondly, we described unique patterns in motor score recovery approaching the definitions of CCS as being more present in the PR groups. As described by Engel-Haber *et al.* in [319], defining CCS has been a controversial research topic and the current definitions may not entirely reflect cases were distal impairments in the upper limb is greater than impairments proximal in the lower limb. It is important to note here that the patterns were observed at recovery (i.e., 26 or 52 weeks after injury), but not at the initial stages following injury, and would therefore note be classified as CCS after injury. However, CCS are known to be a type of SCI that will recover better than injury not exhibiting this specific injury pattern [320]. Additionally, CCS is only defined for cervical injuries, since injuries lower in the spinal cord will not affect the upper limbs. However, it could be of interest to further investigate individuals with greater distal impairment versus proximal motor scores, in the lower limbs specifically. Finally, based on the data collected as part of the Sygen clinical trial, we were able to test hypothesis related to the prescription of antibiotics following SCI and their association with recovery. While we could observe that some antibiotics such as vancomycin or ceftriaxone were prescribed more often in the PR versus comparator groups, the findings were not consistent throughout the difference definitions of PR applied. We could therefore not identify significant differences in the antibiotics prescribed between individuals experiencing a PR and the ones who did not.

7.5.1 Limitations

A number of limitations should be acknowledged related to this work. Firstly, we were not able to study medications in the EMSCI cohort since the information is not collected as part of the protocol for this observational study. It would however be interesting to further study medications in complementary cohorts such as the SCIRehab cohort [46]. On the topic of medications, we focused here on antibiotics but further explorations are warranted to investigate potential differences in, for example, pain management. The

areas of medications studied could then be used as proxy to extent to comparisons in incidence of complications such as neuropathic pain, or compromised bowel and bladder management, which are important concerns for individuals living with SCI. Furthermore, the investigation of antibiotics depicted here would benefit from refinements, relating the prescriptions to their indications (e.g., surgery, prophylaxis, infections) and further considerations towards the combinations of antibiotics prescribed, and clusters that those drugs might form rather than consider each antibiotic as a unique, independent compound. Secondly, definition of outliers, or here PR, comes with inherent challenges. The small sample size described in the clinical definition of PR prevented us from performing meaningful statistical comparisons. Merging, or considering additional data sources would be particularly valuable when specifically focusing on defining rare events.

7.5.2 Conclusions

This study is the first of its kind exploring individuals with sci who experience a recovery beyond what is clinically explainable and expected. Being able to identify such individuals hold the promise to both inform potential avenues in the search for interventions to improve recovery following SCI and better identify those individuals to improve the downstream analysis of data collected in the context of clinical trials.

Declarations

Ethics approval and consent to participate

The study was performed in accordance with the Declaration of Helsinki and was approved by all responsible institutional review boards. The original Sygen clinical trial [117, 118, 33] also received ethical approval, but was conducted before clinical trials were required to be registered (i.e., no clinicaltrial.gov identifier available). All experiments were performed in accordance with relevant guidelines and regulations. Informed consent to participate was obtained from all of the participants as part of the original study.

Availability of data and materials

Anonymized data used in this study is available upon request to the corresponding author and in compliance with the European general data protection regulation (EU GDPR). The code describing the analysis can be accessed on our GitHub repository [3].

Competing interests

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Funding

This study was supported by the Swiss National Science Foundation (Ambizione Grant [PZ00P3_186101, Jutzeler]). LB was supported by an ICORD student exchange award provided by the Praxis Spinal Cord Injury Foundation. SB was supported by the Botnar Research Centre for Child Health Postdoctoral Excellence Programme (#PEP-2021-1008). The funding sources of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Authors' contribution

Lucie Bourguignon: design of the study, analysis, visualisation and interpretation of the data; writing of the manuscript
Louis P. Lukas: substantial contributions to the design and interpretation of the study;

[3] https://github.com/lbourguignon/SCI-phenomenal-recovery

revising the manuscript critically for important intellectual content
Sarah C. Brüningk: substantial contributions to the design and interpretation of the study; revising the manuscript critically for important intellectual content
Fred H. Geisler: revising the manuscript critically for important intellectual content
EMSCI study group: access to data
Catherine R. Jutzeler: substantial contributions to the design and interpretation of the study; revising the manuscript critically for important intellectual content
John L. K. Kramer: design of the study, interpretation of the data; revising the manuscript critically for important intellectual content

Acknowledgements

Parts of this project were conducted on the traditional and unceded territory of the Cayuse, Umatilla and Walla Walla, S'ólh Téméxw (Stó:lō), Hul'qumi'num Treaty Group, səl̓ilwətaʔɬ təməxʷ (Tsleil-Waututh), šxʷməθkʷəy̓əmaʔɬ təməxʷ (Musqueam), Skwxwú7mesh-ulh Temíx̱w (Squamish), and Stz'uminus.

Part C in a nutshell

Part C gave the opportunity to exhibit data science methods tailored to SCI data and research based on this data. Importantly, the adapted methods presented do not only improve the quality of the present research outcomes but also contribute to setting best practices within the field that can be applied in future projects. Adhering to common best practices would enable the field to better compare, reproduce and interpret results across studies and research groups. However, methodological advances are not the only leverage point to achieve reproducibility and transparency. Data visualization, and in particular user-interactive tools, plays a pivotal role when dealing with sensitive data, such as medical and identifiable information.

Part D

Effectively conveying results through interactive data visualization

Introduction

Data visualization is a constitutive part of research communication (e.g., figures published alongside a manuscript, oral and poster presentations), and is often the point of initial access to a project. Accordingly, an impactful visual representation will contribute to the enhanced promotion of the results displayed, especially when addressed to a multidisciplinary audience. However, visualization is also largely restricted by the current means of research communication (e.g., publications, conferences, social media), not only in the amount that can be shared but also in the lack of adaptability.

In parallel, data transparency has been advocated as one of the most important aspects of medical research, particularly in clinical trials and studies, as the way to inform trustworthy evidence-based clinical decisions [321]. The initiatives towards increased data transparency in clinical research involved protocol and trial design disclosures, publication plans to reduce selective publication and independent data analysis [321].

In **Part D** and its corresponding **Chapter 8**, we present interactive, user-driven data visualization as an additional measure towards the promotion of data transparency and better science communication, thus directly benefiting medical research in general and clinical studies in particular.

Chapter 8

The interactive manuscript: From tabular to interactive result presentation and data visualization

---

Adapted from:

Lucie Bourguignon, Jaimie J. Lee*[1], Ryan Loke*, John L. K. Kramer, Catherine R. Jutzeler

Manuscript in preparation, 2024

Publication:
Web site: https://lbourguignon.shinyapps.io/Shiny-perspective/
GitHub repository: https://github.com/lbourguignon/shiny-perspective

Lucie Bourguignon was responsible for the writing of the manuscript and data visualization, on an original idea shared with Catherine R. Jutzeler. Further details can be found in **Section 8.5**.

[1]* indicates equal contribution

8.1 Abstract

Introduction

Advances in computer science and technology have led to unprecedented new possibilities in science, engineering, and business. Data visualization is one notable field which emerged from such advances. Taken in the context of research, it allows for better and further use of the data and proves to be an additional tool to promote data sharing and transparency, especially when dealing with sensitive data.

Areas covered

This perspective covers the use of new technologies and tools for improved data visualisation in research. It specifically focuses on software allowing for dynamic, interactive, and user-controlled plotting in comparison with traditional fixed visualisations and tabular format.

Expert opinion

Interactive data visualization offers an avenue of new possibilities in all steps of data-driven research: data exploration, hypothesis formulation, and result outputs. Owing to the simplicity of its implementation and use, one should advocate for larger deployment of interactive data visualization in research.

8.2 Introduction

In the last decades, an immense amount of data has been generated, affecting all aspects of our lives. The emergence of biobanks and open-source data sets into the research landscape is one illustration of the phenomenon applied to biomedical research. This increase in data availability drove the development of scientific fields, such as computer and data sciences, and new analysis techniques, known under the denomination of machine learning (ML) and artifical intelligence (AI). Because they can disentangle meaningful recurring patterns from complex, multimodal data, those techniques are nowadays prominently featured in our daily lives. Consequently, it is not rare to encounter discussions about ML and AI not only in the scientific literature, but also in mainstream media.

Along with the development of data sciences, a related field emerged: the field of data visualization. It commonly refers to the transfer of data into visual displays, such as plots, charts, or graphs [322]. Visualization is an essential tool for exploring and communicating findings in medical research, and especially in epidemiological surveillance. Notable improvements have indeed been made in the direction of more effective scientific communication. Such advances include the access to extensive information on the effective use of colors [323], or to free online web tools. BioRender [2] is one example of a widely used software when it comes to illustration of biological phenomena across scales, from molecular to specimen level. Additionally, softwares like the `Shiny` R package [3] [90] or Python libraries Bokeh [4] and plotly [5] allow researchers to build interfaces to publish user-friendly and interactive representations of the data used. This possibility marks a significant breakthrough for transmission of our research outcomes as it allows us to overcome limits inherited from the area of printed scientific journals such as the restricted use of colors or number of figures to be presented. It equally increases transparency as it enables the public to not only access carefully selected representations of the data but rather a diversity of representations. This transparency represents a unique opportunity for the reader to draw their own interpretations and potentially formulate new hypotheses. Displaying data in an interactive fashion can also help researchers and policymakers to identify and understand trends that could be overlooked if the data were reviewed solely in tabular form. Besides their advantages regarding visualization, the web applications created can also perform instant analyses, effectively multiplying their impact.

In this article, we will first discuss the importance of data visualization for medical

[2] https://www.biorender.com/
[3] https://shiny.rstudio.com/
[4] https://docs.bokeh.org/en/latest/index.html
[5] https://plotly.com/python/

research in more detail. We will then demonstrate and illustrate the power of interactive data visualization via a living example of a Shiny webpage comparing tabular and interactive visualizations [6]. Finally, we will present our views on how to best take advantage of the new data visualization tools available to make our research outcomes more impactful and meaningful.

8.3 The importance of data visualization in medical research

Data visualization in the context of medical research is a critical tool to better understand complex phenomena and drive new research insights, which would ultimately contribute to better care. Hence, data visualization became crucial not only in the early stages of data exploration, but also in reaching healthcare professionals, policy-makers, and the general public (**Figure 8.1**).

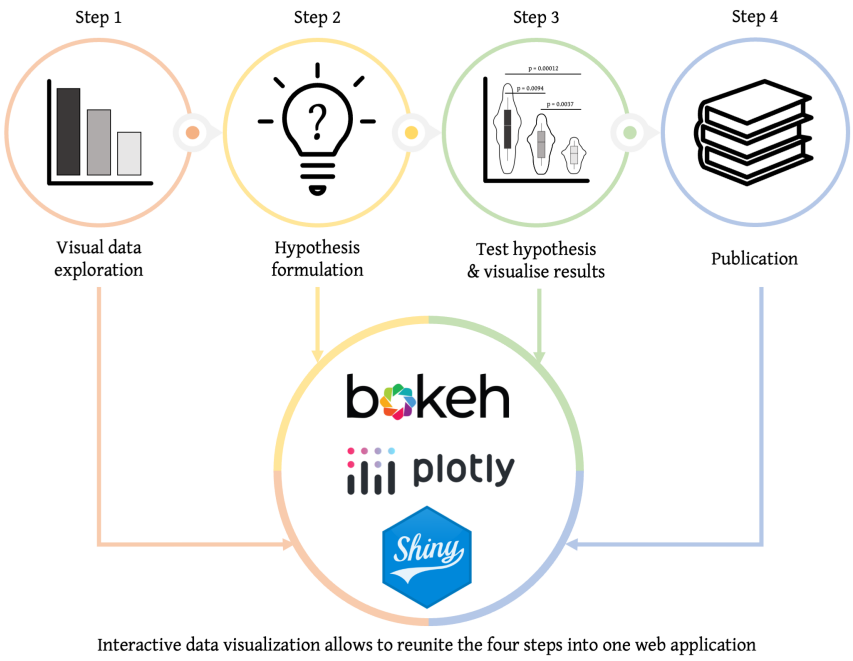


Figure 8.1: Data visualization in the context of research. Interactive data visualization allows to reunite the four steps into one web application.

When first exploring large amounts of data, the initial objective is to better grasp the nature of the data, and reach both a global and granular understanding of the data. While visualization may appear as a qualitative approach, it is an effective way of capturing

[6] https://lbourguignon.shinyapps.io/Shiny-perspective/

the presence of outliers (e.g., impossible values such as oxygen saturation of about 10%) or errors (e.g., temperatures being reported in a mixture of Fahrenheit and Celsius, or female being encoded as "female" or "woman" in the same dataset), which will influence the following data cleaning steps. In helping to identify previously unseen trends, visualization of the data eventually leads to formulating new hypotheses and research questions, providing the most effective type of chart for the data at hand. It is important to note here that an understanding of what makes a good visualization is crucial as it contributes to shaping the downstream analyses.

Once the relevant analyses have been performed comes the time to visualize the results obtained. At that stage, the visualization strategy will depend on the intended message and targeted audience. Figures presented in publications are intended for a specialist audience. The emphasis is put on precision and accuracy of the representation of the results, which might lead to more complex figures. They are often the first window for readers to approach a publication and will determine whether the reader will invest time into the entire manuscript. As such, those figures would not only need to be clear and straightforward, but also stand alone figures. They should be the reflection of the main findings presented, even taken outside of the context of the publication. Although the target audience might be similar, the same topic presented in an oral or poster presentation would require adjustment in its visual support. In this scenario, the reading time would be restricted but guided by the author, allowing for a layer-based representation. On the other hand, visualization in the context of outreach activities such as presentation to a lay audience requires readable and concise information. Similarly as one would adapt the language used, the type of representation should also be adapted to a simpler but accurate format in order to prevent misinterpretation.

The use of interactive visualization tools, such as Shiny web applications, allows to reconcile visualizing data for exploration or presenting research results in diverse contexts (e.g., publications, conferences, grant applications). While the interactive representations were created with a precise goal in mind (e.g., illustrating the publication that it was created for), its adaptability to user inputs present the opportunity to formulate new research questions to be further explored. The degree of complexity can equally be adapted and guiding steps can be provided.

8.4 From tabular to interactive data visualization

To illustrate the use of interactive data visualization, we built a R-based webpage (https://lbourguignon.shi
perspective/) based on the demo database [324] from the MIMIC-IV project [325, 326].
The webpage allows for comparison between components of a standard table display

(often included as Table 1 to describe the study population, (**Figure 8.2**) left panel) and the added information from the corresponding visual representations (right panel). The user can control multiple parameters by selecting only a subset of the data or stratifying the data by groups (left sidebar). Additionally, the structure given in the app, separating demographics, diagnosis and vital information, eases the workflow for a user who would not be familiar with the data (categories of interest, top panel in **Figure 8.2**). Compared to presenting the data in tabular form only, the corresponding plots help to identify outliers as illustrated by the oxygen saturation data (vitals), where one can easily detect an outlier in the adult subgroup. Similarly, visualizing temperatures (vitals) makes it easier to notice that this variable is most likely reported in two distinct scales (Fahrenheit and Celsius) depending on the entries. Based on this interactive presentation of the data, one can get familiar with the data and formulate hypotheses or research questions to be explored. In our example, one would be encouraged to explore group differences based on sex and age categories. Presenting an interactive app as part of a grant application would also allow the reviewer to asses on their own the feasability of the proposed project based on the data presented, hence strengthening the application.

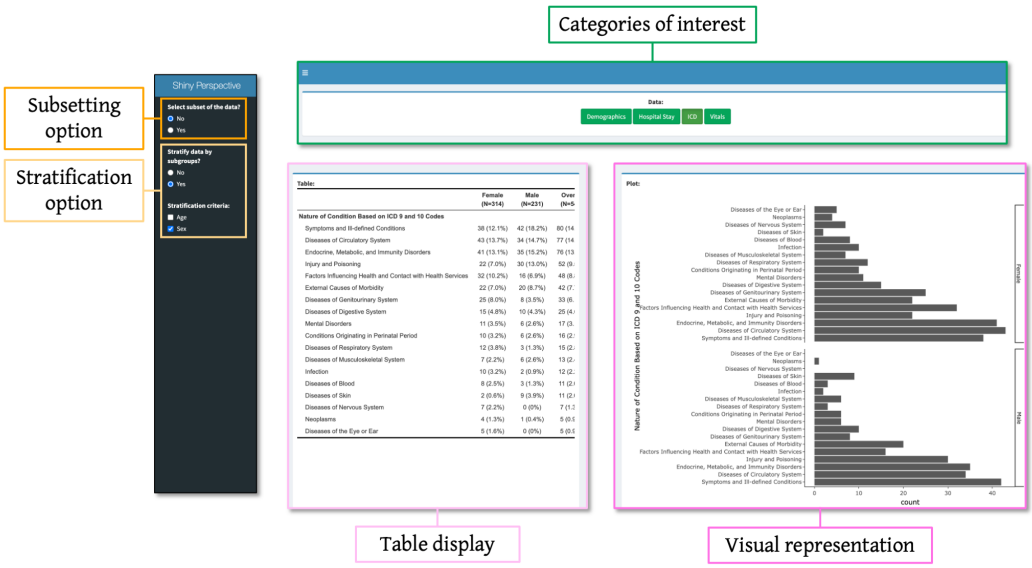


Figure 8.2: Example Shiny app interface.

A key component of web applications, such as the Shiny app presented alongside this perspective, resides in the unique user experience that it allows. The user is not required to have any coding skills to be in control of the data displayed. This holds true in the limit of the flexibility given by the coding team behind the app built. However, a web page, contrary to a manuscript, can be modified and updated even after initial publication to continuously incorporate new features or data, as required or desired by the users.

8.5 Implications for future research

As illustrated by our use-case example, interactive plotting of the data allows for a comprehensive overview of the data at hand, which is currently not the standard format of scientific publication. By contrast, the re-evaluation of clinical trials regularly reveals a lack of transparency, particularly in selective reporting of the results [327]. Introducing new tools such as interactive data visualization would be a straightforward step towards increased transparency, which aligns with the current guidelines for better research reporting, in particular in the case of clinical trials [328]. Ultimately, it would contribute to decreasing bias and improving the overall quality in reporting research outcomes. Additionally, it should be highlighted that this form of data sharing is particularly valuable for medical applications, where data sharing may come with data security and safety concerns. With an online data visualization tool, one can initially circumvent sharing original data and apply certain restrictions (e.g. only display subgroups based on age range, and not on precise age) to ensure anonymity.

We want to reiterate that interactive data visualization holds multiple advantages, while being accessible with a basic programming background. It should therefore be further promoted, encouraged and valued by all actors of the scientific research community, from advisors, to regulatory instances and publishers.

Declarations

Availability of data and materials

Data and code related to the example Shiny app built can be accessed on our GitHub repository [7].

Competing interests

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Funding

This study was supported by the Swiss National Science Foundation (Ambizione Grant [PZ00P3_186101, Jutzeler]). LB was supported by an ICORD student exchange award provided by the Praxis Spinal Cord Injury Foundation.

[7] https://github.com/lbourguignon/shiny-perspective

Authors' contributions

Lucie Bourguignon: design of the study, acquisition, cleaning and visualisation of data, writing of the manuscript
Jaimie J. Lee: cleaning and visualisation of data, revising the manuscript critically for important intellectual content
Ryan Loke: cleaning and visualisation of data, revising the manuscript critically for important intellectual content
John L. K. Kramer: substantial contributions to the design of the study, revising the manuscript critically for important intellectual content
Catherine R. Jutzeler: design of the study, revising the manuscript critically for important intellectual content

Acknowledgements

Parts of this project were conducted on the traditional and unceded territory of the Cayuse, Umatilla and Walla Walla, S'ólh Téméxw (Stó:lō), Hul'qumi'num Treaty Group, səl̓ilwətaʔɬ təməxʷ (Tsleil-Waututh), šxʷməθkʷəy̓əmaʔɬ təməxʷ (Musqueam), Skwxwú7mesh-ulh Temíx̱w (Squamish), and Stz'uminus.

Part D in a nutshell

Part D showcases the importance of data visualization, particularly when data science is applied as the mean to inform other fields, such as medical applications or policy-making. By allowing user-oriented and tailored experiences, data visualization allows for a better understanding of problems at hand. Therefore, visualization tools hold the promise to enhance not only research outcomes themselves, but also their use into clinical practice, in political decision making processes and in communication to the general public, including individuals affected by the conditions studied.

Part E

General discussion

The aim of this thesis was to leverage the potential of data science to maximize clinical impact in SCI research. To achieve this goal, I first utilized existing data sources to benchmark the natural course unfolding following SCI (**Part B**). Here, I investigated the progression of neurological recovery over the last two decades (**Chapter 1**), and of serological markers after injury (**Chapter 2**). Following this, I studied medication prescriptions (**Chapter 3**) and their potential impact on neurological recovery (**Chapter 4**). Subsequently, I applied and modified statistical and ML tools in an attempt to predict neurological recovery (**Chapter 5**). My investigation revealed that such prediction tasks are constrained by the presence of missing data and unexpected recovery patterns, which are specific to SCI data. This led to the development of guidelines on handling missing data in the context of SCI data (**Chapter 6**) and exploration of the unexpected recovery patterns (**Chapter 7**).

While the findings of each individual study have been outlined in the preceding sections, we will now discuss their collective impact on the SCI research domain as a whole, and explore future perspectives that emerge from the remaining limitations.

6 Leveraging data to enhance SCI clinical trials

As exposed previously in this thesis, one of the main concerns of the SCI research community is the search for an intervention benefiting recovery following a SCI. Randomized clinical trials (RCTs) are the gold-standard for testing the safety and effectiveness of an intervention, and many of them have been conducted in the context of SCI research [28]. However, as reported by Martin *et al.*, RCTs come with a high cost, ranging for phase-III clinical trial, testing for safety and effectiveness of an intervention, from 7 to 74 millions US dollars [329]. At times, those high financial stakes may contribute to the early termination of a trial [330].

Beyond financial concerns, early termination of a RCT may be caused by the difficulty in recruiting participants in the trial. Recruitment is particularly critical when studying rare conditions such as SCI and came in the forefront during the worldwide coronavirus disease 2019 (COVID-19) pandemic. This challenge, however, can be tackled by enriching the placebo arm, which is not subject to the intervention tested. This enrichment can be achieved using information from patients whose data have been collected as part of previous observational studies and who were, by definition, not subject to any intervention beyond standard of care. My study showing unchanged neurological recovery patterns over the last two decades (**Chapter 1**) revealed this strategy to be viable. This finding has tremendous impact on both the (clinical) SCI research community and SCI patients as it underlines the stalled progress in the field to provide effective treatments. Moreover, it

provides an opportunity to utilize existing data from individuals who did not receive the intervention of interest for comparison. In the context of current and future RCTs, this enables the maximization of exposure to promising interventions.

Similarly, the design of RCT protocols will be influenced by the study conducted on routine serological markers. These latter are a cornerstone of patient surveillance and play a pivotal role in RCTs. Indeed, those markers are particularly important in scrutinizing the effects of the intervention tested beyond the recovery of the injury itself. Any systematic deviation from the norm may be interpreted as a safety concern and similarly, lead to early termination of a trial. However, detection of deviations from the norm assumes precise knowledge of the norm, adapted to the individuals tested. Establishing this point of reference for individuals with acute SCI was the target of the study reported in **Chapter 2**. We showed that routinely collected serological markers may lay outside of the norm as established for able-bodied individuals in the acute injury phase. This finding allows for adapted and therefore more pertinent comparisons with individuals undergoing a new intervention of interest. This highlights a specific scenario where fluctuations in liver enzymes might be ascribed solely to the tested compound, overlooking variations stemming from the injury itself. Such a misinterpretation could potentially prompt premature termination of the trial. Hence, we provide another avenue to mitigate the limitation of current SCI RCTs.

Finally, as described by Lammertse in [331], promising preclinical findings have to date failed to be translated to human populations. Despite leaving no effective treatment to improve recovery SCI, lessons can, however, also be learnt from these failures. Contrary to the animals in preclinical studies, real-live SCI patients are not only exposed to the treatment of interest. As part of the standard of care, many SCI patients receive antibiotics, to prevent or during infectious episodes, painkillers etc. Reviewing the variety of those medications was the object of **Chapter 3**. We uncovered a vast polypharmacy that effectively makes preclinical and clinical studies incomparable. With exposure to up to 43 unique medications per day, the likelihood of pharmacological interactions is indisputable. As the SCI community recently acknowledged that the answer to enhancing recovery will unlikely lay in a single compound or intervention [28], the characterisation of polypharmacy represents a unique opportunity for drug repurposing. **Chapter 4** outlines my contribution to addressing the challenge of drug repurposing for SCI through a systematic review of nearly 400 publications. Specifically, the publications reviewed discussed the effects on neurological recovery of at least one compound identified as being prescribed after SCI. This review revealed promising candidates for drug repurposing, such as metformin, indicated to manage high blood sugar levels in type 2 diabetes [233].

7 Developing best practices for data analysis in SCI research

The first step towards the use of data science is the access to data. However, data alone is not sufficient. Successful inference from the data requires quality data, appropriate preprocessing and analysis. For an analysis to be relevant to the problem at hand, it relies on detailed knowledge of the methods taken in combination with the peculiarities of the problem setting. In the context of medical research, including SCI research, integration of the clinical background is crucial. **Chapter 5** is a prime example of how considering clinical subgroups in fitting ML models improves the prediction performance overall. We took advantage of clinical observations, reflected in the data, showing that individuals starting with an injury scored with an LEMS of zero in the acute stage after injury are less likely to recover motor functions, compared to individuals whose injury was scored with an LEMS above zero in the same timeframe. This project also depicted how basic approaches in applying ML models for prediction hold numerous limitations, namely the presence of missing data and of individuals recovering beyond clinical expectations. These limitations were the object of ensuing projects of this thesis.

Missing data is the core reason for reduced sample size in analyses relying on complete case data. Given the sparsity and heterogeneity of SCI data, it is key to maximize data usage implying effective handling of missing data. **Chapter 6** addressed this important research question. The specific imputation strategy (or lack thereof) influences the results reported in a downstream analysis. It is therefore necessary to accurately describe and understand the effects of missing data in the context of SCI recovery prediction. Here we performed a simulation study of different patterns of data missingness in SCI registries. By exploiting the characteristics of SCI recovery trajectories, we demonstrated that using imputation based on the last observation carried forward (LOCF) is a viable option in the context of SCI longitudinal data. This finding is in contrast to other research fields where LOCF imputation may not be appropriate, such as in anesthesiology [309] or clinical trials with spaced follow-ups [310]. More specifically, when an observational is missing for an assessment expected around 52 weeks after injury, values evaluated 26 weeks after injury can be used to replace the missing later time point. This reflects the natural recovery trajectory following SCI, where the majority of the recovery will occur in the first six months after initial trauma to reach a plateau thereafter [17]. Our simulation study formally supports this imputation method based on data-driven evidence. It additionally suggested that carrying information forward from earlier time points such as 16 weeks after injury would not introduce more bias compared to other imputation methods. This study will contribute to the augmentation of cohort sizes to be studied in future studies

and provides guidelines to handle missing data. It is therefore an essential building block of best practices in data-driven SCI research.

Aiming to further improve prediction performance, the inspection of individuals for whom predictions were not successful uncovered individuals with recovery trajectories deviating from clinical expectation (e.g., an injury scored with an LEMS of zero in the acute stage after injury undergoing an unexpected improvement). **Chapter 7** identified and described such individuals, referred to as individuals presenting a phenomenal recovery. Our study focused on describing those individuals according to their motor recovery, examined on the myotome level, and their antibiotics prescriptions. We found that individuals exhibiting a phenomenal recovery tend to present with greater impairment in the distal myotomes of the upper limbs compared to the proximal myotomes of the lower limbs. This surprising pattern could relate to previous definitions given of CCS, and further investigations beyond motor scores are warranted. Identifying patterns associated with greater recovery could inform researchers in the development of interventions aiming to improve recovery. Studying prescriptions of antibiotics aligned with our objective of developing drug repurposing in the context of SCI. We hypothesized that differences in recovery could be driven by higher antibiotics intake, reducing inflammation and therefore contributing to better preservation of the tissues affected by the initial trauma [332]. Although the data studied did not support our hypothesis, this study contributes to establishing hypotheses related to drug repurposing as part of the general landscape of data-driven SCI research.

Finally, best practices in research also include research communication and outreach, especially when conducting research at the intersection of multiple fields. Throughout this thesis, the projects conducted led to the development of multiple webpages to visualize raw and processed data and provide users (i.e., clinical partners, individuals affected, and policy-makers) with an interactive experience. **Chapter 8** illustrates how scientific communication with online tools can complement traditional scientific publications, inherently limited in the number and static nature of their visualizations. Specifically, we show how Shiny apps contribute to making data safely accessible and allow for targeted exploration and analysis according to the user's needs. As a result, such webpages promote both transparency and collaboration, with the overarching goal of enhanced research outcomes.

8 Limitations

As mentioned throughout the chapters presented, a number of limitations still need to be overcome. Firstly, the size of SCI data available, while growing, remains small

in comparison with other domains of application of statistical and ML methods. This phenomenon is well illustrated in the context of drug repurposing. ML models have been developed to emulate clinical trials based on electronic health records (EHRs) [333, 334]. These models rely on the combination of causal and deep learning approaches. However, they more importantly rely on big data, including over one hundred million individuals [333] to test around 250 drugs of interest. In SCI, the balance is reversed as illustrated in **Chapter 3**: we identified around 775 drugs prescribed as standard of care across 2040 individuals. For this specific task, the collection of additional data, while commendable, would not suffice considering that SCI is a rare condition. Additionally, the current methods still focus on the effect of unique compounds, which might offer limited applicability to SCI owing to the heterogeneity and complexity of processes occurring following trauma. Hence, alternatives remain to be found.

Related to the size of cohorts collected is the sparsity of data collection. Assessments of bladder, bowel, and sexual function are particularly critical in that regard: Anderson *et al.* showed in [18] that these are ranked higher than walking by individuals living with SCI as major factors to improve their quality of life. However, information about bladder, bowel, and sexual function is often only sparsely reported. In the EMSCI cohort, for example, bowel and bladder functions are assessed through subscores of the SCIM, which is one of variables with highest proportions of missing data at all assessment time points (see **Chapter 1, Additional file 1**), while there is no assessment reflecting sexual function. This absence of relevant data naturally prevents their exploration and calls for further actions in collecting data to address the main areas of interest identified by individuals living with SCI.

Owing to the larger data availability, the focus of this thesis was on neurological recovery. However, the variables at hand also present limitations. Scores such as the LEMS and UEMS can be mathematically identified as aggregates of ordered categorical variables. Indeed, LEMS is constructed from the sum of motor scores of 10 myotomes in the lower extremities, each being scores from zero to five, with 0 being the most severe degree of impairment (see **Table 1**). The grading of individual myotomes is however unequally spaced: for an individual to transition from a myotome graded as zero to one would clinically be different from an individual improving from grade four to five. The latter is expected over the natural course of recovery, while the former is more unlikely to occur. This relevance of this subtlety becomes crucial when appropriately modeling SCI. For simplicity, the studies presented here considered aggregate scores such as LEMS as continuous variables, inherently assuming a linear improvement for an equal increment in the score. Additionally, the aggregate nature of the scores tends to mask clinically relevant distinctions between individuals. This phenomenon was particularly

illustrated in **Chapter 7** and Lukas *et al.* (unpublished), where comparable aggregate scores present significant differences when examining the details at the myotome level. Hence, further work beyond aggregate scores is warranted to better represent the reality of highly heterogeneous clinical presentations.

9 Future perspectives

This thesis illustrates how combining clinical knowledge with existing methods can improve the application of data-driven methods to SCI research. However, numerous clinical aspects remain to be integrated. As described above, the clinical measurements collected as part of daily clinical practice, such as the LEMS or UEMS, are summing scores from the evaluation of individual motor scores, effectively losing levels of details and the inherent dependence between the individual motor scores. A natural next step would be to model motor score sequences in structures such as graphs and build prediction tasks around this elevated, more accurate representation of the motor function. The newly created representation of the data can then be analyzed with more complex models such as graph neural networks.

Going further, an improved description of an individual would require additional information. This can be achieved by combining multiple data modalities, and expanding modeling to combine neurological scores (i.e., motor and sensory scores and their aggregates) and demographics with imaging, or electrophysiology measurements for example. Insight into the way how clinical trials are conducted provides an overview of all modalities employed to monitor individuals after injury. The NISCI clinical trial is a good example in that regard. Conducted from 2019 to 2023 in Switzerland and across Europe, NISCI tested the effect of antibodies against nogo-A, which had been suggested as being beneficial for recovery after SCI [41, 335, 336, 40]. While the primary outcome relied on changes in UEMS, secondary outcomes did not only capture various measures of bowel and bladder functions, including but not restricted to SCIM, but also measurements of nerve conducting velocity, and somatosensory evoked potentials [8]. The latter two have shown to be associated with long-term recovery after SCI [337, 338]. Consistently gathering multiple data modalities across individuals therefore promises to more accurately represent the state of individuals at a personal level. Taking the example of other medical fields such as oncology, the field is moving towards the collection of not only larger but also wider cohorts, gathering multi-model inputs [339]. The resultant investigation of dedicated models allowing for multi modality integration suggest favourable development for SCI research in the next decade.

[8] ClinicalTrials.gov: NCT03935321

10 Conclusion

In search of an intervention that would benefit individuals with SCI in recovering from their injury, data science tools offer a new opening. The integration of clinical knowledge within data-driven approaches guides the SCI research community towards a better understanding of the condition and of the research conducted to date, as illustrated through **Part B** in this thesis. **Part C** further elaborates on adapting data science tools to specific characteristics of SCI, such as plateaued recovery and high heterogeneity. Finally, any research only becomes valuable when communicated between fields, from clinical data collection, to data analyses in a research context, and back to the bedside. Those transitions can equally be supported by effective data visualization as demonstrated in **Part D**. Taken together, the contributions of this thesis participate in revising the approaches employed to discover interventions to improve recovery following SCI.

Appendices

11 Supplementary materials

11.1 Chapter 1

Supplementary material was made publicly available with the corresponding manuscript.

11.2 Chapter 2

Supplementary material was made publicly available with the corresponding manuscript.

11.3 Chapter 3

Supplementary material was made publicly available with the corresponding manuscript.

11.4 Chapter 4

Table 8.1: List of drugs included in analysis.

Drug(s) tested	Number of publications
acetylcysteine	5
acetylsalicylic acid	1
albumin	3
aluminum	1
amiloride	4
amphetamine	2
atorvastatin	9
azithromycin	2
baclofen	1
botulinum toxin	1
bupivacaine	1
buspirone	2
calcitriol	2
carbidopa levodopa	1
carvedilol	2
ceftriaxone	2
ceftriaxone + acetylcysteine	*1*
celecoxib	1
chlorpromazine	1
citalopram	1
clonidine	1
clopidogrel	1
cyproheptadine	3
dantrolene	4
dapsone	1
darbepoetin	1
dexamethasone	15
dexamethasone + estrogen	*1*

Continued on next page

Table 8.1: List of drugs included in analysis. (Continued)

Drug(s) tested	Number of publications
dexamethasone + melatonin	*1*
dexmedetomidine	2
diclofenac	1
epinephrine	1
epinephrine + nitroprusside	*1*
epoetin	4
epoietin	2
escitalopram	2
estradiol	18
estradiol + testosterone	1
estrogen	3
ethanol	2
ethanol + isoflurane	*1*
ethanol + ketamine + pentobarbital	*1*
etomidate	1
etomidate + epoietin	*1*
etomidate + methylprednisolone	*1*
ezetimibe	1
ezetimibe + simvastatin	*1*
fenofibrate	1
fentanyl + nitrous oxide	*1*
fentanyl + nitrous oxide + naloxone	*1*
fluoxetine	4
fluoxetine + vitamin c	*1*
folic acid	2
folic acid + nitrous oxide	*1*
gabapentin	1
glibenclamide	1

Continued on next page

Table 8.1: List of drugs included in analysis. (Continued)

Drug(s) tested	Number of publications
glucosamine	1
glutamine	2
heparin	2
hydralazine	1
ibuprofen	4
immune globulin	3
indomethacin	3
ketoprofen	1
levocarnitine	1
levodopa	2
lidocaine	2
liothyronine	1
lithium	8
magnesium	2
magnesium + methylprednisolone	*1*
magnesium chloride + polyethylene glycol	*3*
magnesium sulfate	5
magnesium sulfate + polyethylene glycol	*2*
mannitol	3
melatonin	21
meloxicam	1
metformin	5
methotrexate	3
methylprednisolone	81
methylprednisolone + acetylcysteine	*1*
methylprednisolone + epoietin	*1*
methylprednisolone + magnesium chloride + polyethylene glycol	*1*
methylprednisolone + magnesium sulfate	*1*

Continued on next page

Table 8.1: List of drugs included in analysis. (Continued)

Drug(s) tested	Number of publications
methylprednisolone + melatonin	*1*
methylprednisolone + methotrexate	*1*
methylprednisolone + mycophenolate	*1*
methylprednisolone + pregabalin	*1*
methylprednisolone + rosuvastatin	*1*
methylprednisolone sodium succinate	23
methylprednisolone sodium succinate + aminocaproic acid	*1*
methylprednisolone sodium succinate + dantrolene	*1*
methylprednisolone sodium succinate + vitamin c	*1*
mexiletine	2
minocycline	22
minocycline + tacrolimus	*1*
modafinil	1
montelukast	2
morphine	6
morphine + minocycline	*1*
morphine sulfate	2
mycophenolate	1
naloxone	24
naltrexone	1
naproxen	2
niacin	1
nicotine	1
nifedipine	1
nitrous oxide	1
omega 3	5
oxandrolone	1
pentobarbital	1

Continued on next page

Table 8.1: List of drugs included in analysis. (Continued)

Drug(s) tested	Number of publications
phenytoin	4
pioglitazone	3
plasma	1
platelets	1
polyethylene glycol	10
prednisolone	1
prednisone	1
pregabalin	2
progesterone	3
progesterone + vitamin d	*1*
propofol	2
selegiline	1
sevoflurane	1
simvastatin	8
sitagliptin	1
tacrolimus	8
tadalafil	1
tamoxifen	8
testosterone	2
theophylline	1
thiopental	1
thiopental + naloxone	*1*
topiramate	3
tramadol	1
trifluoperazine	1
ubiquinone	1
valproic acid	10
vitamin c	3

Continued on next page

Table 8.1: List of drugs included in analysis. (Continued)

Drug(s) tested	Number of publications
vitamin c e	1
vitamin d	2
vitamin e	2
zinc	4

Rows in *italic* highlight drugs tested in combination.

Table 8.2: Neurological and functional outcomes for animal studies included in the review.

Category	Harmonised assessment name	Assessment name as reported in literature
locomotion	BBB	Basso Beatie Brenahan (BBB) locomotor scale
		Basso Beattie and Bresnahan (BBB) rating scale
		Basso Beattie Brenahan (BBB) locomotor scale
		Basso-Beatie-and Bresnahan (BBB) scale
		Basso-Beatie-Bresnahan (BBB) scale
		BBB
		BBB hind limb locomotor rating scale
		BBB lcoomotor score
		BBB locomoter scale
		BBB locomotor rating scale
		BBB locomotor scale
		BBB locomotor scale (canine)
		BBB locomotor scale (modified)
		BBB locomotor scale (mouse version adapted to local protocol)
		BBB locomotor scale (mouse version)
		BBB locomotor score
		BBB Locomotor test
		BBB locomotor test

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		BBB method
		BBB rating scale
		BBB scale
		BBB score
		BBB scoring
		BBB scoring scale
		BBB scoring system
		BBB subscores
		BBB subscoring
		BBB test
		modified BBB hindlimb locomotor scale
		modified murine BBB hindlimb locomotor rating scale
		modified murine BBB hindlimb locomotor-rating scale
		modified murine BBB scale
		straight alley BBB
	BMS	Basso mouse scale
		Basso Mouse Scale (BMS)
		Basso Mouse scale (BMS)
		Basso mouse scale (BMS)
		BBB locomotor scale (mouse version adapted to local protocol)
		BMS

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		BMS scale
		BMS score
	beam walk test	beam walk
		beam walk test
		beam walk tests
		Beam walking test
		narrow beam crossing test
		narrow beam test
		narrow beam test
		narrow beam-crossing test
		tapered beam test
		tapered beam walk test
	footprint analysis	foot print analysis (fine motor control)
		footprint analysis
		footprint analysis (fine motor control)
		footprint recording
	gait analysis	2D hindlimb kinematics during weight-supported treadmill locomotion
		3D kinemtic data
		angulograms (quality and range of motion)
		base of support
		catwalk gait analysis
		CatWalk gait analysis

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		Catwalk-automated quantitative gait analysis
		Gait analysis
		gait analysis (DigiGait)
		gait analysis with CatWalk XT 10.6 multivariate system
		gait recording
		hind limb gait
		kinematic analysis with the CatWalk gait analysis system
		kinematic profile
		locomotion analsyis with MotoRater apparatus
		locomotor analysis with MotoRater apparatus
		toe spread index
	grid walking test	grid walk test
		gridwalk test
		grid walking test
		grid-walking test
		horizontal grid walking
		ability to traverse wire grid
		horizontal grid
		grid footfalls
		grid walking

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
	inclined plane test	angled plane score
		incline plane score (IPS)
		incline plane test method
		inclined plane
		inclined plane assessemnt
		inclined plane assessment
		inclined plane method
		inclined plane method of Rivlin and Tator
		inclined plane score
		inclined plane score (IPS)
		inclined plane task
		inclined plane technique
		inclined plane test
		inclined plane test (modified Rivlin's method)
		inclined plane test method
		inclined plane tests
		inclined plate test
		inclined test
		rivlin and tator's inclined plane test
		Rivlin inclined plane test
	ladder walk test	45 degrees ladder walk test
		footfalls
		horizontal ladder

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		horizontal ladder crossing test
		horizontal ladder task
		horizontal ladder test
		horizontal ladder test (adapted to local protocol)
		horizontal ladder walk test
		horizontal ladder walk tests
		ladder walk
		ladder walk test
		ladder walk tests
		walk on ladder
	locomotor (other)	activity box
		activity box test (ABT)
		activity measures
		categorisation of walking ability (paraplegia/poor walker/walker)
		clinical grading
		clinical motor exam (Drummond and Moore)
		Drummond and Moore criteria
		Drummond and Moore motor function score

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		Eugene D Means and Douglas K Anderson's motility score
		Forelimb locomotor scale
		grading of motor disturbance (Drummond and Moore scale)
		gross motor activity (activity box)
		hind limb motor function score (MFS)
		motor capacity
		motor deficit index
		motor function
		motor function scale
		motor function scale (Farooque)
		motor performance on rotarod
		neurological function (walking status)
		neurological scores (locomotor status)
		Open field test
		open field test
		porcine thoracic behavior scale

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		presence/absence of hind-limb paralysis
		recovery index (mobility)
		rotarod
		rotarod locomotor function test
		spontaneous movement
		unprompted walking motor score
	swimming	swimming performance
		swimming test
	Tarlov scale	five-point modified Tarlov scale
		hind limb motor function (modified Tarlov)
		hind-limb motor-function according to Tarlov
		modified five-point scale developed by Tarlov
		modified Tarlov method
		modified Tarlov rating system
		modified Tarlov scale
		modified Tarlov scale
		modified Tarlov score
		modified Tarlov scoring system

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		modified tarlov's grading scale
		modified Tarlov's motor scale
		motor function (modified Tarlov scale)
		Tarlov motor scale
		Tarlov scale
		Tarlov scoring
		Tarlov scoring system
		Tarlov's scoring system
		Tarlow scale
forelimb function	grip strength	grip strength meter
		grip strength task
	reaching or retrieval	directed forepaw reaching (DFR)
		grasping test (food retrieval)
		modified Montoya's staircase test
		Montoya staircase reaching
		staircase test
		vermicelli handling test
	rearing	cylinder rearing test
		cylinder test (forelimb assymetry)

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		open field test (rearing)
		paw placement
		rearing
sensory and pain	mechanical reactivity	cutaneus trunchi muscle reflex
		cutaneus trunci muscle (CTM) reflex
		foot withdrawal under mechanical stimuli
		girdle test
		localisation reflex
		mechanical reactivity
		mechanical reactivity (von Frey)
		mechanical sensitivity
		mechanical sensitivity (von Frey filaments)
		proprioception
		proprioceptive placing response
		response to mechanical stimuli
		sensory function (paw withdrawal)
		sensory function (von Frey filaments)

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		sensory testing (forelimb withdrawal under mechanical stimulation)
		tactile capacity
		tactile reactivity
		tactile reactivity (girdle test)
		tactile sensory test with Von Frey filaments
		tape sensing and removal test
		touch-evoked agitation
		vocal/sensory score
		vocalization threshold to mechanical pressure
		Von Frey test
		von Frey test
		Von Frey testing
	other reflexes	physiological reflexes
		test of hindlimb reflexes
	pain	gross neurologic examination
		hindpaw pinprick sensory threshold test
		hindpaw pinprick sensory treshold test

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		hindpaw withdrawal threshold for mechanical allodynia
		hindpaw withdrawal treshold for mechanical allodynia
		mechanical allodynia
		mechanical allodynia testing
		painful stimulus by pinching of rat tail
		paraplegia status (tail pinch)
		pinprick
		purposeful response to paw pinch
		Rat Grimace Scale
		response to noxious stimulation
		sensitivity to pain
		Von Frey test of mechanical allodynia/hyperalgesia
	thermal reactivity	acetone drop test
		Hargreave's test
		hot-water test
		neuropathic pain evaluation [acetone drop test and thermal hyperalgesia]

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		nociceptive reactivity (thermal shock threshold tested through tail-flick test)
		nociceptive reactivity (thermal)
		sensory blockade (heat)
		sensory function (hot plate/cold stimulation)
		tail flick test
		tail-flick test
		thermal hyperalgesia
		thermal reacitivity (standard hot-plate test/cold stimulation)
		thermal reactivity
		thermal sensitivity
		Thermal Sensitivity
		thermal sensitivity (tail flick)
	toe spread test	toe spread test
		toe-spread test
		toe spread tests
		toe spread
		toe spread reflex
electrophysiology	electrophysiology (other)	activity in hemidiaphragm and phrenic nerve ipsilateral to hemisection

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		assessment of H-reflex
		compound action potential (CAP) recording
		compound action potentials
		EMG recordings
		frequency dependent depression (FDD) of H-reflex
		H-reflex analysis
		sciatic nerve stimulation
	motor evoked potentials	corticomotor evoked potentials
		corticomotor evoked potentials (CMEPs)
		evoked muscle responses (EMR)
		evoked potential test (MEP)
		motor evoked potential
		motor evoked potential (MEP)
		motor evoked potentials
		motor evoked potentials (MEP)
		motor evoked potentials recording
		Motor-evoked potential (MEP)

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		motor-evoked potential (MEPs)
		motor-evoked potentials (MEPs)
		rubrospinal motor evoked potentials (rMEP)
		spinal motor–evoked potentials (sMEPs)
	somatosensory evoked potentials	cortical somatosensory evoked potentials
		cortico somatosensory evoked potentials (CSEP)
		evoked potentials measured
		SEPs
		somatosensory evoked potential (SEP)
		somatosensory evoked potential (SEPs)
		somatosensory evoked potential (SSEP)
		somatosensory evoked potentials
		somatosensory evoked potentials (SEP)
		somatosensory evoked potentials (SEPs)

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
		somatosensory evoked potentials (SSEP)
		somatosensory evoked potentials (SSEPs)
		somatosensory evoked responses (SER)
		somatosensory-evoked potential (SEPs)
		somatosensory-evoked potentials (SEPs)
		somotosensory evoked potentials (SSEP)
		SSEP
		SSEPs
	spinal cord evoked potentials	spinal cord evoked potential recording
		spinal cord evoked potentials
		spinal cord evoked potentials (SCEPs)
		spinal evoked potentials (SEP)
other	composite scores	motor sensory deficit index (MSDI)
		neurologic scores (motor and sensory deficit)
		sensory and motor evaluations (paraplegia status)

Continued on next page

Table 8.2: Neurological and functional outcomes for animal studies included in the review. (Continued)

Category	Harmonised assessment name	Assessment name as reported in literature
	Gale scale	combined behavioral score (Gale scale/CBS)
		functional deficits scoring
		Gale scale
		gale scale
		modified Gale scale
		motor function scale (modified Gale)
		motor function scale according to Gale et al. (1985)
		motor function score (modified Gale)
		overall hindlimb impairment (modified CBS)
	hindfoot bar grab test	hindfoot bar grab test
		hindfoot bar grab tests
	micturition	bladder function
		micturition (voiding behaviour)
	spinal cord blood flow	spinal cord blood flow
		spinal cord blood flow (SCBF)

Table 8.3: Neurological and functional outcomes for human studies included in the review

Category	Assessment name reported	Assessment
neurological	neurological (motor and sensory)	marked recovery (combination of improvement in AIS grade and walking function)
		pinprick, light touch, motor function scale
		ASIA motor and sensory scores
		ASIA scale: motor and sensory composites
		ASIA motor score, ASIA sensory score
		motor score; light touch (LT) and pin prick (PP) scores
	neurological (motor)	ASIA motor score
		ASIA Motor score
		ISNCSCI motor score
		strength
		discharge motor score
	neurological (other)	improvement to level of injury (change in segment to more caudal location)
	injury severity	marked recovery (combination of improvement in AIS grade and walking function)
		improvement in ASIA scale
		ASIA impairment score and grade
		ASIA grade
functional	functional (general)	Spinal Cord Independence Measure

Continued on next page

Table 8.3: Neurological and functional outcomes for human studies included in the review (Continued)

Category	Assessment name reported	Assessment
		Functional Independence Measure
		London Handicap scale
		Short Form 36 Questionnaire
		FIM discharge score
	functional (mobility)	FIM motor score
	functional (mobility and general)	Walking Index for SCI II (WISCI II), Spinal Cord Independence Measure II (SCIM II)
	functional (mobility and spasticity)	overground walking performance; treadmill walking performance
		spastic reflexes (modified Ashworth scale); walking function
electrophysiology		EMG

American spinal injury association (ASIA) impairment scale (AIS), light touch (LT), pin prick (PP), international standards for neurological classification of spinal cord injury (ISNCSCI), functional independence measure (FIM), walking index for spinal cord injury (WISCI), spinal cord independence measure (SCIM), electromyography (EMG)

Table 8.4: Variables extracted from studies included for analysis.

Category	Variable extracted	Details
General information	Person in charge	Person in charge of the data extraction
General information	Authors	First author et al
General information	Year	Year of publication
General information	Title	Full title
General information	DOI or PMID	Unique identifier
General information	Language	Language of the main text
Inclusion/exclusion	Included/excluded	Included or excluded
Inclusion/exclusion	Reason for exclusion	Primary reason of exclusion
Inclusion/exclusion	Reason for exclusion	Reason of exclusion if primary reason of exclusion is "out of scope"
Inclusion/exclusion	Reason for exclusion (description)	Description of the reason of exclusion
Classification	Data collection	Prospective or retrospective (human studies only)
Classification	Analysis	Prospective or retrospective (human studies only)
Study population	Species	Species studied among humans, mice, rats, dogs, cats, fish, lampreys, sheep, rabbits, guinea pigs, others
Study population	Species information	Information about subspecies used
Study population	Count, n	Total number of subjects reported

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Study population	Count, n control group	Number of subjects in control group (included in analysis)
Study population	Count, n died in control group	Number of subjects assigned to control group not included in analysis due to premature death
Study population	Count, n excluded in control group	Number of subjects assigned to control group not included in analysis for other reasons
Study population	Count, n treatment group	Number of subjects in treatment group (included in analysis)
Study population	Count, n died in treatment group	Number of subjects assigned to treatment group not included in analysis due to premature death
Study population	Count, n excluded in treatment group	Number of subjects assigned to treatment group not included in analysis for other reasons
Study population	Comment on counts	Details on counts, especially when total control + total treament do not add to total n
Study population	Sex (n, ratio, percentage)	Information about sex of subjects as reported in the publication

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Study population	Sex	One option among female, male, mixed and not reported
Study population	Sex (%, male)	% male included in the study
Study population	Age [days, months, years]	Information about age of subjects as reported in the publication
Study population	Age (mean)	Mean age (when applicable)
Study population	Age (SD)	SD age (when applicable)
Study population	Age (min)	Minimum age (when age range reported)
Study population	Age (max)	Maximum age (when age range reported)
Study population	Age (units)	Age units used among days, weeks, months and years
Study population	Age (comments)	Comment on age information, one option among not reported, adult, young, for publication not reporting precise age included
Study population	Weight [g, kg, pounds]	Information about weight of subjects as reported in the publication
Study population	Weight (mean)	Mean weight (when applicable)
Study population	Weight (SD)	SD weight (when applicable)

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Study population	Weight (min)	Minimum weight (when weight range reported)
Study population	Weight (max)	Maximum weight (when weight range reported)
Study population	Weight (unit)	Weight units among g, kg, pounds
Study population	Weight (comments)	Comment on weight information (e.g., not reported)
Study population	Injury characteristics (level, severity)	Information about injury characteristics included level and severity as reported in the publication
Study population	Injury level	Level of injury (unique level or range for animal studies, number of subject per level or category for human studies)
Study population	Injury severity	Injury severity among moderate, mild, severe, complete, incomplete, paraplegia, tetraplegia, not reported, mixed and moderate-severe

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Study population	Injury mechanism	Injury mechanism among contusion, compression, distraction, dislocation, transection, ischemia, trauma and others. Note this classification mainly applies for animal models, injury mechanism reported may differ in human studies
Study population	Injury mechanism (details)	Details on injury mechanisms (e.g., height and weight used in contusion injuries, time before repurfusion in ischmic injuries etc)
Study population	Duration of SCI	Duration of SCI before euthanasia (animals) or duration of SCI before inclusion in study (human)
Drug information	Drug(s)	Drug(s) studied in the publication
Drug information	Drug name harmonized	Drug name harmonized based on [Bourguignon et al., 2022]

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Drug information	MP used as main drug?	Yes or no, for publication investigation methylprednisolone and methylprednisolone sodium succinate only (assess if the drug was the main drug of interest or used as positve control)
Drug information	Dose (absolute dose or mg/kg)	Dose given
Drug information	Time (minutes pre-injury, minutes post-injury)	Timing of start of treatment compared to injury
Drug information	Duration of treatment	Duration of treatment
Drug information	Timing (e.g., BID, PID)	Frequency of treatment
Drug information	Route	Route used for drug administration
Drug information	Route (comment when multiple)	Comments on the route used
Neurological and functional assessment	What was assessed? (e.g., neurological, functional recovery, spasticity, walking function, electrophysiology)	Type of neurological/functional assessment (broad categories)
Neurological and functional assessment	Name/type of asessement	Neurological/functional assessments as named in the publication
Neurological and functional assessment	Name of assessment harmonised	Neurological/functional assessments' names harmonised as described in Table S3

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Neurological and functional assessment	Details on assessement	Details on assessments as described in the publications
Neurological and functional assessment	Timing of assessement	Time of assessment with respect to the injury
Neurological and functional assessment	Assessment on day 28 (yes/no)	Whether subjects were assessed at day 28 after injury (applies to experiments testing methylprednisolone and methylprednisolone sodium succinate only)
Neurological and functional assessment	Was observer blinded?	Options among no, yes and not reported
Neurological and functional assessment	Drug effect on functional assessment	Options qualifying effects among positive, negative, no effect, mixed (assessment), mixed (dosage), mixed (timing), mixed (regime), no stats, mixed (stats/no stats), mixed (assessment) + mixed (timing), not reported, mixed (dosage) + mixed (timing), mixed (dosage) + mixed (assessment), mixed (dosage) + mixed (regime), mixed (assessment) + mixed (regime)

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Neurological and functional assessment	Drug effect on functional assessment (details)	Details on the effects reported allowing to categorize the effects in the previous column
Neuroanatomical assessments	What was assessed? (e.g., histological measures, cavitity measures, ect)	Type of histological assessment (broad categories)
Neuroanatomical assessments	Name/type of asessement	Histology assessments as named in the publication
Neuroanatomical assessments	Timing of assessement	Time of assessment with respect to the injury
Neuroanatomical assessments	Was observer blinded?	Options among no, yes and not reported
Neuroanatomical assessments	Drug effect on neuroanatomical assessment	Options qualifying effects among positive, negative, no effect, mixed (assessment), mixed (dosage), mixed (timing), mixed (timing of assessment), no stats, mixed (stats/no stats), mixed (assessment) + mixed (dosage), not reported, and mixed (assessment) + mixed (timing)
Neuroanatomical assessments	Drug effect on neuroanatomical assessment (details)	Details on the effects reported allowing to categorize the effects in the previous column

Continued on next page

Table 8.4: Variables extracted from studies included for analysis. (Continued)

Category	Variable extracted	Details
Conclusions and others	Drugs given to treat infections/pain ect.	Other drugs given to subjects according to the study protocol (e.g., pain relief plan, infection treatment or prophylaxis, anesthesia)
Conclusions and others	Conclusion of study	Conclusions as reported in the publication
Conclusions and others	Limitations	Limitations mentioned in the publication
Conclusions and others	Remarks/Comments	Personal remark or comments following extractions
Conclusions and others	Combination of drugs tested	Options among no, yes (drug of interest + drug of interest), and yes (drug of interest + drug not of interest)
Conclusions and others	Contradictions present in the results	Yes or no, flags contradictions between text and figures presented in a given manuscript

Table 8.5: Details on the bias classification for animal experiments.

Domain of bias	Classified as “unclear risk of bias”	Classified “high risk of bias”
Dose	No precise dose reported, including “high dose”	Not reported
Species	Subspecies not reported	Not reported
Route	-	Not reported
Level of injury	No precise level or range reported, including “cervical”, “mid-thoracic, ‘thoracic”, “lumbar-sacral”	Not reported
Treatment time	-	Not reported
Results	Mixed results due to lack of statistics reported, including "mixed (stats/no stats)", "mixed effects (assessment) + mixed (stats/no stats)", "no stats"	Not reported
Sample size	Sample size reported as range or bounded	Not reported
Sex	Mixed population (male/female) with ratio not reported	Not reported
Blinding	Not reported	No blinding applied
Age	Reported as “adult”, “young” with no precise age reported	Not reported

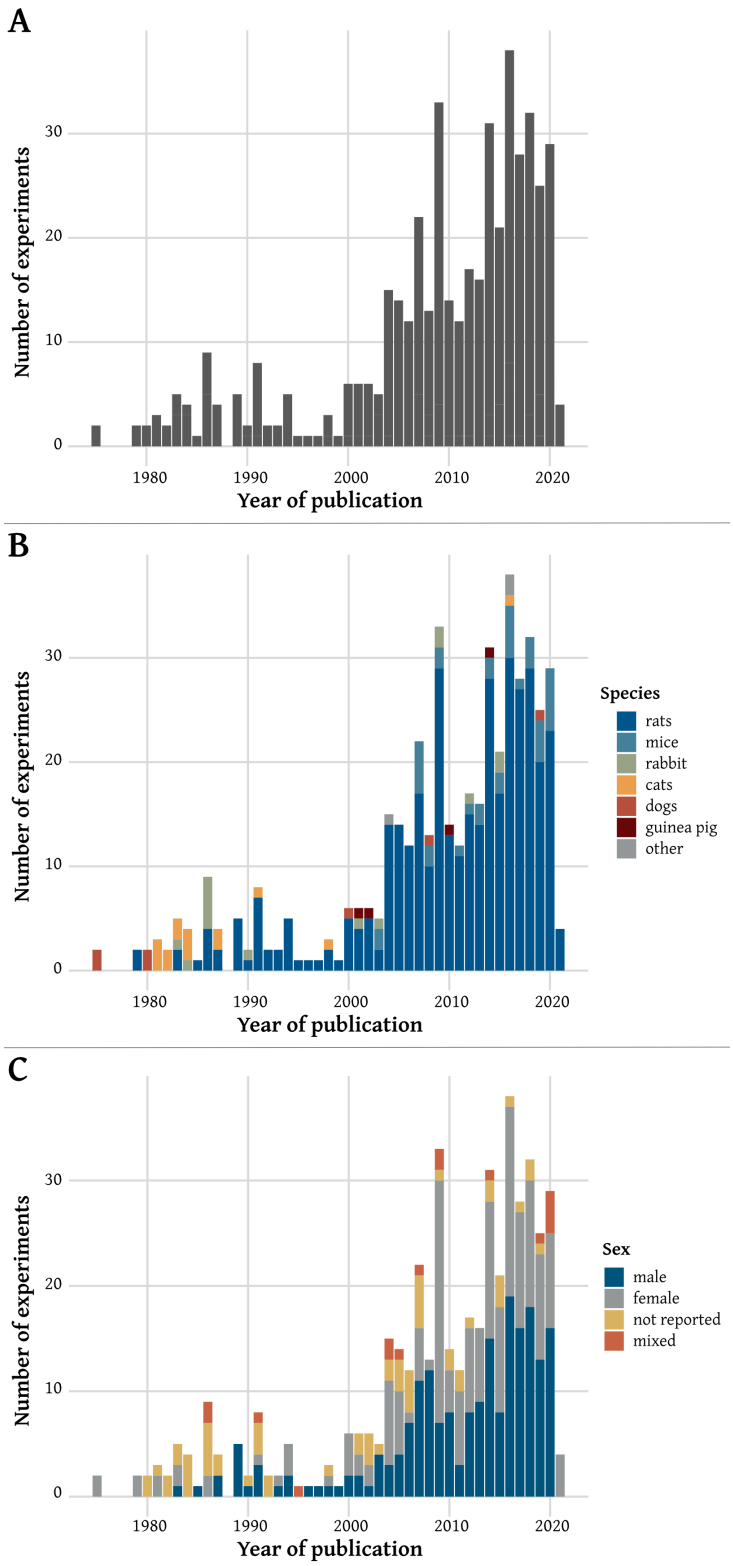


Figure 8.3: Publication trends over time. A. General overview of the number of experiments included per year of publication; **B.** Details of repartition of species used in animal models over time; **C.** Details of repartition of sex in animal models over time.

Table 8.6: Details on the bias classification for animal experiments.

Species	Mean	SD	Median	Q1	Q3
cats	26.53	16.50	24.00	16.50	31.00
dogs	33.43	25.13	26.00	22.00	32.00
guinea pig	21.00	7.55	20.00	17.00	24.50
mice	120.62	83.33	96.00	50.25	176.50
other	31.33	16.17	22.00	22.00	36.00
rabbit	69.36	50.57	47.00	28.50	133.00
rats	61.60	46.71	48.00	32.00	79.50

“Other” include Yucatan miniature pigs ($n = 2$) yellow eel Anguilla anguilla L. ($n = 1$); standard deviation (SD), first quartile (Q1), third quartile (Q3)

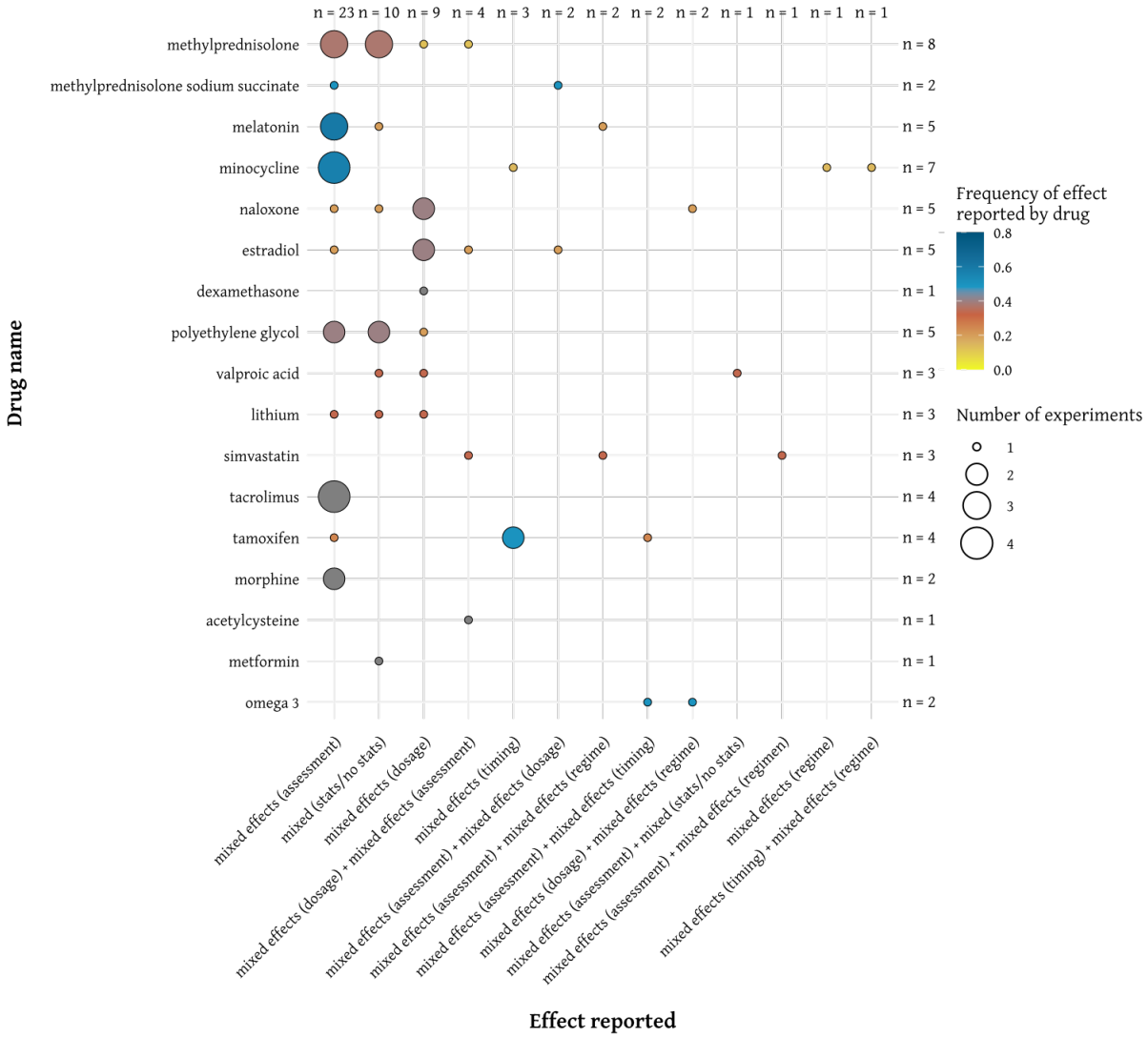


Figure 8.4: Details of the mixed drug effects reported for drugs studied in at least five experiments. Circle size is proportional to the number of experiments reporting the effect of interest. Circles are colored proportionally to the frequency that the effect of interest represents among all experiments studying the drug of interest.

Table 8.7: Bias assessment by animal experiment.

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Pinzon et al. (2008, minocycline)	0	0	0	1	0	0	0	0	0	0	1
Sharp et al (2013, ibuprofen)	0	0	0	0	0	0	0	0	0	0	0
Liu et al (2015, omega 3)	0	1	0	1	0	0	0	0	0	0	2
Bimbova et al (2018, atorvastatin)	0	0	0	1	0	0	0	0	0	1	2
Liu et al (2017, omega 3)	0	1	0	1	0	0	0	0	0	0	2
Yang et al (2016, niacin)	0	0	0	1	0	0	0	0	0	0	1
Jiang et al (2004, MP)	0	1	0	1	0	0	0	0	0	0	2
Halt et al (1992, ethanol + isoflurane)	1	0	2	2	0	0	0	0	0	0	5
Halt et al (1992, ethanol + ketamine + pentobarbital)	1	0	2	2	0	0	0	0	0	0	5
Durham-Lee et al (2011, amiloride)	0	0	0	1	0	0	0	0	0	1	2
Imai et al (2018, amiloride)	0	1	0	0	0	0	0	0	0	1	2
Krisa et al (2012, amphetamine)	0	0	0	1	0	0	0	0	0	0	1
Hook et al (2011, morphine)	0	0	0	0	0	0	0	0	0	0	0
Gao et al (2014, MP)	0	0	0	1	0	0	0	0	0	1	2
Baiyila et al (2018, MP)	0	0	2	1	0	0	2	0	0	1	6
Bilginer et al (2009, MP)	0	0	0	1	0	0	0	0	0	1	2
Bilginer et al (2009, mycophenolate)	0	0	0	1	0	0	0	0	0	1	2
Bilginer et al (2009, MP + mycophenolate)	0	0	0	1	0	0	0	0	0	1	2
Hong et al (2020, vitamin c)	0	1	0	0	0	0	0	0	0	0	1
Martins et al (2018, dantrolene)	0	0	0	0	0	0	0	0	0	0	0
Gao et al (2016, atorvastatin)	0	0	0	0	0	0	0	0	0	1	1
Déry et al (2009, atorvastatin)	0	0	0	1	0	0	0	0	0	1	2
Yeng et al (2016, estradiol)	0	0	0	1	0	0	0	0	0	0	1
Genovese et al (2005, melatonin)	0	0	0	2	0	0	0	2	0	1	5
Pannu et al (2005, atorvastatin)	0	1	0	2	0	0	0	0	0	0	3
Nash et al (2002, MP)	0	0	2	1	0	0	0	0	0	0	3
Zhang et al (2015, azithromycin)	0	1	0	0	0	0	0	0	0	0	1
Faden et al (1981, naloxone)	1	0	0	1	0	0	0	0	0	0	2
Giulian et al (1990, dexamethasone)	0	1	2	2	1	0	0	0	0	1	7
Salzman et al (1991, cyproheptadine)	0	0	0	2	0	0	0	0	0	0	2
Siriphorn et al (2012, estradiol)	0	0	0	1	0	0	0	0	0	0	1
Mohammadshirazi et al (2019, lithium)	0	0	0	1	0	0	0	0	0	1	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Rabchevsky et al (2002, MPSS)	0	0	0	1	0	0	0	0	0	0	1
Borgens et al (2001, PEG)	1	0	2	1	1	0	0	0	0	1	6
Ditor et al (2007, PEG)	0	0	0	2	0	0	0	0	0	0	2
Ditor et al (2007, magnesium sulfate)	0	0	0	2	0	0	0	0	0	0	2
Ditor et al (2007, magnesium sulfate + PEG)	0	0	0	2	0	0	0	0	0	0	2
Liu et al (2015, carvedilol)	0	0	0	0	0	0	0	0	0	0	0
Diaz-Ruiz et al (2011, dapsone)	0	0	0	1	0	0	0	0	0	0	1
Krityakiarana et al (2016, melatonin)	1	0	0	0	0	0	0	0	0	1	2
Vanicky et al (2002, MPSS)	0	0	0	2	0	0	0	0	0	0	2
Behrmann et al (1994, MPSS)	0	0	0	2	0	0	0	0	0	0	2
Sadanaga et al (1989, chlorpromazine)	0	1	0	2	0	0	0	0	0	0	3
Gueye et al (2015, vitamin d)	0	0	0	0	0	0	0	0	0	1	1
Guth et al (1994, indomethacin)	0	0	0	2	0	0	2	0	2	0	6
Nazemi et al (2020, minocycline)	0	0	0	1	0	0	0	0	0	1	2
Lopez et al (2004, bupivacaine)	0	0	0	0	0	0	0	0	0	0	0
Namjoo et al (2018, estradiol) - rats	0	0	0	1	0	0	0	0	0	0	1
Çavus et al (2014, MP)	0	0	0	0	0	0	0	0	0	0	0
Çavus et al (2014, acetylcysteine)	0	0	0	0	0	0	0	0	0	0	0
Çavus et al (2014, MP + acetylcysteine)	0	0	0	0	0	0	0	0	0	0	0
Kang et al (2017, estradiol)	0	0	0	0	2	0	0	0	0	1	3
Baltin et al (2021, MPSS)	0	0	0	2	0	0	0	0	0	1	3
Chen et al (2018, MP)	0	1	0	1	0	0	0	0	0	1	3
Caliskan et al (2016, etomidate)	0	0	0	1	0	0	0	0	0	0	1
Caliskan et al (2016, epoietin)	0	0	0	1	0	0	0	0	0	0	1
Caliskan et al (2016, etomidate + epoietin)	0	0	0	1	0	0	0	0	0	0	1
Cayli et al (2004, MP)	0	0	0	1	0	0	0	0	0	1	2
Cayli et al (2004, melatonin)	0	0	0	1	0	0	0	0	0	1	2
Cayli et al (2004, MP + melatonin)	0	0	0	1	0	0	0	0	0	1	2
Cayli et al (2004, ethanol)	0	0	0	1	0	0	0	0	0	1	2
Cetin et al (2006, MP)	0	0	2	2	0	0	0	0	0	1	5
Cetin et al (2006, epoietin)	0	0	2	2	0	0	0	0	0	1	5
Cetin et al (2006, MP + epoietin)	0	0	2	2	0	0	0	0	0	1	5

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Ha et al. (2008, pregabalin)	0	0	0	1	0	0	0	0	0	1	2
Ha et al. (2008, MP)	0	0	0	1	0	0	0	0	2	1	4
Ha et al. (2008, minocycline)	0	0	0	1	0	0	0	0	2	1	4
Aslan et al (2009, dexmedetomidine)	0	0	1	2	0	0	0	0	2	0	5
Aslan et al (2009, dantrolene)	0	0	1	2	0	0	2	0	0	0	5
Colón et al (2018, tamoxifen)	0	1	0	0	0	0	0	0	0	0	1
Xu et al (2009, dexamethasone)	0	0	0	1	0	0	0	0	0	1	2
Saganová et al (2009, tacrolimus)	0	0	0	2	0	0	0	0	0	1	3
Fabela-Sánchez et al (2018, albumin)	0	0	0	2	0	0	0	0	0	0	2
Darvishi et al (2014, valproic acid)	0	1	0	2	0	0	0	0	1	1	5
Torres et al (2018, dantrolene)	0	0	0	0	0	0	0	0	0	1	1
Guo et al (2018, metformin)	0	0	0	2	0	0	2	0	0	0	4
Chio et al (2021, immune globulin)	0	0	0	1	0	0	0	0	0	0	1
Kopper et al (2019, azithromycin)	0	0	0	0	0	0	0	0	0	0	0
Afshary et al. (2020, minocycline)	0	0	0	2	0	0	0	0	0	0	2
Zhang et al. (2017, metformin) - rats	1	0	1	0	1	0	0	0	0	0	2
Liu et al. (2017, lithium)	0	1	0	0	0	0	0	0	0	1	2
Jin et al. (2021, buspirone)	0	0	0	0	0	0	0	0	0	0	0
Jin et al. (2021, fluoxetine)	0	0	0	0	0	0	0	0	0	0	0
Brandoli et al. (2001, dexamethasone)	0	0	0	2	0	0	0	0	0	0	2
Faden et al (1984, naloxone)	0	0	2	2	2	0	0	0	0	0	6
Hashimoto et al. (1991, naloxone)	0	0	0	2	0	0	0	0	0	0	2
Winkler et al (1994, naloxone)	0	0	0	2	0	0	0	0	0	1	3
Faden et al (1983, naloxone) - cats	0	0	0	2	0	0	0	0	0	0	2
Faden et al (1983, naloxone) - rats	0	0	0	2	0	0	0	0	1	0	3
Faden et al (1983, naloxone) - rabbit	0	0	2	2	1	0	0	0	1	0	6
Chen et al. (2020, ezetimibe)	0	0	0	1	0	0	0	0	0	0	1
Oslau et al (2014, selegiline)	0	1	0	2	0	0	0	0	1	1	5
Salem et al. (2017, MPSS)	0	0	0	1	0	0	0	0	0	0	1
Salem et al. (2017, vitamin c)	0	0	0	1	0	0	0	0	0	0	1
Salem et al. (2017, MPSS + vitamin c)	0	0	0	1	0	0	0	0	0	0	1
Abdanipour et al. (2012, valproic acid)	0	0	0	2	0	0	0	0	0	0	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Teixeira et al. (2018, MP)	0	0	0	0	0	0	0	0	0	0	0
Tong et al. (2018, lithium)	0	1	0	1	2	0	0	0	0	1	5
Karatas et al. (2015, carvedilol)	0	0	2	2	0	2	2	2	0	1	11
Papa et al. (2016, minocycline)	0	1	2	2	0	0	0	0	0	0	5
Pourheydar et al. (2018, ubiquinone)	0	0	0	2	0	0	0	0	0	1	3
Pourheydar et al. (2018, vitamin c)	0	0	0	2	0	0	0	0	0	1	3
Wang et al. (2017, minocycline)	0	0	0	2	0	0	0	0	0	0	2
Wang et al. (2019, minocycline)	0	0	0	1	0	0	0	0	0	1	2
Khoshsirat et al. (2018, MP)	0	0	0	1	0	0	2	2	0	0	5
Fee et al. (2007, progesterone)	0	0	0	0	0	0	0	2	0	0	2
Ritz et al. (2008, estradiol)	0	0	0	1	0	0	0	0	0	0	1
Means et al. (1981, MPSS)	0	0	0	2	0	0	0	0	0	0	2
Holtz et al. (1990, MP)	0	0	0	2	0	0	0	0	1	1	4
Korkmaz et al. (2015, montelukast)	0	0	0	2	1	0	0	2	0	0	5
Haghighi et al. (1987, naloxone)	0	0	2	1	0	0	0	0	0	1	4
Arias (1985, naloxone)	0	0	0	2	0	0	0	0	0	1	3
Ross et al. (1993, MP)	0	0	0	2	0	0	0	0	0	0	2
Gerber et al. (1980, phenytoin)	0	0	2	2	0	0	0	0	0	0	4
Gerber et al. (1980, dexamethasone)	0	0	2	2	0	0	0	0	0	0	4
Silva et al. (2008, prednisone)	0	0	0	0	0	0	0	0	0	1	1
Pan et al. (2006, tacrolimus)	0	0	0	1	0	0	0	0	0	0	1
Liu et al. (2017, MP)	0	0	0	2	0	0	0	0	0	1	3
Liu et al. (2017, methotrexate)	0	0	0	2	0	0	2	0	0	1	5
Liu et al. (2017, MP + methotrexate)	0	0	0	2	0	0	0	0	0	1	3
Ahmad et al. (2016, minocycline)	0	0	0	1	0	0	0	0	0	0	1
Ahmad et al. (2016, tacrolimus)	0	0	0	1	0	0	0	0	0	0	1
Ahmad et al. (2016, minocycline + tacrolimus)	0	0	0	1	0	0	0	0	0	0	1
Meng et al. (2011, MP)	0	1	2	1	0	0	0	0	1	0	5
Shen et al. (2019, levocarnitine)	0	0	0	0	0	0	0	0	0	1	1
Cristante et al. (2013, fluoxetine)	0	0	0	0	0	0	2	0	1	0	3
Zhou et al. (2016, calcitriol)	0	0	0	1	0	0	0	0	0	0	1
Nantwi et al. (1998, theophylline)	1	0	0	1	0	0	0	0	1	1	4

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Genovese et al. (2007, dexamethasone) - mice	1	0	2	2	0	0	0	0	0	1	6
Genovese et al. (2007, melatonin)	1	0	2	2	0	0	0	0	0	1	6
Genovese et al. (2007, dexamethasone + melatonin)	1	0	2	2	0	0	0	0	0	1	6
Farsi et al. (2015, MP)	1	0	0	1	0	0	0	0	0	0	2
Farsi et al. (2015, magnesium sulfate)	1	0	0	1	0	0	0	0	0	0	2
Farsi et al. (2015, MP + magnesium sulfate)	1	0	0	1	0	0	0	0	0	0	2
Yin et al. (2013, MP)	0	0	0	1	0	0	0	0	0	0	1
Lu et al. (2016, MP)	0	0	0	0	0	0	0	0	0	0	0
Li et al. (2016, MP)	0	0	0	1	0	0	0	0	0	0	1
Hou et al. (2015, celecoxib)	0	0	0	0	0	0	2	0	0	1	3
Qinxuan et al. (2020, dexamethasone + estrogen)	0	0	0	0	0	0	0	0	0	0	0
Qinxuan et al. (2020, dexamethasone)	0	0	0	0	0	0	0	0	0	0	0
Letaif et al. (2015, estradiol)	0	0	0	0	0	0	0	0	0	0	0
Hains et al. (2004, phenytoin)	0	0	0	1	0	0	0	0	0	1	2
Mann et al. (2008, epoetin)	0	0	0	2	0	0	0	0	0	0	2
Mann et al. (2008, darbepoetin)	0	0	0	2	0	0	0	0	0	0	2
Liao et al. (2014, MP)	0	0	1	1	0	0	0	0	0	1	3
Li et al. (2019, MP)	0	0	0	0	0	0	2	2	0	1	5
Wu et al. (2019, MP)	0	0	0	1	0	0	2	0	0	0	3
Rong et al. (2018, methotrexate)	0	0	0	0	0	0	2	0	0	1	3
Wong et al. (2012, amphetamine)	0	0	0	0	0	0	0	0	0	0	0
Lima et al. (2020, citalopram)	0	0	0	0	0	0	0	0	0	0	0
Li et al. (2014, MP)	0	0	0	1	0	0	0	0	0	0	1
Chen et al. (2014, vitamin c e)	0	0	0	1	0	0	2	0	0	0	3
Akdemir et al. (1993, MP)	0	0	0	2	0	0	0	0	0	0	2
Genovese et al. (2008, montelukast)	0	0	0	1	0	0	0	0	0	1	2
Chen et al. (2018, plasma)	0	1	0	0	0	2	0	0	0	0	3
Chen et al. (2018, platelets)	0	1	0	0	0	0	0	0	0	0	1
Kim et al. (2004, MP)	0	0	0	1	0	0	0	0	0	0	1
Mbori et al. (2016, MP)	0	0	0	0	0	0	0	0	0	0	0
Wiseman et al. (2009, MP)	0	0	0	1	0	0	0	0	1	0	2
Wiseman et al. (2009, magnesium)	0	0	0	1	0	0	0	0	0	0	1

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Wiseman et al. (2009, magnesium +MP)	0	0	0	1	0	0	0	0	1	0	2
Ates et al. (2007, mexiletine)	0	0	0	1	0	0	0	0	1	1	3
Ates et al. (2007, phenytoin)	0	0	0	1	0	0	0	0	1	1	3
Serarslan et al. (2010, MP)	0	0	0	2	0	0	0	0	0	1	3
Serarslan et al. (2010, tadalafil)	0	0	0	2	0	0	0	0	0	1	3
Hara et al. (2000, MPSS)	0	0	0	2	0	0	0	0	0	0	2
Zendedel et al. (2018, estradiol)	0	1	0	0	0	0	0	0	0	0	1
Braughler et al. (1987, MPSS)	0	0	2	1	0	0	0	0	0	0	3
Robertson et al. (1986, thiopental)	0	0	2	2	1	0	0	0	0	1	6
Robertson et al. (1986, magnesium sulfate)	0	0	2	2	1	0	0	0	0	1	6
Robertson et al. (1986, lidocaine)	0	0	2	2	1	0	0	0	0	1	6
Robertson et al. (1986, naloxone)	0	0	2	2	1	0	0	0	0	1	6
Robertson et al. (1986, thiopental + naloxone)	0	0	2	2	1	0	0	0	0	1	6
Kobrine et al. (1984, lidocaine)	1	0	2	1	0	0	0	0	1	1	6
Hallenbeck et al. (1983, naloxone)	0	0	2	2	0	0	0	0	0	0	4
Watanabe et al. (2012, minocycline)	0	0	2	2	1	0	0	0	0	0	5
Yücel et al. (2006, MP)	0	0	0	1	0	0	0	0	0	0	1
Gürkan et al. (2020, MP)	0	0	0	2	0	0	0	0	1	1	4
Schwartz et al. (2001, phenytoin)	0	0	0	1	0	0	0	0	0	1	2
Tator et al. (1983, liothyronine)	0	0	0	2	0	0	2	0	0	0	4
Young et al. (1982, MPSS)	1	0	2	1	0	0	0	0	1	1	6
Saganova et al. (2008, minocycline)	0	0	0	2	0	0	0	0	0	1	3
Rivlin et al. (1979, epinephrine)	0	0	0	2	0	0	0	0	0	0	2
Rivlin et al. (1979, epinephrine + nitroprusside)	0	0	0	2	0	0	0	0	0	0	2
Zhang et al. (2020, MP)	0	0	0	0	0	0	2	0	0	0	2
Zhang et al. (2020, metformin)	0	0	0	0	0	0	2	0	0	0	2
Genovese et al. (2007, dexamethasone) - mice	0	0	0	1	0	0	0	0	0	1	2
Wu et al. (2017, sevoflurane)	0	1	0	1	0	0	0	0	0	1	3
Lee et al. (2016, fluoxetine + vitamin c)	0	1	0	1	0	0	0	0	0	0	2
de Figueiredo et al. (2018, tramadol)	0	0	0	0	0	0	2	0	0	0	2
Vasconcelos et al. (2016, magnesium chloride + PEG)	0	0	0	0	0	0	0	0	0	0	0
Miranpuri et al. (2017, folic acid)	0	1	0	1	0	0	0	0	0	0	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Gül et al. (2005, MP)	0	0	2	2	0	0	0	0	1	1	6
Gül et al. (2005, melatonin)	0	0	2	2	0	0	0	0	1	1	6
Fu et al. (2007, naproxen)	0	0	0	2	0	0	0	0	0	0	2
Fu et al. (2007, ibuprofen)	0	0	0	2	0	0	0	0	0	0	2
Cheng et al. (2016, estradiol)	0	0	0	2	0	0	2	0	0	0	4
Hu et al. (2012, estradiol)	0	1	0	2	0	0	0	0	0	0	3
Sun et al. (2020, gabapentin)	0	1	1	0	0	0	0	0	0	0	2
McCreedy et al. (2018, diclofenac)	0	1	0	0	0	0	0	0	0	0	1
Tajkey et al. (2015, ceftriaxone)	0	0	0	0	2	0	2	0	0	1	5
Zheng et al. (2011, heparin)	0	0	0	1	0	0	0	0	0	1	2
Nguyen et al. (2012, immune globulin)	0	0	0	2	0	0	0	0	0	0	2
Ueno et al. (2011, minocycline)	0	0	2	0	0	0	0	0	0	0	2
Wang et al. (2009, ibuprofen) - rats	0	0	0	0	0	0	0	0	0	0	0
Wang et al. (2009, naproxen)	0	0	0	0	0	0	0	0	0	0	0
Wang et al. (2009, ibuprofen) - mice	0	0	0	0	0	0	0	0	0	0	0
Ozkunt et al. (2017, MP)	0	0	0	1	0	0	0	0	1	0	2
Ozkunt et al. (2017, epoetin)	0	0	0	1	0	0	0	0	1	0	2
Zakeri et al. (2014, lithium)	0	1	0	2	0	0	0	0	0	0	3
Teng et al. (2004, minocycline)	0	1	0	2	0	0	0	0	0	0	3
Wu et al. (2010, MP)	0	2	0	0	0	0	0	0	0	0	2
Huang et al. (2009, epoetin)	0	0	0	2	0	0	0	0	0	1	3
Lee et al. (2003, minocycline)	0	1	0	2	0	0	0	0	0	0	3
Lin et al. (2016, estradiol)	0	0	0	1	0	1	0	0	0	0	2
Faden et al. (1981, naloxone)	1	0	2	1	0	0	0	0	0	0	4
Holtz et al. (1991, MP)	0	0	0	2	0	0	0	2	0	1	5
Gorio et al. (2007, MPSS)	0	1	2	1	0	0	0	0	0	0	4
Ravikumar et al. (2005, nicotine)	0	0	0	0	0	0	0	0	0	0	0
Know et al. (2009, MP)	0	0	0	2	0	0	0	0	0	0	2
Know et al. (2009, PEG)	0	0	0	2	0	0	0	0	0	0	2
Know et al. (2009, magnesium sulfate)	0	0	0	2	0	0	0	0	0	0	2
Know et al. (2009, magnesium sulfate + PEG)	0	0	0	2	0	0	0	0	0	0	2
Know et al. (2009, magnesium chloride + PEG)	0	0	0	2	0	0	0	0	0	0	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Know et al. (2009, MP + magnesium chloride + PEG)	0	0	0	2	0	2	0	0	0	0	4
Kachadroka et al. (2010, estradiol)	0	0	0	0	0	0	0	0	0	0	0
Roman et al. (2011, PEG)	0	0	0	0	0	0	0	0	1	1	2
Bu et al. (2018, estradiol)	0	0	2	0	0	0	2	0	1	1	6
Fakhri et al. (2020, melatonin)	0	0	0	1	0	0	0	0	0	0	1
Hook et al. (2009, morphine sulfate)	0	0	0	0	0	0	0	0	0	1	1
Garcia-Ovejero et al. (2014, progesterone)	0	0	0	0	0	0	0	0	0	1	1
Erol et al. (2016, MP)	0	0	0	2	0	0	0	0	0	1	3
Erol et al. (2016, topiramate)	0	0	0	2	0	0	0	0	0	1	3
Streijger et al. (2016, magnesium chloride + PEG)	0	0	0	2	0	0	0	0	0	1	3
Streijger et al. (2016, magnesium sulfate)	0	0	0	2	0	0	0	0	0	1	3
Ji et al. (2005, MP)	0	1	0	0	0	0	0	0	0	1	2
Doyle et al. (2004, levodopa)	0	0	2	2	2	0	0	0	1	1	8
Ibarra et al. (2004, MPSS)	0	2	0	0	0	0	0	0	0	0	2
Kuroiwa et al. (2014, amiloride)	0	1	0	0	0	0	0	0	0	0	1
Wells et al. (2003, MP)	0	0	0	0	0	0	0	0	0	1	1
Wells et al. (2003, minocycline)	0	0	0	0	0	0	0	0	0	1	1
Guizar-Sahagun et al. (2009, MPSS)	0	0	0	1	0	0	0	0	0	1	2
Guizar-Sahagun et al. (2009, melatonin)	0	0	0	1	0	0	0	0	0	1	2
Lee et al. (2010, minocycline)	0	0	2	2	0	0	0	0	0	0	4
Lee et al. (2010, simvastatin)	0	0	2	2	0	0	0	0	0	0	4
Zeman et al. (2009, oxandrolone)	0	1	0	1	0	0	0	0	0	0	2
Cole et al. (1989, fentanyl + nitrous oxide)	0	0	0	2	0	0	0	0	1	0	3
Cole et al. (1989, fentanyl + nitrous oxide + naloxone)	0	0	0	2	0	0	0	0	1	0	3
Kuchner et al. (2000, dexamethasone)	0	0	0	1	0	0	0	0	0	0	1
Luo et al. (2013, MP)	0	0	0	0	0	0	0	0	0	1	1
Thomas et al. (1999, progesterone)	0	0	0	1	0	0	0	0	0	1	2
Stewart et al. (2019, folic acid)	0	0	0	0	0	0	0	0	0	0	0
Stewart et al. (2019, nitrous oxide)	0	0	0	0	0	0	0	0	0	0	0
Stewart et al. (2019, folic acid + nitrous oxide)	0	0	0	0	0	0	0	0	0	0	0
Gok et al. (2007, MP)	0	0	0	1	0	0	0	0	0	0	1
Dinc et al. (2013, MP)	0	0	0	2	0	0	2	0	1	1	6

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Lee et al. (2010, magnesium)	0	0	0	2	0	0	0	0	0	1	3
Sonmez et al. (2013, minocycline)	0	0	0	1	0	0	0	0	0	1	2
Cuzzocrea et al. (2008, estradiol)	0	0	0	1	0	0	0	0	1	1	3
Ren et al. (2017, PEG)	1	0	0	2	0	0	0	0	1	1	5
Faden et al. (1984, dexamethasone)	0	0	2	2	0	0	0	2	0	0	6
Faden et al. (1984, MP)	0	0	2	2	0	0	0	0	0	0	4
Xu et al. (2019, melatonin)	0	2	2	2	0	0	0	0	1	0	7
Li et al. (2019, melatonin)	0	0	0	0	0	0	0	0	0	0	0
Yang et al. (2020, melatonin)	0	0	0	0	0	0	0	0	0	1	1
Piao et al. (2014, melatonin)	0	0	0	0	0	0	0	0	0	1	1
Zhang et al. (2019, melatonin)	0	0	0	1	0	0	0	0	0	0	1
Shen et al. (2017, melatonin)	0	1	0	1	0	0	2	0	0	0	4
Esposito et al. (2009, melatonin)	1	0	2	2	0	0	0	2	0	1	8
Jing et al. (2019, melatonin)	0	0	0	1	0	0	2	0	0	0	3
Fee et al. (2010, melatonin)	0	1	0	0	0	0	0	0	0	1	2
Jeffrey-Gauthier et al. (2018, buspirone)	0	0	0	2	0	0	0	0	0	1	3
Holtz et al. (1989, naloxone)	0	0	0	2	0	0	0	0	0	1	3
Park et al. (2012, melatonin)	0	0	0	0	0	0	0	0	0	0	0
Ates et al. (2006, MP)	0	0	0	0	0	0	0	0	1	0	1
Ates et al. (2006, ethanol)	0	0	0	0	0	0	0	0	1	0	1
Yingli et al. (2014, melatonin)	0	0	0	1	0	0	2	0	0	1	4
Yune et al. (2007, minocycline)	0	0	0	1	0	0	0	0	0	0	1
Yune et al. (2007, MP)	0	0	0	1	0	0	0	0	0	0	1
Zhang et al. (2017, metformin) - rats	0	1	0	1	0	0	0	0	0	0	2
Park et al. (2014, hydralazine)	0	1	0	2	0	0	0	0	0	1	4
Stirling et al. (2004, minocycline)	0	0	2	1	0	0	0	0	0	0	3
Weaver et al. (2005, MP)	0	0	0	2	0	0	0	0	0	0	2
Moutaery et al. (2000, aluminum)	0	0	0	1	0	0	0	0	0	0	1
de Mesquita Coutinho et al. (2016, tacrolimus)	0	0	0	1	0	0	0	2	0	1	4
Takami et al. (2002, MP)	0	0	0	1	0	0	0	0	0	0	1
Chikawa et al. (2001, MP)	0	0	0	0	0	0	0	0	0	0	0
Aceves et al. (2019, morphine)	0	0	0	0	0	0	0	0	0	1	1

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Aceves et al. (2019, minocycline)	0	0	0	0	0	0	0	0	0	1	1
Aceves et al. (2019, morphine + minocycline)	0	0	0	0	0	0	0	0	0	1	1
Woller et al. (2014, morphine)	0	0	0	0	0	0	0	0	0	1	1
de la Torre Valdovino et al. (2016, tamoxifen)	1	0	0	1	0	0	0	0	0	1	3
Guo et al. (2015, acetylcysteine)	0	1	0	1	0	0	0	0	1	1	4
Black et al. (1991, naloxone)	0	0	1	2	0	0	0	0	0	0	3
Black et al. (1986, naloxone) - rats	0	0	0	0	0	0	0	0	0	0	0
Black et al. (1986, naloxone) - rats	0	0	0	2	0	0	0	0	0	0	2
Wang et al. (2020, metformin)	0	0	0	2	0	0	0	0	1	0	3
Lin et al. (2019, MP)	0	1	0	0	0	2	0	0	1	0	4
Lin et al. (2019, MPSS)	0	1	0	0	0	2	0	0	1	0	4
Koyanagi, Tator (1997, MP)	0	0	0	2	0	0	0	0	0	1	3
Hook et al. (2017, morphine)	0	0	0	0	0	0	0	0	0	1	1
Wu et al. (2016, botulinum toxin)	0	0	0	1	0	0	0	0	1	0	2
Guth et al. (1994, indomethacin)	0	0	0	2	0	0	2	0	0	0	4
Lee et al. (2012, fluoxetine)	0	1	0	1	0	0	0	0	0	0	2
Gao et al. (2020, melatonin)	0	0	0	2	0	0	0	0	0	0	2
Gorio et al. (2005, MPSS)	0	1	2	1	0	0	2	0	0	0	6
Scali et al. (2013, fluoxetine)	0	1	0	0	0	0	0	0	0	1	2
Dixit et al. (2018, clonidine)	0	0	0	1	0	0	0	0	0	1	2
Zhang et al. (2014, MPSS)	0	0	2	2	0	2	2	2	1	1	12
Nazli et al. (2015, atorvastatin)	0	0	2	2	1	0	0	0	0	0	5
Li et al. (2014, atorvastatin)	0	0	0	1	0	0	0	0	0	1	2
Bharne et al. (2013, MP)	0	0	0	1	0	0	0	0	0	1	2
Cayli et al. (2006, etomidate +MP)	0	0	0	1	0	0	0	0	0	0	1
Cong, Chen (2016, dexamethasone)	0	0	0	1	0	0	0	0	0	0	1
Tan et al. (2015, MP)	0	0	0	1	0	0	0	0	0	1	2
Cakir et al. (2003, acetylcysteine)	0	0	2	2	1	0	0	2	0	1	8
Gao et al. (2015, simvastatin)	0	0	0	1	0	0	0	0	0	0	1
Hou et al. (2016, MP)	0	0	0	0	0	0	0	0	0	1	1
Wang et al. (2014, MP)	0	1	0	0	0	0	0	0	0	0	1
Sozbilen et al. (2018, MP)	0	0	0	1	0	0	0	0	0	1	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Yilmaz et al. (2015, clopidogrel)	0	0	0	2	0	0	0	0	0	0	2
Chen et al. (2015, MP)	0	2	0	2	0	0	0	0	0	0	4
Ok et al. (2012, MP)	0	0	0	1	0	0	0	0	1	0	2
Kazanci et al. (2017, MP + pregabalin)	0	0	2	1	1	0	0	0	0	1	5
Kahveci et al. (2014, MP + rosuvastatin)	0	0	0	1	0	0	0	0	0	0	1
Xian-Hui et al. (2016, MP)	0	0	0	1	0	1	0	2	0	0	4
Kouhzaei et al. (2013, PEG)	0	1	0	1	0	0	0	0	0	1	3
Aceves et al. (2016, morphine)	0	0	0	0	0	0	0	0	0	1	1
Guizar-Sahagun et al. (2005, MPSS)	0	0	0	1	0	0	0	0	0	1	2
De La Torre et al. (1975, mannitol)	0	0	0	2	0	0	0	0	1	1	4
De La Torre et al. (1975, dexamethasone)	0	0	0	2	0	0	0	0	1	1	4
Yates et al. (2014, MP)	0	1	0	2	0	0	0	0	0	1	4
Flamm et al. (1982, naloxone)	0	0	2	2	0	0	0	0	1	1	6
Wallace, Tator (1986, naloxone) - rats	0	0	0	2	0	0	0	0	0	1	3
Wallace, Tator (1986, naloxone) - rats	0	0	0	2	0	0	0	0	0	1	3
Cho et al. (2010, glucosamine)	0	0	0	1	1	2	0	0	0	1	5
Zadeh-Ardabili et al. (2017, vitamin e)	0	0	0	0	0	0	2	0	0	1	3
Gok et al. (2009, albumin)	0	0	0	1	0	0	0	0	0	0	1
Gok et al. (2009, immune globulin)	0	0	0	1	0	0	0	0	0	0	1
Khajoueinejad et al. (2019, calcitriol)	0	0	0	0	0	0	0	0	0	0	0
Lim et al. (2013, omega 3)	0	0	0	2	0	0	0	0	0	0	2
Popovich et al. (2012, glibenclamide)	0	1	0	0	0	0	0	0	0	0	1
Pukos, McTigue (2020, tamoxifen)	0	0	0	0	0	0	0	0	0	0	0
Durham-Lee et al. (2012, amiloride)	0	1	0	2	0	0	0	0	0	1	4
Perez-Espejo et al. (1996, MP)	0	0	0	2	0	0	0	0	1	0	3
Patel et al. (2017, pioglitazone)	0	0	0	2	0	0	0	0	0	1	3
Nash et al. (2002, MP)	0	0	2	1	0	0	0	0	0	1	4
Lankhorst et al. (2000, MP)	0	0	0	0	0	0	0	0	0	0	0
Liu et al. (2010, carbidopa levodopa)	0	0	0	1	0	0	2	0	0	1	4
Yang et al. (2020, glutamine)	0	0	0	0	0	0	0	0	0	0	0
Pannu et al. (2007, atorvastatin)	0	0	0	2	0	0	0	0	0	0	2
Mann et al. (2010, atorvastatin)	0	0	0	2	0	0	0	0	0	0	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Mann et al. (2010, simvastatin)	0	0	0	2	0	0	0	0	0	0	2
King et al. (2006, omega 3) - rats	0	0	0	1	0	0	0	0	0	1	2
King et al. (2006, omega 3) - rats	0	0	0	1	2	0	0	0	0	1	4
Fujimoto et al. (2000, melatonin)	0	0	0	2	0	0	0	0	0	0	2
Schiaveto-de-Souza et al. (2013, melatonin)	0	0	0	2	0	0	0	0	0	1	3
Karami et al. (2013, ketoprofen)	0	0	0	2	0	0	0	0	0	1	3
Tan et al. (2020, estrogen)	0	0	0	0	0	0	0	0	0	0	0
Wang et al. (2015, propofol)	0	0	0	0	0	0	0	0	0	1	1
Zhang et al. (2020, mannitol)	0	0	0	0	0	0	0	0	0	0	0
Yates et al. (2009, modafinil)	0	0	0	1	0	0	0	0	0	1	2
Iwasa et al. (1989, vitamin e)	0	0	0	2	0	0	0	0	0	1	3
Sengelaub et al. (2018, estradiol)	0	0	0	1	0	0	0	0	0	0	1
Sengelaub et al. (2018, testosterone)	0	0	0	1	0	0	0	0	0	0	1
Sengelaub et al. (2018, estradiol + testosterone)	0	0	0	1	0	0	0	0	0	0	1
Patel et al. (2014, acetylcysteine)	0	0	0	2	0	0	0	0	0	1	3
Osuna-Carrasco et al. (2016, tamoxifen)	0	0	0	1	0	0	0	0	0	1	2
Ren et al. (2019, PEG)	0	0	0	2	0	0	0	0	0	1	3
Kaptanoglu et al. (2005, MP)	0	0	0	1	0	0	0	0	0	0	1
Kaptanoglu et al. (2005, mexiletine)	0	0	0	1	0	0	0	0	0	0	1
Xing et al (2016, morphine)	0	0	0	1	0	0	0	0	0	0	1
Mu et al (2000, MP)	0	0	0	1	0	0	0	0	0	0	1
Kazama et al (2001, pentobarbital)	0	0	2	2	1	0	0	0	0	0	5
Genovese et al (2007, dexamethasone)	0	2	0	1	0	0	0	0	0	1	4
Pan et al (2013, tacrolimus)	0	0	0	1	0	0	0	0	0	0	1
Pereira et al (2009, MPSS)	0	0	0	1	0	0	0	0	0	1	2
Cain et al (2007, albumin)	0	0	0	2	0	0	0	0	0	0	2
Liang et al (2019, simvastatin)	0	0	0	0	0	0	0	0	0	1	1
Liang et al (2019, ezetimibe + simvastatin)	0	0	0	0	0	0	0	0	0	1	1
Gao et al (2016, simvastatin)	0	1	0	1	0	0	0	0	0	0	2
Han et al (2012, simvastatin)	0	0	0	2	0	0	0	0	0	0	2
Han et al (2011, simvastatin)	0	0	0	1	0	0	0	0	0	1	2
Han et al (2020, sitagliptin)	0	1	0	0	0	0	0	0	1	0	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
He et al (2016, propofol)	0	0	0	0	0	0	0	0	0	1	1
Holmberg et al (2008, simvastatin)	0	1	0	1	0	1	2	0	0	1	6
Zhang et al (2018, lithium)	0	0	0	1	0	0	0	0	0	0	1
Tedeshi et al (2016, pregabalin)	0	1	0	0	0	0	0	0	0	1	2
Kim et al (2017, lithium)	0	0	0	0	0	0	0	0	0	0	0
Sanli et al (2012, MPSS)	0	0	0	1	0	0	0	0	0	0	1
Salimi et al (2020, ceftriaxone)	0	0	0	0	0	0	2	0	0	0	2
Salimi et al (2020, acetylcysteine)	0	0	0	0	0	0	2	0	0	0	2
Salimi et al (2020, ceftriaxone + acetylcysteine)	0	0	0	0	0	0	2	0	0	0	2
Ni et al (2018, estrogen)	0	1	0	1	0	0	0	0	0	0	2
Xiao Jianru et al (1998, naloxone)	0	0	0	2	0	0	0	2	0	1	5
Baffour et al (1995, MPSS)	0	0	1	2	0	0	0	0	1	0	4
Qi et al (2017, MP)	0	0	0	0	0	0	0	0	0	1	1
Yune et al (2004, estradiol)	0	1	0	1	0	0	0	0	0	0	2
Nacar et al (2014, PEG)	0	0	0	2	0	0	0	0	1	0	3
Nacar et al (2014, atorvastatin)	0	0	0	2	0	0	0	0	0	0	2
Baptiste et al (2009, PEG)	0	0	0	1	0	0	0	0	0	0	1
Mallei et al (2005, prednisolone)	0	1	0	1	0	0	0	0	0	0	2
Madsen et al (1998, tacrolimus)	0	1	2	2	0	0	0	0	0	0	5
Colón et al (2016, tamoxifen)	0	0	0	1	0	0	0	0	0	0	1
Mosquera et al (2014, estradiol)	0	1	0	1	0	0	0	0	0	0	2
Mosquera et al (2014, tamoxifen)	0	1	0	1	0	0	0	0	0	0	2
Tian et al (2009, tamoxifen)	0	0	0	1	0	0	0	0	0	0	1
Kitchen et al (2020, trifluoperazine)	0	1	0	0	0	0	0	0	0	0	1
Namjoo et al (2018, estradiol) - rats	0	0	0	1	0	0	0	0	0	1	2
Borgens et al (2002, PEG)	1	0	2	1	0	0	0	0	0	0	4
Hao et al (1991, naltrexone)	0	0	0	2	0	0	0	0	0	1	3
Ruhollah Hosseini et al (2017, dexamethasone)	0	0	0	1	0	0	0	0	0	1	2
Pedram et al (2018, meloxicam)	0	0	0	0	0	0	0	0	0	0	0
Sharma et al (2004, MPSS)	0	0	1	2	1	0	0	2	1	1	8
Sharma et al (2004, dexamethasone)	0	0	1	2	1	0	0	2	1	1	8
Guptarak et al (2014, tamoxifen)	0	0	0	2	0	0	0	0	0	0	2

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Kermani et al (2016, acetylsalicylic acid)	0	0	0	1	0	0	0	0	0	0	1
Sayin et al (2013, MPSS)	0	0	0	1	0	0	0	0	0	0	1
Baysefer et al (2003, mannitol)	0	0	0	2	0	0	0	0	0	1	3
Farooque et al (1994, MPSS)	0	0	0	2	0	0	0	0	0	1	3
Golding et al (2006, glutamine)	0	0	0	0	0	0	0	0	0	1	1
Abdanipour et al (2019, lithium)	1	0	0	0	0	0	0	0	0	1	2
Charn et al (2011, minocycline)	0	0	0	2	0	0	0	0	0	0	2
Gul et al (2010, MP)	0	0	0	1	0	0	0	0	2	1	4
Gul et al (2010, dexmedetomidine)	0	0	0	1	0	0	0	0	2	1	4
Lang-Lazdunski et al (2001, tacrolimus)	0	0	0	2	1	0	0	0	0	1	4
Rosado et al (2014, MPSS)	0	0	0	0	0	0	0	0	0	0	0
Rosado et al (2014, dantrolene)	0	0	0	0	0	0	0	0	0	0	0
Rosado et al (2014, MPSS + dantrolene)	0	0	0	0	0	0	0	0	0	0	0
Boran et al (2005, MP)	0	0	0	2	0	0	0	0	0	0	2
Boran et al (2005, epoetin)	0	0	0	2	0	0	0	0	0	0	2
Hook et al (2007, morphine sulfate)	0	0	0	0	0	0	0	0	0	0	0
Simpson et al (1991, nifedipine)	0	0	2	2	1	0	0	0	0	1	6
Simpson et al (1991, indomethacin)	0	0	2	2	1	0	0	0	0	1	6
He et al (2017, lithium)	0	1	0	1	0	0	0	0	1	1	4
Almad et al (2011, fenofibrate)	0	1	0	1	0	0	0	0	0	0	2
McTigue et al (2007, pioglitazone)	0	0	0	1	0	0	0	0	0	0	1
Ko et al (2006, minocycline)	0	0	2	0	0	0	0	0	0	0	2
Çelik et al (2015, vitamin d)	0	0	2	2	0	0	0	0	1	1	6
Park et al (2007, pioglitazone)	0	1	2	1	0	0	0	0	0	0	4
Afhami et al (2016, estradiol)	0	0	0	1	0	0	0	0	0	1	2
Gezici et al (2017, methotrexate)	0	0	0	1	0	0	0	0	0	0	1
Narin et al (2017, topiramate)	0	0	0	0	0	0	0	0	0	0	0
Gensel et al (2012, topiramate)	0	0	0	0	0	0	0	0	0	1	1
Yoshizaki et al (2019, heparin)	0	0	0	0	0	0	0	0	0	0	0
Arias (1987, naloxone)	0	1	0	2	0	0	0	0	1	1	5
Arias (1987, dexamethasone)	0	1	0	2	0	0	0	0	1	1	5
Naftchi et al (1991, MPSS + aminocaproic acid)	0	0	2	2	0	0	0	0	2	1	7

Table 8.7 – continued from previous page

Experiment	Species	Sample size	Sex	Age	Level of injury	Dose	Treatment time	Route	Results	Blinding	Total bias score
Romero-Ramírez et al (2020, MP)	0	0	0	0	0	0	0	0	1	1	2
Zhang et al (2009, tacrolimus)	0	0	0	2	0	0	0	0	0	1	3
Zhang et al (2014, MP)	0	0	2	0	0	0	0	0	0	0	2
Rabinowitz et al (2008, MP)	0	0	0	1	0	0	0	0	0	1	2
Penas et al (2011, valproic acid)	0	0	0	0	0	0	0	0	1	0	1
Chu et al (2015, valproic acid)	0	0	0	1	0	0	0	0	0	0	1
Lee et al (2012, valproic acid)	0	1	0	1	0	0	0	0	0	0	2
Lu et al (2013, valproic acid)	0	0	0	1	0	0	0	0	0	0	1
Lv et al (2012, valproic acid)	0	1	0	2	0	0	0	0	0	0	3
Lv et al (2011, valproic acid)	0	1	0	2	0	0	0	0	0	0	3
Hao et al (2013, valproic acid)	0	1	0	1	0	0	0	0	0	0	2
Wang et al (2020, valproic acid)	0	0	0	1	0	0	0	0	0	1	2
Li et al (2019, zinc)	0	0	0	1	0	0	0	0	0	1	2
Lin et al (2020, zinc) - mice	0	1	0	0	0	0	2	0	0	0	3
Li et al (2020, zinc)	0	1	0	0	0	0	0	0	0	0	1
Lin et al (2020, zinc) - mice	0	1	0	0	0	0	0	0	0	0	1

11.5 Chapter 5

Table 8.8: Serological markers studied and their normal range

Serological marker	Normal range
Erythrocytes	3.80-5.90 _/pl
Hemoglobin	12.0-18.0 g/dl
Hematocrit	35.0-52.0 %
MCHC	32-36 g/dl
MCV	80-100 fl
Thrombocytes	140-440 tsd/ul
Leucocytes	4.3-10.8 _/nl
Hemoglobin per erythrocyte	27-34 pg
Alkaline phosphatase	35-171 U/l
ASAT	0-35 U/l
ALAT	0-45
Total bilirubin	0.0-1.1 mg/dl
Gamma-GT	0-65 U/l
Lactate dehydrogenase	0-248 U/l
Calcium	2.22-2.66 mmol/l
Creatinine	0.5-1.0 mg/dl
Total proteins	5.70-8.2 g/dl
Blood urea nitrogen	1.70-8.30 mmol/l
Potassium	3.50-5.10 mmol/l
Sodium	136-148 mmol/l
Cholinesterase	5320-12920 U/l
Amylase	0-115 U/l
Lipase	0-80 U/l
Glucose	4.1-5.9 mmol/l
INR	0.8-1.1
Partial thromboplastin time	26-40 s
CRP	0.0-0.5 mg/dl
Quick test	80-127%

Continued on next page

Table 8.8: Serological markers studied and their normal range (Continued)

Serological marker	Normal range

mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), aspartate aminotransferase (ASAT), alanine aminotransferase (ALAT), γ-glutamyl transferase (Gamma-GT), international normalized ratio (INR), C-reactive protein (CRP)

Table 8.9: Parameter grid of the regression models used in the prediction task

Hyperparameter	Values
LASSO	
alpha	np.arange (0.5, 1500, 0.5)
Ridge regression	
alpha	np.arange (0.5, 1500, 0.5)
RF regressor	
n_estimators	10, 25, 50, 100
max_features	'sqrt', 'log2', None
max_depth	3, 6, 9
max_leaf_nodes	3, 6, 9
SVM regressor with linear kernel	
epsilon	np.arange(0, 1.5, 0.1)
GBR	
n_estimators	10, 25, 50, 100
learning rate	0.001, 0.01, 0.05
subsample	0.5, 0.7, 0.8
max_depth	3, 5, 7
min_samples_split	8, 10, 15
min_samples_leaf	5, 8, 10
XGBoost	
n_estimators	10, 25, 50, 100
max_depth	3, 5, 6
eta	0.001, 0.01, 0.05
subsample	0.3, 0.5, 0.9
colsample_bytree	0.5, 0.9
gamma	0.2, 0.3, 0.4, 0.5
LightGBM	
learning rate	0.0.5, 0.1
max_depth	1, 2, 3

Continued on next page

Table 8.9: Parameter grid of the regression models used in the prediction task (Continued)

Hyperparameter	Values
num_leaves	2, 3
metric	'l2', 'l1', 'poisson'
min_child_samples	10
LightGBM (stratified parameters)	
learning rate	0.0.5, 0.1
max_depth	1, 2
num_leaves	2, 3
metric	'l2', 'l1', 'poisson'
min_child_samples	5

Unless specified the same parameters have been used for both approaches (non stratified and stratified cohort). Scoring for all GridSearchCV was done with the negative root mean squared error. least absolute shrinkage and selection operator (LASSO); random forest (RF); support vector machines (SVM); gradient boosting regressor (GBR); extreme gradient boosting (XGBoost); light gradient boosting machine (LightGBM)

Table 8.10: Frequency of the features across all 50 seed iterations. Values can range from zero, always omitted due to high correlation, to 50, always present across all seeds.

	Mean	Median	Minimu	Maximu	Range	Sampling frequency
Erythrocytes	50	50	50	50	50	50
Hemoglobin	0	0	0	0	0	0
Hematocrit	0	0	0	0	0	0
MCHC	50	50	50	50	50	0
MCV	50	50	50	50	50	0
Thrombocytes	50	50	50	50	50	0
Leucocytes	50	50	50	50	50	0
Hemoglobin per erythrocyte	0	0	4	0	50	0
Alkaline phosphatase	50	50	50	50	50	0
ASAT	50	50	50	50	50	0
ALAT	41	40	41	49	50	0
Total bilirubin	50	50	50	50	50	3
Gamma-GT	49	50	50	17	0	0
Lactate dehydrogenase	47	41	50	50	50	0
Calcium	50	50	49	50	50	0
Creatinine	50	50	50	50	50	0
Total proteins	0	1	50	1	0	
Blood urea nitrogen	13	19	29	44	50	0
Potassium	50	50	50	50	50	0
Sodium	50	50	50	50	50	0
Cholinesterase	43	33	13	50	0	0
Amylase	50	50	50	50	50	0
Lipase	26	13	50	39	37	0
Glucose	50	50	50	50	50	0
INR	50	50	50	50	50	0

Continued on next page

Table 8.10: Frequency of the features across all 50 seed iterations. Values can range from zero, always omitted due to high correlation, to 50, always present across all seeds. (Continued)

	Mean	Median	Minimu	Maximu	Range	Sampling frequency
Partial thromboplastin time	50	50	50	50	50	0
CRP	50	50	50	50	50	0
Quick test	0	0	50	50	10	0
Age	50	50	50	50	50	50
Very acute LEMS	50	50	50	50	50	50

mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), aspartate aminotransferase (ASAT), alanine aminotransferase (ALAT), γ-glutamyl transferase (Gamma-GT), international normalized ratio (INR), C-reactive protein (CRP), lower extremity motor score (LEMS)

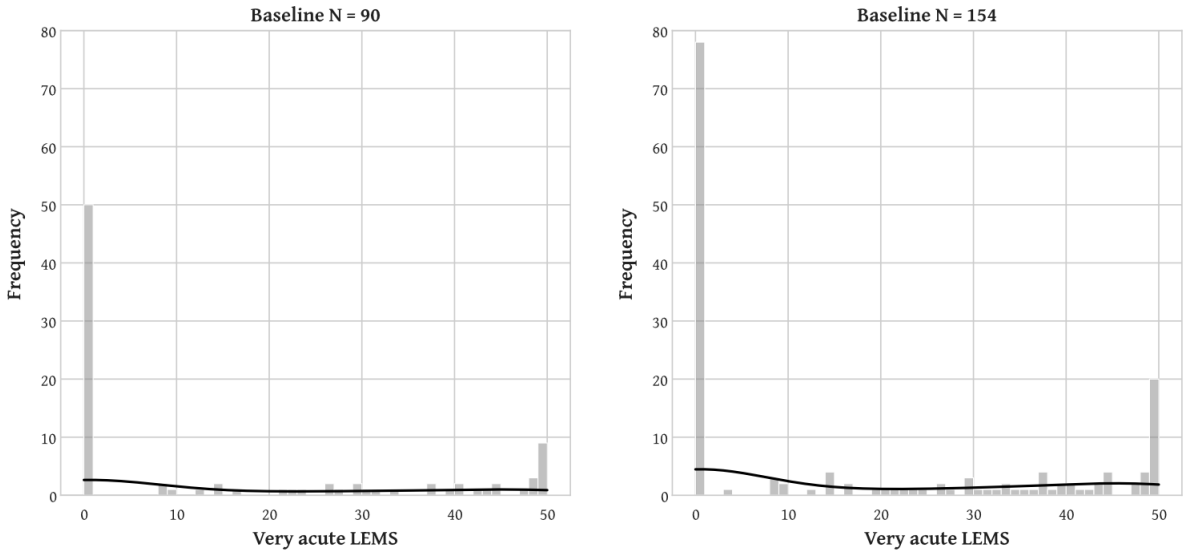


Figure 8.5: Lower extremity motor score (LEMS) distribution at the very acute stage. lower extremity motor score (LEMS)

Table 8.11: Murnau cohorts before and after preprocessing.

	Murnau included	**LOCF included**	**Very acute LEMS included**
Subject characteristics			
Total, *n*	118	136	154
Age in years at injury			
Mean±SD	47.83±18.31	47.57±18.27	47.64±18.52
Sex, n (%)			0.1629
Male	96 (81)	112 (82)	126 (82)
Female	22 (19)	24 (18)	28 (18)
LEMS score, mean±SD			
Very acute	19.65±20.50	19.67±20.48	17.37±20.25
	$n = 118$	$n = 136$	$n = 154$
After 26/52 weeks	29.58±21.60	29.10±21.66	25.69±22.40
	$n = 118$	$n = 136$	$n = 154$
AIS grade at very acute stage, *n* (%)			
A	43 (36.4)	48 (35.3)	48 (31.2)
B	13 (11.0)	17 (12.5)	17 (11.0)
C	12 (10.2)	14 (10.3)	14 (9.1)
D	0 (39.8)	54 (39.7)	54 (35.1)
E	3 (0.6)	0 (0)	0 (0)
NA	3 (2.5)	3 (2.2)	21 (13.6)

The Murnau included column represents the included patients that meet the prerequisites without imputation. Last observation carried forward (LOCF) included and very acute LEMS included columns represent patients included after LOCF and very acute LEMS imputation, respectively. last observation carried forward (LOCF), lower extremity motor score (LEMS), standard deviation (SD), American spinal injury association (ASIA) impairment scale (AIS), not available (NA)

Table 8.12: Description of the different cohorts.

	Cohort $n = 90$	**Cohort** $n = 154$	**Test statistic**	**p value**
Age in years at injury				
Mean±SD	47.11±19.09	47.64±18.52	-0.2594	0.7953
Sex, n (%)			0.1215	0.7274
Male	76 (84)	126 (82)	126 (82)	
Female	14 (16)	28 (18)	28 (18)	
LEMS score, mean±SD				
Very acute	15.37±19.63	17.37±20.25	-0.7407	0.4589
After 26/52 weeks	24.59±22.62	25.69±22.40	-0.3177	0.7507
AIS grade at very acute stage, *n* (%)			0.9243	0.8196
A	28 (31.1)	48 (31.2)		
B	12 (13.3)	17 (11.0)		
C	10 (11.1)	14 (9.1)		
D	27 (30.0)	54 (35.1)		
E	3 (0.6)	0 (0)		
NA	13 (14.4)	21 (13.6)		

standard deviation (SD), lower extremity motor score (LEMS), American spinal injury association (ASIA) impairment scale (AIS), not available (NA)

Table 8.13: Mean and standard deviation (SD) of the different preprocessing approaches for non-stratified regression models.

MAE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
Linear	14.53 ± 3.70	11.58 ± 2.71	11.59 ± 2.64	11.14 ± 2.59	12.64 ± 3.10	11.84 ± 2.63	10.83 ± 1.84	9.18 ± 1.31	9.25 ± 1.31	24.38 ± 4.59	9.29 ± 1.33
LASSO	11.12 ± 2.04	11.38 ± 1.70	11.55 ± 1.75	10.94 ± 1.92	11.75 ± 1.75	11.36 ± 2.00	11.06 ± 1.73	9.39 ± 1.26	9.47 ± 1.25	9.77 ± 1.22	9.48 ± 1.24
Ridge	12.36 ± 3.31	12.19 ± 2.03	12.13 ± 1.93	11.58 ± 2.10	12.94 ± 2.35	12.09 ± 2.27	10.89 ± 1.81	9.22 ± 1.30	9.30 ± 1.30	11.84 ± 1.72	9.33 ± 1.32
RF	**10.03 ± 2.01**	**10.36 ± 1.97**	10.13 ± 2.11	**9.37 ± 2.18**	**10.53 ± 1.84**	**10.64 ± 1.87**	9.89 ± 1.92	8.89 ± 1.72	9.01 ± 1.62	9.14 ± 1.59	8.55 ± 1.61
SVM	11.12 ± 2.99	11.12 ± 2.39	10.82 ± 2.18	10.42 ± 2.49	11.43 ± 2.87	11.10 ± 2.56	**9.69 ± 2.88**	**8.11 ± 2.04**	8.09 ± 2.04	11.56 ± 1.78	**8.15 ± 2.23**
XGBoost	10.38 ± 2.26	10.73 ± 2.17	10.38 ± 2.28	9.62 ± 1.18	12.17 ± 2.04	11.86 ± 2.18	11.49 ± 1.92	9.31 ± 1.51	9.44 ± 1.49	9.60 ± 1.35	9.35 ± 1.35
GBR	10.14 ± 2.18	10.38 ± 1.98	**9.69 ± 2.12**	9.72 ± 2.24	11.33 ± 1.73	11.33 ± 1.75	10.37 ± 1.74	9.07 ± 1.42	9.16 ± 1.41	**9.12 ± 1.50**	8.95 ± 1.41

Continued on next page

Table 8.13: Mean and standard deviation (SD) of the different preprocessing approaches for non-stratified regression models. (Continued)

MAE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
LightGBM	10.56 ± 2.11	10.67 ± 1.84	10.25 ± 2.11	9.91 ± 2.14	11.01 ± 1.78	10.91 ± 1.87	9.98 ± 1.80	8.88 ± 1.24	8.96 ± 1.24	9.68 ± 1.19	8.83 ± 1.3
RMSE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
Linear	20.97 ± 4.17	16.86 ± 3.46	16.70 ± 3.24	15.89 ± 3.14	18.05 ± 3.61	16.90 ± 3.17	14.46 ± 2.91	12.74 ± 2.26	12.82 ± 2.26	32.72 ± 3.99	12.78 ± 2.27
LASSO	14.93 ± 2.92	14.66 ± 2.45	15.03 ± 2.41	14.08 ± 2.65	15.05 ± 2.32	14.81 ± 2.69	14.49 ± 2.69	12.80 ± 2.13	12.86 ± 2.10	13.08 ± 1.99	12.84 ± 2.12
Ridge	17.49 ± 4.12	16.31 ± 2.46	16.11 ± 2.42	14.93 ± 2.67	16.78 ± 2.71	15.95 ± 2.65	14.49 ± 2.86	12.76 ± 2.24	12.85 ± 2.23	15.92 ± 2.04	12.80 ± 2.25
RF	14.38 ± 3.65	14.60 ± 3.54	14.27 ± 3.60	**13.63 ± 3.63**	14.53 ± 3.35	14.69 ± 3.44	14.37 ± 3.72	12.85 ± 2.79	12.79 ± 2.61	**12.98 ± 2.67**	12.64 ± 2.67

Continued on next page

Table 8.13: Mean and standard deviation (SD) of the different preprocessing approaches for non-stratified regression models. (Continued)

MAE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
SVM	15.79 ± 3.83	15.90 ± 3.19	15.93 ± 2.92	15.06 ± 3.39	16.07 ± 3.44	15.58 ± 3.21	16.17 ± 4.19	13.73 ± 3.21	13.77 ± 3.20	15.92 ± 2.05	14.35 ± 3.22
XGBoost	14.67 ± 3.58	15.14 ± 3.37	14.51 ± 3.18	13.87 ± 3.28	16.12 ± 3.21	15.96 ± 3.30	15.33 ± 3.04	13.32 ± 2.35	13.26 ± 2.48	13.50 ± 2.26	12.80 ± 2.32
GBR	**14.29 ± 3.56**	14.54 ± 3.19	**13.51 ± 3.39**	13.77 ± 3.35	15.15 ± 2.96	15.01 ± 3.00	14.39 ± 3.35	13.80 ± 2.51	12.84 ± 2.48	13.16 ± 2.48	12.75 ± 2.48
LightGBM	14.39 ± 3.32	**14.50 ± 3.06**	13.78 ± 3.11	13.82 ± 3.34	**14.47 ± 3.02**	**14.46 ± 3.26**	**13.98 ± 3.49**	**12.54 ± 2.44**	12.56 ± 2.43	13.00 ± 2.24	**12.43 ± 2.50**

Highlighted in **bold** is the best score of the cohort. mean absolute error (MAE), root mean squared error (RMSE), least absolute shrinkage and selection operator (LASSO), random forest (RF), support vector machines (SVM), extreme gradient boosting (XGBoost), gradient boosting regressor (GBR), light gradient boosting machine (LightGBM)

Table 8.14: Mean and standard deviation (SD) of the different preprocessing approaches for stratified regression models.

MAE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
Linear	11.51 ± 3.00	13.46 ± 3.20	13.07 ± 3.26	12.88 ± 3.34	14.79 ± 2.62	12.70 ± 3.12	9.38 ± 1.89	8.40 ± 1.45	8.49 ± 1.46	18.16 ± 5.25	8.46 ± 1.41
LASSO	9.45 ± 2.11	9.39 ± 1.92	9.75 ± 1.95	8.90 ± 2.13	9.53 ± 2.06	9.80 ± 1.98	9.33 ± 1.91	8.35 ± 1.38	8.36 ± 1.39	8.45 ± 1.34	8.35 ± 1.37
Ridge	10.23 ± 2.37	9.77 ± 1.91	9.98 ± 1.96	9.72 ± 2.05	9.77 ± 1.98	9.96 ± 2.04	9.35 ± 1.91	8.40 ± 1.39	8.43 ± 1.40	8.94 ± 1.33	8.38 ± 1.38
RF	9.97 ± 2.02	10.09 ± 2.12	9.97 ± 2.16	9.42 ± 2.17	10.71 ± 2.03	10.34 ± 1.76	9.43 ± 2.07	8.54 ± 1.63	8.55 ± 1.42	8.58 ± 1.48	8.31 ± 1.64
SVM	8.97 ± 2.76	**8.91** ± **2.73**	9.10 ± 2.61	8.82 ± 2.77	**9.18** ± **2.69**	**8.69** ± **2.70**	**7.41** ± **3.04**	**6.59** ± **2.14**	**6.62** ± **2.13**	9.63 ± 1.76	**6.59** ± **2.14**
XGBoost	**8.97** ± **2.57**	9.06 ± 2.59	**8.77** ± **2.78**	**8.52** ± **2.71**	9.30 ± 2.82	9.27 ± 2.82	8.30 ± 2.67	7.88 ± 1.75	7.67 ± 1.78	**7.67** ± **1.84**	7.73 ± 1.67
GBR	9.64 ± 2.08	9.60 ± 2.00	9.53 ± 2.02	9.39 ± 2.13	9.93 ± 1.98	10.01 ± 1.96	9.59 ± 1.86	8.42 ± 1.41	8.45 ± 1.38	8.45 ± 1.36	8.31 ± 1.37

Continued on next page

Table 8.14: Mean and standard deviation (SD) of the different preprocessing approaches for stratified regression models. (Continued)

MAE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
LightGBM	9.89 ± 2.15	10.21 ± 2.00	9.62 ± 2.24	8.94 ± 2.41	10.27 ± 2.17	10.10 ± 2.17	9.57 ± 1.92	8.23 ± 1.38	8.26 ± 1.39	8.69 ± 1.47	8.11 ± 1.40
RMSE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
Linear	17.44 ± 3.67	20.35 ± 3.44	19.48 ± 3.73	18.85 ± 3.47	21.72 ± 3.79	19.11 ± 3.71	13.74 ± 3.59	12.33 ± 2.69	12.47 ± 2.70	27.05 ± 5.21	12.24 ± 2.64
LASSO	13.99 ± 3.71	**13.51 ± 3.63**	14.17 ± 3.56	**13.12 ± 3.70**	**13.63 ± 3.73**	**13.95 ± 3.64**	**13.55 ± 3.62**	**12.13 ± 2.61**	**12.13 ± 2.61**	**12.18 ± 2.62**	**12.09 ± 2.60**
Ridge	14.78 ± 3.77	13.76 ± 3.60	14.25 ± 3.36	13.82 ± 3.55	13.73 ± 3.63	14.08 ± 3.56	13.57 ± 3.61	12.18 ± 2.61	12.19 ± 2.61	12.55 ± 2.57	12.13 ± 2.60
RF	14.33 ± 3.57	14.45 ± 3.57	14.13 ± 3.59	13.83 ± 3.64	14.77 ± 3.48	14.44 ± 3.27	14.38 ± 3.76	12.81 ± 2.87	12.66 ± 2.66	12.63 ± 2.70	12.84 ± 2.87

Continued on next page

Table 8.14: Mean and standard deviation (SD) of the different preprocessing approaches for stratified regression models. (Continued)

MAE	All	Mean	Median	Minimum	Maximum	Range	Baseline $n = 90$	Sampling frequency	Sampling frequency + noise	Encoded	Baseline $n = 154$
SVM	14.38 ± 4.31	15.24 ± 4.42	15.54 ± 4.09	14.66 ± 3.93	15.36 ± 4.29	14.99 ± 4.31	14.16 ± 4.94	12.72 ± 3.52	12.74 ± 3.50	14.03 ± 2.42	12.72 ± 3.52
XGBoost	14.32 ± 4.43	14.49 ± 4.51	14.04 ± 4.53	13.64 ± 4.33	14.77 ± 4.54	14.64 ± 4.51	13.94 ± 4.71	12.75 ± 3.13	12.66 ± 3.18	12.59 ± 3.19	12.44 ± 3.02
GBR	**13.92 ± 3.60**	13.84 ± 3.64	**13.66 ± 3.56**	13.58 ± 3.58	14.20 ± 3.56	14.08 ± 3.60	13.87 ± 3.60	12.31 ± 2.60	12.30 ± 2.61	12.33 ± 2.65	12.32 ± 2.57
LightGBM	14.65 ± 3.35	15.23 ± 3.28	14.17 ± 3.44	13.32 ± 3.80	14.87 ± 3.49	14.64 ± 3.70	14.21 ± 3.65	12.37 ± 2.68	12.39 ± 2.70	13.09 ± 2.60	12.16 ± 2.67

Highlighted in **bold** is the best score of the cohort. mean absolute error (MAE), root mean squared error (RMSE), least absolute shrinkage and selection operator (LASSO), random forest (RF), support vector machines (SVM), extreme gradient boosting (XGBoost), gradient boosting regressor (GBR), light gradient boosting machine (LightGBM)

Table 8.15: Pairwise comparison of the non stratified vs stratified model based on very acute lower extremity motor score (LEMS)

Model	RMSE t value	RMSE p value	MAE t value	MAE p value
All features				
LR	7.7657	**4.37E-10**	6.9807	**7.11E-09**
LASSO regression	3.4875	1.04E-03	8.3738	**5.15E-11**
Ridge regression	6.2187	**1.08E-07**	6.1836	**1.22E-07**
RF regressor	0.3343	7.40E-01	0.4458	6.58E-01
SVM regressor	5.2293	**3.51E-06**	8.6326	**2.09E-11**
XGBoost	1.2082	2.33E-01	6.2638	**9.17E-08**
GBR	2.1423	3.72E-02	3.4218	1.26E-03
LightGBM	-1.7027	9.50E-02	5.5915	**9.90E-07**
Mean cohort				
LR	-9.4788	**1.15E-12**	-6.0014	**2.33E-07**
LASSO regression	4.5166	**3.97E-05**	11.8027	**6.21E-16**
Ridge regression	6.7795	**1.46E-08**	9.6696	**6.07E-13**
RF regressor	0.9083	3.68E-01	2.0552	4.52E-02
SVM regressor	2.0632	4.44E-02	9.3646	**1.70E-12**
XGBoost	1.9575	5.60E-02	7.8661	**3.06E-10**

Continued on next page

Table 8.15: Pairwise comparison of the non stratified vs stratified model based on very acute lower extremity motor score (LEMS) (Continued)

Model	RMSE t value	RMSE p value	MAE t value	MAE p value
GBR	4.1944	**1.14E-04**	5.8198	**4.43E-07**
LightGBM	-4.855	**1.27E-05**	3.646	6.44E-04
Median cohort				
LR	-7.7014	**5.48E-10**	-5.5733	**1.06E-06**
LASSO regression	3.3568	1.53E-03	10.2929	**7.65E-14**
Ridge regression	6.6432	**2.37E-08**	10.6194	**2.64E-14**
RF regressor	0.8914	3.77E-01	1.1779	2.45E-01
SVM regressor	1.6333	1.09E-01	9.1117	**4.02E-12**
XGBoost	1.3465	1.84E-01	6.2279	**1.04E-07**
GBR	-1.0527	2.98E-01	1.283	2.06E-01
LightGBM	-2.804	7.21E-03	4.1131	**1.49E-04**
Minimum cohort				
LR	-7.0735	**5.10E-09**	-4.5311	**3.78E-05**
LASSO regression	3.753	4.63E-04	10.8661	**1.19E-14**
Ridge regression	3.7399	4.83E-04	7.446	**1.36E-09**
RF regressor	-1.0306	3.08E-01	-0.3602	7.20E-01

Continued on next page

Table 8.15: Pairwise comparison of the non stratified vs stratified model based on very acute lower extremity motor score (LEMS) (Continued)

Model	RMSE t value	RMSE p value	MAE t value	MAE p value
SVM regressor	1.3524	1.82E-01	5.8821	**3.55E-07**
XGBoost	0.8234	4.14E-01	5.6448	**8.21E-07**
GBR	1.1821	2.43E-01	2.4351	1.86E-02
LightGBM	2.5622	1.35E-02	5.3808	**2.07E-06**
Maximum cohort				
LR	-8.0921	**1.38E-10**	-4.8532	**1.28E-05**
LASSO regression	4.8256	**1.40E-05**	11.0114	**7.46E-15**
Ridge regression	7.0753	**5.07E-09**	10.3547	**6.25E-14**
RF regressor	-1.786	8.03E-02	-1.3263	1.91E-01
SVM regressor	1.8293	7.35E-02	7.9337	**2.41E-10**
XGBoost	4.7509	**1.81E-05**	11.8678	**5.08E-16**
GBR	4.9263	**9.96E-06**	8.5262	**3.03E-11**
LightGBM	-2.7547	8.22E-03	4.6006	**3.00E-05**
Range cohort				
LR	-5.148	**4.65E-06**	-2.3987	2.03E-02
LASSO regression	3.4933	1.02E-03	7.5393	**9.74E-10**

Continued on next page

Table 8.15: Pairwise comparison of the non stratified vs stratified model based on very acute lower extremity motor score (LEMS) (Continued)

Model	RMSE t value	RMSE p value	MAE t value	MAE p value
Ridge regression	4.9236	**1.00E-05**	8.3309	**5.99E-11**
RF regressor	1.6123	1.13E-01	2.4869	1.63E-02
SVM regressor	1.6824	9.88E-02	9.343	**1.83E-12**
XGBoost	4.3087	**7.88E-05**	9.8684	**3.12E-13**
GBR	4.955	**9.03E-06**	8.2515	**7.90E-11**
LightGBM	-0.9095	3.68E-01	4.7695	**1.70E-05**
Sampling frequency cohort				
LR	3.5346	9.02E-04	9.7821	**4.16E-13**
LASSO regression	5.0276	**7.04E-06**	10.6571	**2.33E-14**
Ridge regression	4.6009	**2.99E-05**	8.7564	**1.36E-11**
RF regressor	0.3518	7.26E-01	2.9259	5.19E-03
SVM regressor	4.3352	**7.22E-05**	8.8164	**1.11E-11**
XGBoost	3.0653	3.53E-03	7.8507	**3.23E-10**
GBR	5.8472	**4.02E-07**	8.095	**1.37E-10**
LightGBM	2.2234	3.08E-02	8.2631	**7.59E-11**
Cohort encoded based on normal range				

Continued on next page

Table 8.15: Pairwise comparison of the non stratified vs stratified model based on very acute lower extremity motor score (LEMS) (Continued)

Model	RMSE t value	RMSE p value	MAE t value	MAE p value
LR	6.9072	**9.24E-09**	6.8781	**1.02E-08**
LASSO regression	6.0577	**1.91E-07**	12.1718	**2.00E-16**
Ridge regression	10.3337	**6.70E-14**	12.0916	**2.55E-16**
RF regressor	2.9279	5.16E-03	6.6165	**2.61E-08**
SVM regressor	6.9328	**8.43E-09**	9.0193	**5.52E-12**
XGBoost	4.2239	**1.04E-04**	10.9961	**7.84E-15**
GBR	6.9537	**7.82E-09**	5.8894	**3.46E-07**
LightGBM	-0.5922	5.56E-01	9.8503	**3.31E-13**
Baseline $n = 90$				
LR	4.6982	**2.16E-05**	13.2762	**7.49E-18**
LASSO regression	5.2092	**3.76E-06**	13.9141	**1.21E-18**
Ridge regression	5.5963	**9.74E-07**	13.2688	**7.66E-18**
RF regressor	-0.1172	9.07E-01	4.3117	**7.80E-05**
SVM regressor	8.9132	**7.94E-12**	12.6071	**5.38E-17**
XGBoost	4.0609	**1.76E-04**	12.6748	**4.40E-17**
GBR	5.513	**1.30E-06**	7.8678	**3.04E-10**

Continued on next page

Table 8.15: Pairwise comparison of the non stratified vs stratified model based on very acute lower extremity motor score (LEMS) (Continued)

Model	RMSE t value	RMSE p value	MAE t value	MAE p value
LightGBM	-2.7204	9.00E-03	4.328	**7.40E-05**
Baseline $n = 154$				
LR	4.6437	**2.59E-05**	9.7457	**4.70E-13**
LASSO regression	5.5601	**1.11E-06**	11.439	**1.93E-15**
Ridge regression	5.297	**2.77E-06**	9.8255	**3.60E-13**
RF regressor	-1.5084	1.38E-01	2.1355	3.77E-02
SVM regressor	6.9439	**8.10E-09**	8.7088	**1.61E-11**
XGBoost	2.05	4.57E-02	11.573	**1.27E-15**
GBR	4.2551	**9.39E-05**	8.41	**4.54E-11**
LightGBM	3.7132	5.24E-04	11.4817	**1.69E-15**

For each cohort the non stratified model was compared to the stratified model with a dependent t test, with both root mean squared error (RMSE) and mean absolute error (MAE) scores. Significant p-values, after Bonferroni correction, are indicated in **bold**.
linear regression (LR), least absolute shrinkage and selection operator (LASSO), random forest (RF), support vector machines (SVM), extreme gradient boosting (XGBoost), gradient boosting regressor (GBR), light gradient boosting machine (LightGBM)

11.6 Chapter 6

Supplementary material was made publicly available with the corresponding manuscript.

11.7 Chapter 7

Table 8.16: Central cord syndrome (CCS) of the phenomenal and comparator groups defined in the Sygen and European multicenter study on human spinal cord injury (EMSCI) cohorts according to the clinical definition

	Sygen cohort		**EMSCI cohort**	
	PR group	Comparator group	PR group	Comparator group
CCS-1	2 (33)	0 (0)	2 (25)	0 (0)
CCS-5	1 (16.7)	0 (0)	1 (12.5)	0 (0)
CCS-10	0 (0)	0 (0)	0 (0)	0 (0)
CCS-19	0 (0)	0 (0)	0 (0)	0 (0)
NLI-based CCS	3 (50)	1 (0.3)	3 (37.5)	0 (0)

central cord syndrome (CCS), neurological level of injury (NLI)

Table 8.17: Antibiotics prescribed and their proportions in the phenomenal recovery (PR) and comparator groups from the clinical definition

	PR group	**Comparator group**
acetic acid	0.17	0.003
amoxicillin	0.17	0.06
amoxicillin and clavulanate potassium	0.17	0.04
ampicillin	0.17	0.10
bacitracin	0.33	0.14
cefazolin	0.50	0.58
cefotaxime	0.17	0.07
ceftazidime	0.33	0.29

Continued on next page

Table 8.17: Antibiotics prescribed and their proportions in the phenomenal recovery (PR) and comparator groups from the clinical definition (Continued)

	PR group	Comparator group
ceftriaxone	0.67	0.17
ciprofloxacin	0.17	0.48
clindamycin	0.33	0.12
gentamicin	0.83	0.46
metronidazole	0.33	0.17
nafcillin	0.17	0.08
ofloxacin	0.33	0.07
piperacillin	0.17	0.14
ticarcillin	0.17	0.14
trimethoprim	0.50	0.43
vancomycin	1.00	0.44

phenomenal recovery (PR)

Table 8.18: Antibiotics prescribed and their proportions in the phenomenal recovery (PR) and comparator groups from the statistical definition

	PR group	Comparator group	p-value
amikacin	0.04	0.03	1
amoxicillin	0.04	0.05	1
amoxicillin and clavulanate potassium	0.04	0.05	1
ampicillin	0.11	0.10	1
ampicillin and sulbactam	0.11	0.05	0.38
bacitracin	0.18	0.08	0.16
cefazolin	0.54	0.45	0.41
cefotaxime	0.18	0.09	0.18
cefoxitin	0.04	0.02	0.49

Continued on next page

Table 8.18: Antibiotics prescribed and their proportions in the phenomenal recovery (PR) and comparator groups from the statistical definition (Continued)

	PR group	Comparator group	p-value
ceftazidime	0.11	0.19	0.41
ceftriaxone	0.32	0.07	< 0.001
cefuroxime	0.04	0.05	1
ciprofloxacin	0.14	0.25	0.32
clindamycin	0.11	0.07	0.42
gentamicin	0.43	0.38	0.66
imipenem	0.11	0.01	0.03
metronidazole	0.04	0.07	0.69
nafcillin	0.07	0.05	0.66
nitrofurantoin	0.04	0.05	1
norfloxacin	0.04	0.03	1
ofloxacin	0.03	0.03	1
penicillin	0.04	0.05	1
piperacillin	0.07	0.09	1
ticarcillin	0.07	0.08	1
tobramycin	0.11	0.08	0.71
trimethoprim	0.29	0.37	0.51
vancomycin	0.36	0.35	1

phenomenal recovery (PR), Note that p-values reported are not Bonferroni-corrected, in **bold** are the p-values significant after multiple testing correction.

12 Use of artifical intelligence (AI) tools

AI-based tool	Use case	Scope	Remarks
DeepL	Translation of text	Résumé	From English to French

Continued on next page

AI-based tool	Use case	Scope	Remarks
		Zusammenfassung	From English to German
ChatGPT 3.0	Refine sentences and wording	Section 4.2	
	Brainstorming ideas	Thesis title	

artifical intelligence (AI)

13 Curriculum vitae

Date of Birth	12^{th} January 1999	**Google Scholar**	Lucie Bourguignon
Nationality	French	**GitHub**	lbourguignon
Email	blucie@ethz.ch	**ORCID**	0000-0001-8049-6461

Education

2020-now PhD candidate in Biomedical Data Science - ETH Zürich, Switzerland

Supervised by Prof. Dr. Catherine Jutzeler
Topic: Harnessing the potential of data science to enhance clinical trials for spinal cord injury

2018-2020 MSc in Computational Biology and Bioinformatics - ETH Zürich, Switzerland

Overall Grade Point Average : 5.59/6

2016-now MD-PhD program - INSERM school, France

Rank : 7/70

2015-2018 BSc in Human Medicine - Université de Bordeaux, France

First Year Rank : 125/1854

Research Experience

Feb 2023 - Jul 2023 **Research exchange**, supervised by Prof. Dr. John Kramer
ICORD, International Collaboration on Repair Discoveries
UBC Faculty of Medicine and VGH Research Institute, British Columbia, Vancouver, Canada

Topic: The concept of positive deviance applied to spinal cord injury recovery: an exploratory analysis

Jun 2020 - now **PhD project**, supervised by Prof. Dr. Catherine Jutzeler
BMDS, Biomedical Data Science
Department of Health Sciences & Technologies, ETH Zürich, Zürich, Switzerland
Topic: Harnessing the potential of data science to enhance clinical trials for spinal cord injury

Nov 2019 - Jun 2020 **Master's thesis**, supervised by Dr. Catherine Jutzeler & Prof. Dr. Karsten Borgwardt
MLCB, Machine Learning and Computational Biology
Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland
Topic: Mortality prediction using self-reported health records and large scale genomic data

Oct 2019 - Nov 2019 **Semester research project**, supervised by Dr. Catherine Jutzeler
MLCB, Machine Learning and Computational Biology
Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland
Topic: Study of blood markers in spinal cord injuries

Jun 2019 - Jul 2019 **Semester research project**, supervised by Prof. Dr. Sebastian Bonhoeffer
Theoretical Biology
Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
Topic: HIV long-term evolution analysis

Jun 2018 - Aug 2018 **Summer research project**, supervised by Dr. Olivier Tenaillon
QEM, Quantitative Evolutionary Microbiology
Infection, antimicrobials, modelling, evolution institute, INSERM UMR 1137, Paris, France
Topic: DNA sequencing of *E. coli*, coalescent simulations

Jun 2017 - Aug 2017 **Summer research project**, supervised by Dr. Olivier Saut
MONC, Mathematical Modeling Applied to Oncology
Mathematical Institute of Bordeaux, INRIA South-West, Bordeaux, France
Topic: Prediction of the evolution of kidney tumors based on MRI images

Presentation & invited lectures

May 2024 Poster presentation - 2024 ASIA Annual Scientific Meeting
The concept of positive deviance applied to spinal cord injury recovery - an analysis of medications received by patients exhibiting a phenomenal recovery

Dec 2023 Oral presentation - 2023 French MD-PhD scientific days
Data-driven approaches to enhance drug discovery in spinal cord injuries

Oct 2023 Poster presentation - 2023 ISCoS Annual Scientific Meeting

The concept of positive deviance applied to spinal cord injury recovery - an analysis of patients exhibiting a "phenomenal recovery"

Oct 2023 Poster presentation - 2023 ISCoS Annual Scientific Meeting

Effects of commonly administered drugs on spinal cord injury - A systematic review

Jul 2023 Invited lecture - 2023 UBC Vancouver Summer Program Pharmacology course

Introduction to the scientific method

Mar 2023 Poster presentation - 2023 ICORD annual research meeting

Analysis of phenomenal recovery after spinal cord injury - a data-driven approach to enhance recovery

Dec 2022 Oral presentation - 2022 French MD-PhD scientific days

Data-driven approaches to inform drug repurposing for spinal cord injury

Oct 2022 Oral presentation - 2022 Swiss MD-PhD program retreat

When clinical data goes missing: challenges and impact of data imputation in the field of spinal cord injury

Sept 2022 Poster presentation - 2022 ISCoS Annual Scientific Meeting

Do commonly administered drugs inadvertently modify the progression of spinal cord injury?

Sept 2022 Oral presentation - 2022 ISCoS Annual Scientific Meeting

Studying missingness in spinal cord injury data: Challenges and impact of data imputation

June 2022 Oral presentation - 2022 Swiss MD-PhD conference

Studying the polypharmacy administered following acute spinal cord injury

May 2022 Oral presentation - 2022 ASIA Annual Scientific Meeting

What the literature tells us about drugs used in acute care following spinal cord injury

Dec 2021 Oral presentation - 2021 ISCoS Annual Scientific Meeting

Trauma-induced perturbations of serological markers

Travel grant: SCI Research Collaboration Grant from the Spinal Research Institute

Oct 2021 Oral presentation - 2021 French MD-PhD scientific days

Informative missingness - The example of data from spinal cord injury clinical research

Sept 2020 Oral presentation - 2020 French MD-PhD scientific days

Can your genome say for how long you will survive?

Honors and awards

Mar 2023 Third prize PhD poster award - CAD$ 250

Annual research meeting of the ICORD Institute

Feb 2023 International exchange award - CAD$ 7'500

International exchange award from the ICORD Institute

Feb 2023 SNF mobility grant - CHF 16'000

Mobility grant awarded for a 6-months research stay in Vancouver, BC, Canada
Under the supervision of Prof. Dr. John Kramer

Sept 2021 Travel grant - £220

SCI Research Collaboration Grant from the Spinal Research Institute

Jan 2020 Master's thesis mobility grant - 1'500€

Ecole de l'INSERM Liliane Bettencourt

Outreach activities

Participant in Business Concept course from Innosuisse:
I was selected to participate in this 12-week acceleration program for early-stage startup ideas. This course aims to teach how to turn an idea into a promising startup, train entrepreneurial thinking and acting and prepare participants theoretically as well as practically for the foundation and management of their own company.

Lead organiser of the 2023 annual French MD-PhD thematic conference:
I proposed the topic *Data analysis and artificial intelligence approaches for biomedical research*, which was accepted by the board of the INSERM school French MD-PhD program. Alongside with a fellow MD-PhD student, we are in charge of building the program and contacting both national and international speakers. Additionally, we will give a 90-minute presentation to the MD-PhD audience, introducing essential concepts in data analysis, statistics and machine learning applied to medical data.

Active member of the Swiss MD-PhD association (SMPA):
Since June 2022, I am in charge of communication of the association on diverse social media platforms (e.g. Twitter, LinkedIn), participate to the redaction of the trimestrial newsletter sent to more than 250 members, represent and promote the association to medical students, and co-organise local events in Zurich. In June 2023, my role was made official and I took the position of Head of Social Media as part of the board of the SMPA.

Scientifica 2023:
Co-organiser of the BMDS lab booth on *Understanding spinal cord injuries better thanks to AI*

Participation in the Kangaroo goes Science project:
The Kangaroo goes science project aims at promoting science and research towards seventh grade female students. My role was to prepare a video presenting my daily life as a female researcher at ETH. You can find the video here.

Scientifica 2021 :
The event, co-organised by ETH Zurich and the University of Zurich and ETH Zurich, aims to bring scientific research and knowledge into the public domain. Every second year, researchers and scientists from both institutions are invited to present their work around a specific topic. In its last edition focusing on "Synthetic naturally", I was involved in the development of a web-based application illustrating the use of machine learning and growing medical data to inform clinical decision. Through simulated examples, we were able to introduce to a broader audience our research for personalised and precision medicine.

Service to the community

Reviewer for scientific journals :

Web of Science ResearcherID ADP-6132-2022

Teaching assistant:

376-1723-00L Big Data Analysis in Biomedical Research

376-1983-00L Foundations of Data Science

Supervision of junior researchers

Garance Jaques : Master student in Computational Biology and Bioinformatics
Studying missingness in serological markers in relationship to recovery after spinal cord injury

Eljas Röllin : Master student in Computational Biology and Bioinformatics
Characterising the importance of genetic relatedness in phenotype prediction

Richard Affolter : Master student in Computational Biology and Bioinformatics
Integrating genetic and environmental data into phenotype prediction

Ufuk Ilgin : Master student in Molecular Health Sciences
Studying missing data and imputation methods in medical data

Mariia Kuleba : Research assistant in the Biomedical Data Science lab
Impact of the COVID-related restrictions on the physical activities of patients with neurological conditions

Jaimie Lee & Ryan Loke : Bachelor and Master students in Pharmacological sciences
The interactive manuscript: from tabular to interactive result presentation and data visualization

Jan Matthias : Master student in Molecular Health Sciences

Machine learning-based recovery prediction in spinal cord injury

Maya Louage : Master student in Computational Biology and Bioinformatics

Studying temporal variations in recovery patterns following spinal cord injury

Skills & Interests

Programming Languages : R (intermediate), Python (intermediate), C++ (beginner)

Languages : French (native), English (fluent), German (B2), Spanish (A2)

Sports : latin dances, running, climbing

Activities : reading, piano, cooking & baking

Acknowledgements

As everyone who went through the same process knows, a PhD is not the work of a unique person, but rather a combination of forces. I would like to dedicate this section to everyone that supported me, one way or another, in the last four years.

I would first like to thank the members of my committee: Prof. Dr. med. Armin Curt, Prof. Dr. Karsten Borgwardt, Prof. Dr. Torsten Hothorn and Prof. Dr. Catherine Jutzeler. Their joint guidance was crucial in developing and conducting the projects that I am reporting here. I would like to particularly express my gratitude to Prof. Dr. Catherine Jutzeler. She is and will remain a role-model for me, and I have been honoured to work by her side for the last five years. She greatly contributed to shape me as a better scientist, but more importantly as a better person. I could never thank her enough for everything she brought to my life.

During the course of my PhD, I had the great privilege to see Catherine shape her own team, full of wonderful colleagues and friends, in particular Louis Lukas, who was there from the start, for my greatest pleasure, and Dr. Sarah Brüningk, who greatly contributed to most of my projects. I would also like to thank all the students that I had the chance to supervise: Garance Jaques, Eljas Röllin, Richard Affolter, Ufuk Ilgin, Mariia Kuleba, Jan Matthias and Maya Louage. It has always been my desire to transmit them what I have learned along the way, but I truly learnt at least as much from them.

I would like to acknowledge that my time as a PhD student and related activities were made possible by several organisations: the Swiss National Science Foundation, the Praxis Spinal Cord Injury Institute, and the Spinal Research Institute, for their financial support; the Ecole de l'INSERM Liliane Bettencourt and Swiss MD-PhD association, which allowed me to keep in touch with my medical vocation; and the Swiss Federal Railways, as I worked on this thesis for quite a few hours in their trains. Additionally, I would like to remember that behind all the data I used stand individuals who have experienced or are experiencing living with a spinal cord injury. Their data was always handled with care and having at heart to advance knowledge to help them now and in the future.

As part of my PhD, I had the great chance to stay for six months as part of Prof. Dr. John Kramer's lab in Vancouver, BC, Canada. I would like to thank him, first for

his warm welcome within the lab, second for his continuous enthusiasm and valuable input and feedback related to my projects, and finally for the spirit that he instigated as part of his team. Numerous people contributed to making my stay unforgettable. I would like to address a special thank you to my colleague, flatmate and friend Dr. Paulina Scheuren; to Ryan Loke, Jaimie Lee and Lorenz Mauerer (aka the cool kids' crew) for our (mainly culinary) adventures; to Bobo Tong, the one necessary person to hold everything together; to Cassie Choles for her moral support; to Anh Khoa Vo, for being Anh (if you know, you know) and all other members of the lab (Dr. Lukas Linde, Himali Bergeron-Vitez, Nicole Bailey and many more). Because life is not only about science, I would also like to address a special note to the people I met at the Dance Vancouver Studio and to Gabriel in particular. They brought me joy when I needed it the most, and for that I thank them very much. My stay in Vancouver was undeniably one of the most enriching parts of my PhD. Knowing how hard it was for me to leave only shows how much I cherish this experience and all the moments we shared. Vancouver and all the above-mentioned friends left an indelible mark on me. From the bottom of my heart, thank you.

Talking about friends, I would like to address a special note to *meine besten Freunde*: Lauriane, Daniela, Thomas, Maurizio and Philippe; but also Giulia and Caroline. Their unwavering support was an essential key to go through the four years of my PhD.

Last but not least, I would like to thank my family, in particular my parents, Karine and Laurent Bourguignon, my brother Julien, my grandparents, Nicole and Marc Baudet [9] and Josette Etcheberigarray. I know that it was not always easy to cope with me, but your continuous love and support made me grow and brought me where I am today. My successes - past, present and future - are all yours.

[9] Je n'oublie pas le New England Journal of Medicine.

Bibliography

[1] Steven Kirshblum, Brittany Snider, Rüdiger Rupp, and Mary Schmidt Read. Updates of the international standards for neurologic classification of spinal cord injury. *Physical Medicine and Rehabilitation Clinics of North America*, 31(3):319–330, August 2020.

[2] Per Brodal. *The central nervous system*. Oxford University Press, New York, NY, 5 edition, July 2016.

[3] Christopher S. Ahuja, Jefferson R. Wilson, Satoshi Nori, Mark R. N. Kotter, Claudia Druschel, Armin Curt, and Michael G. Fehlings. Traumatic spinal cord injury. *Nature Reviews Disease Primers*, 3(1), April 2017.

[4] Seyed Behnam Jazayeri, Seyed Farzad Maroufi, Esmaeil Mohammadi, Mohammad Amin Dabbagh Ohadi, Ellen-Merete Hagen, Maryam Chalangari, Seyed Behzad Jazayeri, Mahdi Safdarian, Shayan Abdollah Zadegan, Zahra Ghodsi, and Vafa Rahimi-Movaghar. Incidence of traumatic spinal cord injury worldwide: A systematic review, data integration, and update. *World Neurosurgery: X*, 18:100171, April 2023.

[5] Michael Fehlings, Anoushka Singh, Lindsay Tetreault, Sukhvinder Kalsi-Ryan, and Aria Nouri. Global prevalence and incidence of traumatic spinal cord injury. *Clinical Epidemiology*, page 309, September 2014.

[6] Nancy P. Thorogood, Vanessa K. Noonan, Xiaozhi Chen, Nader Fallah, Suzanne Humphreys, Nicolas Dea, Brian K. Kwon, and Marcel F. Dvorak. Incidence and prevalence of traumatic spinal cord injury in canada using health administrative data. *Frontiers in Neurology*, 14, July 2023.

[7] R A Cripps, B B Lee, P Wing, E Weerts, J Mackay, and D Brown. A global map for traumatic spinal cord injury epidemiology: towards a living data repository for injury prevention. *Spinal Cord*, 49(4):493–501, November 2010.

[8] Claudio Barbiellini Amidei, Laura Salmaso, Stefania Bellio, and Mario Saia. Epidemiology of traumatic spinal cord injury: a large population-based study. *Spinal Cord*, 60(9):812–819, April 2022.

[9] Yuying Chen, Ying Tang, Victoria Allen, and Michael J. DeVivo. Aging and spinal cord injury: External causes of injury and implications for prevention. *Topics in Spinal Cord Injury Rehabilitation*, 21(3):218–226, August 2015.

[10] A Curt and V Dietz. Electrophysiological recordings in patients with spinal cord injury: significance for predicting outcome. *Spinal Cord*, 37(3):157–165, March 1999.

[11] T N Bryce, F Biering-Sørensen, N B Finnerup, D D Cardenas, R Defrin, T Lundeberg, C Norrbrink, J S Richards, P Siddall, T Stripling, R-D Treede, S G Waxman, E Widerström-Noga, R P Yezierski, and M Dijkers. International spinal cord injury pain classification: part i. background and description. *Spinal Cord*, 50(6):413–417, December 2011.

[12] Patrick Freund, Maryam Seif, Nikolaus Weiskopf, Karl Friston, Michael G Fehlings, Alan J Thompson, and Armin Curt. Mri in traumatic spinal cord injury: from clinical assessment to neuroimaging biomarkers. *The Lancet Neurology*, 18(12):1123–1135, December 2019.

[13] Sukhvinder Kalsi-Ryan, Colin Chan, Mary Verrier, Armin Curt, Michael Fehlings, Marc Bolliger, and Inge-Marie Velstra. The graded redefined assessment of strength sensibility and prehension version 2 (gv2): Psychometric properties. *The Journal of Spinal Cord Medicine*, 42(sup1):149–157, September 2019.

[14] Dorland. *Dorland's illustrated medical dictionary*. Dorland's Medical Dictionary. W B Saunders, London, England, 32 edition, May 2011.

[15] Rüdiger Rupp, Fin Biering-Sørensen, Stephen P. Burns, Daniel E. Graves, James Guest, Linda Jones, Mary Schmidt Read, Gianna M. Rodriguez, Christian Schuld, Keith E. Tansey-MD, Kristen Walden, and Steven Kirshblum. International standards for neurological classification of spinal cord injury. *Topics in Spinal Cord Injury Rehabilitation*, 27(2):1–22, March 2021.

[16] Steven C. Kirshblum, Stephen P. Burns, Fin Biering-Sorensen, William Donovan, Daniel E. Graves, Amitabh Jha, Mark Johansen, Linda Jones, Andrei Krassioukov, M.J. Mulcahey, Mary Schmidt-Read, and William Waring. International standards for neurological classification of spinal cord injury (revised 2011). *The Journal of Spinal Cord Medicine*, 34(6):535–546, November 2011.

[17] Steven Kirshblum, Brittany Snider, Fatma Eren, and James Guest. Characterizing natural recovery after traumatic spinal cord injury. *Journal of Neurotrauma*, 38(9):1267–1284, May 2021.

[18] Kim D. Anderson. Targeting recovery: Priorities of the spinal cord-injured population. *Journal of Neurotrauma*, 21(10):1371–1383, October 2004.

[19] M Itzkovich, H Shefler, L Front, R Gur-Pollack, K Elkayam, V Bluvshtein, I Gelernter, and A Catz. SCIM III (spinal cord independence measure version III): reliability of assessment by interview and comparison with assessment by observation. *Spinal Cord*, 56(1):46–51, September 2017.

[20] M. Itzkovich, I. Gelernter, F. Biering-Sorensen, C. Weeks, M. T. Laramee, B. C. Craven, M. Tonack, S. L. Hitzig, E. Glaser, G. Zeilig, S. Aito, G. Scivoletto, M. Mecci, R. J. Chadwick, W. S. El Masry, A. Osman, C. A. Glass, P. Silva, B. M. Soni, B. P. Gardner, G. Savic, E. M. Bergström, V. Bluvshtein, J. Ronen, and A. Catz. The spinal cord independence measure (scim) version iii: Reliability and validity in a multi-center international study. *Disability and Rehabilitation*, 29(24):1926–1933, January 2007.

[21] National Academies Press, January 1988.

[22] NIH's Definition of a Clinical Trial | grants.nih.gov — grants.nih.gov. `https://grants.nih.gov/policy/clinical-trials/definition.htm`. [Accessed 30-01-2024].

[23] Christopher J. Pannucci and Edwin G. Wilkins. Identifying and avoiding bias in research. *Plastic and Reconstructive Surgery*, 126(2):619–625, August 2010.

[24] Andrea Skelly, Joseph Dettori, and Erika Brodt. Assessing bias: the importance of considering confounding. *Evidence-Based Spine-Care Journal*, 3(01):9–12, February 2012.

[25] Vance W Berger and Derek V Exner. Detecting selection bias in randomized clinical trials. *Controlled Clinical Trials*, 20(4):319–327, August 1999.

[26] Giovanni Tripepi, Kitty J. Jager, Friedo W. Dekker, and Carmine Zoccali. Selection bias and information bias in clinical research. *Nephron Clinical Practice*, 115(2):c94–c99, April 2010.

[27] Eduardo Hariton and Joseph J Locascio. Randomised controlled trials – the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG: An International Journal of Obstetrics amp; Gynaecology*, 125(13):1716–1716, June 2018.

[28] Valerie A. Dietz, Nolan Roberts, Katelyn Knox, Sherilynne Moore, Michael Pitonak, Chris Barr, Jesus Centeno, Scott Leininger, Kent C. New, Peter Nowell, Matthew Rodreick, Cedric G. Geoffroy, Argyrios Stampas, and Jennifer N. Dulin. Fighting for recovery on multiple fronts: The past, present, and future of clinical trials for spinal cord injury. *Frontiers in Cellular Neuroscience*, 16, September 2022.

[29] Michael B. Bracken, Mary Jo Shepard, William F. Collins, Theodore R. Holford, Wise Young, David S. Baskin, Howard M. Eisenberg, Eugene Flamm, Linda Leo-Summers, Joseph Maroon, Lawrence F. Marshall, Phanor L. Perot, Joseph Piepmeier, Volker K.H. Sonntag, Franklin C. Wagner, Jack E. Wilberger, and H. Richard Winn. A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal-cord injury: Results of the second national acute spinal cord injury study. *New England Journal of Medicine*, 322(20):1405–1411, May 1990.

[30] Shannon Hextrum and Stephanie Bennett. A critical examination of subgroup analyses: The national acute spinal cord injury studies and beyond. *Frontiers in Neurology*, 9, February 2018.

[31] R. John Hurlbert. Methylprednisolone for acute spinal cord injury: an inappropriate standard of care. *Journal of Neurosurgery: Spine*, 93(1):1–7, July 2000.

[32] Michael G. Fehlings, Jefferson R. Wilson, Lindsay A. Tetreault, Bizhan Aarabi, Paul Anderson, Paul M. Arnold, Darrel S. Brodke, Anthony S. Burns, Kazuhiro Chiba, Joseph R. Dettori, Julio C. Furlan, Gregory Hawryluk, Langston T. Holly, Susan Howley, Tara Jeji, Sukhvinder Kalsi-Ryan, Mark Kotter, Shekar Kurpad, Brian K. Kwon, Ralph J. Marino, Allan R. Martin, Eric Massicotte, Geno Merli, James W. Middleton, Hiroaki Nakashima, Narihito Nagoshi, Katherine Palmieri, Andrea C. Skelly, Anoushka Singh, Eve C. Tsai, Alexander Vaccaro, Albert Yee, and James S. Harrop. A clinical practice guideline for the management of patients with acute spinal cord injury: Recommendations on the use of methylprednisolone sodium succinate. *Global Spine Journal*, $7(3_supplement) : 203S - -211S, September 2017$.

[33] Fred H. Geisler, William P. Coleman, Giacinto Grieco, and Devinder Poonian. The sygen® multicenter acute spinal cord injury study. *Spine*, 26(Supplement):S87–S98, December 2001.

[34] Fred H. Geisler, Frank C. Dorsey, and William P. Coleman. Recovery of motor function after spinal-cord injury — a randomized, placebo-controlled trial with GM-1 ganglioside. *New England Journal of Medicine*, 324(26):1829–1838, June 1991.

[35] Thanuja Dharmadasa and Matthew C Kiernan. Riluzole, disease stage and survival in als. *The Lancet Neurology*, 17(5):385–386, May 2018.

[36] Michael G. Fehlings, Jefferson R. Wilson, Ralph F. Frankowski, Elizabeth G. Toups, Bizhan Aarabi, James S. Harrop, Christopher I. Shaffrey, Susan J. Harkema, James D. Guest, Charles H. Tator, Keith D. Burau, Michele W. Johnson, and Robert G. Grossman. Riluzole for the treatment of acute traumatic spinal cord injury: rationale for and design of the nactn phase i clinical trial. *Journal of Neurosurgery: Spine*, 17(Suppl1):151–156, September 2012.

[37] Robert G. Grossman, Michael G. Fehlings, Ralph F. Frankowski, Keith D. Burau, Diana S.L. Chow, Charles Tator, Angela Teng, Elizabeth G. Toups, James S. Harrop, Bizhan Aarabi, Christopher I. Shaffrey, Michele M. Johnson, Susan J. Harkema, Maxwell Boakye, James D. Guest, and Jefferson R. Wilson. A prospective, multicenter, phase i matched-comparison group trial of safety, pharmacokinetics, and preliminary efficacy of riluzole in patients with traumatic spinal cord injury. *Journal of Neurotrauma*, 31(3):239–255, February 2014.

[38] Michael Fehlings, Ali Moghaddamjou, James Harrop, Jonathon Ball, Paul Arnold, James Guest, Shekar Kurpad, James Schuster, Ahmad Nassr, Jefferson Wilson, Darrel Brodke, Wilson Ray, and Branko Kopjar. 69. safety and efficacy of riluzole in acute spinal cord injury (riscis): a multicenter, randomized, placebo-controlled, double-blinded trial. *The Spine Journal*, 23(9):S36, September 2023.

[39] Michael Fehlings, Branko Kopjar, and Robert Grossman. Efficacy and safety of riluzole in acute spinal cord injury (sci). rationale and design of aospine phase iii multi-center double blinded randomized controlled trial. (riscis). *Global Spine Journal*, $6(1_{s}uppl)$: $s--0036--1582944--s--0036--1582944, April2016$.

[40] Thomas Liebscher, Lisa Schnell, Dina Schnell, Jeannette Scholl, Regula Schneider, Mirjam Gullo, Karim Fouad, Anis Mir, Martin Rausch, Diana Kindler, Frank P. T. Hamers, and Martin E. Schwab. Nogo-a antibody improves regeneration and locomotion of spinal cord–injured rats. *Annals of Neurology*, 58(5):706–719, September 2005.

[41] Klaus Kucher, Donald Johns, Doris Maier, Rainer Abel, Andreas Badke, Hagen Baron, Roland Thietje, Steven Casha, Renate Meindl, Baltazar Gomez-Mancilla, Christian Pfister, Rüdiger Rupp, Norbert Weidner, Anis Mir, Martin E. Schwab, and Armin Curt. First-in-man intrathecal application of neurite growth-promoting anti-nogo-a antibodies in acute spinal cord injury. *Neurorehabilitation and Neural Repair*, 32(6–7):578–589, June 2018.

[42] *A Dictionary of Epidemiology*. Oxford University Press, January 2014.

[43] Jacquelyn J. Cragg, Bobo Tong, Catherine R. Jutzeler, Freda M. Warner, Neil Cashman, Fred Geisler, and John L. K. Kramer. A longitudinal study of the neurologic safety of acute baclofen use after spinal cord injury. *Neurotherapeutics*, 16(3):858–867, February 2019.

[44] Bobo Tong, Catherine R. Jutzeler, Jacquelyn J. Cragg, Lukas Grassner, Jan M. Schwab, Steve Casha, Fred Geisler, and John L. K. Kramer. Serum albumin predicts long-term neurological outcomes after acute spinal cord injury. *Neurorehabilitation and Neural Repair*, 32(1):7–17, December 2017.

[45] Sagun Tuli, Jayshree Tuli, William P. Coleman, Fred H. Geisler, and Andrei Krassioukov. Hemodynamic parameters and timing of surgical decompression in acute cervical spinal cord injury. *The Journal of Spinal Cord Medicine*, 30(5):482–490, January 2007.

[46] Gale Whiteneck, Julie Cassaway, Marcel Dijkers, and Amitabh Jha. New approach to study the contents and outcomes of spinal cord injury rehabilitation: The SCIRehab project. *The Journal of Spinal Cord Medicine*, 32(3):251–259, January 2009.

[47] Teresa Foy, Ginger Perritt, Deepa Thimmaiah, Lauren Heisler, Jennifer Lookingbill Offutt, Kara Cantoni, Ching-Hui Hseih, Julie Gassaway, Rebecca Ozelie, and Deborah Backus. Occupational therapy treatment time during inpatient spinal cord injury rehabilitation. *The Journal of Spinal Cord Medicine*, 34(2):162–175, March 2011.

[48] Seyed Behnam Jazayeri, Samuel Berchi Kankam, Ali Golestani, Parnian Shobeiri, Morteza Gholami, Mohammad Amin Dabbagh Ohadi, Seyed Farzad Maroufi, Mohammad Reza Fattahi, Hamid Malekzadeh, Seyed Behzad Jazayeri, Zahra Ghodsi, Seyed Mohammad Ghodsi, and Vafa Rahimi-Movaghar. A systematic review and meta-analysis of the global epidemiology of pediatric traumatic spinal cord injuries. *European Journal of Pediatrics*, 182(12):5245–5257, October 2023.

[49] Yi Kang, Han Ding, Hengxing Zhou, Zhijian Wei, Lu Liu, Dayu Pan, and Shiqing Feng. Epidemiology of worldwide spinal cord injury: a literature review. *Jnanabha*, 6:1–9, December 2017.

[50] Muller Andreas C and Sarah Guido. *Introduction to machine learning with python*. O'Reilly Media, September 2016.

[51] Nam K Tran, Samer Albahra, Larissa May, Sarah Waldman, Scott Crabtree, Scott Bainbridge, and Hooman Rashidi. Evolving applications of artificial intelligence and machine learning in infectious diseases testing. *Clinical Chemistry*, 68(1):125–133, December 2021.

[52] Scott Mayer McKinney, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, Greg S. Corrado, Ara Darzi, Mozziyar

Etemadi, Florencia Garcia-Vicente, Fiona J. Gilbert, Mark Halling-Brown, Demis Hassabis, Sunny Jansen, Alan Karthikesalingam, Christopher J. Kelly, Dominic King, Joseph R. Ledsam, David Melnick, Hormuz Mostofi, Lily Peng, Joshua Jay Reicher, Bernardino Romera-Paredes, Richard Sidebottom, Mustafa Suleyman, Daniel Tse, Kenneth C. Young, Jeffrey De Fauw, and Shravya Shetty. International evaluation of an ai system for breast cancer screening. *Nature*, 577(7788):89–94, January 2020.

[53] Debraj Sen and Anusree Majumder. Artificial intelligence in cancer diagnostics and therapy: current perspectives. *Indian Journal of Cancer*, 0(0):0, 2021.

[54] T Rowland, L Ohno-Machado, and A Ohrn. Comparison of multiple prediction models for ambulation following spinal cord injury. *Proc. AMIA Symp.*, pages 528–532, 1998.

[55] Timothy Belliveau, Alan M. Jette, Subramani Seetharama, Jeffrey Axt, David Rosenblum, Daniel Larose, Bethlyn Houlihan, Mary Slavin, and Chantal Larose. Developing artificial neural network models to predict functioning one year after traumatic spinal cord injury. *Archives of Physical Medicine and Rehabilitation*, 97(10):1663–1668.e3, October 2016.

[56] P.R.S.S. Venkatapathiraju, Dr.V.Asanambigai, and Dr. Suresh Babu Mudunuri. Modelling a deep kernel-basel learning approach for spinal cord injury prediction. 2022.

[57] Dhruv Kapoor and Clark Xu. Spinal cord injury ais predictions using machine learning. *eneuro*, 10(1):ENEURO.0149–22.2022, December 2022.

[58] Zachary DeVries, Mohamad Hoda, Carly S Rivers, Audrey Maher, Eugene Wai, Dita Moravek, Alexandra Stratton, Stephen Kingwell, Nader Fallah, Jérôme Paquet, and Philippe Phan. Development of an unsupervised machine learning algorithm for the prognostication of walking ability in spinal cord injury patients. *The Spine Journal*, 20(2):213–224, February 2020.

[59] Faisal AlHuthaifi, Joseph Krzak, Timothy Hanke, and Lawrence C. Vogel. Predictors of functional outcomes in adults with traumatic spinal cord injury following inpatient rehabilitation: A systematic review. *The Journal of Spinal Cord Medicine*, 40(3):282–294, November 2016.

[60] Chihiro Kato, Osamu Uemura, Yasunori Sato, and Tetsuya Tsuji. Functional outcome prediction after spinal cord injury using ensemble machine learning. *Archives of Physical Medicine and Rehabilitation*, 105(1):95–100, January 2024.

[61] Adrian Cathomen, Laura Sirucek, Tim Killeen, Rainer Abel, Doris Maier, Norbert Weidner, Rüdiger Rupp, Torsten Hothorn, John D. Steeves, Armin Curt, and Marc Bolliger. Inclusive trial designs in acute spinal cord injuries: Prediction–based stratification of clinical

walking outcome and projected enrolment frequencies. *Neurorehabilitation and Neural Repair*, 36(4–5):274–285, February 2022.

[62] Zhan Sizheng, Huang Boxuan, Xue Feng, and Zhang Dianying. A functional outcome prediction model of acute traumatic spinal cord injury based on extreme gradient boost. *Journal of Orthopaedic Surgery and Research*, 17(1), October 2022.

[63] Chang Ho Yoon, Robert Torrance, and Naomi Scheinerman. Machine learning in medicine: should the pursuit of enhanced interpretability be abandoned? *Journal of Medical Ethics*, 48(9):581–585, May 2021.

[64] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artif. Intell.*, 267:1–38, 2017.

[65] Timothy L Lash, Tyler J VanderWeele, Sebastien Haneuse, and Kenneth J Rothman. *Modern epidemiology*. Lippincott Williams and Wilkins, Philadelphia, PA, 4 edition, January 2021.

[66] V. Dietz and K. Fouad. Restoration of sensorimotor functions after spinal cord injury. *Brain*, 137(3):654–667, October 2013.

[67] Charles H. Tator. Update on the pathophysiology and pathology of acute spinal cord injury. *Brain Pathology*, 5(4):407–413, October 1995.

[68] Leanne M Ramer, Matt S Ramer, and Elizabeth J Bradbury. Restoring function after spinal cord injury: towards clinical translation of experimental strategies. *The Lancet Neurology*, 13(12):1241–1256, December 2014.

[69] M H Rabadi, S K Mayanna, and A S Vincent. Predictors of mortality in veterans with traumatic spinal cord injury. *Spinal Cord*, 51(10):784–788, July 2013.

[70] Spinal cord injury facts and figures at a glance. *The Journal of Spinal Cord Medicine*, 36(1):1–2, January 2013.

[71] Andre M. Samuel, Daniel D. Bohl, Bryce A. Basques, Pablo J. Diaz-Collado, Adam M. Lukasiewicz, Matthew L. Webb, and Jonathan N. Grauer. Analysis of delays to surgery for cervical spinal cord injuries. *Spine*, 40(13):992–1000, July 2015.

[72] Fernando L. Vale, Jennifer Burns, Amie B. Jackson, and Mark N. Hadley. Combined medical and surgical treatment after acute spinal cord injury: results of a prospective pilot study to assess the merits of aggressive medical resuscitation and blood pressure management. *Journal of Neurosurgery*, 87(2):239–246, August 1997.

[73] M Saliba, D Saadeh, F Bouchand, B Davido, C Duran, B Clair, C Lawrence, D Annane, P Denys, J Salomon, L Bernard, and A Dinh. Outcome of bloodstream infections among spinal cord injury patients and impact of multidrug-resistant organisms. *Spinal Cord*, 55(2):148–154, December 2016.

[74] Christiana L. Cheng, Tova Plashkes, Tian Shen, Nader Fallah, Suzanne Humphreys, Colleen O'Connell, A. Gary Linassi, Chester Ho, Christine Short, Karen Ethans, Rebecca Charbonneau, Jérôme Paquet, and Vanessa K. Noonan and. Does specialized inpatient rehabilitation affect whether or not people with traumatic spinal cord injury return home? *Journal of Neurotrauma*, 34(20):2867–2876, October 2017.

[75] Michael J. De Vivo, J. Stuart Krause, and Daniel P. Lammertse. Recent trends in mortality and causes of death among persons with spinal cord injury. *Archives of Physical Medicine and Rehabilitation*, 80(11):1411–1419, November 1999.

[76] Nanna B Finnerup, Nadine Attal, Simon Haroutounian, Ewan McNicol, Ralf Baron, Robert H Dworkin, Ian Gilron, Maija Haanpää, Per Hansson, Troels S Jensen, Peter R Kamerman, Karen Lund, Andrew Moore, Srinivasa N Raja, Andrew S C Rice, Michael Rowbotham, Emily Sena, Philip Siddall, Blair H Smith, and Mark Wallace. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. *The Lancet Neurology*, 14(2):162–173, February 2015.

[77] M J DeVivo. Epidemiology of traumatic spinal cord injury: trends and future implications. *Spinal Cord*, 50(5):365–372, January 2012.

[78] E M Hagen, G E Eide, T Rekand, N E Gilhus, and M Gronning. A 50-year follow-up of the incidence of traumatic spinal cord injuries in western norway. *Spinal Cord*, 48(4):313–318, October 2009.

[79] Mitsunori Toda, Eiji Nakatani, Kaoru Omae, Masanori Fukushima, and Takaaki Chin. Age-specific characterization of spinal cord injuries over a 19-year period at a japanese rehabilitation center. *PLOS ONE*, 13(3):e0195120, March 2018.

[80] Erik von Elm, Douglas G. Altman, Matthias Egger, Stuart J. Pocock, Peter C. Gøtzsche, and Jan P. Vandenbroucke. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, 61(4):344–349, April 2008.

[81] Giorgio Scivoletto, Monica Torre, Alessia Mammone, Doris D. Maier, Norbert Weidner, Martin Schubert, Ruediger Rupp, Rainer Abel, Kalke Yorck-Bernhard, Kriz Jiri, Armin Curt, and Marco Molinari. Acute traumatic and ischemic spinal cord injuries have a

comparable course of recovery. *Neurorehabilitation and Neural Repair*, 34(8):723–732, July 2020.

[82] Christian Schuld, Julia Wiese, Andreas Hug, Cornelia Putz, Hubertus J.A. van Hedel, Martina R. Spiess, EM-SCI Study Group Norbert Weidner, and Rüdiger Rupp. Computer implementation of the international standards for neurological classification of spinal cord injury for consistent and efficient derivation of its subscores including handling of data from not testable segments. *Journal of Neurotrauma*, 29(3):453–461, February 2012.

[83] Flavia Stein Amiram Catz, Malka Itzkovich. The catz-itzkovich SCIM: a revised version of the spinal cord independence measure. *Disability and Rehabilitation*, 23(6):263–268, January 2001.

[84] J F Ditunno, P L Ditunno, G Scivoletto, M Patrick, M Dijkers, H Barbeau, A S Burns, R J Marino, and M Schmidt-Read. The walking index for spinal cord injury (WISCI/WISCI II): nature, metric properties, use and misuse. *Spinal Cord*, 51(5):346–355, March 2013.

[85] Hubertus J. van Hedel, Markus Wirz, and Volker Dietz. Assessing walking ability in subjects with spinal cord injury: Validity and reliability of 3 walking tests. *Archives of Physical Medicine and Rehabilitation*, 86(2):190–196, February 2005.

[86] Ralph J. Marino, Tarcisio Barros, Fin Biering-Sorensen, Stephen P. Burns, William H. Donovan, Daniel E. Graves, Michael Haak, Lesley M. Hudson, and Michael Priebe. International standards for neurological classification of spinal cord injury. *The Journal of Spinal Cord Medicine*, 26(sup1):S50–S56, January 2003.

[87] Peter Diggle, Patrick Heagerty, Kung-Yee Liang, and Scott Zeger. *Analysis of Longitudinal Data*, chapter Chapter 12. Time-dependent covariates. Oxford Statistical Science Series, 2002.

[88] Sean L. Simpson, Lloyd J. Edwards, Keith E. Muller, Pranab K. Sen, and Martin A. Styner. A linear exponent AR(1) family of correlation structures. *Statistics in Medicine*, 29(17):1825–1838, April 2010.

[89] Andrew Gelman, Jennifer Hill, and Masanao Yajima. Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5(2):189–211, April 2012.

[90] Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. *shiny: Web Application Framework for R*, 2023. https://shiny.posit.co/, https://github.com/rstudio/shiny.

[91] Winston Chang and Barbara Borges Ribeiro. *Package "ShinyDashboard": create dashboards with "Shiny"*, 2023. https://github.com/rstudio/shinydashboard.

[92] Michael J. DeVivo and Yuying Chen. Trends in new injuries, prevalent cases, and aging with spinal cord injury. *Archives of Physical Medicine and Rehabilitation*, 92(3):332–338, March 2011.

[93] Yuying Chen, Michael J. DeVivo, J. Scott Richards, and Theresa B. SanAgustin. Spinal cord injury model systems: Review of program and national database from 1970 to 2015. *Archives of Physical Medicine and Rehabilitation*, 97(10):1797–1804, October 2016.

[94] Jonviea D. Chamberlain, Olivier Deriaz, Margret Hund-Georgiadis, Sonja Meier, Anke Scheel-Sailer, Martin Schubert, Gerold Stucki, and Martin WG Brinkhof. Epidemiology and contemporary risk profile of traumatic spinal cord injury in switzerland. *Injury Epidemiology*, 2(1), November 2015.

[95] E S Lapteva, M R Tsutsunava, G M Podoprigora, and D S Diachkova-Gertseva. Falls in the elderly and senior age prevention perspectives. *Adv. Gerontol.*, 32(3):469–476, 2019.

[96] GEORGE St.J. PERROTT and DOROTHY F. HOLLAND. Population trends and problems of public health. *Milbank Quarterly*, 83(4):569–608, November 2005.

[97] Marco Franceschini, Humberto Cerrel Bazo, Fulvio Lauretani, Maurizio Agosti, and M. Cristina Pagliacci. Age influences rehabilitative outcomes in patients with spinal cord injury (SCI). *Aging Clinical and Experimental Research*, 23(3):202–208, June 2011.

[98] M P Jensen, A R Truitt, K G Schomer, K M Yorkston, C Baylor, and I R Molton. Frequency and age effects of secondary health conditions in individuals with spinal cord injury: a scoping review. *Spinal Cord*, 51(12):882–892, October 2013.

[99] S Knútsdóttir, H Thórisdóttir, K Sigvaldason, H Jónsson, A Björnsson, and P Ingvarsson. Epidemiology of traumatic spinal cord injuries in iceland from 1975 to 2009. *Spinal Cord*, 50(2):123–126, September 2011.

[100] Armin Curt, Hubertus J.A. Van Hedel, Daniel Klaus, and Volker Dietz and. Recovery from a spinal cord injury: Significance of compensation, neural plasticity, and repair. *Journal of Neurotrauma*, 25(6):677–685, June 2008.

[101] Brian A. Lee, Benjamin E. Leiby, and Ralph J. Marino. Neurological and functional recovery after thoracic spinal cord injury. *The Journal of Spinal Cord Medicine*, 39(1):67–76, December 2014.

[102] John L. K. Kramer, Daniel P. Lammertse, Martin Schubert, Armin Curt, and John D. Steeves. Relationship between motor recovery and independence after sensorimotor-complete cervical spinal cord injury. *Neurorehabilitation and Neural Repair*, 26(9):1064–1071, May 2012.

[103] Freda M. Warner, Jacquelyn J. Cragg, Catherine R. Jutzeler, Frank Röhrich, Norbert Weidner, Marion Saur, Doris D. Maier, Christian Schuld, Armin Curt, and John K. Kramer. Early administration of gabapentinoids improves motor recovery after human spinal cord injury. *Cell Reports*, 18(7):1614–1618, February 2017.

[104] Freda M. Warner, Catherine R. Jutzeler, Jacquelyn J. Cragg, Bobo Tong, Lukas Grassner, Frank Bradke, Fred Geisler, and John K. Kramer. The effect of non-gabapentinoid anticonvulsants on sensorimotor recovery after human spinal cord injury. *CNS Drugs*, 33(5):503–511, April 2019.

[105] Vieri Failli, Marcel A. Kopp, Christine Gericke, Peter Martus, Susann Klingbeil, Benedikt Brommer, Inês Laginha, Yuying Chen, Michael J. DeVivo, Ulrich Dirnagl, and Jan M. Schwab. Functional neurological recovery after spinal cord injury is impaired in patients with infections. *Brain*, 135(11):3238–3250, October 2012.

[106] Andrea L Behrman, Mark G Bowden, and Preeti M Nair. Neuroplasticity after spinal cord injury and training: An emerging paradigm shift in rehabilitation and walking recovery. *Physical Therapy*, 86(10):1406–1425, October 2006.

[107] Blessing N.R. Jaja, Jetan Badhiwala, James Guest, James Harrop, Chris Shaffrey, Max Boakye, Shekar Kurpad, Robert Grossman, Elizabeth Toups, Fred Geisler, Brian Kwon, Bizhan Aarabi, Mark Kotter, Michael G. Fehlings, and Jefferson R. Wilson. Trajectory-based classification of recovery in sensorimotor complete traumatic cervical spinal cord injury. *Neurology*, 96(22):e2736–e2748, April 2021.

[108] Jefferson R. Wilson, Blessing N.R. Jaja, Brian K. Kwon, James D. Guest, James S. Harrop, Bizhan Aarabi, Christopher I. Shaffrey, Jetan H. Badhiwala, Elizabeth G. Toups, Robert G. Grossman, and Michael G. Fehlings. Natural history, predictors of outcome, and effects of treatment in thoracic spinal cord injury: A multi-center cohort study from the north american clinical trials network. *Journal of Neurotrauma*, 35(21):2554–2560, November 2018.

[109] Steven M. Schwartz, Kevin Wildenhaus, Amy Bucher, and Brigid Byrd. Digital twins and the emerging science of self: Implications for digital health experience design and "small" data. *Frontiers in Computer Science*, 2, October 2020.

[110] Sreeram V. Ramagopalan, Alex Simpson, and Cormac Sammon. Can real-world data really replace randomised clinical trials? *BMC Medicine*, 18(1), January 2020.

[111] Kert Viele, Scott Berry, Beat Neuenschwander, Billy Amzal, Fang Chen, Nathan Enas, Brian Hobbs, Joseph G. Ibrahim, Nelson Kinnersley, Stacy Lindborg, Sandrine Micallef, Satrajit Roychoudhury, and Laura Thompson. Use of historical control data for assessing treatment effects in clinical trials. *Pharmaceutical Statistics*, 13(1):41–54, August 2013.

[112] M. J. Mulcahey, Linda A. T. Jones, Frank Rockhold, Rediger Rupp, John L. K. Kramer, Steven Kirshblum, Andrew Blight, Daniel Lammertse, James D. Guest, and John D. Steeves. Adaptive trial designs for spinal cord injury clinical trials directed to the central nervous system. *Spinal Cord*, 58(12):1235–1248, September 2020.

[113] Deborah M. Stein, Jay Menaker, Karen McQuillan, Christopher Handley, Bizhan Aarabi, and Thomas M. Scalea. Risk factors for organ dysfunction and failure in patients with acute traumatic cervical spinal cord injury. *Neurocritical Care*, 13(1):29–39, April 2010.

[114] Brian K. Kwon, Ona Bloom, Ina-Beate Wanner, Armin Curt, Jan M. Schwab, James Fawcett, and Kevin K. Wang. Neurochemical biomarkers in spinal cord injury. *Spinal Cord*, 57(10):819–831, July 2019.

[115] Brian K. Kwon, Femke Streijger, Nader Fallah, Vanessa K. Noonan, Lise M. Bélanger, Leanna Ritchie, Scott J. Paquette, Tamir Ailon, Michael C. Boyd, John Street, Charles G. Fisher, and Marcel F. Dvorak. Cerebrospinal fluid biomarkers to stratify injury severity and predict outcome in human traumatic spinal cord injury. *Journal of Neurotrauma*, 34(3):567–580, February 2017.

[116] Jetan H. Badhiwala, Jefferson R. Wilson, Brian K. Kwon, Steven Casha, and Michael G. Fehlings. A review of clinical trials in spinal cord injury including biomarkers. *Journal of Neurotrauma*, 35(16):1906–1917, August 2018.

[117] Fred H. Geisler, William P. Coleman, Giacinto Grieco, and Devinder Poonian. Recruitment and early treatment in a multicenter study of acute spinal cord injury. *Spine*, 26(Supplement):S58–S67, December 2001.

[118] Fred H. Geisler, William P. Coleman, Giacinto Grieco, and Devinder Poonian. Measurements and recovery patterns in a multicenter study of acute spinal cord injury. *Spine*, 26(Supplement):S68–S86, December 2001.

[119] Steven C. Kirshblum, William Waring, Fin Biering-Sorensen, Stephen P. Burns, Mark Johansen, Mary Schmidt-Read, William Donovan, Daniel E. Graves, Amitabh Jha, Linda Jones, M. J. Mulcahey, and Andrei Krassioukov. Reference for the 2011 revision of the

international standards for neurological classification of spinal cord injury. *The Journal of Spinal Cord Medicine*, 34(6):547–554, November 2011.

[120] Hans-Christoph Pape, Rolf Lefering, Nerida Butcher, Andrew Peitzman, Luke Leenen, Ingo Marzi, Philip Lichte, Christoph Josten, Bertil Bouillon, Uli Schmucker, Philip Stahel, Peter Giannoudis, and Zsolt Balogh. The definition of polytrauma revisited. *Journal of Trauma and Acute Care Surgery*, 77(5):780–786, November 2014.

[121] Anh K. Vo, Fred Geisler, Lukas Grassner, Jan Schwab, Gale Whiteneck, Catherine Jutzeler, and John L. K. Kramer. Serum albumin as a predictor of neurological recovery after spinal cord injury: a replication study. *Spinal Cord*, 59(3):282–290, August 2020.

[122] Gerardo Avila-Martin, Iriana Galan-Arriero, Julio Gómez-Soriano, and Julian Taylor. Treatment of rat spinal cord injury with the neurotrophic factor albumin-oleic acid: Translational application for paralysis, spasticity and pain. *PLoS ONE*, 6(10):e26107, October 2011.

[123] Steve Casha, Tiffany Rice, David P. Stirling, Claudia Silva, Sharmilee Gnanapavan, Gavin Giovannoni, R. John Hurlbert, and V. Wee Yong. Cerebrospinal fluid biomarkers in human spinal cord injury from a phase II minocycline trial. *Journal of Neurotrauma*, 35(16):1918–1928, August 2018.

[124] Senda Ajroud-Driss, Mohammad Saeed, Humaira Khan, Nailah Siddique, W. Y. Hung, Robert Sufit, Scott Heller, Jennifer Armstrong, Pat Casey, Teepu Siddique, and Thomas J. Lukas. Riluzole metabolism and CYP1a1/2 polymorphisms in patients with ALS. *Amyotrophic Lateral Sclerosis*, 8(5):305–309, January 2007.

[125] René Bruno, Nicole Vivier, Guy Montay, Aim éLe Liboux, Larry K. Powe, Jean-Christophe Delumeau, and Gerald R. Rhodes. Population pharmacokinetics of riluzole in patients with amyotrophic lateral sclerosis. *Clinical Pharmacology &amp Therapeutics*, 62(5):518–526, November 1997.

[126] H. J. Nelis and A. P. De Leenheer. Metabolism of minocycline in humans. *Drug Metab Dispos*, 10(2):142–146, 1982.

[127] S. Saivin and G. Houin. Clinical pharmacokinetics of doxycycline and minocycline. *Clinical Pharmacokinetics*, 15(6):355–366, December 1988.

[128] Food, Drug Administration, et al. Drug-induced liver injury: premarketing clinical evaluation. *Guidance for industry*, 2009.

[129] P. B. Watkins, M. Merz, M. I. Avigan, N. Kaplowitz, A. Regev, and J. R. Senior. The clinical liver safety assessment best practices workshop: rationale, goals, accomplishments and the future. *Drug Saf*, 37 Suppl 1(Suppl 1):1–7, Nov 2014.

[130] United States Food and Drug Administration. Guidance for industry: premarketing risk assessment. 2005.

[131] C H Hulme, S J Brown, H R Fuller, J Riddell, A Osman, J Chowdhury, N Kumar, W E Johnson, and K T Wright. The developing landscape of diagnostic and prognostic biomarkers for spinal cord injury in cerebrospinal fluid and blood. *Spinal Cord*, 55(2):114–125, December 2016.

[132] Anthony S. Burns, Bum Suk Lee, John F. Ditunno, and Alan Tessler. Patient selection for clinical trials: The reliability of the early spinal cord injury examination. *Journal of Neurotrauma*, 20(5):477–482, May 2003.

[133] M. B. Bracken, W. F. Collins, D. F. Freeman, M. J. Shepard, F. W. Wagner, R. M. Silten, K. G. Hellenbrand, J. Ransohoff, W. E. Hunt, and P. L. Perot. Efficacy of methylprednisolone in acute spinal cord injury. *JAMA*, 251(1):45–52, Jan 1984.

[134] M. B. Bracken, M. J. Shepard, T. R. Holford, L. Leo-Summers, E. F. Aldrich, M. Fazl, M. Fehlings, D. L. Herr, P. W. Hitchon, L. F. Marshall, R. P. Nockels, V. Pascale, P. L. Perot, J. Piepmeier, V. K. Sonntag, F. Wagner, J. E. Wilberger, H. R. Winn, and W. Young. Administration of methylprednisolone for 24 or 48 hours or tirilazad mesylate for 48 hours in the treatment of acute spinal cord injury. Results of the Third National Acute Spinal Cord Injury Randomized Controlled Trial. National Acute Spinal Cord Injury Study. *JAMA*, 277(20):1597–1604, May 1997.

[135] Krzysztof Gutkowski, Alina Chwist, and Marek Hartleb. Liver injury induced by high-dose methylprednisolone therapy: A case report and brief review of the literature. *Hepatitis Monthly*, 11(8):656–661, August 2011.

[136] Mohamed A. H. Kadle and Natalia V. Mazurchik. Hepatotoxicity induced by high dose of methylprednisolone therapy in a patient with multiple sclerosis: A case report and brief review of literature. *Open Journal of Gastroenterology*, 06(05):146–150, 2016.

[137] Y. Shoenfeld, Y. Gurewich, L.A. Gallant, and J. Pinkhas. Prednisone-induced leukocytosis. *The American Journal of Medicine*, 71(5):773–778, November 1981.

[138] B. Vásárhelyi and L. A. Debreczeni. Lab Test Findings in the Elderly. *EJIFCC*, 28(4):328–332, Dec 2017.

[139] Callum G. Fraser. Age-related changes in laboratory test results. *Drugs &amp Aging*, 3(3):246–257, 1993.

[140] R A Glennie, V K Noonan, N Fallah, S E Park, N P Thorogood, A Cheung, C G Fisher, M F Dvorak, and J T Street. Reliability of the spine adverse events severity system (SAVES) for individuals with traumatic spinal cord injury. *Spinal Cord*, 52(10):758–763, July 2014.

[141] Travis E. Marion, Carly S. Rivers, Dilnur Kurban, Christiana L. Cheng, Nader Fallah, Juliet Batke, Marcel F. Dvorak, Charles G. Fisher, Brian K. Kwon, Vanessa K. Noonan, and John T. Street. Previously identified common post-injury adverse events in traumatic spinal cord injury—validation of existing literature and relation to selected potentially modifiable comorbidities: A prospective canadian cohort study. *Journal of Neurotrauma*, 34(20):2883–2891, October 2017.

[142] J. L. Segal, B. F. Maltby, M. I. Langdorf, R. Jacobson, S. R. Brunnemann, and W. J. Jusko. Methylprednisolone disposition kinetics in patients with acute spinal cord injury. *Pharmacotherapy*, 18(1):16–22, 1998.

[143] Jacquelyn J. Cragg, Jenny Haefeli, Catherine R. Jutzeler, Frank Röhrich, Norbert Weidner, Marion Saur, Doris D. Maier, Yorck B. Kalke, Christian Schuld, Armin Curt, and John K. Kramer. Effects of pain and pain management on motor recovery of spinal cord–injured patients. *Neurorehabilitation and Neural Repair*, 30(8):753–761, July 2016.

[144] Wenjing Sun, Molly J.E. Larson, Conrad M. Kiyoshi, Alexander J. Annett, William A. Stalker, Juan Peng, and Andrea Tedeschi. Gabapentinoid treatment promotes corticospinal plasticity and regeneration following murine spinal cord injury. *Journal of Clinical Investigation*, 130(1):345–358, December 2019.

[145] Andrea Tedeschi, Sebastian Dupraz, Claudia J. Laskowski, Jia Xue, Thomas Ulas, Marc Beyer, Joachim L. Schultze, and Frank Bradke. The calcium channel subunit alpha2delta2 suppresses axon regeneration in the adult CNS. *Neuron*, 92(2):419–434, October 2016.

[146] Jan P. Vandenbroucke, Erik von Elm, Douglas G. Altman, Peter C. Gøtzsche, Cynthia D. Mulrow, Stuart J. Pocock, Charles Poole, James J. Schlesselman, and Matthias Egger. Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. *International Journal of Surgery*, 12(12):1500–1524, December 2014.

[147] Mark N. Hadley and Beverly C. Walters. Introduction to the guidelines for the management of acute cervical spine and spinal cord injuries. *Neurosurgery*, 72(supplement 2):5–16, March 2013.

[148] M. N. Hadley, B. C. Walters, P. A. Grabb, N. M. Oyesiku, G. J. Przybylski, D. K. Resnick, T. C. Ryken, and D. H. Mielke. Guidelines for the management of acute cervical spine and spinal cord injuries. *Clin Neurosurg*, 49:407–498, 2002.

[149] J W Fawcett, A Curt, J D Steeves, W P Coleman, M H Tuszynski, D Lammertse, P F Bartlett, A R Blight, V Dietz, J Ditunno, B H Dobkin, L A Havton, P H Ellaway, M G Fehlings, A Privat, R Grossman, J D Guest, N Kleitman, M Nakamura, M Gaviria, and D Short. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. *Spinal Cord*, 45(3):190–205, December 2006.

[150] J D Steeves, , J K Kramer, J W Fawcett, J Cragg, D P Lammertse, A R Blight, R J Marino, J F Ditunno, W P Coleman, F H Geisler, J Guest, L Jones, S Burns, M Schubert, H J A van Hedel, and A Curt. Extent of spontaneous motor recovery after traumatic cervical sensorimotor complete spinal cord injury. *Spinal Cord*, 49(2):257–265, August 2010.

[151] Amylou C. Dueck, Tito R. Mendoza, Sandra A. Mitchell, Bryce B. Reeve, Kathleen M. Castro, Lauren J. Rogak, Thomas M. Atkinson, Antonia V. Bennett, Andrea M. Denicoff, Ann M. O'Mara, Yuelin Li, Steven B. Clauser, Donna M. Bryant, James D. Bearden, Theresa A. Gillis, Jay K. Harness, Robert D. Siegel, Diane B. Paul, Charles S. Cleeland, Deborah Schrag, Jeff A. Sloan, Amy P. Abernethy, Deborah W. Bruner, Lori M. Minasian, and Ethan Basch. Validity and reliability of the US national cancer institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *JAMA Oncology*, 1(8):1051, November 2015.

[152] *R: A language and environment for statistical computing : Reference index*. 2010.

[153] Elinor Ben-Menachem. Pregabalin pharmacology and its relevance to clinical practice. *Epilepsia*, 45(s6):13–18, August 2004.

[154] D. Wu, Y. S. Kang, U. Bickel, and W. M. Pardridge. Blood-brain barrier permeability to morphine-6-glucuronide is markedly reduced compared with morphine. *Drug Metab Dispos*, 25(6):768–771, Jun 1997.

[155] Kirsten K. Viktil, Hege S. Blix, Tron A. Moger, and Aasmund Reikvam. Polypharmacy as commonly defined is an indicator of limited value in the assessment of drug-related problems. *British Journal of Clinical Pharmacology*, 63(2):187–195, February 2007.

[156] Danijela Gnjidic, Sarah N. Hilmer, Fiona M. Blyth, Vasi Naganathan, Louise Waite, Markus J. Seibel, Andrew J. McLachlan, Robert G. Cumming, David J. Handelsman, and David G. Le Couteur. Polypharmacy cutoff and outcomes: five or more medicines were used to identify

community-dwelling older men at risk of different adverse outcomes. *Journal of Clinical Epidemiology*, 65(9):989–995, September 2012.

[157] Argyrios Stampas, Claudia Pedroza, Jennifer N. Bush, Adam R. Ferguson, John L. Kipling Kramer, and Michelle Hook. The first 24 h: opioid administration in people with spinal cord injury and neurologic recovery. *Spinal Cord*, 58(10):1080–1089, May 2020.

[158] Michelle A. Hook, Sarah A. Woller, Eric Bancroft, Miriam Aceves, Mary Katherine Funk, John Hartman, and Sandra M. Garraway. Neurobiological effects of morphine after spinal cord injury. *Journal of Neurotrauma*, 34(3):632–644, February 2017.

[159] Michelle A. Hook, Georgina Moreno, Sarah Woller, Denise Puga, Kevin Hoy, Robyn Balden, and James W. Grau. Intrathecal morphine attenuates recovery of function after a spinal cord injury. *Journal of Neurotrauma*, 26(5):741–752, May 2009.

[160] Michael B. Bracken, Mary Jo Shepard, William F. Collins, Theodore R. Holford, David S. Baskin, Howard M. Eisenberg, Eugene Flamm, Linda Leo-Summers, Joseph C. Maroon, Lawrence F. Marshall, Phanor L. Perot, Joseph Piepmeier, Volker K. H. Sonntag, Franklin C. Wagner, James L. Wilberger, H. Richard Winn, and Wise Young. Methylprednisolone or naloxone treatment after acute spinal cord injury: 1-year follow-up data. *Journal of Neurosurgery*, 76(1):23–31, January 1992.

[161] C. Michael Goplen, Wesley Verbeek, Sung Hyun Kang, C. Allyson Jones, Donald C. Voaklander, Thomas A. Churchill, and Lauren A. Beaupre. Preoperative opioid use is associated with worse patient outcomes after total joint arthroplasty: a systematic review and meta-analysis. *BMC Musculoskeletal Disorders*, 20(1), May 2019.

[162] H Mestre, T Alkon, S Salazar, and A Ibarra. Spinal cord injury sequelae alter drug pharmacokinetics: an overview. *Spinal Cord*, 49(9):955–960, May 2011.

[163] P García-López, A Martínez-Cruz, G Guízar-Sahagún, and G Castañeda-Hernández. Acute spinal cord injury changes the disposition of some, but not all drugs given intravenously. *Spinal Cord*, 45(9):603–608, December 2006.

[164] Jack L. Segal, Sherry R. Brunnemann, Ibrahim M. Eltorai, and Michael Vulpe. Decreased systemic clearance of lorazepam in humans with spinal cord injury. *The Journal of Clinical Pharmacology*, 31(7):651–656, July 1991.

[165] Antonio Ibarra, Gabriel Guízar-Sahagún, Dolores Correa, Roberto Kretschmer, Israel Grijalva, Francisco J. Flores-Murrieta, Gilberto Castañeda-Hernández, Alberto Odor, Rosa M.

López, Rebecca Franco-Bourland, Ana L. Espitia, Hermelinda Salgado-Ceballos, and Ignacio Madrazo. Alteration of cyclosporin-a pharmacokinetics after experimental spinal cord injury. *Journal of Neurotrauma*, 13(5):267–272, May 1996.

[166] Lucie Bourguignon, Bobo Tong, Fred Geisler, Martin Schubert, Frank Röhrich, Marion Saur, Norbert Weidner, Rüdiger Rupp, Yorck-Bernhard B. Kalke, Rainer Abel, Doris Maier, Lukas Grassner, Harvinder S. Chhabra, Thomas Liebscher, Jacquelyn J. Cragg, John Kramer, Armin Curt, and Catherine R. Jutzeler and. International surveillance study in acute spinal cord injury confirms viability of multinational clinical trials. *BMC Medicine*, 20(1), June 2022.

[167] Gary McLean, John V. Hindle, Bruce Guthrie, and Stewart W. Mercer. Co-morbidity and polypharmacy in parkinson's disease: insights from a large scottish primary care database. *BMC Neurology*, 17(1), July 2017.

[168] Seyed-Mohammad Fereshtehnejad, Kristina Johnell, and Maria Eriksdotter. Anti-dementia drugs and co-medication among patients with alzheimer's disease. *Drugs &amp Aging*, 31(3):215–224, February 2014.

[169] Joanie M Thelen, Sharon G Lynch, Amanda S Bruce, Laura M Hancock, and Jared M Bruce. Polypharmacy in multiple sclerosis: Relationship with fatigue, perceived cognition, and objective cognitive performance. *Journal of Psychosomatic Research*, 76(5):400–404, May 2014.

[170] Giorgia Cosano, Manuela Giangreco, Silvia Ussai, Tullio Giorgini, Emanuele Biasutti, Fabio Barbone, and Federica Edith Pisa. Polypharmacy and the use of medications in inpatients with acquired brain injury during post-acute rehabilitation: A cross-sectional study. *Brain Injury*, 30(3):353–362, February 2016.

[171] Samir H Haddad and Yaseen M Arabi. Critical care management of severe traumatic brain injury in adults. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 20(1):12, 2012.

[172] Katharina A. Kierner, Dietmar Weixler, Eva K. Masel, Verena Gartner, and Herbert H. Watzke. Polypharmacy in the terminal stage of cancer. *Supportive Care in Cancer*, 24(5):2067–2074, November 2015.

[173] T.A. Mcallister. Polypharmacy in sepsis. *The Lancet*, 301(7818):1520, June 1973.

[174] Elio Scarpini, Daniela Galimberti, and Laura Ghezzi. Disease-modifying drugs in alzheimer&amp#39s disease. *Drug Design, Development and Therapy*, page 1471, December 2013.

[175] J. Cummings and N. Fox. Defining disease modifying therapy for alzheimer's disease. *The Journal of Prevention of Alzheimer's Disease*, pages 1–7, 2017.

[176] Anthony E. Lang and Alberto J. Espay. Disease modification in parkinson's disease: Current approaches, challenges, and future considerations. *Movement Disorders*, 33(5):660–677, April 2018.

[177] T. Schulenborg, O. Schmidt, A. van Hall, H. E. Meyer, M. Hamacher, and K. Marcus. Proteomics in neurodegeneration – disease driven approaches. *Journal of Neural Transmission*, 113(8):1055–1073, July 2006.

[178] Ting Tian, Sensen Zhang, and Maojun Yang. Recent progress and challenges in the treatment of spinal cord injury. *Protein amp; Cell*, 14(9):635–652, February 2023.

[179] Jayne Donovan and Steven Kirshblum. Clinical trials in traumatic spinal cord injury. *Neurotherapeutics*, 15(3):654–668, July 2018.

[180] Catherine R. Jutzeler, Lucie Bourguignon, Bobo Tong, Elias Ronca, Eric Bailey, Noam Y. Harel, Fred Geisler, Adam R. Ferguson, Brian K. Kwon, Jacquelyn J. Cragg, Lukas Grassner, and John L. K. Kramer. Pharmacological management of acute spinal cord injury: a longitudinal multi-cohort observational study. *Scientific Reports*, 13(1), April 2023.

[181] Lauren Cadel, Amanda C. Everall, Sander L. Hitzig, Tanya L. Packer, Tejal Patel, Aisha Lofters, and Sara J. T. Guilcher. Spinal cord injury and polypharmacy: a scoping review. *Disability and Rehabilitation*, 42(26):3858–3870, May 2019.

[182] Nurdan Paker, Yelda Soluk Özdemir, Derya Buğdaycı, Berna undefinedelik, and Yasemin Bölükbaş. Prevalence and predictors of polypharmacy among community-based individuals with traumatic spinal cord injury. *The Journal of Spinal Cord Medicine*, 46(6):958–963, December 2021.

[183] Di Zhang, Jun Xuan, Bin-bin Zheng, Yu-long Zhou, Yan Lin, Yao-sen Wu, Yi-fei Zhou, Yi-xing Huang, Quan Wang, Li-yan Shen, Cong Mao, Yan Wu, Xiang-yang Wang, Nai-feng Tian, Hua-Zi Xu, and Xiao-lei Zhang. Metformin improves functional recovery after spinal cord injury via autophagy flux stimulation. *Molecular Neurobiology*, 54(5):3327–3341, May 2016.

[184] Mark J. Lambrechts and James L. Cook. Nonsteroidal anti-inflammatory drugs and their neuroprotective role after an acute spinal cord injury: A systematic review of animal models. *Global Spine Journal*, 11(3):365–377, January 2020.

[185] Xingxing Wang, Stephane Budel, Kenneth Baughman, Grahame Gould, Kang-Ho Song, and Stephen M. Strittmatter. Ibuprofen enhances recovery from spinal cord injury by limiting tissue loss and stimulating axonal growth. *Journal of Neurotrauma*, 26(1):81–95, January 2009.

[186] Haoli Wang, Zhilong Zheng, Wen Han, Yuan Yuan, Yao Li, Kailiang Zhou, Qingqing Wang, Ling Xie, Ke Xu, Hongyu Zhang, Huazi Xu, Yanqing Wu, and Jian Xiao. Metformin promotes axon regeneration after spinal cord injury through inhibiting oxidative stress and stabilizing microtubule. *Oxidative Medicine and Cellular Longevity*, 2020:1–20, January 2020.

[187] Michael B. Bracken, Mary Jo Shepard, Karen G. Hellenbrand, William F. Collins, Linda S. Leo, Daniel F. Freeman, Franklin C. Wagner, Eugene S. Flamm, Howard M. Eisenberg, Joseph H. Goodman, Phanor L. Perot, Barth A. Green, Robert G. Grossman, John N. Meagher, Wise Young, Boguslav Fischer, Guy L. Clifton, William E. Hunt, and Nathan Rifkinson. Methylprednisolone and neurological function 1 year after spinal cord injury: Results of the national acute spinal cord injury study. *Journal of Neurosurgery*, 63(5):704–713, November 1985.

[188] Jose A. Canseco, Brian A. Karamian, Daniel R. Bowles, Michael P. Markowitz, Stephen L. DiMaria, Nicholas C. Semenza, Mark R. Leibensperger, Michael L. Smith, and Alexander R. Vaccaro. Updated review: The steroid controversy for management of spinal cord injury. *World Neurosurgery*, 150:1–8, June 2021.

[189] L. Shamseer, D. Moher, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: elaboration and explanation. *BMJ*, 349(jan02 1):g7647–g7647, January 2015.

[190] Anthony J Viera and Joanne M Garrett. Understanding interobserver agreement: the kappa statistic. *Fam. Med.*, 37(5):360–363, May 2005.

[191] Jonathan A C Sterne, Jelena Savović, Matthew J Page, Roy G Elbers, Natalie S Blencowe, Isabelle Boutron, Christopher J Cates, Hung-Yuan Cheng, Mark S Corbett, Sandra M Eldridge, Jonathan R Emberson, Miguel A Hernán, Sally Hopewell, Asbjørn Hróbjartsson, Daniela R Junqueira, Peter Jüni, Jamie J Kirkham, Toby Lasserson, Tianjing Li, Alexandra McAleenan, Barnaby C Reeves, Sasha Shepperd, Ian Shrier, Lesley A Stewart, Kate Tilling, Ian R White, Penny F Whiting, and Julian P T Higgins. Rob 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, page l4898, August 2019.

[192] Jonathan AC Sterne, Miguel A Hernán, Barnaby C Reeves, Jelena Savović, Nancy D Berkman, Meera Viswanathan, David Henry, Douglas G Altman, Mohammed T Ansari, Isabelle

Boutron, James R Carpenter, An-Wen Chan, Rachel Churchill, Jonathan J Deeks, Asbjørn Hróbjartsson, Jamie Kirkham, Peter Jüni, Yoon K Loke, Theresa D Pigott, Craig R Ramsay, Deborah Regidor, Hannah R Rothstein, Lakhbir Sandhu, Pasqualina L Santaguida, Holger J Schünemann, Beverly Shea, Ian Shrier, Peter Tugwell, Lucy Turner, Jeffrey C Valentine, Hugh Waddington, Elizabeth Waters, George A Wells, Penny F Whiting, and Julian PT Higgins. Robins-i: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, page i4919, October 2016.

[193] Luke A. McGuinness and Julian P. T. Higgins. Risk-of-bias visualization (robvis): An r package and shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, 12(1):55–61, May 2020.

[194] N. Eric Naftchi. Treatment of mammalian spinal cord injury with antioxidants. *International Journal of Developmental Neuroscience*, 9(2):113–126, January 1991.

[195] Ramazan Kahveci, Emre Cemal Gökçe, Bora Gürer, Aysun Gökçe, Uçler Kisa, Duran Berker Cemil, Mustafa Fevzi Sargon, Fatih Ozan Kahveci, Nurkan Aksoy, and Bülent Erdoğan. Neuroprotective effects of rosuvastatin against traumatic spinal cord injury in rats. *European Journal of Pharmacology*, 741:45–54, October 2014.

[196] Brian K. Kwon, Josee Roy, Jae H.T. Lee, Elena Okon, Hongbin Zhang, Jeffrey C. Marx, and Mark S. Kindy. Magnesium chloride in a polyethylene glycol formulation as a neuroprotective therapy for acute spinal cord injury: Preclinical refinement and optimization. *Journal of Neurotrauma*, 26(8):1379–1393, August 2009.

[197] Femke Streijger, Jae H.T. Lee, Neda Manouchehri, Elena B. Okon, Seth Tigchelaar, Lisa M. Anderson, Greg A. Dekaban, David A. Rudko, Ravi S. Menon, Jennifer F. Iaci, Donald C. Button, Andrea M. Vecchione, Andrey Konovalov, Patrick D. Sarmiere, Chi Ung, Anthony O. Caggiano, and Brian K. Kwon. The evaluation of magnesium chloride within a polyethylene glycol formulation in a porcine model of acute spinal cord injury. *Journal of Neurotrauma*, 33(24):2202–2216, December 2016.

[198] P. S. Halt, R. A. Swanson, and A. I. Faden. Alcohol exacerbates behavioral and neurochemical effects of rat spinal cord trauma. *Archives of Neurology*, 49(11):1178–1184, November 1992.

[199] Alex S. Rivlin and Charles H. Tator. Effect of vasodilators and myelotomy on recovery after acute spinal cord injury in rats. *Journal of Neurosurgery*, 50(3):349–352, March 1979.

[200] T Cheriyan, D J Ryan, J H Weinreb, J Cheriyan, J C Paul, V Lafage, T Kirsch, and T J Errico. Spinal cord injury models: a review. *Spinal Cord*, 52(8):588–595, June 2014.

[201] Sandrine Lopez, Alain Privat, Nathalie Bernard, Freddy Ohanna, Christine Vergnes, and Xavier Capdevila. Intrathecal bupivacaine protects against extension of lesions in an acute photochemical spinal cord injury model. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 51(4):364–372, April 2004.

[202] J.-X. Hao, X.-J. Xu, H. Aldskogius, Å. Seiger, and Z. Wiesenfeld-Hallin. Beneficial effect of the opioid receptor antagonist naltrexone on hyper-sensitivity induced by spinal cord ischemia in rats: disassociation with mk-801. *Restorative Neurology and Neuroscience*, 3(5):257–266, 1991.

[203] Joseph R. Madsen, Paul MacDonald, Nina Irwin, David E. Goldberg, Gui-Lan Yao, Karina F. Meiri, Ilonna J. Rimm, Philip E. Stieg, and Larry I. Benowitz. Tacrolimus (fk506) increases neuronal expression of gap-43 and improves functional recovery after spinal cord injury in rats. *Experimental Neurology*, 154(2):673–683, December 1998.

[204] Min-Sheng Piao, Jung-Kil Lee, Jae-Won Jang, Hyuk Hur, Shin-Seok Lee, LuWei Xiao, and Hyung-Seok Kim. Melatonin improves functional outcome via inhibition of matrix metalloproteinases-9 after photothrombotic spinal cord injury in rats. *Acta Neurochirurgica*, 156(11):2173–2182, May 2014.

[205] Mina Afhami, Fatemeh Abbaszadeh, Elham Saghaei, Kobra Naseri, Mohammad Javan, and Masoumeh Jorjani. The demyelination and altered motor performance following electrolytic lesion in the ventrolateral white matter of spinal cord in male rats: Benefit of post-injury administration of estradiol. 20:157–171, 09 2016.

[206] D. MICHELE BASSO, MICHAEL S. BEATTIE, and JACQUELINE C. BRESNAHAN. A sensitive and reliable locomotor rating scale for open field testing in rats. *Journal of Neurotrauma*, 12(1):1–21, February 1995.

[207] D. Michele Basso, Lesley C. Fisher, Aileen J. Anderson, Lyn B. Jakeman, Dana M. Mctigue, and Phillip G. Popovich. Basso mouse scale for locomotion detects differences in recovery after spinal cord injury in five common mouse strains. *Journal of Neurotrauma*, 23(5):635–659, May 2006.

[208] Rachel B. Song, D. Michele Basso, Ronaldo C. da Costa, Lesley C. Fisher, Xiaokui Mo, and Sarah A. Moore. Adaptation of the basso–beattie–bresnahan locomotor rating scale for use in a clinical model of spinal cord injury in dogs. *Journal of Neuroscience Methods*, 268:117–124, August 2016.

[209] Kristan A. Leech, Catherine R. Kinnaird, and T. George Hornby. Effects of serotonergic medications on locomotor performance in humans with incomplete spinal cord injury. *Journal of Neurotrauma*, 31(15):1334–1342, August 2014.

[210] Nathan Evaniew, Vanessa K. Noonan, Nader Fallah, Brian K. Kwon, Carly S. Rivers, Henry Ahn, Christopher S. Bailey, Sean D. Christie, Daryl R. Fourney, R. John Hurlbert, A.G. Linassi, Michael G. Fehlings, and Marcel F. Dvorak. Methylprednisolone for the treatment of patients with acute spinal cord injuries: A propensity score-matched cohort study from a canadian multi-center spinal cord injury registry. *Journal of Neurotrauma*, 32(21):1674–1683, November 2015.

[211] Ha Seong Kim, Kil-Byung Lim, Jiyong Kim, Joongmo Kang, Hojin Lee, Sang Wan Lee, and Jeehyun Yoo. Epidemiology of spinal cord injury: Changes to its cause amid aging population, a single center study. *Annals of Rehabilitation Medicine*, 45(1):7–15, February 2021.

[212] Yasuo Ito, Yoshihisa Sugimoto, Masao Tomioka, Nobuo Kai, and Masato Tanaka. Does high dose methylprednisolone sodium succinate really improve neurological status in patient with acute cervical cord injury?: a prospective study about neurological recovery and early complications. *Spine*, 34(20):2121–2124, September 2009.

[213] Guang-Zhi Ning, Qiang Wu, Yu-Lin Li, and Shi-Qing Feng. Epidemiology of traumatic spinal cord injury in asia: a systematic review. *J. Spinal Cord Med.*, 35(4):229–239, July 2012.

[214] Vafa Rahimi-Movaghar, Mohammad Kazem Sayyah, Hesam Akbari, Reza Khorramirouz, Mohammad R Rasouli, Maziar Moradi-Lakeh, Farhad Shokraneh, and Alexander R Vaccaro. Epidemiology of traumatic spinal cord injury in developing countries: a systematic review. *Neuroepidemiology*, 41(2):65–85, June 2013.

[215] Christopher K Thompson and T George Hornby. Divergent modulation of clinical measures of volitional and reflexive motor behaviors following serotonergic medications in human incomplete spinal cord injury. *J. Neurotrauma*, 30(6):498–502, March 2013.

[216] M B Bracken and T R Holford. Effects of timing of methylprednisolone or naloxone administration on recovery of segmental and long-tract neurological function in NASCIS 2. *J. Neurosurg.*, 79(4):500–507, October 1993.

[217] Steven Casha, David Zygun, M Dan McGowan, Ish Bains, V Wee Yong, and R John Hurlbert. Results of a phase II placebo-controlled randomized trial of minocycline in acute spinal cord injury. *Brain*, 135(Pt 4):1224–1236, April 2012.

[218] Oliver Maric, Björn Zörner, and Volker Dietz. Levodopa therapy in incomplete spinal cord injury. *J. Neurotrauma*, 25(11):1303–1307, November 2008.

[219] Mary J Clark, Gregory F Petroski, Micah O Mazurek, Kristofer J Hagglund, Ashley K Sherman, Andrew B Lammy, Martin K Childers, and Michael E Acuff. Testosterone replacement

therapy and motor function in men with spinal cord injury: a retrospective analysis. *Am. J. Phys. Med. Rehabil.*, 87(4):281–284, April 2008.

[220] Bahram Aminmansour, Ali Asnaashari, Majid Rezvani, Fariborz Ghaffarpasand, Seyed Mohammad Amin Noorian, Masih Saboori, and Parisa Abdollahzadeh. Effects of progesterone and vitamin D on outcome of patients with acute traumatic spinal cord injury; a randomized, double-blind, placebo controlled study. *J. Spinal Cord Med.*, 39(3):272–280, May 2016.

[221] Jacob E Sunshine, Armagan Dagal, Stephen P Burns, Richard J Bransford, Fangyi Zhang, Shu-Fang Newman, Bala G Nair, and Sam R Sharar. Methylprednisolone therapy in acute traumatic spinal cord injury: Analysis of a regional spinal cord model systems database. *Anesth. Analg.*, 124(4):1200–1205, April 2017.

[222] Peter Felleiter, Nicole Müller, Frederik Schumann, Olga Felix, and Peter Lierz. Changes in the use of the methylprednisolone protocol for traumatic spinal cord injury in switzerland. *Spine*, 37(11):953–956, May 2012.

[223] Satoshi Tsutsumi, Takayoshi Ueta, Keiichiro Shiba, Shunsaku Yamamoto, and Kenji Takagishi. Effects of the second national acute spinal cord injury study of high-dose methylprednisolone therapy on acute cervical spinal cord injury-results in spinal injuries center. *Spine*, 31(26):2992–6; discussion 2997, December 2006.

[224] Osman Arikan Nacar, Hakan Eroglu, Nuri Eralp Cetinalp, Guner Menekse, Ali Erdem Yildirim, Ozhan Merzuk Uckun, Ergun Daglioglu, Omer Faruk Turkoglu, and Ahmet Deniz Belen. Systemic administration of atorvastatin improves locomotor functions and hyperacute-acute response after experimental spinal cord injury: an ultrastructural and biochemical analysis. *Turk. Neurosurg.*, 24(3):337–343, 2014.

[225] Duo Zhang, Fang Wang, Xu Zhai, Xiao-Hui Li, and Xi-Jing He. Lithium promotes recovery of neurological function after spinal cord injury by inducing autophagy. *Neural Regeneration Res.*, 13(12):2191–2199, December 2018.

[226] Hai-Hu Hao, Li Wang, Zhi-Jian Guo, Lang Bai, Rui-Ping Zhang, Wei-Bing Shuang, Yi-Jia Jia, Jie Wang, Xiao-Yu Li, and Qiang Liu. Valproic acid reduces autophagy and promotes functional recovery after spinal cord injury in rats. *Neurosci. Bull.*, 29(4):484–492, August 2013.

[227] Yan Zhang, Zongjian Liu, Wenxiu Zhang, Qichao Wu, Yanjun Zhang, Yadong Liu, Yun Guan, and Xueming Chen. Melatonin improves functional recovery in female rats after acute spinal cord injury by modulating polarization of spinal microglial/macrophages. *J. Neurosci. Res.*, 97(7):733–743, July 2019.

[228] Sarah A Woller. *Analgesia Or Addiction: Implications for Morphine Use After Spinal Cord Injury*. Texas A & M University, 2012.

[229] A I Faden, T P Jacobs, and J W Holaday. Opiate antagonist improves neurologic recovery after spinal injury. *Science*, 211(4481):493–494, January 1981.

[230] A I Faden, T P Jacobs, and J W Holaday. Comparison of early and late naloxone treatment in experimental spinal injury. *Neurology*, 32(6):677–681, June 1982.

[231] G Hubsher, M Haider, and M S Okun. Amantadine: the journey from fighting flu to treating parkinson disease. *Neurology*, 78(14):1096–1099, April 2012.

[232] Anne Corbett, James Pickett, Alistair Burns, Jonathan Corcoran, Stephen B Dunnett, Paul Edison, Jim J Hagan, Clive Holmes, Emma Jones, Cornelius Katona, Ian Kearns, Patrick Kehoe, Amrit Mudher, Anthony Passmore, Nicola Shepherd, Frank Walsh, and Clive Ballard. Drug repositioning for alzheimer's disease. *Nat. Rev. Drug Discov.*, 11(11):833–846, November 2012.

[233] Lilian Beatriz Aguayo Rojas and Marilia Brito Gomes. Metformin: an old but still the best treatment for type 2 diabetes. *Diabetol. Metab. Syndr.*, 5(1):6, February 2013.

[234] Yue Guo, Fang Wang, Haopeng Li, Hui Liang, Yuhuan Li, Zhengchao Gao, and Xijing He. Metformin protects against spinal cord injury by regulating autophagy via the mTOR signaling pathway. *Neurochem. Res.*, 43(5):1111–1117, May 2018.

[235] Tao Zhang, Fang Wang, Kang Li, Chengwei Lv, Kai Gao, and Chaoliang Lv. Therapeutic effect of metformin on inflammation and apoptosis after spinal cord injury in rats through the Wnt/β-catenin signaling pathway. *Neurosci. Lett.*, 739:135440, November 2020.

[236] Michelle A Hook, Grace T Liu, Stephanie N Washburn, Adam R Ferguson, Anne C Bopp, John R Huie, and James W Grau. The impact of morphine after a spinal cord injury. *Behav. Brain Res.*, 179(2):281–293, May 2007.

[237] Michelle A Hook, Stephanie N Washburn, Georgina Moreno, Sarah A Woller, Denise Puga, Kuan H Lee, and James W Grau. An IL-1 receptor antagonist blocks a morphine-induced attenuation of locomotor recovery after spinal cord injury. *Brain Behav. Immun.*, 25(2):349–359, February 2011.

[238] Nicola Nosengo. Can you teach old drugs new tricks? *Nature*, 534(7607):314–316, June 2016.

[239] Brian K Kwon, Lesley J J Soril, Mark Bacon, Michael S Beattie, Armin Blesch, Jacqueline C Bresnahan, Mary Bartlett Bunge, Sarah A Dunlop, Michael G Fehlings, Adam R Ferguson,

Caitlin E Hill, Soheila Karimi-Abdolrezaee, Paul Lu, John W McDonald, Hans W Müller, Martin Oudega, Ephron S Rosenzweig, Paul J Reier, Jerry Silver, Eva Sykova, Xiao-Ming Xu, James D Guest, and Wolfram Tetzlaff. Demonstrating efficacy in preclinical studies of cellular therapies for spinal cord injury - how much is enough? *Exp. Neurol.*, 248:30–44, October 2013.

[240] Phillip G Popovich, Stanley Lemeshow, John C Gensel, and C Amy Tovar. Independent evaluation of the effects of glibenclamide on reducing progressive hemorrhagic necrosis after cervical spinal cord injury. *Exp. Neurol.*, 233(2):615–622, February 2012.

[241] Nelson Adami Andreollo, Elisvânia Freitas dos Santos, Marina Rachel Araújo, and Luiz Roberto Lopes. Rat's age versus human's age: what is the relationship? *Arq. Bras. Cir. Dig.*, 25(1):49–51, 2012.

[242] Sulagna Dutta and Pallav Sengupta. Men and mice: Relating their ages. *Life Sci.*, 152:244–248, May 2016.

[243] Rix Cambridge. Modeling the age difference between humans and felis catus biological cat age: The actual age of a cat. *J. Anim. Res. Nutr.*, 02(02), 2017.

[244] Pallav Sengupta and Sulagna Dutta. Mapping the age of laboratory rabbit strains to human. *Int. J. Prev. Med.*, 11:194, December 2020.

[245] James Guest, Nilanjana Datta, George Jimsheleishvili, and David R Gater, Jr. Pathophysiology, classification and comorbidities after traumatic spinal cord injury. *J Pers Med*, 12(7), July 2022.

[246] T Ueno, Y Ohori, J Ito, S Hoshikawa, S Yamamoto, K Nakamura, S Tanaka, M Akai, Y Tobimatsu, and T Ogata. Hyperphosphorylated neurofilament NF-H as a biomarker of the efficacy of minocycline therapy for spinal cord injury. *Spinal Cord*, 49(3):333–336, March 2011.

[247] Jee Y Lee, Hwang S Kim, Hye Y Choi, Tae H Oh, and Tae Y Yune. Fluoxetine inhibits matrix metalloprotease activation and prevents disruption of blood-spinal cord barrier after spinal cord injury. *Brain*, 135(Pt 8):2375–2389, August 2012.

[248] Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, December 1986.

[249] Tammy Jiang, Jaimie L. Gradus, and Anthony J. Rosellini. Supervised machine learning: A brief primer. *Behavior Therapy*, 51(5):675–687, September 2020.

[250] Teus H. Kappen, Wilton A. van Klei, Leo van Wolfswinkel, Cor J. Kalkman, Yvonne Vergouwe, and Karel G. M. Moons. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagnostic and Prognostic Research*, 2(1), June 2018.

[251] Sharon E Davis, Robert A Greevy, Thomas A Lasko, Colin G Walsh, and Michael E Matheny. Comparison of prediction model performance updating protocols: Using a data-driven testing procedure to guide updating. *AMIA Annu. Symp. Proc.*, 2019:1002–1010, 2019.

[252] Jenny Yang, Andrew A. S. Soltan, and David A. Clifton. Machine learning generalizability across healthcare settings: insights from multi-site covid-19 screening. *npj Digital Medicine*, 5(1), June 2022.

[253] Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machine learning. 2017.

[254] Mohor Biplab Sengupta, Mahashweta Basu, Sourav Iswarari, Kiran Kumar Mukhopadhyay, Krishna Pada Sardar, Biplab Acharyya, Pradeep K. Mohanty, and Debashis Mukhopadhyay. Csf proteomics of secondary phase spinal cord injury in human subjects: Perturbed molecular pathways post injury. *PLoS ONE*, 9(10):e110885, October 2014.

[255] Ahmed A. Albayar, Abigail Roche, Przemyslaw Swiatkowski, Sarah Antar, Nouran Ouda, Eman Emara, Douglas H. Smith, Ali K. Ozturk, and Basem I. Awad. Biomarkers in spinal cord injury: Prognostic insights and future potentials. *Frontiers in Neurology*, 10, January 2019.

[256] Sophie Stukas, Jennifer Cooper, Jasmine Gill, Nader Fallah, Michael A. Skinnider, Lise Belanger, Leanna Ritchie, Angela Tsang, Kevin Dong, Femke Streijger, John Street, Scott Paquette, Tamir Ailon, Nicolas Dea, Raphaele Charest-Morin, Charles G. Fisher, Christopher S. Bailey, Sanjay Dhall, Jean-Marc Mac-Thiong, Jefferson R. Wilson, Sean Christie, Marcel F. Dvorak, Cheryl L. Wellington, and Brian K. Kwon. Association of csf and serum neurofilament light and glial fibrillary acidic protein, injury severity, and outcome in spinal cord injury. *Neurology*, 100(12), March 2023.

[257] Iris Leister, Barbara Altendorfer, Doris Maier, Orpheus Mach, Christof Wutte, Andreas Grillhösl, Angel Arevalo-Martin, Daniel Garcia-Ovejero, Ludwig Aigner, and Lukas Grassner. Serum levels of glial fibrillary acidic protein and neurofilament light protein are related to the neurological impairment and spinal edema after traumatic spinal cord injury. *Journal of Neurotrauma*, 38(24):3431–3439, December 2021.

[258] Lucie Bourguignon, Anh Khoa Vo, Bobo Tong, Fred Geisler, Orpheus Mach, Doris Maier, John L.K. Kramer, Lukas Grassner, and Catherine R. Jutzeler. Natural progression of routine

laboratory markers after spinal trauma: A longitudinal, multi-cohort study. *Journal of Neurotrauma*, May 2021.

[259] Iris Leister, Lukas D. Linde, Anh Khoa Vo, Thomas Haider, Georg Mattiassich, Lukas Grassner, Wolfgang Schaden, Herbert Resch, Catherine R. Jutzeler, Fred H. Geisler, John L. K. Kramer, and Ludwig Aigner. Routine blood chemistry predicts functional recovery after traumatic spinal cord injury: A post hoc analysis. *Neurorehabilitation and Neural Repair*, 35(4):321–333, February 2021.

[260] John F. Ditunno, Hugues Barbeau, Bruce H. Dobkin, Robert Elashoff, Susan Harkema, Ralph J. Marino, Walter W. Hauck, David Apple, D. Michele Basso, Andrea Behrman, Daniel Deforge, Lisa Fugate, Michael Saulino, Michael Scott, and Joanie Chung. Validity of the walking scale for spinal cord injury and other domains of function in a multicenter clinical trial. *Neurorehabilitation and Neural Repair*, 21(6):539–550, March 2007.

[261] Lucie Bourguignon, Louis P. Lukas, James D. Guest, Fred H. Geisler, Vanessa Noonan, Armin Curt, Sarah C. Brüningk, and Catherine R. Jutzeler. Studying missingness in spinal cord injury data: challenges and impact of data imputation. *BMC Medical Research Methodology*, 24(1), January 2024.

[262] Arsalan Alizadeh, Scott Matthew Dyck, and Soheila Karimi-Abdolrezaee. Traumatic spinal cord injury: An overview of pathophysiology, models and acute injury mechanisms. *Frontiers in Neurology*, 10, March 2019.

[263] M M Mukaka. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.*, 24(3):69–71, September 2012.

[264] Shivani Gupta and Atul Gupta. Dealing with noise problem in machine learning data-sets: A systematic review. *Procedia Computer Science*, 161:466–474, 2019.

[265] Xingquan Zhu and Xindong Wu. Class noise vs. attribute noise: A quantitative study. *Artificial Intelligence Review*, 22(3):177–210, November 2004.

[266] Joshua Chuah, Uwe Kruger, Ge Wang, Pingkun Yan, and Juergen Hahn. Framework for testing robustness of machine learning-based classifiers. *Journal of Personalized Medicine*, 12(8):1314, August 2022.

[267] Joel L. Horowitz and Charles F. Manski. Identification and robustness with contaminated and corrupted data. *Econometrica*, 63(2):281, March 1995.

[268] Akira Komori, Hiroki Iriyama, Takako Kainoh, Makoto Aoki, Toshio Naito, and Toshikazu Abe. The impact of infection complications after trauma differs according to trauma severity. *Scientific Reports*, 11(1), July 2021.

[269] Jineta Banerjee, Jaclyn N. Taroni, Robert J. Allaway, Deepashree Venkatesh Prasad, Justin Guinney, and Casey Greene. Machine learning in rare disease. *Nature Methods*, 20(6):803–814, May 2023.

[270] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty. Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8):1509–1515, November 2004.

[271] G Scivoletto, F Tamburella, L Laurenza, and M Molinari. Distribution-based estimates of clinically significant changes in the international standards for neurological classification of spinal cord injury motor and sensory scores. *Eur. J. Phys. Rehabil. Med.*, 49(3):373–384, June 2013.

[272] Donald B. Rubin. Inference and missing data. *Biometrika*, 63(3):581–592, 1976.

[273] Daniel A. Newman. Missing data: Five practical guidelines. *Organizational Research Methods*, 17(4):372–411, September 2014.

[274] Rachael A Hughes, Jon Heron, Jonathan A C Sterne, and Kate Tilling. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *International Journal of Epidemiology*, 48(4):1294–1304, March 2019.

[275] James D Dziura, Lori A Post, Qing Zhao, Zhixuan Fu, and Peter Peduzzi. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J. Biol. Med.*, 86(3):343–358, September 2013.

[276] Roderick Little and Donald Rubin. *Statistical Analysis with Missing Data, Third Edition*. Wiley, April 2019.

[277] Ana-Maria Simundic. Bias in research. *Biochemia Medica*, page 12–15, 2013.

[278] Alma Pedersen, Ellen Mikkelsen, Deirdre Cronin-Fenton, Nickolaj Kristensen, Tra My Pham, Lars Pedersen, and Irene Petersen. Missing data and multiple imputation in clinical epidemiological research. *Clinical Epidemiology*, Volume 9:157–166, March 2017.

[279] Jiang Li, Xiaowei S. Yan, Durgesh Chaudhary, Venkatesh Avula, Satish Mudiganti, Hannah Husby, Shima Shahjouei, Ardavan Afshar, Walter F. Stewart, Mohammed Yeasin, Ramin Zand, and Vida Abedi. Imputation of missing values for electronic health record laboratory data. *npj Digital Medicine*, 4(1), October 2021.

[280] Andrew C. Smith, Stephanie R. Albin, Denise R. O'Dell, Jeffrey C. Berliner, David Dungan, Mitch Sevigny, Christina Draganich, James M. Elliott, and Kenneth A. Weber II. Axial

mri biomarkers of spinal cord damage to predict future walking and motor function: a retrospective study. *Spinal Cord*, 59(6):693–699, October 2020.

[281] Marianne Riksheim Stavseth, Thomas Clausen, and Jo Røislien. How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data. *SAGE Open Medicine*, 7:205031211882291, January 2019.

[282] Marc H. Gorelick. Bias arising from missing data in predictive models. *Journal of Clinical Epidemiology*, 59(10):1115–1123, October 2006.

[283] Peng Li, Elizabeth A. Stuart, and David B. Allison. Multiple imputation: A flexible tool for handling missing data. *JAMA*, 314(18):1966, November 2015.

[284] Marjan Javanbakht, Johnny Lin, Amy Ragsdale, Soyeon Kim, Suzanne Siminski, and Pamina Gorbach. Comparing single and multiple imputation strategies for harmonizing substance use data across hiv-related cohort studies. *BMC Medical Research Methodology*, 22(1), April 2022.

[285] Wei Ye, Ling Zhang, Wenqing Zhang, Xiaojiao Wu, Dong Yi, and Yazhou Wu. A comparison of single imputation and multiple imputation methods for missing data in different oncogene expression profiles. *Biostatistics amp; Epidemiology*, 6(1):113–127, January 2022.

[286] Henrik S. Jørgensen, Hirofumi Nakayama, Hans O. Raaschou, and Tom S. Olsen. Recovery of walking function in stroke patients: The copenhagen stroke study. *Archives of Physical Medicine and Rehabilitation*, 76(1):27–32, jan 1995.

[287] Ellen L. Carroll, Joanne G. Outtrim, Faye Forsyth, Anne E. Manktelow, Peter J. A. Hutchinson, Olli Tenovuo, Jussi P. Posti, Lindsay Wilson, Barbara J. Sahakian, David K. Menon, and Virginia F. J. Newcombe. Mild traumatic brain injury recovery: a growth curve modelling analysis over 2 years. *Journal of Neurology*, 267(11):3223–3234, June 2020.

[288] MA Leone, S. Bonissoni, L. Collimedaglia, F. Tesser, S. Calzoni, A. Stecco, P. Naldi, and F. Monaco. Factors predicting incomplete recovery from relapses in multiple sclerosis: a prospective study. *Multiple Sclerosis Journal*, 14(4):485–493, January 2008.

[289] Timothy T. Roberts, Garrett R. Leonard, and Daniel J. Cepela. Classifications in brief: American spinal injury association (asia) impairment scale. *Clinical Orthopaedics amp; Related Research*, 475(5):1499–1504, May 2017.

[290] Chin Wan Yoke and Zarina Mohd Khalid. Comparison of multiple imputation and complete-case in a simulated longitudinal data with missing covariate. In *AIP Conference Proceedings*. AIP Publishing LLC, 2014.

[291] Christian Schuld, Steffen Franz, Karin Brüggemann, Laura Heutehaus, Norbert Weidner, Steven C. Kirshblum, and Rüdiger Rupp. International standards for neurological classification of spinal cord injury: impact of the revised worksheet (revision 02/13) on classification performance. *The Journal of Spinal Cord Medicine*, 39(5):504–512, June 2016.

[292] A. Rogier T. Donders, Geert J.M.G. van der Heijden, Theo Stijnen, and Karel G.M. Moons. Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, 59(10):1087–1091, October 2006.

[293] Jun Shao and Bob Zhong. Last observation carry-forward and last observation analysis. *Statistics in Medicine*, 22(15):2429–2441, July 2003.

[294] Sanford Weisberg. *Applied Linear Regression*. Wiley, January 2005.

[295] Leif Peterson. K-nearest neighbor. *Scholarpedia*, 4(2):1883, 2009.

[296] *Support Vector Machines*. Springer New York, 2008.

[297] Leo Breiman. *Machine Learning*, 45(1):5–32, 2001.

[298] Rich Caruana and Alexandru Niculescu-Mizil. An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning - ICML '06*, ICML '06. ACM Press, 2006.

[299] John W. Graham, Allison E. Olchowski, and Tamika D. Gilreath. How many imputations are really needed? some practical clarifications of multiple imputation theory. *Prevention Science*, 8(3):206–213, June 2007.

[300] Stef van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations inr. *Journal of Statistical Software*, 45(3), 2011.

[301] Andrew Miles. Obtaining predictions from models fit to multiply imputed data. *Sociological Methods amp; Research*, 45(1):175–185, October 2015.

[302] page 283–287. Springer New York.

[303] Roderick J. A. Little. A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404):1198–1202, December 1988.

[304] Nicholas J Tierney and Dianne H Cook. Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations. 2018.

[305] Olawale F. Ayilara, Lisa Zhang, Tolulope T. Sajobi, Richard Sawatzky, Eric Bohm, and Lisa M. Lix. Impact of missing data on bias and precision when estimating change in patient-reported outcomes from a clinical registry. *Health and Quality of Life Outcomes*, 17(1), June 2019.

[306] Simon B Goldberg, Daniel M Bolt, and Richard J Davidson. Data missing not at random in mobile health research: Assessment of the problem and a case for sensitivity analyses. *Journal of Medical Internet Research*, 23(6):e26749, June 2021.

[307] Abel Torres-Espín, Jenny Haefeli, Reza Ehsanian, Dolores Torres, Carlos A Almeida, J Russell Huie, Austin Chou, Dmitriy Morozov, Nicole Sanderson, Benjamin Dirlikov, Catherine G Suen, Jessica L Nielson, Nikos Kyritsis, Debra D Hemmerle, Jason F Talbott, Geoffrey T Manley, Sanjay S Dhall, William D Whetstone, Jacqueline C Bresnahan, Michael S Beattie, Stephen L McKenna, Jonathan Z Pan, and Adam R Ferguson. Topological network analysis of patient similarity for precision management of acute blood pressure in spinal cord injury. *eLife*, 10, November 2021.

[308] Guoxin Fan, Sheng Yang, Huaqing Liu, Ningze Xu, Yuyong Chen, Jie He, Xiuyun Su, Mao Pang, Bin Liu, Lanqing Han, and Limin Rong. Machine learning-based prediction of prolonged intensive care unit stay for critical patients with spinal cord injury. *Spine*, 47(9):E390–E398, October 2021.

[309] Hyun Kang. The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, 64(5):402, 2013.

[310] John M Lachin. Fallacies of last observation carried forward analyses. *Clinical Trials*, 13(2):161–168, September 2015.

[311] Joseph G. Ibrahim and Geert Molenberghs. Missing data methods in longitudinal studies: a review. *TEST*, 18(1):1–43, February 2009.

[312] Yue Wu, Terry J. Lyons, and Kate E. A. Saunders. Deriving information from missing data: implications for mood prediction. 2020.

[313] Susan M. Fox-Wasylyshyn and Maher M. El-Masri. Handling missing data in self-report measures. *Research in Nursing amp; Health*, 28(6):488–495, 2005.

[314] Stef van Buuren. *Flexible Imputation of Missing Data, Second Edition*. Chapman and Hall/CRC, July 2018.

[315] V K Noonan, B K Kwon, L Soril, M G Fehlings, R J Hurlbert, A Townson, M Johnson, and M F Dvorak. The rick hansen spinal cord injury registry (rhscir): a national patient-registry. *Spinal Cord*, 50(1):22–27, November 2011.

[316] John K. Yue, Mary J. Vassar, Hester F. Lingsma, Shelly R. Cooper, David O. Okonkwo, Alex B. Valadka, Wayne A. Gordon, Andrew I. R. Maas, Pratik Mukherjee, Esther L. Yuh, Ava M. Puccio, David M. Schnyer, Manley, Geoffrey T. TRACK-TBI Investigators, Scott S. Casey, Maxwell Cheong, Kristen Dams-O'Connor, Allison J. Hricik, Emily E. Knight, Edwin S. Kulubya, David K. Menon, Diane J. Morabito, Jennifer L. Pacheco, and Tuhin K. Sinha. Transforming research and clinical knowledge in traumatic brain injury pilot: Multicenter implementation of the common data elements for traumatic brain injury. *Journal of Neurotrauma*, 30(22):1831–1844, November 2013.

[317] David R Marsh, Dirk G Schroeder, Kirk A Dearden, Jerry Sternin, and Monique Sternin. The power of positive deviance. *BMJ*, 329(7475):1177–1179, November 2004.

[318] Warren L Berggren and Joe D Wray. Positive deviant behavior and nutrition education. *Food Nutr. Bull.*, 23(4 Suppl):9–10, December 2002.

[319] Einat Engel-Haber, Brittany Snider, and Steven Kirshblum. Central cord syndrome definitions, variations and limitations. *Spinal Cord*, 61(11):579–586, April 2023.

[320] Srikanth N. Divi, Gregory D. Schroeder, John J. Mangan, Madeline Tadley, Wyatt L. Ramey, Jetan H. Badhiwala, Michael G. Fehlings, F. Cumhur Oner, Frank Kandziora, Lorin M. Benneker, Emiliano N. Vialle, Shanmuganathan Rajasekaran, Jens R. Chapman, and Alexander R. Vaccaro. Management of acute traumatic central cord syndrome: A narrative review. *Global Spine Journal*, $9(1_suppl) : 89S - -97S, May2019$.

[321] Mohit Joshi and Payal Bhardwaj. Impact of data transparency: Scientific publications. *Perspectives in Clinical Research*, 9(1):31, 2018.

[322] Qi Li. *Overview of Data Visualization*, page 17–47. Springer Singapore, 2020.

[323] Fabio Crameri, Grace E. Shephard, and Philip J. Heron. The misuse of colour in science communication. *Nature Communications*, 11(1), October 2020.

[324] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. Mimic-iv clinical database demo. 2023.

[325] Alistair E. W. Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, Steven Horng, Tom J. Pollard, Sicheng Hao, Benjamin Moody, Brian Gow, Li-wei H. Lehman, Leo A. Celi, and Roger G. Mark. Mimic-iv, a freely accessible electronic health record dataset. *Scientific Data*, 10(1), January 2023.

[326] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. Mimic-iv. 2023.

[327] Kerry Dwan, Douglas G. Altman, Mike Clarke, Carrol Gamble, Julian P. T. Higgins, Jonathan A. C. Sterne, Paula R. Williamson, and Jamie J. Kirkham. Evidence for the selective reporting of analyses and discrepancies in clinical trials: A systematic review of cohort studies of clinical trials. *PLOS Medicine*, 11(6):1–22, 06 2014.

[328] *Guidelines for Reporting Health Research: A User's Manual*. Wiley, July 2014.

[329] Linda Martin, Melissa Hutchens, Conrad Hawkins, and Alaina Radnov. How much do clinical trials cost? *Nature Reviews Drug Discovery*, 16(6):381–382, May 2017.

[330] Task Force of the Working Group on Arrhythmias of the European Society of Cardiology. The early termination of clinical trials: causes, consequences, and control. with special reference to trials in the field of arrhythmias and sudden death. task force of the working group on arrhythmias of the european society of cardiology. *Circulation*, 89(6):2892–2907, June 1994.

[331] D P Lammertse. Clinical trials in spinal cord injury: lessons learned on the path to translation. the 2011 international spinal cord society sir ludwig guttmann lecture. *Spinal Cord*, 51(1):2–9, November 2012.

[332] Khashayar Afshari, Nazanin Momeni Roudsari, Naser-Aldin Lashgari, Nazgol-Sadat Haddadi, Arvin Haj-Mirzaian, Malihe Hassan Nejad, Hamed Shafaroodi, Mehdi Ghasemi, Ahmad Reza Dehpour, and Amir Hossein Abdolghaffari. Antibiotics with therapeutic effects on spinal cord injury: a review. *Fundamental amp; Clinical Pharmacology*, 35(2):277–304, October 2020.

[333] Michal Ozery-Flato, Yaara Goldschmidt, Oded Shaham, Sivan Ravid, and Chen Yanover. Framework for identifying drug repurposing candidates from observational healthcare data. *JAMIA Open*, 3(4):536–544, December 2020.

[334] Ruoqi Liu, Lai Wei, and Ping Zhang. A deep learning framework for drug repurposing via emulating clinical trials on real-world patient data. *Nat. Mach. Intell.*, 3(1):68–75, January 2021.

[335] Raihan Mohammed, Kaesi Opara, Rahul Lall, Utkarsh Ojha, and Jinpo Xiang. Evaluating the effectiveness of anti-nogo treatment in spinal cord injuries. *Neural Development*, 15(1), January 2020.

[336] Björn Zörner and Martin E. Schwab. Anti-nogo on the go: from animal models to a clinical trial. *Annals of the New York Academy of Sciences*, 1198(s1), June 2010.

[337] Armin Curt and Volker Dietz. Nerve conduction study in cervical spinal cord injury: significance for hand function. *NeuroRehabilitation*, 7(3):165–173, December 1996.

[338] Roop Singh, Jitendra Wadhwani, Vijay Singh Meena, Pankaj Sharma, Kiranpreet Kaur, and Svareen. Electrophysiological study in acute spinal cord injury patients: Its correlation to neurological deficit and subsequent recovery assessment by asia score. *Indian Journal of Orthopaedics*, 54(5):678–686, April 2020.

[339] Jana Lipkova, Richard J. Chen, Bowen Chen, Ming Y. Lu, Matteo Barbieri, Daniel Shao, Anurag J. Vaidya, Chengkuan Chen, Luoting Zhuang, Drew F.K. Williamson, Muhammad Shaban, Tiffany Y. Chen, and Faisal Mahmood. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell*, 40(10):1095–1110, October 2022.