Prospective cohort study clavicle and rib fractures

A Data Management Plan created using DMPonline

Creators: Arthur Sweet, Divisie HS

Affiliation: UMC Utrecht

Funder: UMC Utrecht

Template: UMC Utrecht DMP

Project abstract:

Both clavicle fractures and rib fractures have been shown to act as a marker of severity of the chest injury and may both independently double the risk of mortality. These injuries combined may even further worsen the outcome. Treatment and outcomes of both isolated injuries have been well described, yet it remains unclear how these injuries combined impact the outcome. Therefore, this study aims to describe the treatment and health-related outcomes of patients with combined clavicle and rib fractures.

ID: 70797

Start date: 03-02-2021

End date: 03-02-2022

Last modified: 28-06-2021

Prospective cohort study clavicle and rib fractures

1. General features

1.1. Please fill in the table below. When not applicable (yet), please fill in N/A.

DMP template version	29 (don't change)
ABR number *(only for human-related research)*	
METC number *(only for human-related research)*	To be determined
DEC number *(only for animal-related research)*	
Acronym/short study title	Clavribs
Name Research Folder	xx-xxx_Clavribs
Name Division	Chirurgische specialismen
Name Department	Traumatologie
Partner Organization	-
Start date study	3-2-2021
Planned end date study	3-2-2021
Name of datamanager consulted*	Dax Steins
Check date by datamanager	18-02-2021

1.2 Select the specifics that are applicable for your research.

- Prospective study
- Monocenter study
- Non-WMO
- Observational study

2. Data Collection

2.1 Give a short description of the research data.

For this prospective cohort study, we include all trauma patients with clavicle and rib factures (aged 16 years or older) who were admitted to the UMC Utrecht (UMCU) between 01-01-2015 and 31-12-2017.

The Traumazorgnetwerk Midden-Nederland (TZMN), which is regional trauma registry, supplies the initial research dataset to the coordinating investigator (CI). This dataset only contains UMCU patients. Consequently, additional health care data may be extracted from the electronic

health records (EHRs; HiX). When the dataset is complete, the CI will be pseudonymize the data for furthere analysis.

Data on long-term outcomes will be prospectively acquired using questionnaires. All eligible patients will be sent a letter with information on the study and an PIF/IC form. After return of the signed IC forms, a new letter will be sent with the questionnaires. After these letters with filed questionnaires are returned, the patients will be included.

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
Human	120	Trauma registry	Excel	Quantitative	.xlsx	<1GB
Human	120	EPD (Hix)	Excel	Quantitative	.xlsx	<1GB
Human	120	EQ-5D-5L	Excel	Quantitative	.xlsx	<1GB
Human	120	QuickDASH	Excel	Quantitative	.xlsx	<1GB

2.2 Do you reuse existing data?

- Yes, please specify

For this study, we shall reuse UMCU patient data between 01-01-2015 and 31-12-2017 from the trauma registry and HiX.

We also acquire new data by filling questionnaires as long-term outcome measures.

2.3 Describe who will have access to which data during your study.

My division datamanager receives the dataset from TZMN via a secure link that contains direct identifying personal data (e.g. date of birth) and research data. The datamanager is authorized to link different datasets of the selected patient group and thus has access to personal data such as patientID. The key table linking study specific IDs to patient IDs is available to the datamanager and members of the research team.

Type of data	Who has access
Direct identifying personal data	Research team, dHS Datamanager
Key table linking study specific IDs to Patient IDs	PI, dHS Datamanager, research team
Pseudonymized data	Research team, dHS Datamanager

2.4 Describe how you will take care of good data quality.

1. All trauma parameters will be checked and filled out by the PI (i.e. a qualified trauma surgeon and the treating physician).

2. Original data will be supplemented out of notes from concerned medical specialities.
3. In advance, clear definitions of each parameter/variable are formulated.
4. The dataset will be checked by the study PI (i.e. the treating physician) and research team for outliers and missing values.
5. Outliers will be manually checked for correctness by the PI (i.e. the treating physician)
6. Missing values will be either manually reconstructed by the PI (i.e. the treating physician): missing EMV can be reconstructed by descriptions of the neurologist (e.g., clear and alert, commotional). Missing, non-retraceable values (e.g., arterial blood gas analysis) will be accepted as such or multiple imputation will be performed.
7. An audit trail will be kept of all patients who were eligible and were sent a letter and all patients that returned the IC form and questionnaires. These data will be added to the complete dataset.

#	Question	Yes	No	N/A
1.	Do you use a certified Data Capture Tool or Electronic Lab Notebook?		x	
2.	Have you built in skips and validation checks?			x
3.	Do you perform repeated measurements?		x	
4.	Are your devices calibrated?			x
5.	Are your data (partially) checked by others (4 eyes principle)?	x		
6.	Are your data fully up to date?	x		
7.	Do you lock your raw data (frozen dataset)	x		
8.	Do you keep a logging (audit trail) of all changes?		x	
9.	Do you have a policy for handling missing data?	x		
10.	Do you have a policy for handling outliers?	x		

2.5 Specify data management costs and how you plan to cover these costs.

#	Type of costs	Division ("overhead")	Funder	Other (specify)
1.	Time of datamanager	X		
2.	Storage	X		
3.	Archiving	X		

No extra costs will be involved in managing and storing data.
The management and storage will be done in existing UMC Utrecht facilities.
No extra salary costs are involved in this process.
No costs are involved in statistics, since Stata will be used of which the UMCU has a license.

2.6 State how ownership of the data and intellectual property rights (IPR) to the data will be managed, and which agreements will be or are made.

UMC Utrecht is and remains the owner of all collected data for this study. The data is collected in a relatively large patient group and is very valuable for further, broader studies in Europe. It may for example be used to find study subjects for future treatment studies. Our data cannot be protected with IPR, but its value will be taken into account when making our data available to others, when setting up Research Collaborations and when drawing up Data Transfer Agreement(s).

3. Personal data (Data Protection Impact Assessment (DPIA) light)

Will you be using personal data (direct or indirect identifying) from the Electronic Patient Dossier (EPD), DNA, body material, images or any other form of personal data?

- Yes, go to next question

I will process personal data. I have consulted the division data manager and I do not have to complete a full DPIA. I therefore fill out this DPIA light and proceed to 3.1.

3.1 Describe which personal data you are collecting and why you need them.

Which personal data?	Why?
1. Patient Demographics 2. ASA-classification 3. Mechanism of injury AIS scores ISS 4. Arterial blood pH and base excess 5. Adverse events 6. Biochemical and radiological investigations	To describe the study population. (so that it is clear to what other populations the results can be extrapolated)
Injury characteristics • Number of rib fractures • Bilateral rib fractures • Location of rib fractures • Displacement of rib fractures • Dorsal rib fracutres • Type clavicle fracture • Displacement of clavicle fracture (primary & secondary) • Intra-thoracic injuries (i.e. pneumothorax, hemothorax, lung contusion) • Sternum fracture • Scapula fracture	To describe the injuries of the cohort.

Surgery-related outcomes • Operative fixation clavicle fracture • Approach of fixation clavicle fractures (anterior/superior) • Osteosynthesis material clavicle • Operative fixation of rib fractures • Approach of ORIF rib fractures • Number of ribs fixated • Duration until surgery	To assess what type operative the patients had

In-hospital outcomes
 	
• HLOS
• Admission to the ICU
• ILOS
• Mechanical ventilation
• Days on mechanical ventilation (DMV)
• Chest tube placement
• Chest tube duration
• Destination of discharge
• Glasgow outcome scale
• Mortality
• Complications
  ◦ Pneumonia
  ◦ Pneumothorax
  ◦ Hemothorax
  ◦ Excess pleural fluid
  ◦ Delirium
  ◦ Urinary tract infection
• Complications clavicle
  ◦ Non-union
• Secundary dislocation
  ◦ Persistent pain complaints (6 and 12 weeks)
  ◦ Reintervention
• Complications ribs
  ◦ Non-union
  ◦ Persistent pain complaints (6 and 12 weeks)
  ◦ Reintervention
	To describe the outcomes of patients with clavicle and rib fractures and to identify if there are certain patients who are at higher risk of adverse outcomes

Long-term outcomes • Quality of Life EQ-5D-5L questionnaire • QuickDASH questionnaire • NRS of the clavicle and ribs	To describe quality of life and functionality after follow up.

3.2 What legal right do you have to process personal data?

- Study-specific informed consent

3.3 Describe how you manage your data to comply to the rights of study participants.

All eligible patients will be sent a letter with an PIF/IC form by the treating physician. After patients returned the letter and signed the IC form, another letter will be sent with the questionnaires. After return of the completed questionnaires, these patients will be pseudonymized and included in the study. All steps will be noted in an audit trail.

The data are pseudonymized and the linking table to personal data is saved. An authorized person manages the linking table, can re-identify study participants when necessary and deliver, correct or delete the data.

In case the patients were diseased before the date of the asking for informed consent, no informed consent and questionnaires can be collected. In-hospital data of all diseased patients will be collected. Our data manager will check whether the patients have previously objected against participation of a study, in which case the patients will be excluded. This study fulfils all conditions to use these data without consent as this information is necessary for proper execution of this study, this study serves a broad medical interest and asking for consent is impossible in diseased patients.

3.4 Describe the tools and procedures that you use to ensure that only authorized persons have access to personal data.

We use the secured Research Folder Structure that ensures that only authorized personnel has access to personal data, including the key table that links personal data to the pseudoID.

3.5 Describe how you ensure secure transport of personal data and what contracts are in place for doing that.

Data will be requested from the TZMN and shared with our division datamanager via SURFdrive. This is a secure institutional cloud storage, to share data from TZMN via a encrypted URL-link. No data will be transported outside the UMCU network drives

4. Data Storage and Backup

4.1 Describe where you will store your data and documentation during the research.

The digital files will be stored in the secured Research Folder Structure of the UMC Utrecht. We will need +/- 5 GB storage space, so the capacity of the network drive will be sufficient.

4.2 Describe your backup strategy or the automated backup strategy of your storage locations.

All (research) data is stored on UMC Utrecht networked drives from which backups are made automatically twice a day by the division IT (dIT).

5. Metadata and Documentation

5.1 Describe the metadata that you will collect and which standards you use.

1. For the data shared by TZMN in Excel, they prepared a codebook. Additional data from the electronic health records, including the outcome of the upper mentioned questionnaires (in 2.1) will be merged with this database. Consequently, the codebook with updated accordingly. We do not use any existing metadata standards for this study.

5.2 Describe your version control and file naming standards.

We will distinguish versions by indicating the version in the filename of the master copy by adding a code after each edit, for example V1.1 (first number for major versions, last for minor versions). The most recent copy at the master location is always used as the source, and before any editing, this file is saved with the new version code in the filename. The file with the highest code number is the most recent version. Every month, we will move minor versions to a folder OLD. The major versions will be listed in a version document (projxVersDoc.txt), stating the distinguishing elements per listed version.

.

6. Data Analysis

6 Describe how you will make the data analysis procedure insightful for peers.

We will be using Stata, version 14, for statistical analysis of the data. The scripts will contain comments, such that every step in the analysis is documented and peers can read why I made certain decisions during the analysis phase.

7. Data Preservation and Archiving

7.1 Describe which data and documents are needed to reproduce your findings.

The data package will contain: the raw data, the study protocol describing the methods and materials, the script to process the data, the scripts leading to tables and figures in the publication, a codebook with explanations on the variable names, and a ‘read_me.txt’ file with an overview of files included and their content and use.

7.2 Describe for how long the data and documents needed for reproducibility will be available.

Data and documentation needed to reproduce findings from this non-WMO study will be stored for at least 15 years.

7.3 Describe which archive or repository (include the link!) you will use for long-term archiving of your data and whether the repository is certified.

After finishing the project, the data package will be stored at the UMC Utrecht Research Folder Structure and is under the responsibility of the Principal Investigator of the research group.

7.4 Give the Persistent Identifier (PID) that you will use as a permanent link to your published dataset.

I will be using a DOI-code and will update this plan as soon as I have the code.

8. Data Sharing Statement

8.1 Describe what reuse of your research data you intend or foresee, and what audience will be interested in your data.

The raw data can be of interest for other researchers or for spin off projects.
To be determined.

8.2 Are there any reasons to make part of the data NOT publicly available or to

restrict access to the data once made publicly available?

- Yes (please specify)

The publication will be open assessable. The study protocol and datamanagementplan can be requested by the researchers but will not be open assessable.
To be determined.

8.3 Describe which metadata will be available with the data and what methods or software tools are needed to reuse the data.

1. After finishing the project the metadata will be available for at least the mandatory time. The original and pseudonymized data will not be available because of
privacy legislation. In the event that peers within the UMC like to reuse our anonymized data this can only be granted if the research question is in line with the
original research question of our research. Every application therefore will be thoroughly screened upon this requirement.
To be determined.

8.4 Describe when and for how long the (meta)data will be available for reuse

- Other (please specify)

For the timeperiod indicated in 7.2

8.5 Describe where you will make your data findable and available to others.

As the data is privacy-sensitive, we will publish metadata in publication(s) in peer-reviewd journals. Nonetheless, original and pseudonymized data will be treated
as aforementioned, which will be within the UMC Utrecht.
To be determined.