Traumabase

MAP 573 Student projects

Julie Josse, Tobias Gauss, Jean-Denis Moyer, Sophie Hamada, Olivier Auliard, Imke Mayer

École Polytechnique; Traumabase® Group, CapGemini Invent

October 22nd, 2019

Introduction

Collaborators

Academia: **Julie Josse** (X, Inria), Jean-Pierre Nadal (ENS, EHESS), Stefan Wager (Stanford), Wei Jiang (X), Nicolas Prost (X), **Imke Mayer** (X)
Traumabase (APHP): **Tobias Gauss, Sophie Hamada, Jean-Denis Moyer**, and others
CapGemini invent: **Olivier Auliard, Estelle Noel, Patrick Lee**, Julien Sauvan, Romain Laveilssière

Context

Major trauma: any injury that endangers the life or the functional integrity of a person. Road traffic accidents, interpersonal violence, self-harm, falls, etc. $\rightarrow$ hemorrhage and traumatic brain injury

Major source of mortality and handicap in France and worldwide (3rd cause of death, 1st cause among 16-45 yo, 2-3rd cause of disability)

$\Rightarrow$ A public health challenge

Context

Major trauma: any injury that endangers the life or the functional integrity of a person. Road traffic accidents, interpersonal violence, self-harm, falls, etc. $\rightarrow$ hemorrhage and traumatic brain injury

Major source of mortality and handicap in France and worldwide (3rd cause of death, 1st cause among 16-45 yo, 2-3rd cause of disability)

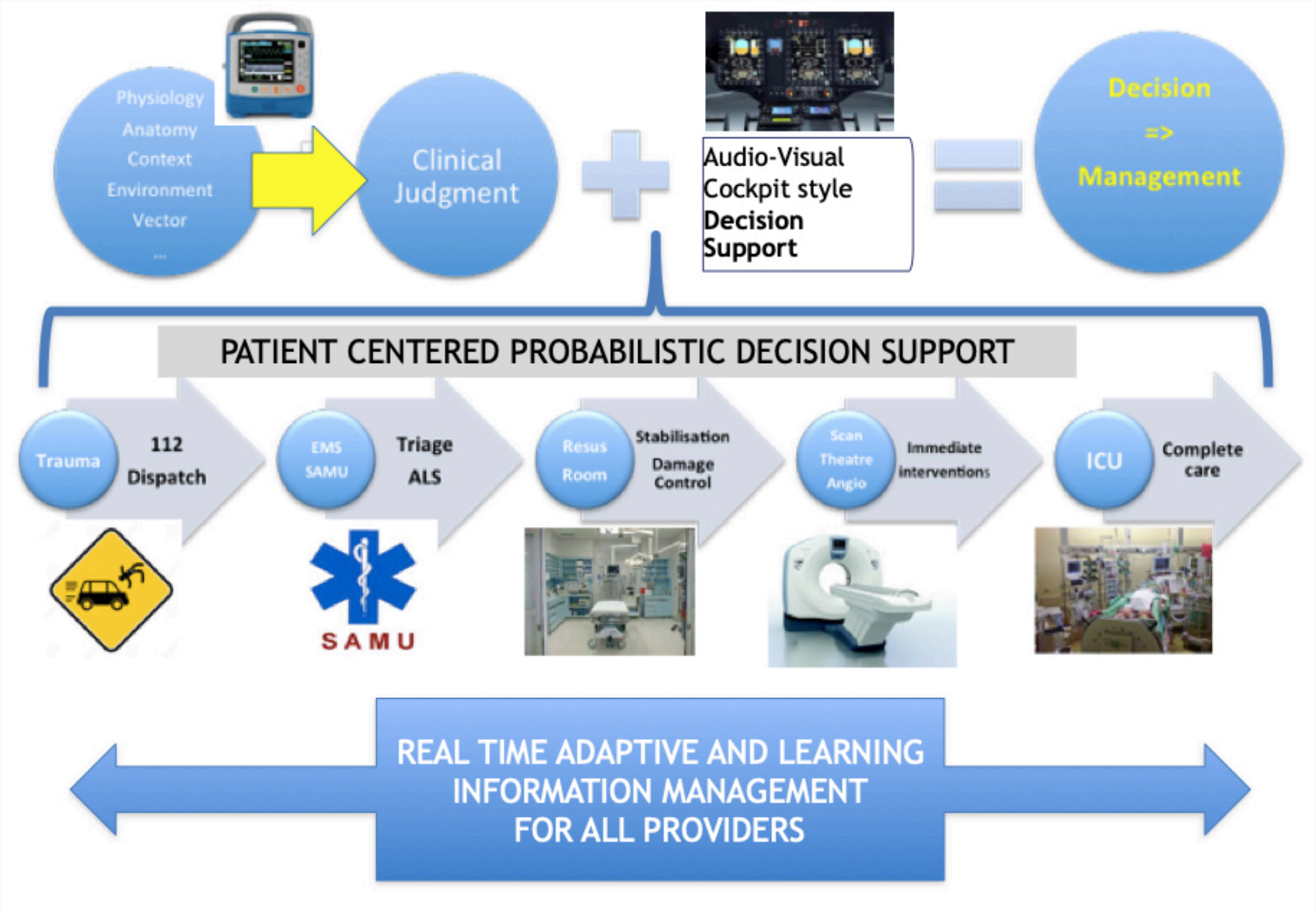
$\Rightarrow$ A public health challenge

Patient prognosis can be improved: **standardized and reproducible procedures** but still **personalized** for the patient and the trauma system.

Trauma decision making: fast and **complex decisions** under **strong time constraint** in a dynamic and multi-player environment (fragmentation: loss or distortion of information) with high levels of uncertainty and **stress**.
Issues: patient management exceeds time frames, diagnostic errors, decisions not reproducible, etc.

$\Rightarrow$ can machine learning help improve the decision process?

Decision support tool for the management of major trauma: TrauMatrix



Decision support tool for the management of major trauma: TrauMatrix

TrauMatrix : an integrative decision support and information management solution to clinicians for the first 24 hours of major trauma management to improve patient care and survival in major trauma.

- Analysis of the Traumabase
- Develop mathematical tools and machine learning models to predict trauma specific outcomes and decisions
  $\Rightarrow$ Scientific and methodological challenges
- Develop a user friendly and ergonomic interface for clinicians
- Test in real-time its impact on clinician decision making and patient outcome

$\Rightarrow$ Trans-disciplinary research and collaboration (medical, cognitive, mathematical, technological)

Traumabase

- $20,000$ patients
- 250 continuous and categorical variables: **heterogeneous**
- 16 hospitals: **multilevel data**
- 4,000 new patients/ year

Center	Accident	Age	Sex	Weight	Lactactes	BP	shock	...
Beaujon	fall	54	m	85	NM	180	yes	
Pitie	gun	26	m	NR	NA	131	no	
Beaujon	moto	63	m	80	3.9	145	yes	
Pitie	moto	30	w	NR	Imp	107	no	
HEGP	knife	16	m	98	2.5	118	no	
⋮								⋱

Traumabase

- $20,000$ patients
- 250 continuous and categorical variables: **heterogeneous**
- 16 hospitals: **multilevel data**
- 4,000 new patients/ year

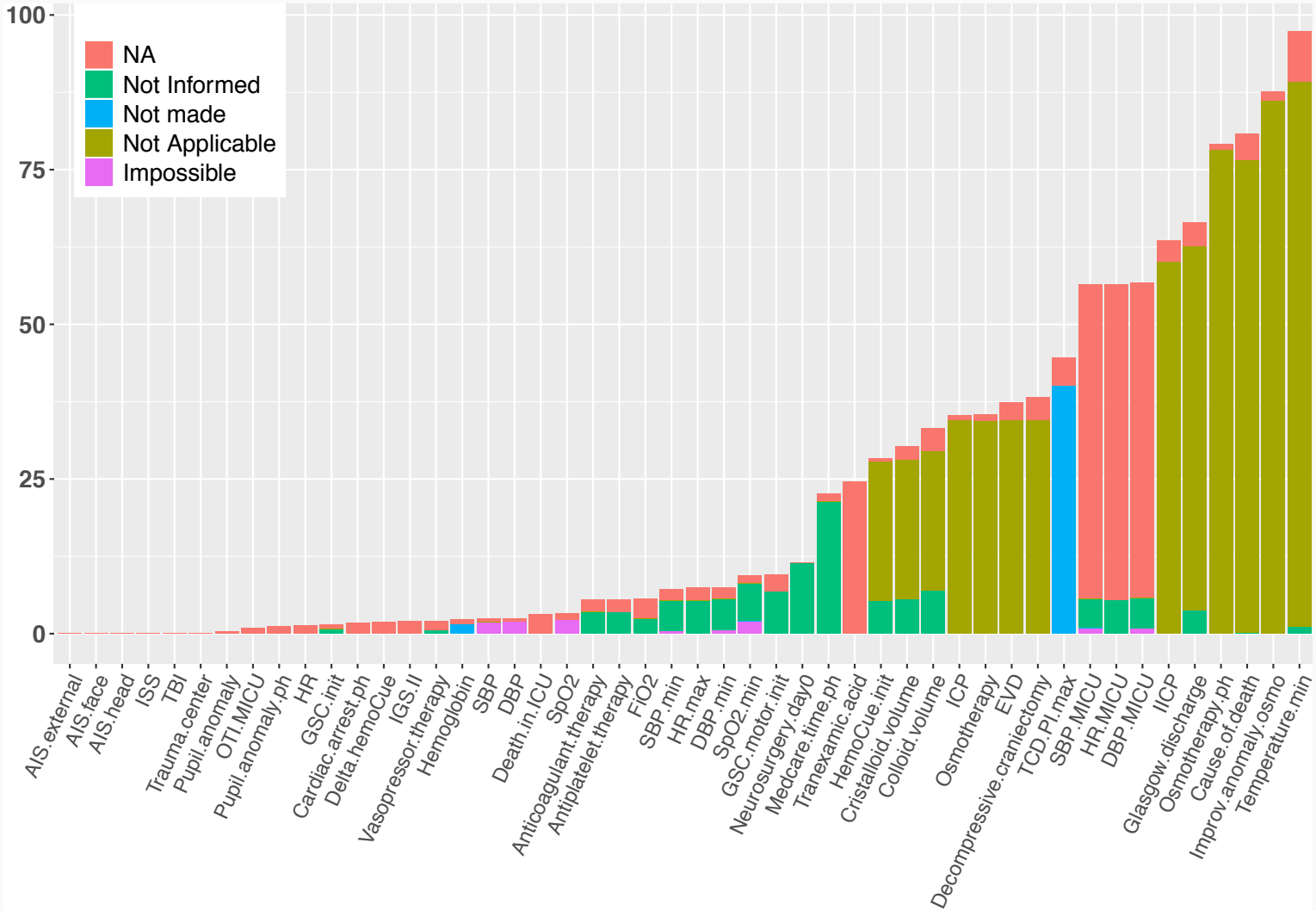
Center	Accident	Age	Sex	Weight	Lactactes	BP	shock	...
Beaujon	fall	54	m	85	NM	180	yes	
Pitie	gun	26	m	NR	NA	131	no	
Beaujon	moto	63	m	80	3.9	145	yes	
Pitie	moto	30	w	NR	Imp	107	no	
HEGP	knife	16	m	98	2.5	118	no	
⋮								⋱

$\Rightarrow$ **Predict** the Glasgow score, whether to start a blood transfusion, to administer fresh frozen plasma, etc...

$\Rightarrow$ **Estimate causal effect**: Administration of the **treatment** "tranexamic acid" (within 3 hours after the accident) on the **outcome** mortality for traumatic brain injury patients

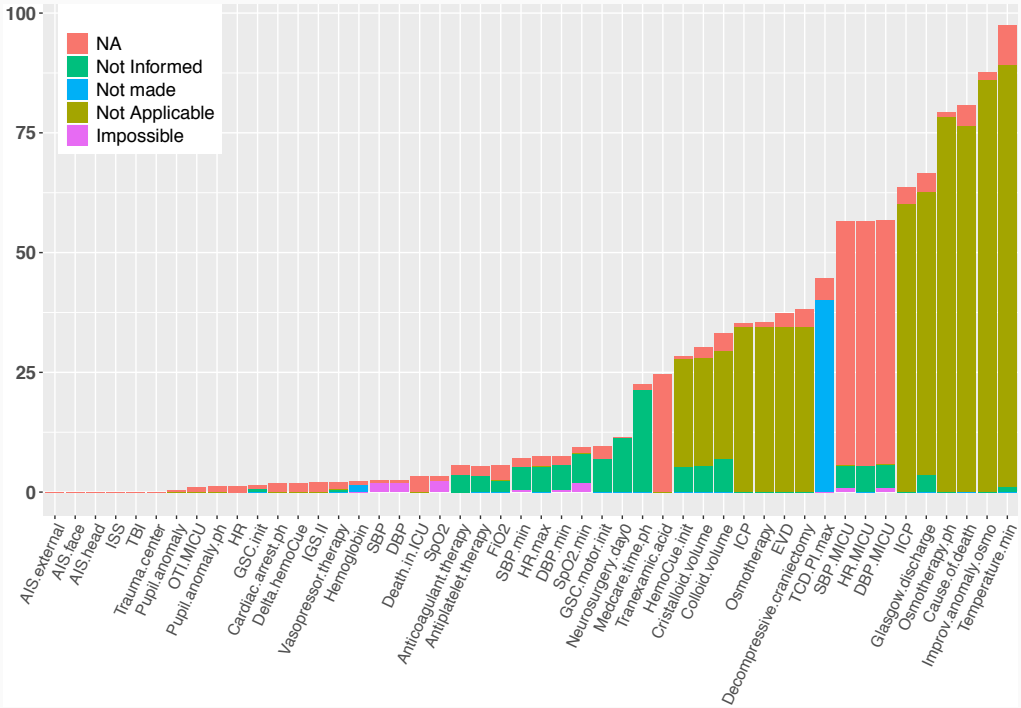
Traumabase

⇒ Missing values



Traumabase

⇒ Missing values



⇒Heterogeneous data (data integration, from different sources. Hospital effect : lack of standardization) - Evolutive data

⇒ High barriers to aggregation of medical data (privacy concerns, proprietary attitude towards data, complexity/size of aggregated data, updates problems). Data stay on each site : distribute computation

Ongoing work

- Logistic regression with missing values. Prediction of hemorrhagic shock. W. Jiang (PhD, X), M. Lavielle (INRIA XPOP), TraumaBase
- On the consistency of supervised learning with missing values. G. Varoquaux (INRIA), E. Scornet (X)
- Causal inference (double robust) methods with missing values. S. Wager (Stanford), I Mayer (PhD, X-EHESS)
- Distributed multilevel matrix completion for medical databases. G. Robin (PhD, X), B. Narasimhan (Stanford), F. Husson (Agrocampus)
- Effect of fibrinogen on mortality in traumatic hemorrhagic shock: a propensity score analysis.TraumaBase, Polytechnique students
- Effect of tranexamic acid on mortality for head trauma patient. I. Mayer (PhD, X-EHESS), J-P Nadal (EHESS), TraumaBase

MAP 573 projects

Step 1: Sign the confidentiality agreement.

MAP 573 projects

Projects 1 (1 group) and 2 (2 groups): Causal inference to assess the effect of a treatment on survival.

With Jean-Denis Moyer (MD).

Context

- Tranexamic acid (or Exacyl): antifibrinolytic agent limiting excessive bleeding
- Traumatic brain injury: any brain lesion visible on a CT scan
- Randomized placebo-controlled trials are very costly and most trials in the past decade did not find treatment effects for patients with TBI.
- Treatment assignment in observational data is not randomized, hence treatment effect estimation is more complicated.

Objectives

- Based on the observational data Traumabase, estimate the effect of tranexamic acid on survival by adjusting for confounding variables.
- Matching: adjust via matching techniques
- HTE: assess whether it is possible to leverage HTE estimation techniques to derive subgroups of patients relative to their benefit from tranexamic acid

First steps

- Methodology bibliography (seminal papers: Rubin (1973), Rosenbaum & Rubin (1983); review (matching): Iacus et al. (2012); review (HTE): Dorie et al. (2019), Wendling et al. (2018); introduction to causal inference: lecture notes from S. Wager; coursera `https://www.coursera.org/learn/crash-course-in-causality`)
- Medical background: CRASH 3 trial `https://crash3.lshtm.ac.uk/` and its results

MAP 573 projects

Project 4 (1 group): Assessing initial brain perfusion status with transcranial Doppler in patients with trauma brain injury: what contribution to predict outcome?

With Sophie Hamada (MD) and Olivier Auliard, Chief Data Scientist Capgemini

Context

- prognostic of a patient with TBI depends on multiple factors: clinical severity of the TBI (GCS score, pupil reactivity, severity of tomodensitometrical lesions), secondary induced brain damages (hypoxia, hypotension), severity of extra-cranial traumatic lesions, **(initial) brain perfusion status**
- prognostic targets: mortality, neurological outcome
- Transcranial Doppler (TCD) and the measured pulsatility index (PI) predictive of secondary damages, hypoperfusion
- TCD: Doppler ultrasonography that measures the velocity of blood flow through the brain's blood vessels by measuring the echoes of ultrasound waves moving transcranially (through the cranium).

Objectives

- Demonstrate that the PI measurement at the arrival of a patient with TBI at the hospital adds complementary prognostic information to the reference scores (IMPACT score and CRASH basal model) for the prediction of mortality at D14.
- Several secondary objectives.

MAP 573 projects

Project 5: Impact of the choice of the definition of a hemorrhagic shock on patient characteristics, consumption of care, patient's outcomes and methodological implications: a systematic review and registry based study

With Sophie Hamada (MD) and Olivier Auliard, Chief Data Scientist Capgemini

Context

- Hemorrhagic shock: one of the main causes of death in major patients, the first preventable cause of death
- Any definition impacts the main endpoint (mostly death but also several outcomes that matter to patients and cares consumptions).
- Definitions can be physiological (based on arterial blood pressure), or pragmatic (based on transfusion, or use of resources) or composite.
- Most used definition: the transfusion of 10 or more units of red blood cells (RBC) within 24 hours of emergency department admission (also known as massive transfusion).
- Problem: excludes many patients.

Objectives

- Identify the most important / frequently used definition of HS
- Test these definitions within the Traumabase to describe the groups of patients triggered
- Try to propose a definition of HS that allow to catch the majority of patients dying of HS, and being the most transfused.

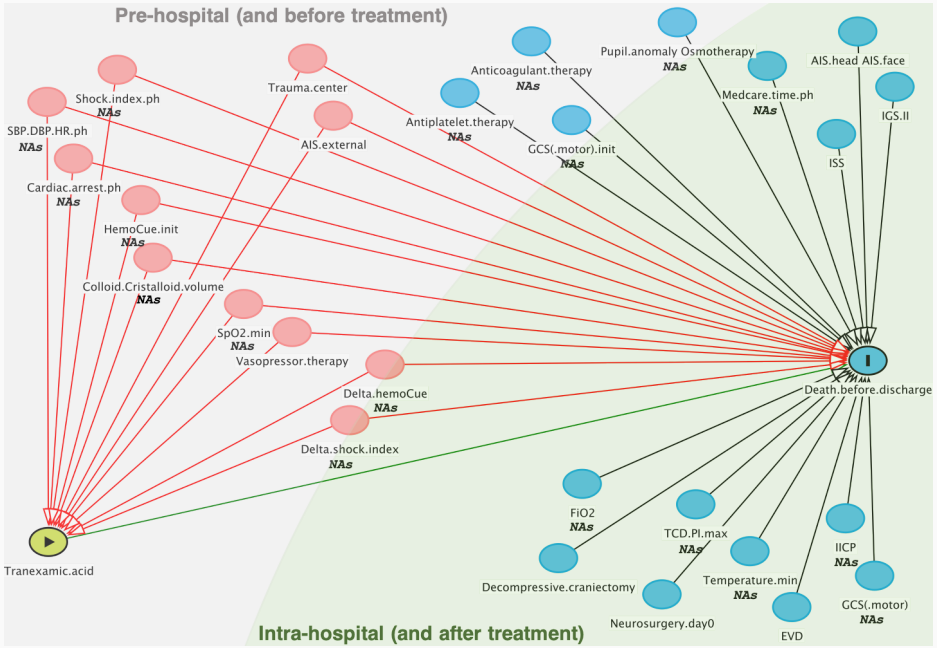
Traumabase catalogue

Short descriptions of all Traumabase variables:

`http://catalogue-data.traumabase.eu`
Access: traumabase2
Password: 5cwYdLABSwfjkp7d

Projects 1 and 2: Tranexamic acid

- Tranexamic acid (or Exacyl): antifibrinolytic agent limiting excessive bleeding
- Traumatic brain injury: any brain lesion visible on a CT scan
- Confounding factors for administration of tranexamic acid: 11 selected pre-treatment variables
- Additional information: 26 variables predictive of the outcome (in-ICU mortality), related to traumatic brain injury



Graph produced with DAGitty (Textor et al. (2011))

- TBI very heterogeneous: the patients injury/physiological profiles differ a lot – symptoms and degrees of severity cover a large spectrum
- MDs think that for certain types of lesion the tranexamic acid is beneficial whereas for others it could be harmful.

Causal inference resources

Stefan Wager (Graduate School of Business, Stanford)
Lecture notes (will be shared via Dropbox)

Susan Athey (Graduate School of Business, Stanford)
Material: `https://drive.google.com/drive/folders/1SEEOMluxBcSAb_tsDYgcLFtOQaeWtkLp`
Videos: `https://www.aeaweb.org/webcasts/2018/machine-learning-and-econometrics-part-1`

Introduction to causal inference (University of Pennsylvania)
`https://www.coursera.org/learn/crash-course-in-causality`

Missing value resource platform

More information and details on missing values: **R-miss-tastic** platform.

→ Theoretical and practical tutorials, popular datasets, bibliography, workflows (in R), active contributors/researchers in the community, etc.

https://rmisstastic.netlify.com

Causal inference groups

Matching
HTE 1
HTE 2