

Trauma Directors Tool Kit (tdtk-package)

A R package for patient data analysis

Eric W. Olle
eric_rprog@pm.me
March 2019
LibrePlanet 2019

[Github tdtk-package link](#)

Initial reasoning and goals

The trauma directors tool kit is an R package designed for/with trauma professionals, but can be applied to any medical registry/EHR data.

Design goals:

- Software base and package should be in the FSF/GPL ecosystem.
- Both the software and package should be scalable to larger data sets as well as have necessary tools for basic local analysis.
- Ability to be used outside the free ecosystem and in commercially available statistical software.
- Easily modifiable by any end user.
- Provide both summary, graphical and accessible ML/AI for the medical professional.
- Blinds the PHI to allow for dissemination and use by the data scientists without the need for additional training.
- Ability to write/generate reports and figures as needed by the end users.

Overall design

R-package

- R is licensed under GPLvX*.
- The tdtk-package is GPLv3.
- Easily accessible (i.e. no gatekeepers) on GitHub, GitLab or equivalent
- Initial functions concentrate on the cleaning and mapping of data.

Reports

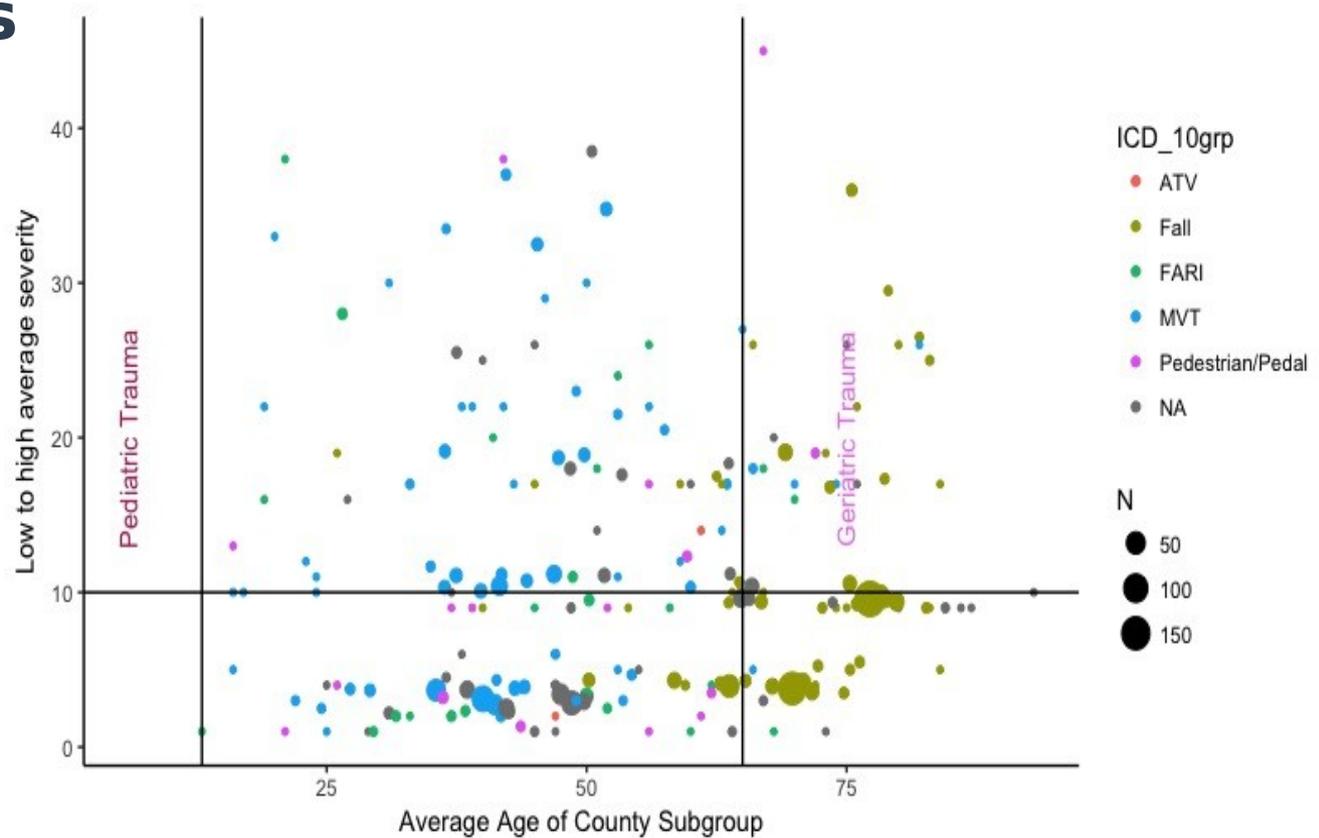
- Initial work done using markdown.
- Should be easily convertible to different formats.
- Easy to modify.
- Are operating system agnostic.

* GPLvX=multiple version see: <https://www.r-project.org/Licenses/>

Summary statistics

Summary Statistics

- Basic column summary statistics (not shown).
- Basic summary statistics on grouped data (not shown).
- Column stats are plotted using standard plots (not shown).
- Additional plots done for visualization of injury prevention on a county level.



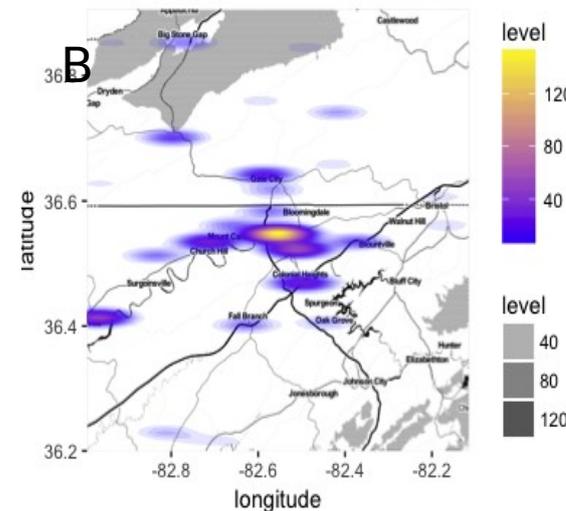
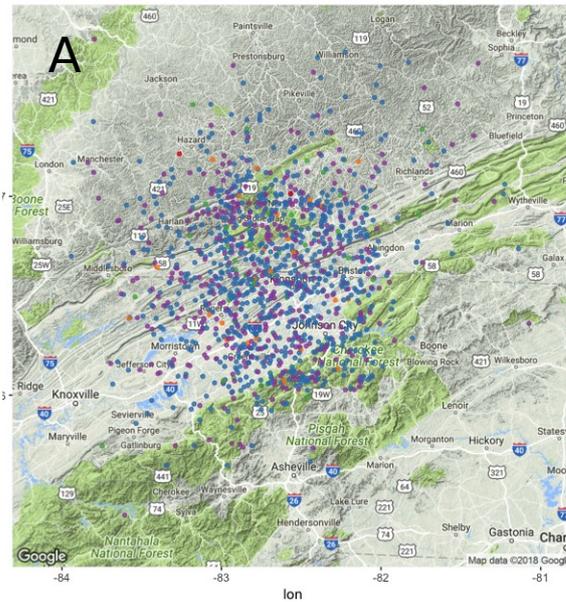
Geo-spatial statistics

Different basic geo-spatial statistics.

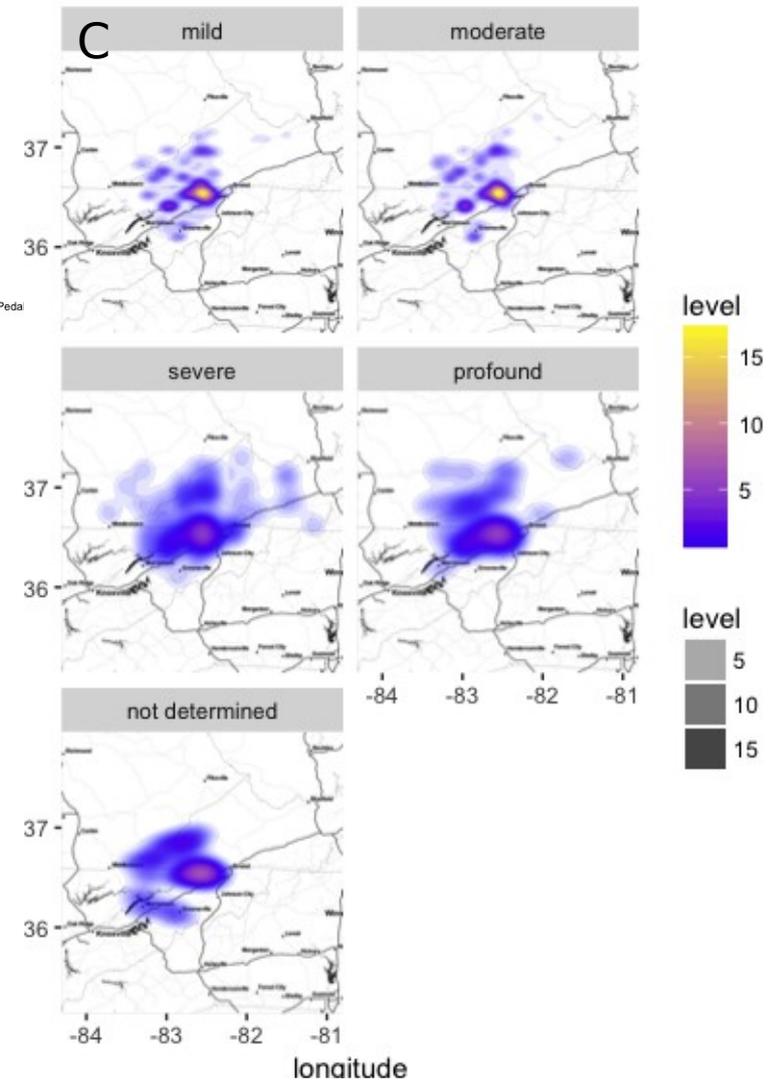
A. Shows individual patients with random noise added to prevent potential id.

B. Shows full cohort 2d statistical density.

C. Shows statistical density of different injury severity categories.



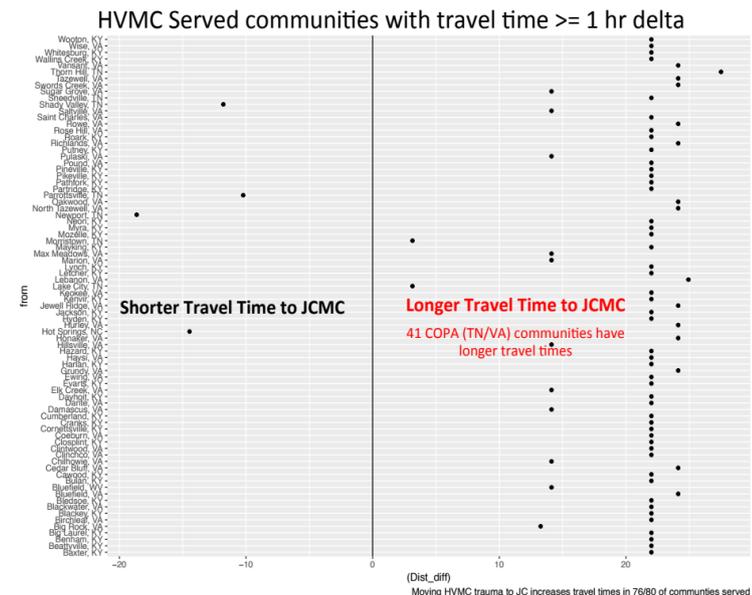
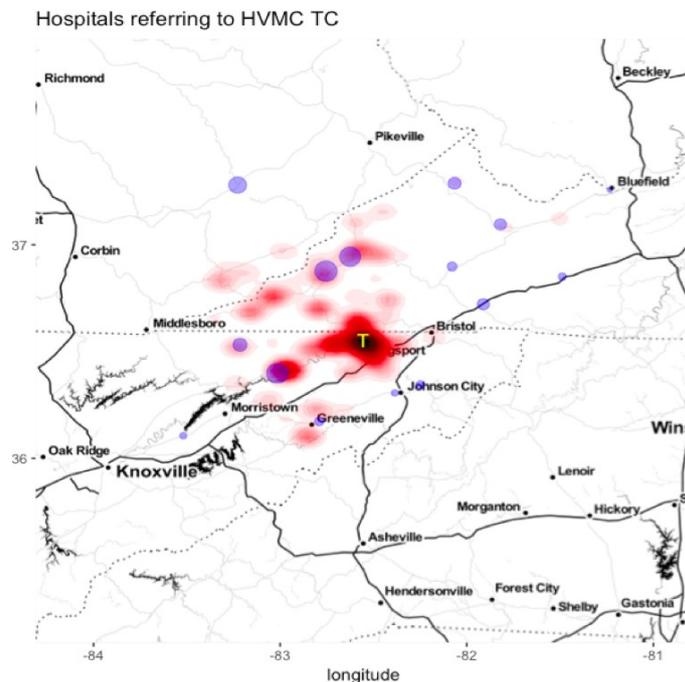
ICD_10grp
 • ATV
 • Fall
 • FARI
 • MVT
 • Pedestrian/Peda
 • NA



Travel times to trauma center

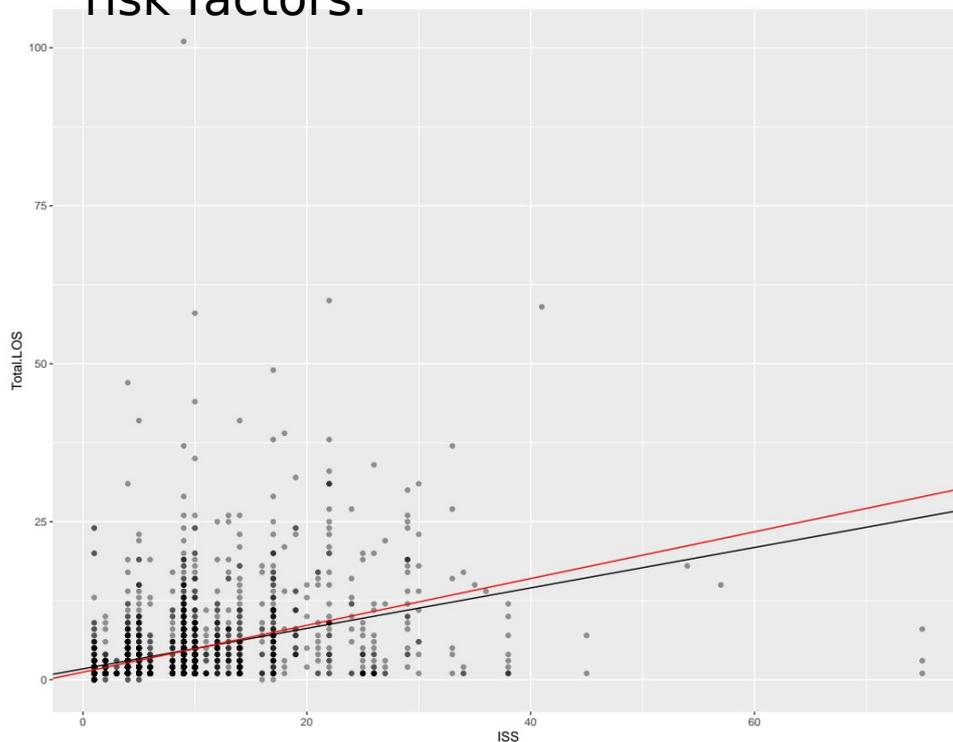
The “Golden Hour” or time from injury to definitive care is a central tenant of trauma care. Using publicly available data, a time difference model between two trauma centers was developed.

There two common ways for patients to enter the trauma service. This can be from a transfer (A) or directly from estimated community of injury (B).



Linear Modeling Length of Stay (LOS)

The model is unadjusted for medical risk factors.



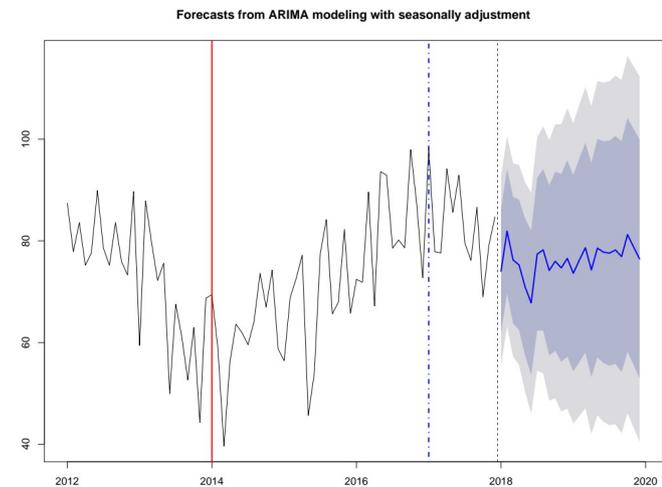
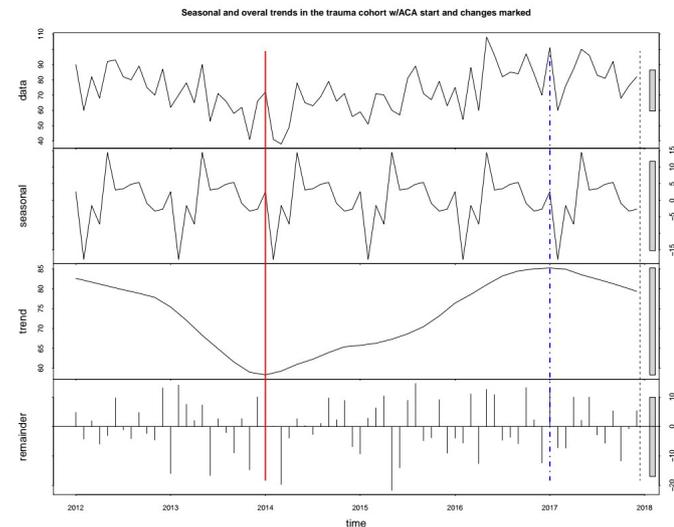
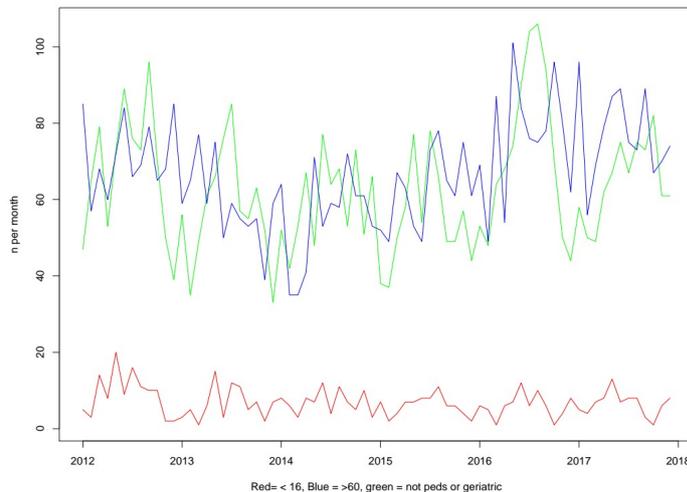
Initial ML model

- Linear modeling is a standard method used to look at the correlation between injury severity and total length of stay (LOS).
- Current model is not adjusted for medical or socioeconomic risk factors.
- Black line is full patient population.
- Red line indicates estimated LOS with mortalities removed from the data set.

Time series modeling of patient arrivals

Using the forecast-package an ARIMA model was developed to help trauma doctors predict the estimated number of patients.

- A. Unadjusted and separated to look at age group drivers.
- B. Effect of seasonal adjustment.
- C. Final model with predictions.



Red line = ACA implementation, Blue dashed = repeal, Black = end of data set.

Ongoing & Future Projects

- Refining of the tdtk-package.*
- Cleaning/continued report development.
- Testing of the report generation.
- Inclusion of PCA, random forest and neural network ML.
- Development of automated blinding/PHI removal by direct db queries.

* “No program survive first contact with the end-user”
-Sun Tzu’s Ancient Art of Programming

Conclusions

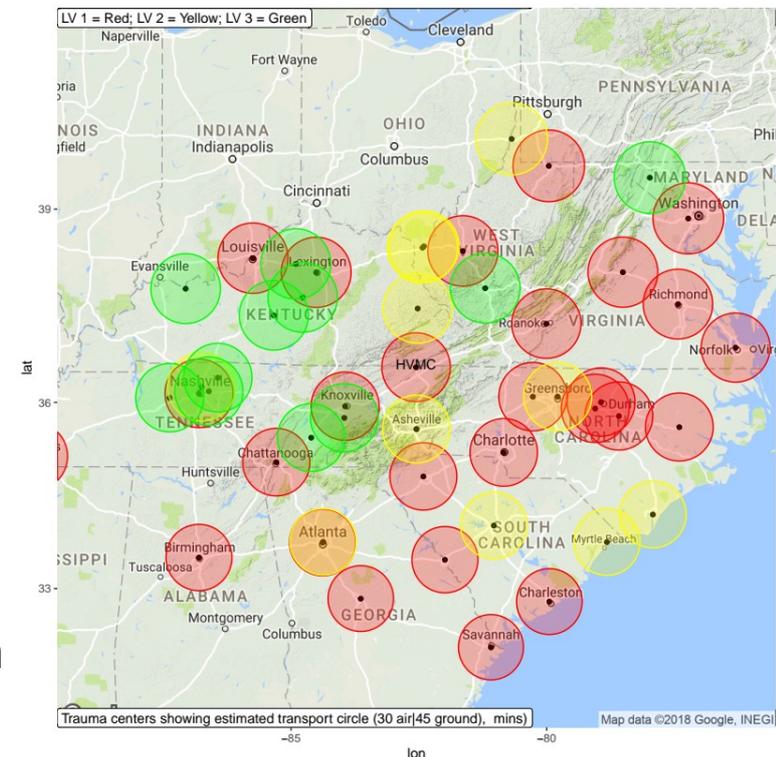
R statistical program along with the tdtk-package is a minimal program that can be used for the analysis and blinding of patient data.

It is operating system independent and R can be used in several commercially available stats software packages.

It provides a set of tools that can be used to target injury prevention.

TDTK provides a framework that can be applied to any patient data set.

Future iterations will include: potential customization reports, ML/AI, calculation of social vulnerability index and direct DB access via dbplyr-package (tidyverse ecosystem)



ACS or state, verified trauma centers with a full complement of advanced practice surgeons.

Acknowledgments

Dr. Tiffany Lasky (Trauma director - ETSU/HVMC)
Vanessa McMellan (Trauma team registrar - HVMC)
Dr. Sarah Robison (Surgery Resident - ETSU)

Free Software Foundation Team
R-Core team

R-Packages:
Tidyverse
ggmap
geosphere
Forecast

All maps were obtained using the google map api and ggmap-package. All google maps are marked with required markings (i.e. © google and data providers)

Thank You!

Questions?