TU Informatics

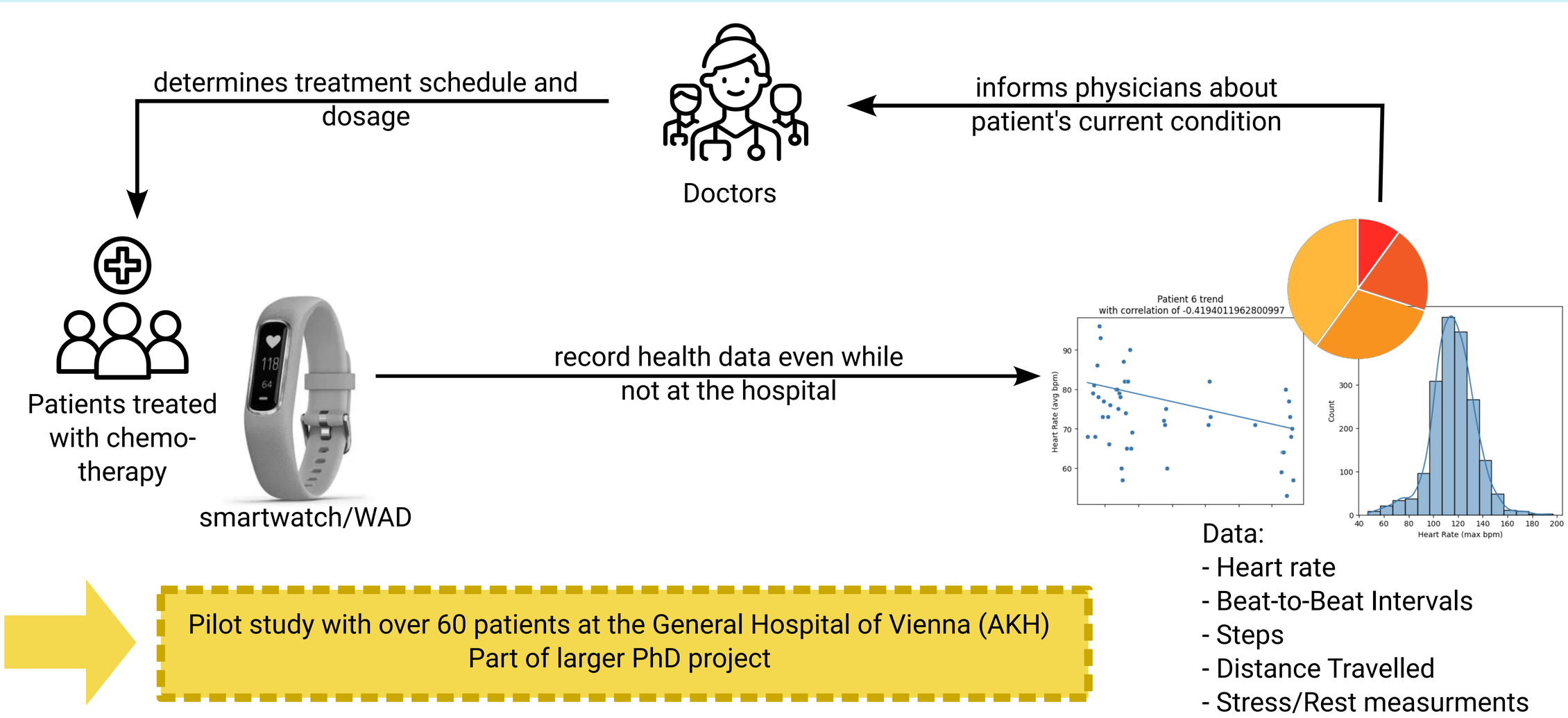
Translational Research Unit

Feasibility study for a patient smartwatch data set in cancer research

Paul Adrian Erpenstein
Data Science

TU Wien Informatics/ Medizinische Universität Wien
Innere Medizin I
Translational Research Unit
Supervisor: Anna Berghoff, Dr. Med.
Contact: e12107369@student.tuwien.ac.at

Can we detect chemotherapy side effects before they occur?



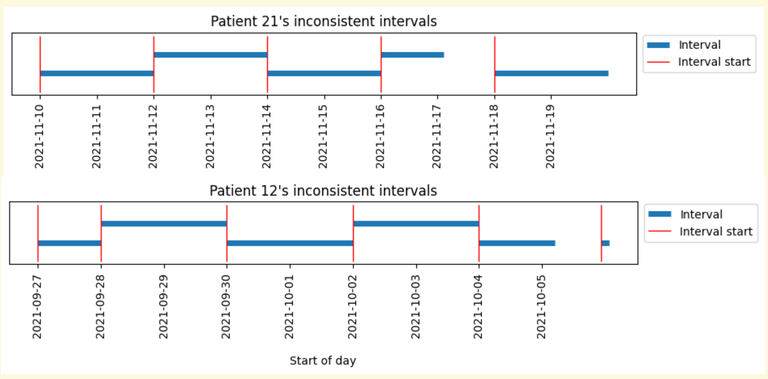
Data:
- Heart rate
- Beat-to-Beat Intervals
- Steps
- Distance Travelled
- Stress/Rest measurments

Pilot study with over 60 patients at the General Hospital of Vienna (AKH)
Part of larger PhD project

3 major challenges regarding data

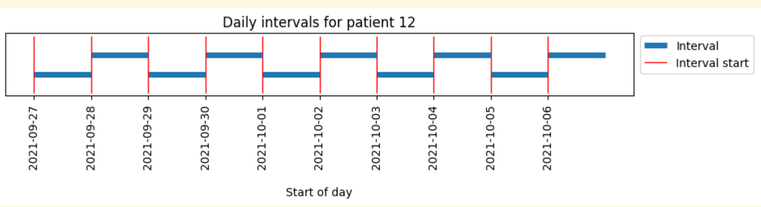
Data structure

Smartwatches recorded data in inconsistent time-intervals



Hard to use and interpret
Minutely and hourly measurements hardly relevant

Create data transformation pipeline to restructure measurement intervals into daily format



Equidistant time setps:
+ easier to read and understand
+ easier to aggregate with other clinical (mostly daily) data

Visualization

How do you communicate a large corpus of time series measurements effectively to the study personnel?

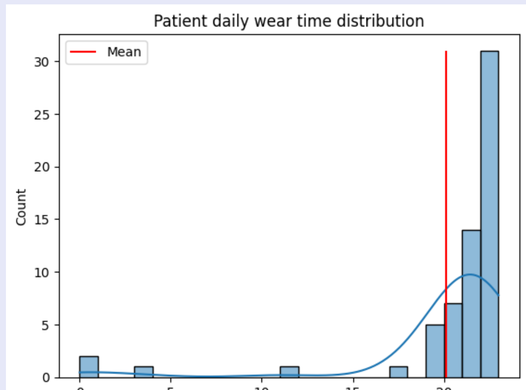
Development of interactive dashboard

Left side: Filter criteria based on medical data and patient profile
Right side: Pie chart describing the number of matching, accepted and rejected patients. Heatmap timeline showing amount of wear time on a per day basis over 6 month period.
Green: Exceeded minimum wear time
Black: No wear time at all

Each pixel respresents a day
Enormous number of days with no data!

Missing data

Data set is **extremely sparse**! Many missing days with absolutely no data. **Why is this?**



On days where the patients do wear the smartwatch they tend to wear them almost a full day!
This is because many patients simply wear the watch until it is empty and forget to recharge it! We can overcome this by utilizing an important concept in chemotherapy treatment.

Treatment cycles: Chemotherapy is administered over relatively long intervals. For example every two weeks.
A new treatment cycle starts on the day the treatment is given.

Only look at cycles, for which enough data is available. Analyze predictive factors that forecast treatment delays or dropouts.

Treatment cycle dates, delays and dropout are currently being collected by a medical student participating in the project. Assisted in developing new data model for entering cycles data.

Column name	Data type
Start date	Date
End date	Date
Type	"Cycle" or "Delay"
ECOG	0, 1, 2, 3, 4, 5
other variables	TBD

Outlook & Next steps:
- assist with data collection and integration of cycles data
- handle beat-to-beat intervals (multiple gigabytes of data)
- start creating a cycles based model that predicts treatment delays and dropout

Could be a great master's thesis!