

SLTC panel

*Clinical Text Mining – Challenges and future visions*

Lilja Øvrelid

University of Oslo

November 09, 2018

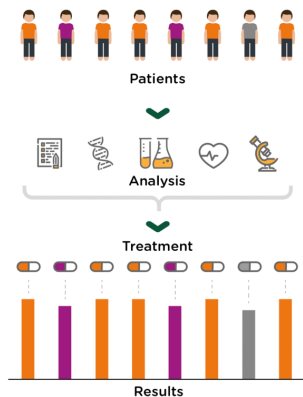
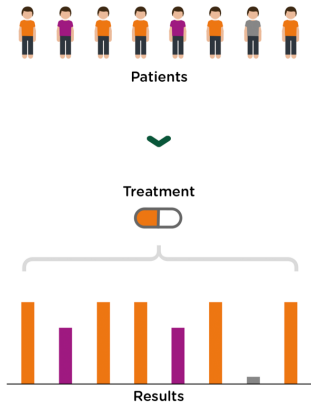


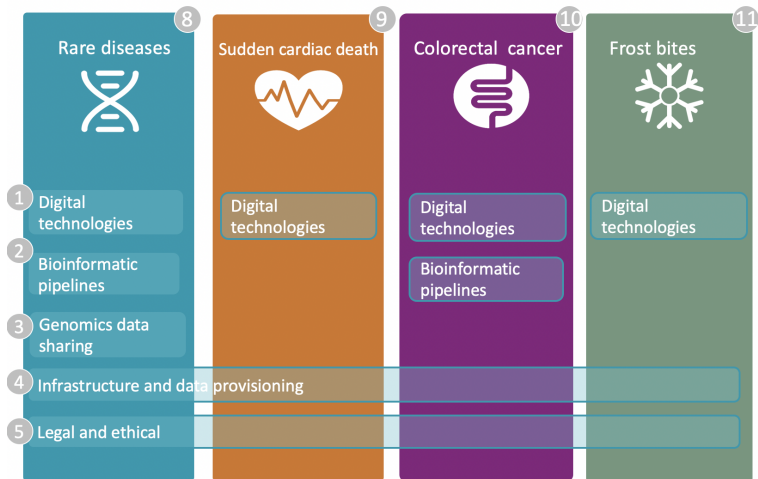


- ▶ Lighthouse project, funded by Norwegian Research Council for 3 years
- ▶ Multi-disciplinary consortium of ICT researchers, biomedical researchers, clinicians and industrial partners
- ▶ Also: law researchers, patient organizations



# Precision medicine







- ▶ Still no data access!
- ▶ Ethical board (REK)  $\Rightarrow$  Hospital information officer  $\Rightarrow$  Hospital ICT
- ▶ Preparing secure processing environment at University
- ▶ De-identification
- ▶ Development of annotation guidelines using **synthetic** data (Rama et. al, 2018)



- ▶ Clinical text mining is a truly **multi-disciplinary** effort:
  - ▶ NLP researchers
  - ▶ clinicians
  - ▶ hospital IT-staff
  - ▶ legal experts
- ▶ **Data access**
  - ▶ difficult due to data privacy regulations
  - ▶ vary from country to country
- ▶ **Data sharing**
  - ▶ few open resources apart from English
  - ▶ start from scratch with every project



- ▶ NLP focus on a variety of text types, domains and languages:
  - ▶ Cross-lingual methods
  - ▶ Domain adaptation, distant supervision
- ▶ Clinical decision support
- ▶ Image and text