Citizen-centered big data for health care, prevention and research

The Personal Health Train

Health-RI Annual Meeting 2019, 15:30 – 16:00

Andre Dekker
MAASTRO Clinic, Maastricht UMC+, Maastricht University
4th HEALTH-RI CONFERENCE - OPENING DOORS TO P4 HEALTH

Accessible research data and resources for predictive, preventive, personalized, participatory healthcare and vital citizens
Predictive & Personalized
Participatory & Preventive
Can we predict a tulip’s color by looking at the bulb?
Predicting the color of a tulip - AUC

AUC
1.00
0.72
0.50
Predicting the survival of lung cancer patients

AUC 1.00
AUC 0.72
AUC 0.50
Lung Cancer
2 year survival
158 patients
5 MDs
Prospective
AUC: 0.56 (0.41-0.62)

Oberije et al., Radiother Oncol. 2014; 112: 37–43.
People’s journey: data in- en output

Intergenerational
Personalized & organized
Where is the real world data?

Source: raconteur.net
Where is the research data?

Source: NFU data4lifesciences
Fragmentation is the keyword

The problem is social, cultural, and organizational.
## Data landscape

### Data elements

<table>
<thead>
<tr>
<th>Patients</th>
<th>Clinical research</th>
<th>Clinical registries</th>
<th>Clinical routine</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>3% of patients</td>
<td>100% of patients</td>
<td>100% of patients</td>
</tr>
<tr>
<td>104%</td>
<td>100% of features</td>
<td>3% of features</td>
<td>100% of features</td>
</tr>
<tr>
<td>285</td>
<td>5% missing</td>
<td>20% missing</td>
<td>80% missing</td>
</tr>
<tr>
<td>data points</td>
<td>240 data points</td>
<td>2000 data points</td>
<td></td>
</tr>
</tbody>
</table>
A different approach

• If sharing is the problem: Don’t share the data

• If you can’t bring the data to the research

• You have to bring the research to the data

• Challenges
  • The research application has to be distributed (trains & track)
  • The data has to be understandable by an application (i.e. not a human) -> FAIR data stations
One of the two original Health-RI pitches

• KNAW Dream Facilities April 2015 (euroCAT)
Kamerbrief over data laten werken voor gezondheid

Minister Bruins (Medische Zorg en Sport) stuurt een brief aan de Tweede Kamer over data laten werken voor gezondheid. Om data te laten werken voor gezondheid is vertrouwen nodig. De patiënt moet erop kunnen vertrouwen dat zijn/haar data wordt omgegaan en de zorgverlener moet erop kunnen vertrouwen dat data-analyses valide resultaten opleveren.
Health-RI implementation: Actions 2019

1. Transition of existing infrastructures
   TraIT, BMMRI, D4LS, PSI
   FAIR, GDPR, Stewardship

2. Transition to federated infrastructures
   Personal Health Train
Predictive & Personalized
Participatory & Preventive
Prevention of toxicities

Korte levensverwachting - geen PCI

Ervaring patient: Decisional Conflict, Control Preference, SDM

Ervaring artsen: SDM

Kosteneffectiviteit zorgpad
Personalized cardiac risk prediction

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC Score for Model with Clinical Features</th>
<th>AUC Score for Model with Lab Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framingham Model</td>
<td>0.4732</td>
<td>0.5973</td>
</tr>
<tr>
<td>Kaplan-Meier</td>
<td>0.6984</td>
<td></td>
</tr>
<tr>
<td>Aalen's Additive model</td>
<td>0.7738</td>
<td>0.7961</td>
</tr>
<tr>
<td>Weibull Hazard Model</td>
<td>0.7989</td>
<td>0.7653</td>
</tr>
<tr>
<td>Microsoft Joint Longitudinal Hazard model</td>
<td>0.8362</td>
<td>0.8102</td>
</tr>
</tbody>
</table>

Source: Microsoft & Apollo Hospitals India
Primary & Secondary Prevention CAD

Regional (Limburg) alliance of Huisartsen, CBS, SanaNet, IDS, MASTRO and MUMC+ for CARRIER - Coronary ARtery disease Risk estimations & Interventions for prevention and EaRly detection

Objective 1: Help GPs & citizens estimate risk of CAD event
Objective 2: Prevention of CAD event through coached self-care

(P1) General population
(P2) Patients at risk of CAD event
(P3) Patients with at least one CAD event

Target population for intervention

IDS: Data governance
MAASTRO: Data stations and data science

Prognostic model
Self-care with eCoach
Predictive model
SANANET

RETROSPECTIVE DATA
- CBS: socio-economic data
- Huisartsen: medical history, lifestyle
- MUMC+: clinical data & treatments

*CAD: coronary artery disease
Prevention of type 2 diabetes

The Maastricht Study: Clinical phenotypic data on 8000 individuals (25% type 2 diabetes)

CBS - Statistics Netherlands: social, economic, and environmental data on 17M+ Dutch citizens
Cancer early detection - SHERLOC
Support the citizens journey

mental health
healthy ageing
primary healthcare
healthy youth
environment & health
prevention
food & health
big data
closing the innovation gap
learning healthcare system
self management
e-health
personalized medicine
value based healthcare
Discussion

• How has data infrastructure supported or made your work/project possible?
  • It is the core of what we do

• What lessons have you learnt that you want to pass on to others with regards to research infrastructure, FAIR, data sharing, ELSI?
  • Just do it, it is NOT difficult to start small – harder to scale up

• What are things that you are unable to do but would like to do in the future when it comes to research infrastructure?
  • Connect all citizens & legal entities holding interesting data

• What could be the next steps? What do you expect from the Health-RI community?
  • Adopt & scale up the PHT for prevention use cases (citizens)
  • Sign the manifesto! -> www.personalhealthtrain.nl
Citizen-centered big data for health care, prevention and research

*The Personal Health Train*

Health-RI Annual Meeting 2019, 15:30 – 16:00