Computational Personalization: Data science methods for personalized health

Maurits Kaptein
“Providing the right treatment to the right patient, at the right dose at the right time”
Outline:

- Defining personalized healthcare
- Analysis of the Randomized Controlled Trial (RCT)
- A computational approach to personalization
Defining personalized healthcare
\{patient, time, treatment, dose\} \rightarrow \text{outcome}. 
The reward, $r$, is a function of the context, $x$, (the characteristics of the patient), and the actions, $a$, (the treatment).
arg \max_a f(x, a)
\[
\sum_{t=1}^{T} \arg \max_{a_t} f(x_t, a_t),
\]

We choose the treatments such that we maximize the reward over all treatments.
Why is this difficult?
High dimensional learning from noisy data
High dimensional learning from noisy data
Learning causal relationships
- High dimensional learning from noisy data
- Learning causal relationships
- Balancing learning and earning
The Randomized Controlled Trial
Advantages:

1. Transparent and understandable
2. Causal effects through randomization
3. Practically appealing
Advantages:
1. Transparent and understandable
2. Causal effects through randomization
3. Practically appealing

Disadvantages:
1. Examines a very small number of options
2. Poor balancing of earning and learning
3. Inability to (re-)use data after trial
A computational approach
Disadvantages:

1. Loss of transparency: black-box
Disadvantages:

1. Loss of transparency: black-box
2. Practical challenges: no deterministic choices
Disadvantages:

1. Loss of transparency: black-box
2. Practical challenges: no deterministic choices
3. Causal effects not guaranteed: need additional analysis
Disadvantages:

1. Loss of transparency: black-box
2. Practical challenges: no deterministic choices
3. Causal effects not guaranteed: need additional analysis
4. Computationally challenging
Why would we want this?
Conclusion