

Participant's thoughts on session posters:

During the Mozfest 90 minute session, participants were invited to interact with 8 posters inspired in the Matrix red pill blue pill analogy. In each poster, the blue pill contained information on the advantages that artificial intelligence, machine learning and open datasets offer to healthcare, and the red pills contained less talked about but very crucial facts that prevent us from seeing the real benefits of healthcare applied technologies.

The 25 session participants shared their thoughts, ideas, and questions in the form of post-its, that were later discussed and implemented as session learnings. The following document contains all participant thoughts on each of the posters. If you wish to replicate the session, you can access all materials at <http://bit.ly/healthmozfest19>.

1. Publicly available genomic data allows researchers from all over the world to study disease and advance science. 81% of participants in genome mapping studies are of European ancestry.
 - Why
 - What is the impact on privacy?
 - Information policies can enrich bias
 - How are we going to treat populations then?
 - More inclusive datasets might help
 - How do we build systems that encourage diverse research?
2. The Framingham risk score predicts the risk for future coronary heart disease events, but only works in a specific cohort of people
 - Possibilities of bias against black people, women, etc.
 - It is important to focus on the development of a new risk score.
 - but we have learned a lot from this study affecting population health
 - The study is still valuable for that group
 - True
 - The study should be corrected so that AI can identify heart disease properly
3. machine-learning has the potential to save thousands of people from skin cancer each year. Most programs are learning on light skin and may underperform on images of lesions in the skin of color.
 - Not an either/or but a spectrum of development?
 - Will skin color change the spectrum?
 - How do we prompt/encourage diversity of data?
 - How do we achieve more diversity without the pitfalls?

- We need more data to train the algorithms
 - We need greater feedback loops
4. AI has introduced the possibility of using healthcare data to produce powerful models that can automate diagnosis, but scientists often lack access to balanced datasets needed to optimally train algorithms and avoid replicating bias
- True
 - Exactly (+5)
 - That is the "benefit" argument we are listening. The risks? All the time
 - Greater inclusivity is clearly needed to see the true benefits of AI
 - Are Algorithms mature enough?
 - What is the trustworthiness of AI? (+1)
 - What is a balanced dataset?
 - Data maintenance will require more people
 - Can we improve the underlying data for AI algorithms?
5. The use of disease surveillance data has become crucial for public health action against new threats, including Dengue and Zika. The potential biases and lack of comparability across countries are limiting the efficient use of the data
- At what expense?
 - Sounds important to focus on, especially in countries with limited data
 - Data can be misused
 - Data management and sharing are not easy. Human resources and infrastructures are needed.
 - We need global open infrastructures
6. A study that tracked cancer mortality in the U.S. Helps public health authorities identify populations at risk. Lack of diverse research subjects is a key reason why the study predicts that black and Hispanic people are more likely to die from cancer than they really are.
- How might we validate?
 - Yes, bias.
 - Bias in datasets is a problem
 - Well being is very subjective
 - This depends on the research
 - Associated data risks must be avoided
 - Is this a "healthy version" of a minority report?
 - Is bias not explicit in the research outputs?
 - This sounds great in theory as long as feedback loops are used (+2)

7. Scientists have created an algorithm that can predict a person's risk of developing Alzheimer's disease. The algorithm only works on native English speakers and misdiagnoses those who have a different accent.
 - Prediction is not always right.
 - Questions about the algorithm (+3)
 - If they predict that I will have Alzheimer's, will I be treated unequally from now on?
 - How can scientists predict this?
 - Is the dataset that informed this diverse enough?
8. Open datasets can help us build more responsible models and achieve good health and well being, but they are often imbalanced; predominantly male and white.
 - Prediction is not the same as explaining and acting about it.
 - What about other genders apart from male and female?
 - This requires responsible governance principles and global legislation
 - This is true of closed data, too.
 - Data can also be used for evil.
 - Open datasets can be dangerous if they go in the wrong hands
 - We need governance structures
 - Open data has risk attached which need to be addressed
 - Health data is super personal
 - Exactly (+4)
 - What about private companies using data?
 - True, but is this sharing free of risks?
 - The workspace must be changed.
 - This is not surprising
 - It makes a lot of sense, this is where the money is.