AVOIDING BIAS

IN MACHINE

LEARNING ALGORITHMS

Karen Gibson

Overview

Algorithms are not always scientific fact!

Often they are:

- decisions made by a humans and embedded into code

...and this means that they could contain:

- the conscious or unconscious bias of the person who made the algorithm

- the biases of the historical data that was used to help the person create the algorithm

Disaggregating data

One size 'fits all' PPE:

- Total fail rate: 16.3%

- Female use fail rate: 18.2%

- Male use fail rate: 9.7%

Disaggregating our data tells a more equitable story.

Table 1 Filtering face piece class 3 (FFP3) respirator fit test results.

			Sex		
			Female	**Male**	**Total**
Test result	Pass	Count (%)	665 (81.8)	213 (90.3)	878 (83.7)
	Fail	Count (%)	148 (18.2)	23 (9.7)	171 (16.3)
Total		Count	813	236	1049

Source:
Respiratory personal protective equipment for healthcare workers: impact of sex differences on respirator fit test results,
British Journal of Anaesthesia:
https://www.bjanaesthesia.org/article/S0007-0912(20)30851-5/fulltext#%20

Ethical challenges

with disaggregated data

- first consider whether it is necessary for the particular findings of your data set:

    - what are you trying to achieve?

    - Is it possible without data disaggregation?

- if is also essential ensure that disaggregated data is used for ethically sound reasons, and not to further marginalise a subset of society

Approaches to these challenges

Define appropriate standards

- Build the algorithm with a training set that is representative of the population being analysed

- Create a set of benchmark criteria

- Explain how an audit should check from data bias

Audit and improve based on outcomes

Conclusion