



“To Err is Human”? Affective dimensions of AI in the clinic

Dr Sarah Chan

Reader in Bioethics, Usher Institute

Associate Director, Centre for Biomedicine, Self & Society

Deputy Director, Mason Institute

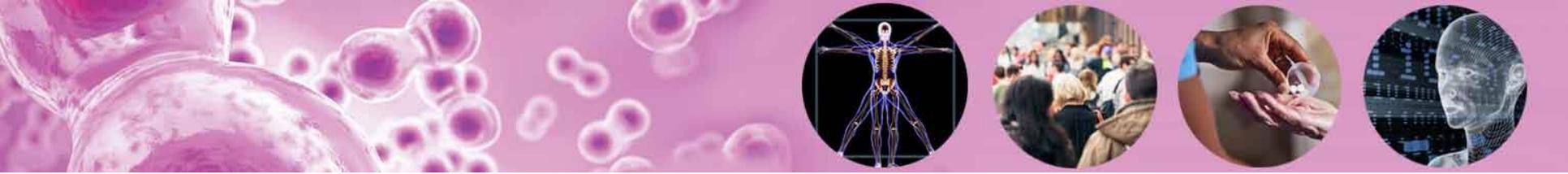


THE UNIVERSITY
of EDINBURGH



Centre for
Biomedicine
Self and Society

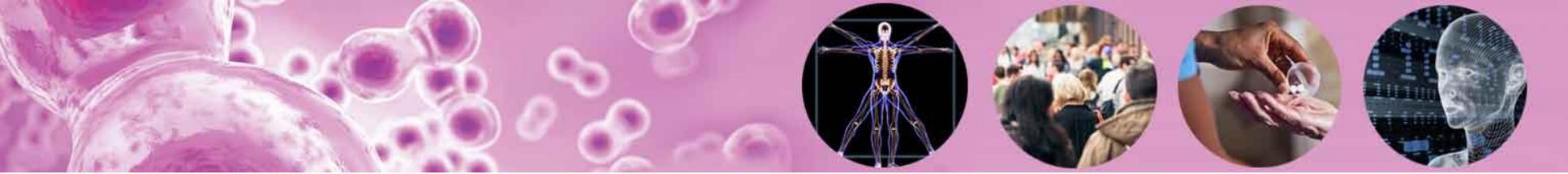




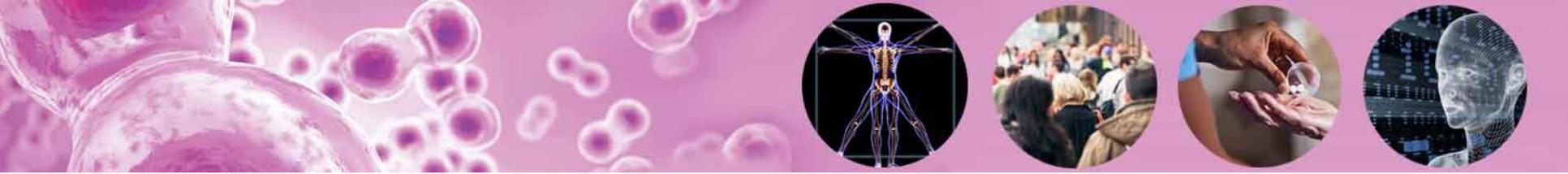
Key points



- The use of AI in health and social care has the potential to introduce new elements to care relationships
- AIs can play an important role in our moral narratives around health care delivery and health outcomes, especially with respect to decision-making, responsibility and communication
- Affective dimensions of AI, particularly physical and embodied aspects, are likely to be significant in mediating this

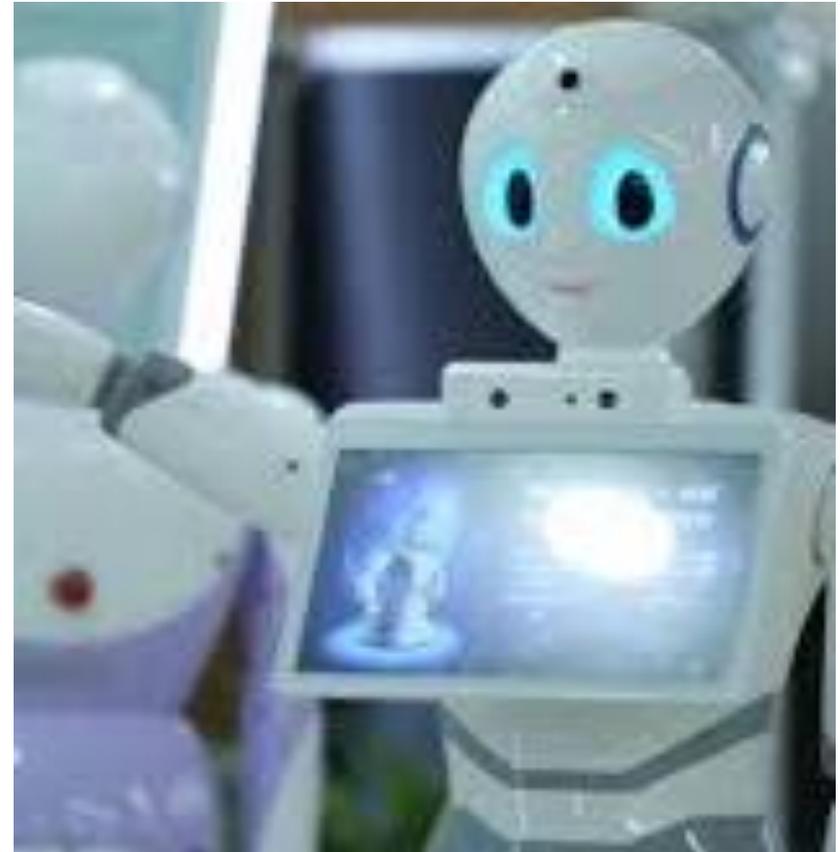


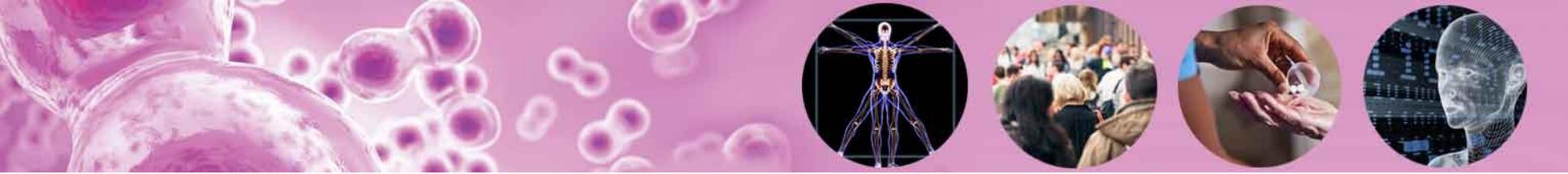
I. AI in the clinic



AI in health care: transformative imaginaries

- Data-driven technologies:
 - Prediction, diagnosis, treatment
 - Drug discovery
 - Health systems optimisation
- Patient-facing applications:
 - Medical devices and wearables
 - Robot carers, robot-assisted surgery, primary-care AI (e.g. Babylon Health)
 - Telemedicine

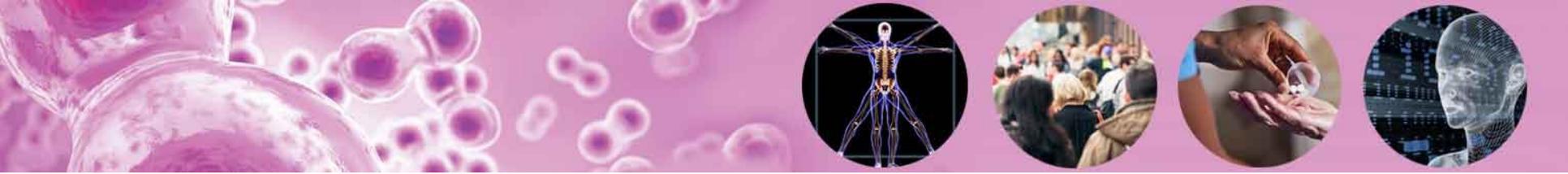




Social robots: no longer just sci-fi



- Robots capable of (stimulating) affective, embodied, emotional responses
- Social robotics: how will these new entities integrate into, influence and alter human social interactions?
- Social robots in care



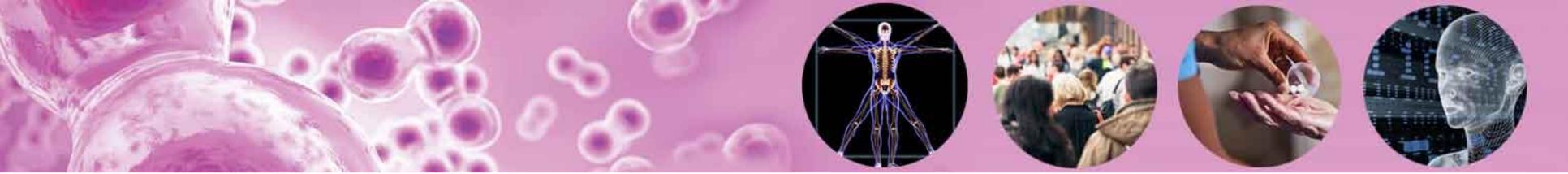
AI in healthcare decision-making

What happens if the AI gets it wrong?

- Moral and legal responsibility for AI(-assisted) medical error

... What happens if the AI gets it *right*?

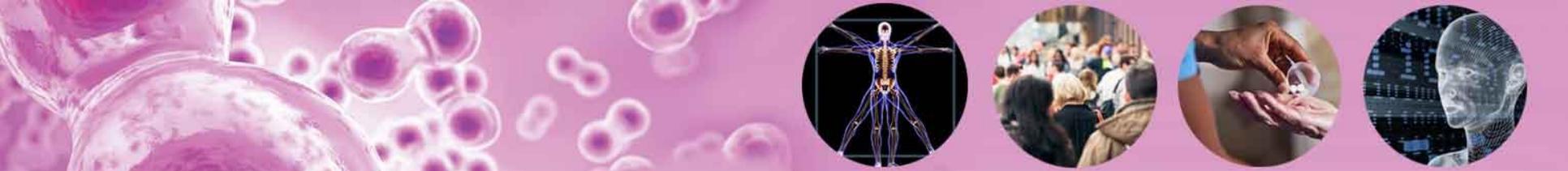
- What are the implications as AI begins to out-perform human doctors?



AI as moral agents?

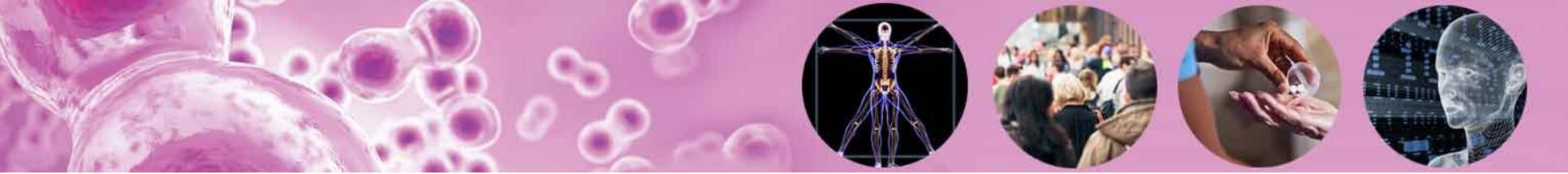
- Weak / narrow versus strong / general AI
- Current state of technology limited to “weak” AI
- Full-fledged “robot doctors” are still a long way off!
- But...
- “Computers As Social Actors” (Nass & Moon, 2000)
- Robots as participants in “dyadic moral interactions” (Swiderska & Küster, 2020)
- AI / ‘social robots’ as actors in our moral narratives





Medical decisions as moral decisions

- Medical decisions are ‘morally salient’
- Not only in terms of consequences, but the nature of the decision & relationships involved
 - Doctors as ideal (or at least better) moral agents
- D-P relationship involves a degree of delegated / substituted moral agency
- In consulting HCPs, we are implicitly asking them to assume (a degree of) moral responsibility for us
 - “What would you do if you were me?”
- Medical decisions: process of joint decision-making in the context of shared moral agency
- AI-assisted medicine: introducing a third agent into the relationship?



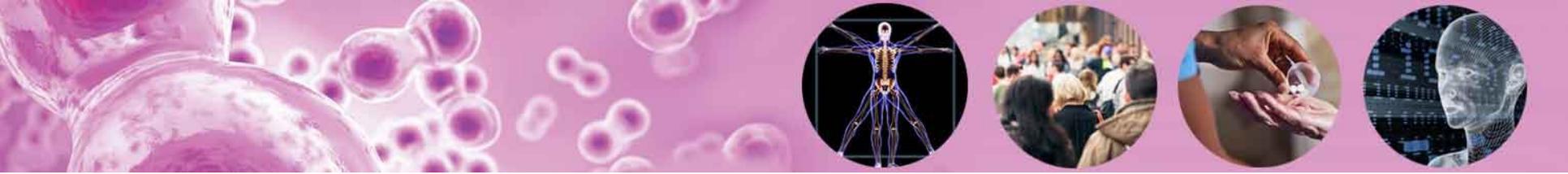
II. 'Getting it wrong'

AI, medical error and moral responsibility



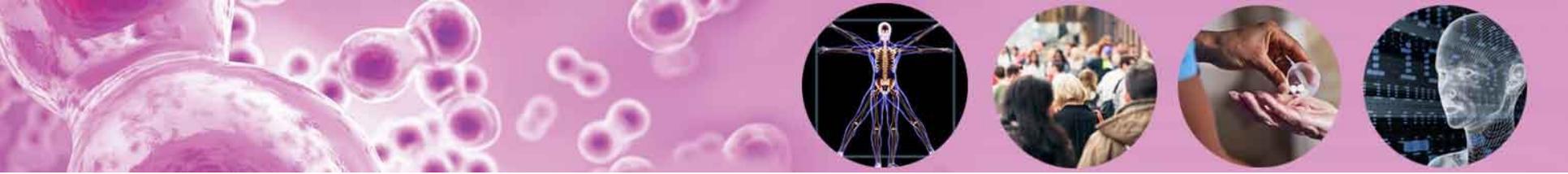
Responsibility for AI healthcare decisions

- Who is (or can be held) responsible for medical error resulting from AI(-assisted) decisions?
- “Should we sue our robot doctors?”
 - (Would it be desirable for us to be able to...?)
- *Pace Tigar*d: AIs do not need to be true moral agents in order to have ‘functional morality’ and to be held responsible
- Are there reasons to attribute responsibility to AI in this way, in a healthcare context?



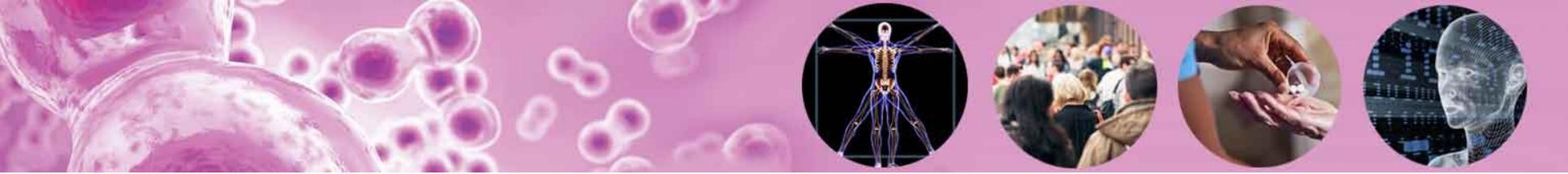
Holding robots responsible

- Basis for legal liability: “duty of care” owed by doctor to patient
- ‘Care’: not only legal but moral concept; relational
- Ability to create ‘relationships’ with AI / robot doctors: (perception) of moral responsibility; expectations of care
- Assigning responsibility: having someone to blame?
 - Satisfaction of ‘moral narrative’, even where not rational / consistent
 - Legal justifications of retribution and satisfaction
- Robots as blame-takers?
 - Function of blame-taking in medical settings (Tigard 2018)
 - Could being able to blame robots serve a useful role?



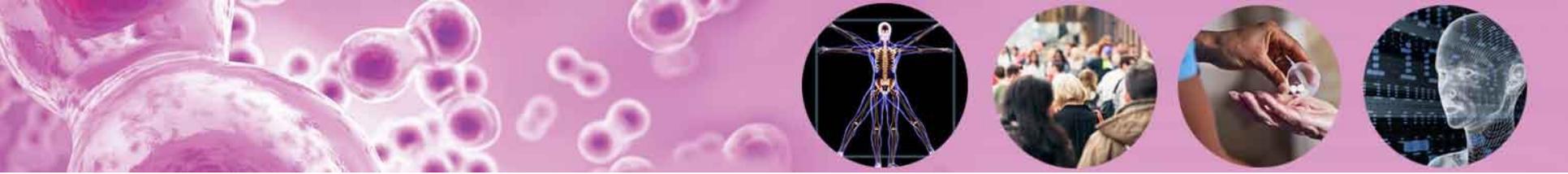
Human factors and AI explanations

- Importance of explanations in AI
 - EC HLEG on AI: principles include ‘explicability’ (trustworthiness + explainability)
 - GDPR “right to explanation”
- *Algorithmic* explainability
- But algorithmic reasons are not moral reasons
- Explanatory role of robots as responsible ‘agents’ within moral narratives
- Is shared humanness a necessary part of making explanations or reasons for moral decisions legible and hence trustworthy?
- If so, what aspect of humanness is at stake?



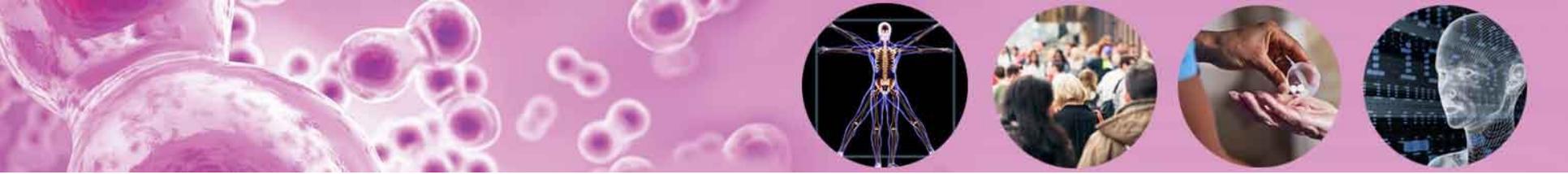
‘Gut instincts’? Embodiment & explainability

- What forms the basis of trust in human doctors and their decisions?
 - Medical expertise is embodied and physical as well as cognitive
- Human doctors and human patients share fundamental aspects of embodiment
- Might decisions made by AI / robot doctors be different because of their different embodiment?
 - Embodiment → values?
- Or might *perceptions* of different embodiment diminish trust in AI decisions?
- Implications for AI and social robot design?



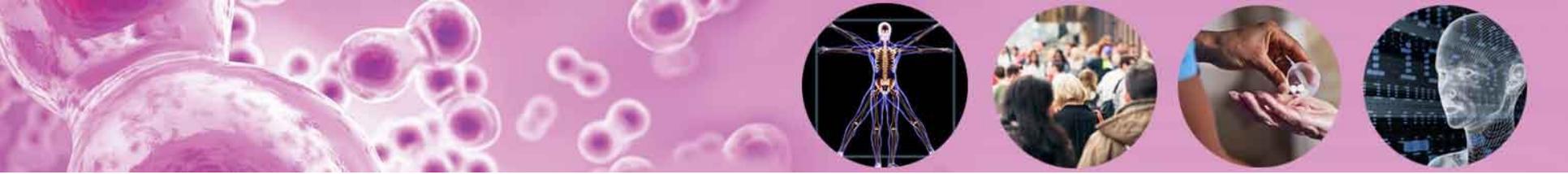
III. 'Getting it right'

when AI is better than human?



Can AI improve on human medicine?

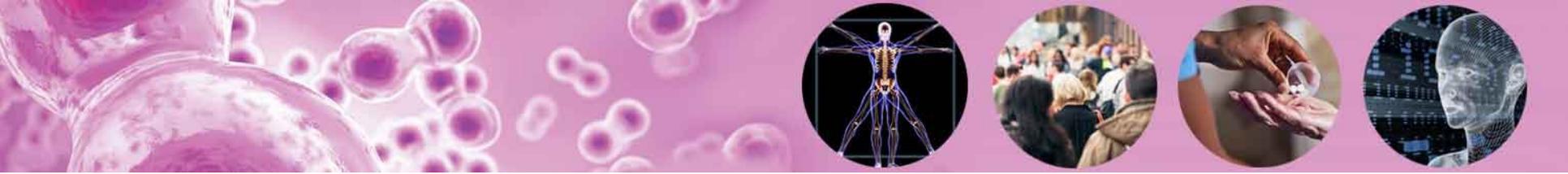
- ‘Big data’, predictive and personalised medicine
 - Could AI be used to predict disease earlier or more accurately? Or prescribe more effective treatments?
- Digital (computational) pathology
 - Algorithmic analysis is beginning to outperform human pathologists in some areas (Acs *et al.* 2020; Cui & Zhang 2021)
- Implications?
 - Some human roles in health care might become redundant
 - Might remove choice about whether to use AI – liability for failure to use ‘gold standard’
 - Would you ever want your doctor to be wrong?



When being wrong is a good thing?



- “Man told he’s going to die by doctor on video-link robot”
(<https://www.bbc.com/news/world-us-canada-47510038>)
 - “I look up and there’s this robot at the door...”
 - (Telemedicine not AI – but robot has presence in the narrative)
 - Merely a matter of communication?



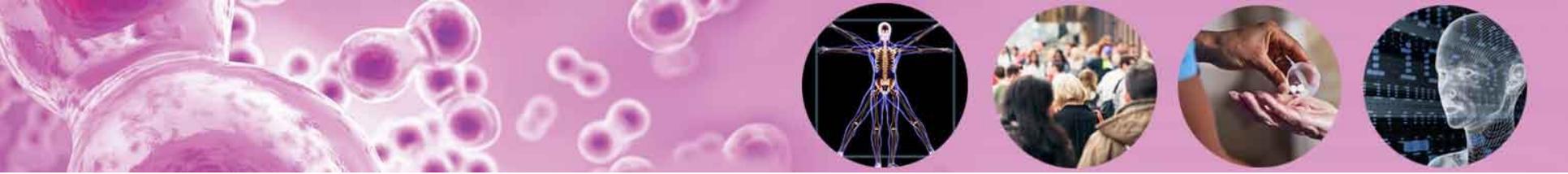
The value of fallibility?

Are you sure?

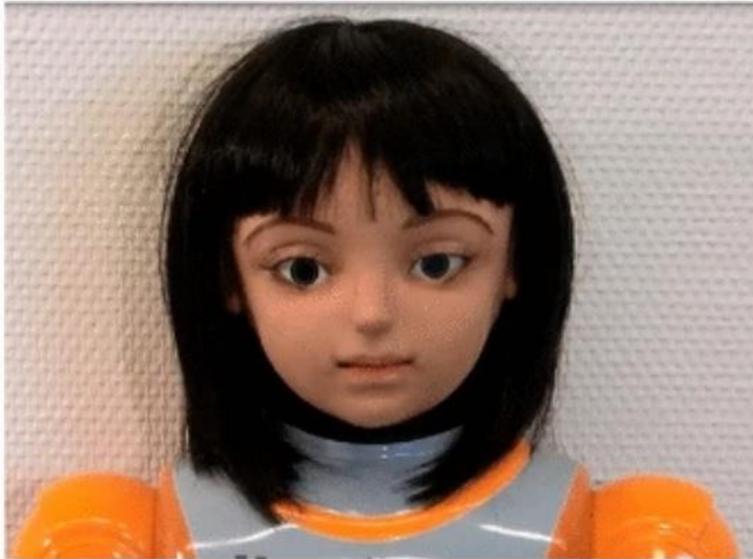
Could it be a mistake?

- How might receiving an adverse diagnosis from an ‘infallible’ AI be different to receiving it from a human doctor?

- Who should break the bad news?
 - Human: affective, emotional... fallible?
 - Shared vulnerability
 - Possibility of error permits narrative negotiation – time to “come to terms”
 - Cf ‘narrative vulnerability’ (Scully)
 - AI / robot: ??

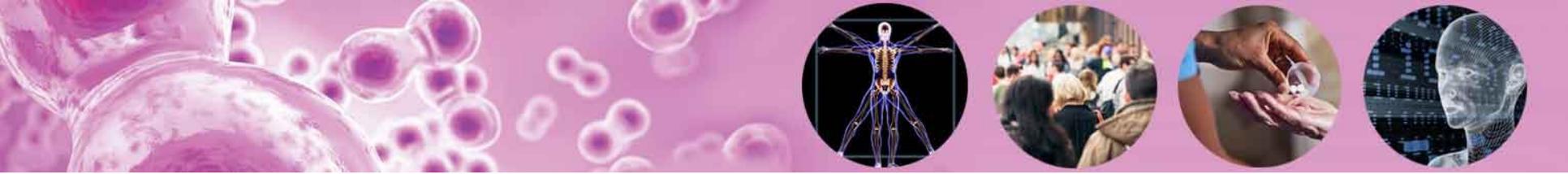


AI, affect and emotion

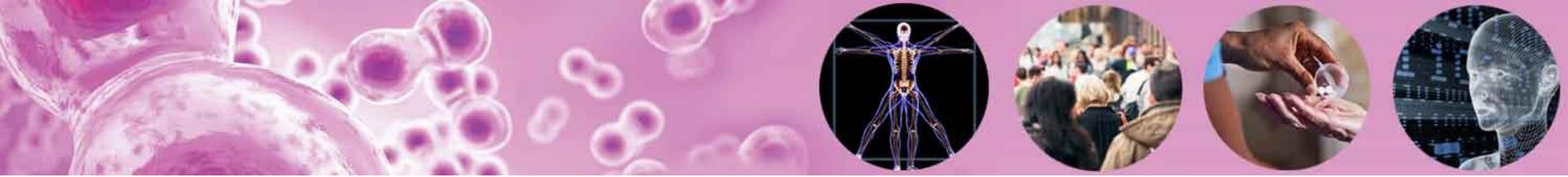


Hanson RoboKind Alice R50 bringing bad health news

- Alternatively: might using robots to communicate bad news allow for affective / emotional privacy?
 - More positive responses to robots bringing bad news compared to human doctors? (Hoorn & Winter, 2017)
- Various aspects of decision-making and communication
 - Potential for AI / robotics to be used in different ways to mediate these

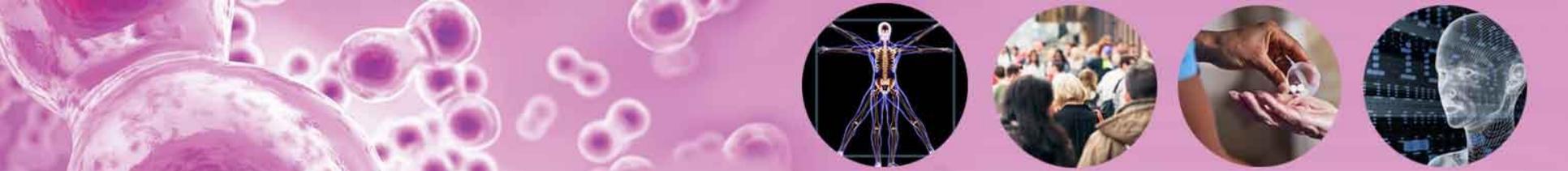


IV. Concluding thoughts



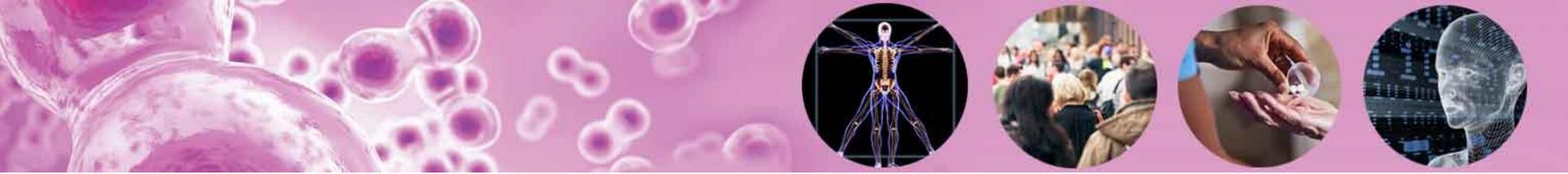
New configurations and relations

- From doctor-patient... to doctor-patient-AI
- “AI technologies and robotics not only re-materialise the boundaries of the human and the machine in affective and relational ways that challenge old distinctions and binaries between the artificial and natural, rational and emotional, and human and non-human, but they do so by augmenting and, indeed, changing human capabilities...” (De Togni *et al*, 2021)



AI decisions, rights & wrongs

- Significance of ‘human factors’ in medicine
 - Cf right to human review of algorithmic decisions
- Why do we want a human element?
 - because we fear machines are fallible?
 - Or because they are not fallible *enough*? (Or in the wrong ways?)
 - Cf concerns about posthumans: “... their values would not be ours...” (Agar, 2011)
- Importance of human(-like) embodiment in decisions?
 - Shared embodiment as component of shared moral values?
- Fallibility as (intrinsically) human?
 - Cf moral enhancement and “freedom to fall”
 - Is the ‘right’ decision the one a human would have made?
- AI as enabling affective, emotional and relational distance?



In favour of invulnerability

- Biological invulnerability: an advantage of robot carers?
- How to negotiate affective and emotional aspects of care relationships?

Global development

RoboDoc: how India's robots are taking on Covid patient care

The pandemic has spurred on robotics companies building machines to perform tasks in hospitals and other industries

- [Coronavirus - latest updates](#)
- [See all our coronavirus coverage](#)

Global development is supported by **BILL & MELINDA GATES foundation**

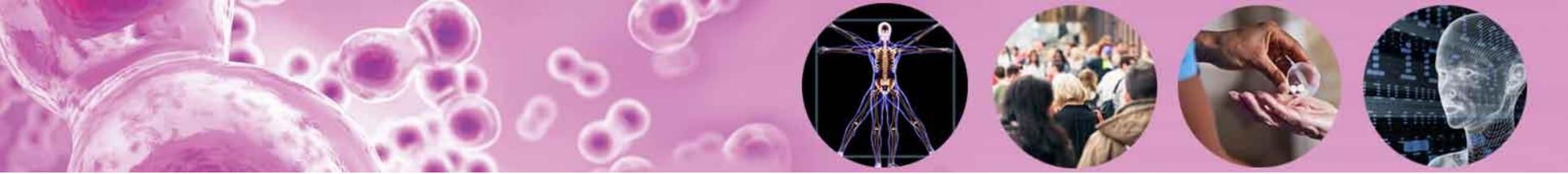
About this content
Kalpana Sunder
Wed 2 Dec 2020 06:00 GMT

217



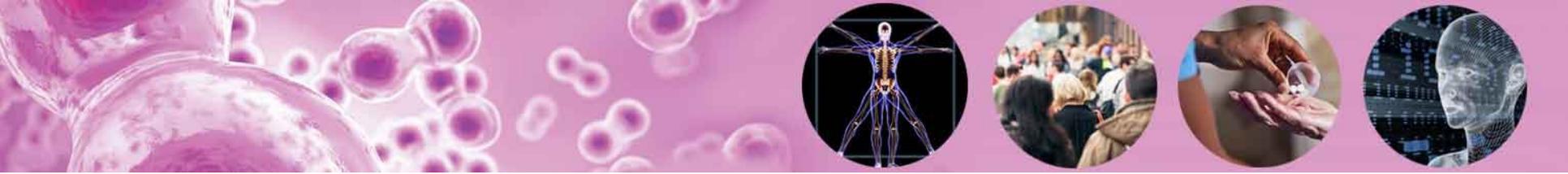
▲ Robots such as Mitra are being used to reduce risk of infection for medical staff. Photograph: Adnan Abidi/Reuters

<https://www.theguardian.com/global-development/2020/dec/02/robodoc-how-india-robots-are-taking-on-covid-patient-care-mitra>



Summary and conclusions

- AIs may have an important role not just in health care, but our moral narratives surrounding it
- Their capacity to fulfil this role will depend, at least partly, on the features that enable us to relate to them in certain ways
- ‘Narrative explainability’:
 - “things happen for a reason”
- ‘Narrative adaptability’:
 - does fallibility and the capacity for error provide space for negotiation?
- Affective, embodied and relational dimensions of AI will mediate this
- Robot doctors and health-related AI as *socio-ethical imaginaries* for (re-)envisioning values in health care



Thank you!



sarah.chan@ed.ac.uk



@sarahwchan

Special thanks to my colleagues Giulia De Togni, Sonja Erikainen and Sarah Cunningham-Burley (Wellcome Trust Seed Award “AI and Health: Exploring Affect and Relationality Across Three Sites of Intelligence and Care” [213643/Z/18/Z])