Dr Spiros Denaxas, Senior Lecturer at Institute of Health Informatics, UCL

Spiros is a leading biomedical data scientist operating at the intersection of computer science, clinical research and healthcare delivery. He leads the UCL Institute of Health Informatics Data Science Lab and his research focuses on applying and evaluating machine learning and data-driven methods for analyzing big biomedical data and answering clinically-significant questions. He has a proven track-record of understanding and translating real-world questions and challenges into statistical learning problems across industries. He has demonstrated success in developing cross- functional collaborations by bringing together stakeholders across engineering and healthcare and succinctly conveying complex technical concepts to non-technical audiences. His research has informed international clinical practice in the prevention and treatment of millions of patients with, or at risk of, cardiovascular disease.

SUMMARY OF EVIDENCE

Good afternoon and thank you for the opportunity to address this All Parliamentary Party Group meeting on Artificial Intelligence on the vision of a data-enabled, AI-driven future.

My name is **Spiros Denaxas** and I am a computer scientist by background. I am currently an Associate Professor of Biomedical Informatics at University College London. I am also an investigator in the newly-created national Health Data Research UK Institute (https://www.hdruk.ac.uk/). HDR UK is a joint investment led by the Medical Research Council, together with the National Institute for Health Research (England), the Chief Scientist Office (Scotland), Health and Care Research Wales, Health and Social Care Research and Development Division (Public Health Agency, Northern Ireland), the Engineering and Physical Sciences Research Council, the Economic and Social Research Council, the British Heart Foundation and Wellcome.

My **lab** (http://denaxaslab.org/) focuses on analyzing electronic health records which are data that are generated during our interactions as patients with the NHS such as for example when you visit your General Practitioner or undergo surgery in hospital or end up in A&E. We analyze these data using, amongst other, machine learning and AI approaches, to do everything from uncovering hidden patterns of diseases, to

building better tools to predicting when somebody is more likely to develop diabetes or to be admitted in hospital with a heart attack.

We have heard from our expert colleagues on the ethical considerations, legal challenges and technical standards so my evidence today comes as someone who is on the ground, building and applying some of the approaches for analyzing large amounts of health data for improving human health and healthcare. Importantly, I also provide evidence today as somebody who is educating the next generation of data-researchers in my capacity as Programme Director for the MSc in Health Data Science at UCL. This next generation of researchers will eventually be the enactors and consumers of these approaches.

What will an AI-driven healthcare system look like in 2052? Unsurprisingly, this is a very hard question to answer…! There are broadly two schools of thought around what an AI future holds, the utopian one and the dystopian one. If I were to paint an image of a utopian NHS, it would be a **rapidly learning** system where every single patient interaction is a learning point. The NHS will be using AI for converting huge amounts of heterogeneous health data into actionable knowledge, enacting this knowledge and rapidly evaluating its impact in terms of treatment, prevention and measurable benefits to patients. This process, which at the moment takes years, should in the future take minutes and enable the delivery of personalized, efficient and safe care to everybody.

Today, I would like to raise just two of the short-term challenges that we face today on our path to a rapidly learning NHS:

1. Data access
At the heart of AI, or any other data-driven approach for that matter, lie data. To achieve a rapidly learning healthcare system, it's critical to have streamlined access to the data itself at a national level. The UK has the unique potential of having high-quality, longitudinal health and social care data from birth to death on a national level, for 65 million citizens. At the moment, these data are siloed which make access to both academia and industry researchers very challenging. Our definition of what these data are is also likely to change with whole genome sequencing becoming a commodity, where all patients get sequenced at birth or with patient's bringing their own data from mobile phones, smartwatches or home sensors and integrating them with the healthcare system.

2. Strengthening the talent pipeline
The skills required to build and evaluate reliable and responsible AI approaches, especially in healthcare, are much broader and diverse than STEM subjects that the current focus is on.T he focus is often on maximizing algorithmic accuracy without considering the implications of deploying the algorithm in routine care early on. Future AI researchers should not only have the solid technical background but also have a solid understanding the underlying healthcare processes that generate the data which in turn would enable them to understand recognize the biases that exist how to address them. Similarly, clinicians should be aware how the data they record and generate during care effectively influence the care they will deliver in the future. We should be encouraging lifelong learning, tackling new emerging challenges using new emerging technologies through a flexible and adaptable curriculum or we risk, as we

say in machine learning, to overfit i.e. end up with a solution that is good now but will not adapt to the future needs and challenges.

To overcome these challenges, my recommendations are:
• The UK should Invest in a robust **national infrastructure which empowers patients** to see what data are recorded, when, by who and how are they being used for their care. Patients should be aware when a decision with regards to their care is based on an algorithm or not. Furthermore, patients should be able to decide who uses what data for research and what are the benefits of such research to their health or the health of others.
• Comparably to genomics, the branch of molecular biology concerned with the structure and evolution of our genetic material, we need to establish a complementary **phenomics discipline** to systematically investigate, refine and redefine disease classification from information about health states (things that are observable, e.g. symptoms), health traits (things that can be measured, e.g. blood pressure) and their variation over time.
• We should **proactively promote capacity building in skills from relevant social sciences** that are required to understand the ethical, social and political challenges of AI in healthcare. In parallel, we should advocate and **reward reproducible research**, enable the sharing of algorithms and the use of open standards in research.
• Finally, we should encourage and reward the next generation of health researchers to go out there and co-design, **co-create AI tools and approaches with clinicians, patients and the public**.

To conclude, the potential of an AI-driven healthcare system is simply fascinating. The NHS will be able in the future to rapidly learn and act using large amounts of heterogeneous data and deliver better, personalized safer care.