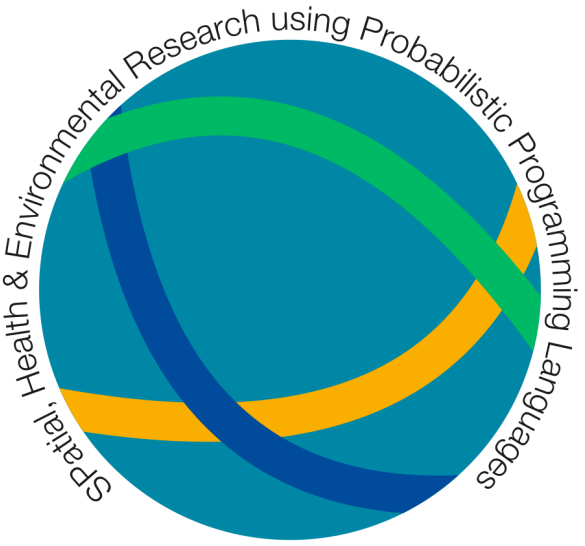
Health Modelling & Forecasting Workshop Report

The Alan Turing Institute – Thursday 23rd January 2025



SPHERE-PPL

Executive Summary

AI and Data Science are at the heart of the UK's plan to enhance the NHS and improve health outcomes for its population. The SPHERE-PPL Team brought together 50 key stakeholders from across the Health sector, including representatives from UKHSA, Cancer Research UK, several NHS Hospitals and ICBs, alongside Data Science experts from The Alan Turing Institute & The Jean Golding Institute. The day workshop saw attendees highlight key challenges within the Health, leading to the design of several forecasting contests that will have direct impact through focussed outputs. Contests include: modelling NHS Capacity & OPEL 4 events using several multi-year datasets, predicting Lyme Disease cases for UK regions, and a narrative competition utilising NHS Urgent & Emergency Care data to find new insights and interpretations to aid evidence-based decision-making. The workshop ended with a discussion on targeted training and support required to supercharge stakeholder analytical capability. The resounding message throughout the event was the clear appetite for stakeholders to provide data for contests and for the forecast predictions to be integrated into decision-making as soon as possible.

Table of Contents

Why Health Forecasting?

AI & Data Science are set to revolutionise the way we understand health & disease, and improve health outcomes. By leveraging increasing numbers of large multi-year datasets, forecasts should be able to provide predictions with greater levels of details and accuracy to more complex problems. A limiting factor in their widespread use is difficulty connecting stakeholders with advanced analytical capability. SPHERE-PPL aims to connect stakeholders with interested analytical teams through workshops and forecasting contests. This democratisation of datasets and methodologies will lead to enhanced predictions and increase the use of forecasting and modelling for data-driven decision-making. Within our health framework, we have three key themes: Disease, Healthcare Operations, and Complex Data.

Key Highlights & Insights

Activity 1 - Key Challenges in Health

The first activity involved teams discussing the challenges they face in their respective fields and how forecasting and modelling might aid in finding solutions. We have synthesised the inputs into challenge statements below, including overarching questions relating to the topics:

Disease

- **Vector-Borne Disease & Climate Change:** How is climate change impacting the spread and severity of vector-borne diseases, and what are the most effective mitigation strategies?
- **Health Data Biases & Disparities:** What are the prevalent biases in health data and how do they perpetuate existing health disparities? How can we ensure data collection and analysis is equitable and representative?
- **Health Zone Boundaries & Disease Control:** How do changes in health zone boundaries impact disease surveillance and control efforts, and what are the optimal strategies for managing these transitions?
- **Antibiotic Resistance Patterns & Treatment Efficacy:** What are the emerging patterns of antibiotic resistance, and how can we develop innovative strategies to combat this growing threat to public health?

Healthcare Operations

- **Predicting Serious System Events & Resource Allocation:** How can we accurately predict serious system events within healthcare settings to proactively allocate resources and prevent adverse outcomes?
- **Public Sector Outcomes Data & Performance vs. Value:** How can we bridge the gap between performance targets and meaningful outcomes data (e.g., value for money, cost efficiency) within the public healthcare sector?
- **Mental Health Concerns & Root Causes:** What are the underlying drivers of mental health concerns, and how can we develop targeted interventions that address these root causes?
- **Hospital Systems & Confounding Factors:** What are the confounding factors within hospital systems that impact patient outcomes, and how can we account for these factors in data analysis and quality improvement initiatives?
- **Intervention Impact & Hospital Collaboration:** How can we effectively scale successful interventions from a single hospital to a network of hospitals, ensuring consistent and impactful results across multiple settings?
- **Maternal Mortality & Black Women:** What are the specific factors contributing to the 2-5x increase in morbidity and mortality for Black women during pregnancy, both within the UK and globally, and what targeted interventions can effectively address this disparity?

Complex Data

- **Data Sharing & Organizational Collaboration:** How can we overcome the inconsistencies in data sharing between healthcare organizations to facilitate seamless collaboration and improve patient care?
- **Systematic Errors & Data Reconciliation:** How can we identify and address systematic errors arising from discrepancies between real-time, disaggregated data from trusts and aggregated linelist data, ensuring data accuracy and reliability?
- **Experimental Design & Data Analysis:** How can we move beyond traditional regression analysis and design innovative experiments that leverage complex healthcare data to generate actionable insights?
- **Communicating Results & Uncertainty:** How can we effectively communicate complex healthcare data results, including uncertainty and potential limitations, to diverse audiences in a clear and impactful manner?

Activity 2 - Forecast Challenge Design

The second activity took the challenges identified by the teams and aimed to produce forecasting contest plans, including an exam question, usable data and output format.

Disease

- **UK Lyme Disease predictions at regional spatial scale**
    - PHE & UKHSA datasets covering multiple decades
    - Baseline output could be a 10 year forecast with additional stratification for different local authorities
- **Species Distribution Modelling of Asian tiger mosquito (Aedes albopictus) to 2050**
    - Global Biodiversity Information Facility datasets of occurrences
    - WorldClim historic climate and weather data
    - Interactive species distribution map showing probabilities

Healthcare Operations

- **Predicting OPEL4 Events in Hospitals**
    - Operational Pressures Escalation Levels (OPEL) Framework
    - Dataset from BNSSG ICB with a binary outcome for multiple years and 3 hospitals
    - Output of a rolling 10 day forecast that can be deployed as an app on live data servers
- **Predicting 111 usage at differing spatial scales (i.e. England, Commissioning Region, ICB)**
    - 5-years of 111 data is freely available and is broken down at different scales.
    - Prediction for multiple horizons (daily, weekly, monthly)
    - Could integrate into other studies looking at demand and capacity of mental health services

Complex Data

- **Unearthing hidden narratives and trends in Urgent and Emergency Care (UEC) datasets**
    - We have access to a very large, multi-variable Urgent and Emergency Care dataset that may hold novel and hidden insights
    - This contest would revolve around building a data story and understanding its impact to acute health services

Activity 3 - Training, Support & Events

We closed the event with a discussion on how SPHERE-PPL can best support the community, through training, events and opportunities. We have split the suggestions out into approximate time horizons for when we will be able to get the ideas up and running.

Implementation Timescale	Opportunities
Short Term (<6 months)	• Integrated training opportunities involving applied projects • Forecasting and modelling pipeline examples • More online resources & events to maximise attendees • Reading group
Medium Term (6-12 months)	• PhD & Postdoc events • Technical Questions Help Desk • Stakeholder engagement sessions • Writing white papers • Multiple choices of training in different PPLs • Lighting Talks based on specific problems/challenges
Long Term (12 months +)	• Mentor Scheme for modelling and forecasting that provides impact for individuals • Cross academia-civil servant training • Horizon scanning / what's coming next? • Careers page for SPHERE-PPL Website

Next Steps

SPHERE-PPL will host a minimum of 3 health-focused forecasting challenges in 2025, covering all three themes of our heath framework. Our plan is to organise a series of online sessions in early March to introduce everyone to the contests and how to take part. We will then run contests from mid-March in 3 month blocks, taking us through to the end of the year!

If there are any questions, please don't hesitate to email us at info@sphere-ppl.org.

Attending Organisations

- Atkins Realis
- Cancer Research UK
- Centre for Cancer Screening, Prevention and Early Diagnosis, Wolfson Institute of Population Health
- Centre of Systems Modelling and Quantitative Biomedicine University of Birmingham
- Imperial College London
- Institute of Health Informatics, University College London
- Jean Golding Institute
- JUNIPER Partnership
- London School of Hygiene & Tropical Medicine
- NHS Bristol, North Somerset & South Gloucestershire Integrated Care Board
- NHS Patient and Public Involvement
- Queen Mary University of London
- The Alan Turing Institute
- Turing-Roche Partnership
- UK Health Security Agency
- University College Hospitals London (NHS Trust)
- University of Bristol
- University of Hertfordshire
- University of Leeds
- University of Liverpool
- University of Oxford