1	1. Optimising data quality for clinical decision support (CDS)	short term (1-2 years)	Our big data analytics platform, iCARE, will provide secure and governed access to patient- level linked longitudinal data to enable research addressing data fragmentation across electronic systems and improvements in EHR data quality. We will link these data with Imperial Health Knowledge and Tissue Banks and will have contributed algorithms and knowledge into the HDRUK Phenotype Library (CALIBER/Phenoflow). We hypothesise that natural language processing (NLP) models can improve EHR data quality for appropriate/timely clinical decision-making, precision medicine and better patient outcomes. We will have tested whether NLP algorithms can derive pertinent information from clinical narratives/structured information fields held across multiple electronic systems (collaboration: SAVANA), and we will

3	3. Increasing real- world clinical trial capacity	short term (1-2 years)	There is a need for randomised trials to be conducted in conditions that are closer to usual clinical practice. Based on studies such as the Salford lung study, we hypothesise that our integrated healthcare records could support a system to facilitate randomised clinical trials and generate more effective real-time safety monitoring and data analytics to support the whole clinical trial lifecycle. NorthWest EHealth Limited (NWEH) are leaders in the field of electronic health records (EHRs) enabled randomised clinical trials (RCTs). They provide a feasibility and recruitment platform (FARSITE) and a clinical trials platform (ConneXon) which enable rapid feasibility and efficient recruitment, capture of EHR data for clinical trials, more effective real-time safety monitoring and data analytics to support the whole clinical trial lifecycle. They have a successful track record in developing the technology behind the Salford Lung study. Working with NWEH and Imperial College Health Partners, we will (1) expand our consultation with patients and citizens on acceptable use of their healthcare records for clinical trials held within the iCARE environment and (2) deploy clinical trial technologies for the design, recruitment and conduct of studies embedded within digital primary and secondary care clinical records.
4	4. Embedding Clinical Decision Support within the Electronic Health Record	short term (1-2 years)	Objective 1 deliverables will have demonstrated the ability to improve Cerner-EHR data quality to create a less fragmented and more usable EHR record. Building on this, we will use EHR-embedded CDS for near-real time patient risk assessment towards a reduction in patient harm events (venous thromboembolism (VTE)/falls) and risky-behaviours (smoking/alcohol). We will develop, test, and have deployed rule-based and machine learning algorithms to deliver (1) accurate and timely identification of patients with unhealthy smoking

5	1. Automated interpretation of continuous cardiotocography (CTG) to improve maternal and foetal outcomes	medium term (2-3 years)	Foetal heart rate traces, cardiotocography (CTG), are challenging for health professionals to interpret visually in a continuous manner during labour. However, overinterpretation can lead to unnecessary invasive intervention (caesarean sections/instrumental deliveries); under interpretation is associated with foetal morbidity and mortality. We hypothesise that a reliable, automated AI tool can recognise and highlight

7	3. Optimising health and social care systems	medium term (2-3 years)	Health and social care systems are struggling to provide timely, safe and efficient services, particularly during a pandemic and through the post-pandemic recovery. We will develop and apply integrated machine learning and optimisation approaches for patient journeys through the healthcare system, with emphasis on improving efficient delivery of high quality and safe care and reducing inequalities.. Our approach will be data-driven and draw from quantitative logistics applications, having been recently successfully applied to hospital care prioritisation. For two exemplar patient groups (Subsets of people with mental ill-health and complex elderly patients), we will (1) use data from our Integrated Care System to complete a descriptive analysis of patient journeys through the health and social care system,

9	2. Interpretable AI applications for addressing priorities in cancer care	long term (4-5 years)	Current priorities in cancer care are driven by rising numbers of cancer patients, increased case complexity and expanding array of treatment options, and the need to improve UK cancer survival rates to be in line with the best-performing countries. We hypothesise that interpretable artificial intelligence (IAI) tools can address these priorities, by supporting (1) early diagnosis of people with suspected cancer, and (2) the identification of complex cancer cases to create adequate time for discussion of cases where it is needed and ensure the most effective use of valuable clinical and diagnostic time, in line with NHS England's guidance for cancer alliances 'Streamlining Multi-Disciplinary Team Meetings' (MDTs). Exploiting our AI testbed in iCARE (objective 2) we will use neuro-symbolic deep learning and natural language processing models to (1) develop IAI technology and tooling to support primary care physicians in making timely referrals for suspected cancer cases; (2) deliver an IAI-based decision-support tool to identify/prioritise complex cancer cases enabling MDT meeting streamlining. We will evaluate the acceptability and feasibility of IAI- driven decision support in cancer care, leading to a randomised controlled trial, in NWL, measuring impact on timeliness of cancer referrals and specialist review in two specific cancer patient cohorts.
10	3. Define approaches to Increasing the efficiency of trials embedded in routine clinical practice through interim monitoring.	long term (4-5 years)	Traditional randomised controlled trials are typically analysed after the last participant has completed their final follow up visit. We know that adaptive trials can increase efficiency by incorporating interim monitoring which can allow the trial to stop early when there is clear benefit, futility or harm. Trials embedded in routine clinical practice have the potential added advantage of being able to undertake this monitoring more frequenteWn9/F1 0 G, 30(d)3(ertak)-3(e)s4 Tali st3(car)9tie