Artificial Intelligence for Multiple Long-term conditions (AIM): A consensus statement from the NIHR AIM consortia

Hajira Dambha-Miller[1], Andrew Farmer[2], Krishnarajah Nirantharakumar[3], Thomas Jackson[4] Christopher Yau[5,6,7], Lauren Walker[8] Iain Buchan[9], Sarah Finer[10], Michael R Barnes[10,] Nick J Reynolds[11], Gyuchan Thomas Jun[12], Satheesh Gangadharan[13], Simon Fraser [14] and Bruce Guthrie[8]

1. Primary Care Research Centre, University of Southampton, Southampton, UK.
2. Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK.
3. Institute of Applied Health Research, University of Birmingham
4. Institute of Inflammation and Ageing, University of Birmingham
5. Nuffield Department of Women's & Reproductive Health, University of Oxford, Oxford, UK
6. Nuffield Department of Population Health, University of Oxford, Oxford, UK
7. Health Data Research, London, UK
8. Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK
9. Institute of Population Health, University of Liverpool, Liverpool, UK
10. Faculty of Medicine and Dentistry, Queen Mary University of London, London, UK
11. Institute of Translational and Clinical Medicine, Newcastle University Medical School, Newcastle upon Tyne, UK
12. School of Design and Creative Arts, Loughborough University, UK
13. Leicestershire Partnership NHS Trust, UK
14. Department of Public Health, University of Southampton
15. Usher Institute, University of Edinburgh

Correspondence: Dr Hajira Dambha-Miller, Primary Care Research Centre, University of Southampton, Southampton, SO16 5ST, Email: H.Dambha-Miller@soton.ac.uk

Abstract

Recent advances in causal machine learning and wider artificial intelligence (AI) methods could provide new insights into the natural histories and potential prevention of clusters of multiple long-term conditions or multimorbidity (MLTC-M). When combined with expertise in clinical practice, applied health research and social science, there is potential to systematically identify and map new clusters of disease, understand the trajectories of patients with these conditions throughout their life course, predict serious adverse outcomes, optimise therapies and consider the influence of wider determinants such as environmental, behavioural and psychosocial factors. The National Institute of Health Research (NIHR) recently funded multidisciplinary consortia to bring together AI specialists, experts in big data and MLTC-M in the first and second waves of this new programme. The so-called AIM consortia of researchers will spearhead the use of artificial intelligence methods and develop insights for the identification and subsequent prevention of MLTC-M. This consensus agreement is aimed at facilitating a community of learning within the AIM consortia, promoting cooperation, transparency and rigour in our approaches while maintaining high methodological standards and consistency in defining and reporting within our research. In bringing together these research collaborations, there is also an opportunity to foster shared learning, synergies and rapidly compare and validate new AI approaches across our respective studies. This step is critical to implementation on the pathway to patient and public benefit.

Scope and aim

This statement was developed by the first wave of the NIHR AIM consortia and received input from the second wave. It includes representatives across thirteen universities from Edinburgh, Birmingham, Oxford, Southampton, Nottingham, Kent, Manchester, St. Andrews, Liverpool, Newcastle, QMUL, Loughborough and UCL. The multidisciplinary collaborations include front-line primary, secondary and social care staff; researchers in primary, secondary, and social care; health informatics and data science (including AI) experts; epidemiologists; qualitative researchers; statisticians; clinical and health services researchers, geographers; health economists; sociologists, human factors design researchers and public contributors. A summary of each study represented within the AIM consortia and their respective aims is included in table 1. The agreements reached are entirely those of the consortia; sponsors and funders have had no role in the development or reporting of this statement.

After initial discussions on the need for such a statement in our individual projects, we met to refine ideas and reach an agreement on item inclusion. We acknowledged variations in aims and purpose of research but found shared interest and overlapping aims with regards to the development of MLTC-M clusters that could then be externally validated across our respective studies. We collectively agreed on the need for a priori consensus on definitions of MLTC-M,

clustering variables, outcomes, managing data requests, as well transparency in reporting and methodological approaches to permit meaningful comparison between findings and validation that in turn, could contribute toward more rapid translation to patient benefit.

Table 1: Summary of studies included within wave 1 and 2 of the AIM consortia

Principal Investigator/s	Hajira Dambha-Miller Andrew Farmer	Thomas Jackson Krishnarajah Nirantharakumar	Bruce Guthrie	Lauren Walker Iain Buchan	Nick Reynolds Michael Barnes	Gyuchan Thomas Jun Satheesh Gangadharan	Simon Fraser Nisreen Alwan
Study title	The development and validation of population clusters for integrating health and social care: A mixed-methods study on Multiple Long-Term Conditions	OPTIMising therapies, disease trajectories and AI-assisted clinical management for patients Living with complex multimorbidity (OPTIMAL study)	Artificial Intelligence and Multimorbidity: Clustering in Individuals, Space and Clinical Context (AIM-CISC)	DynAIRx: AIs for dynamic prescribing optimisation and care integration in multimorbidity	AI-Multiply - Using artificial intelligence (AI) to characterize the dynamic inter-relationships between MUltiple Long-term condiTIons and PoLYpharmacy and across diverse UK populations and inform health care pathways	Data-driven machine-learning aided stratification and management of multiple long-term conditions in adults with intellectual disabilities	Multidisciplinary Ecosystem to study Lifecourse Determinants and Prevention of Early-onset Burdensome Multimorbidity
Study Acronym	AIM-Cluster	OPTIMAL	AIM-CISC	DynAIRx	AI-Multiply	DECODE	MELD-B
Abbreviated study aim	To develop and validate population clusters that consider health and social care determinants and subsequent need for people with MLTC-M using data-driven AI methods compared to expert-driven approaches, followed by evaluation of cluster trajectories and	To integrate multimodal primary care, secondary care, qualitative, and prescribing data to develop an AI tool with the aim of optimising clinical decision making in patients with cMM.	To use artificial intelligence and state-of-the-art data science, social science and health service research methods to understand clustering of morbidities within individuals, within communities, and in key clinical contexts.	To extract information from scattered clinical records on how health and medication change over time, apply AI and statistical approaches to predict risk of poor outcomes and develop visualisations that overlay these combined data to enable easier	To characterise MLTC-M and polypharmacy (MLTC-M-PP) trajectories and define the interrelationships between MLTC-M clusters, polypharmacy, inequalities and healthcare outcomes over the life course	To apply machine learning approaches to identify clusters and trajectories of MLTCs in people with intellectual disabilities and to utilise them to develop actionable insights for effective care coordination to improve the health and wellbeing of people	To safely deliver an Artificial Intelligence (AI)-enhanced epidemiological analytic system in which optimal lifecourse time points and targets for prevention of early-onset, burdensome multimorbidity are identified through multidisciplinary

	their association with health outcomes and costs.			review of medications. We shall pilot this approach in GP prescribing audit and feedback systems, and co-design future medicines optimisation decision support systems with patients and prescribers.		with intellectual disabilities	synthesis and analysis of birth cohorts and electronic health records, and disseminated to key stakeholders
Study population	Adults aged over 18 years	Adults aged over 18 years	Adults aged over 18 years	We will focus on three groups of people at high risk of rapidly worsening health from multimorbidity: 1. Older people with frailty. 2. People with mental and physical health problems. 3. Other people with four or more long-term health conditions taking ten or more drugs.	Adults aged over 18 years	Adults aged over 18 years with learning disabilities	From birth to age 65
Study design	AI methods and Mixed methods;	AI methods and Mixed methods	Objective 1: Unsupervised and supervised clustering of morbidities within individuals, primarily cross-sectional but exploring longitudinal trajectories. Objective 2: Genetic associations of morbidity clusters, and causal pathway analysis. Objective 3: Natural Language Processing (NLP) to enrich neighbourhood data, and multilevel and machine learning analysis of multimorbidity	AI methods and qualitative engagement	AI methods and Mixed methods, ,	AI methods and mixed methods	Qualitative evidence synthesis, Delphi study, AI methods, causal inference methods and modelling using birth cohort data and routine electronic health records

			clustering in places. Objective 4: NLP enrichment of routine clinical data; predictive modelling of serious adverse events using knowledge graphs and Multi-Layer Networks; process mining of care pathways; complex intervention development and early feasibilty testing.				
Follow-up period	Up to 10 years	Up to 10 years	Objective 1: Up to 10 years. Objective 2: Cross-sectional (although genetics is lifelong). Objective 3: Cross-sectional. Objective 4: Prediction is short time horizon (likely 3 years or less), but data will cover 5+ years (but include COVID periods which are a hard interruption of process mining for example).	Up to 12 years	Up to 10 years	Up to20 years	Birth to age 65 for the primary analyses
AI methods	Semi-automated based on shallow machine learning clustering using expert crafted features calculated from raw data; Fully-automated based on deep artificial neural networks and explainable AI techniques	Embedding approaches for learning numerical latent representations of structured electronic health record data that account for heterogeneous data modality, informative observational processes and non-random missingness. Latent class analysis for unsupervised population	Objective 1: Multiple clustering methods. For cross-sectional data includes hierarchical and non-hierarchical clustering, probabilistic latent class analysis, clustering in the projected space, fuzzy partitioning, restricted latent class analysis, Poisson matrix factorisation, and biclustering analysis.	Structured clinical data and narrative processing leading to a minimum core dataset, further extended by information extracted from, and tracked across, clinical narratives using NLP contextualised language models such as BERT. Patient histories will be modelled as temporal	a novel transfer of Bursty Dynamics theory from complex systems research to characterise healthcare event structure; Association Rules Mining to illuminate event sequence; topological data analysis as a novel approach to understanding complex MLTC-M-PP networks; market-	Computational clustering method Feature selection Dynamic Time Warping (unsupervised clustering algorithm) DeepCare approach (Long Short Term Memory deep learning neural network) Transition-based medical recommendation model with interactions (TMR4I)	Multiple clustering methods based on the agreed condition list, but also including other conditions and other burdensome features identified in the evidence synthesis and Delphi study: dimensionality reduction (UMAP), distance measures for clustering mixed data types (Gower transformation),

		stratification with clinical relevance. Prediction models which incorporate causal inference techniques to allow for counterfactual predictions. Recommendation algorithms for medications.	For longitudinal, will explore probabilistic methods such as a finite mixture of hidden Markov models. Objective 2: adapted existing genetic analysis methods. Objective 3: NLP of media and social media free-text for enhanced geoprofiling; multilevel modelling and will explore additional value of non-linear regression methods including neural networks and random forest, and extended observation models, such as beta and Dirichlet Objective 4: NLP of clinical free-text for enhanced serious adverse event ascertainment; Process Mining Project Methodology approaches; knowledge graphs with machine learning methods adapted to graph data such as relational learning. Multi-layer networks are a common method we will explore in objectives 1, 3 and 4. Explainability is a common theme.	graphs capturing clinical events (diagnoses, prescriptions etc.) in their timeline and extracted using 3D convolutional neural networks. This will exploit recent advances in video and time-series classification to discover temporal patterns and not just sequences of events (as with recurrent neural networks). The output will be a time-series of clinical feature vectors, which can be used to predict outcomes or to define clusters of typical patient trajectories. We shall use soft, temporal clustering algorithms to track a patient's membership of each cluster over their recorded history.	basket analysis with topic modelling to define latent associated conditions; ethnographic analysis of interdisciplinary AI-in-health working practices. Quantitative methods will be optimised using well-curated data (UK-Biobank and CPRD), and then validated on routinely collected electronic health record (EHR) NHS data from the North-East of England and East London, with particular focus on intersectional contrasts, including ethnicity and deprivation.		clustering (Agglomerative, Hierarchical and K-prototypes), clustering evaluation (Elbow, Silhouette and Gap methods), modelling and interpreting clusters (Decision Trees and XGBoost-SHAP), topological data analysis (supported by visualisation tools e.g. Kepler Mapper, persistent homology), graph- based techniques (e.g. spectral clustering, community detection and clustering inference on annotated networks). Based on overlapping features between birth cohort and routine data, train a classification model to identify those who are within burdensome clusters (using XGBoost or CatBoost).Use a semi-supervised learning algorithm e.g. SemiBoost Use a dynamical approach to infer causal relationships based on time-series data of lifecourse trajectories. Graph neural networks for inference and comparison to clustering approaches,

Non-AI methods	Semi-structured qualitative interview study, Delphi and Cohort analysis	Semi-structured qualitative interviews and discreet choice experiment	We will use epidemiological/ biostatistical methods in some analyses (and compare findings and explainability with machine learning). Objective 4 has an intervention development element which is mixed qualitative and quantitative methods.	Semi-structured interviews, focus groups, workshops and think aloud studies with healthcare practitioners and patients	Outcomes from AI methods will be compared to and refined further using epidemiology methods and trial emulation. A central assumption of this NIHR call is that bringing together clinicians, data scientists and social scientists will enable more comprehensive understandings of MLTC-M trajectories. WP4 will explore how interdisciplinary AI-in-healthcare collaborations proceed and what ensures success. Using an ethnographic approach and qualitative methods, we will map the trajectory of the co-constructed collaboration and technology. .	Outputs of AI analysis will be visualised for two different audience: i) people with intellectual disabilities and carers (visual narratives, animations and infographics); ii) health and social care professionals (interactive dashboard) Reflective semi-structured interviews Stakeholder workshops World Café approach	Qualitative evidence synthesis.Delphi study with patients, carers and MLTC-M experts to identify burdensomeness / complexity indicators. Directed acyclic graph (DAG) methods to explore associations between early life exposures and burdensome, early-onset MLTC-M clusters. G-methods to estimate the effect of intervening on single or a combination of early-life determinants at the critical time points. Stakeholder engagement to co-produce public health implementation recommendations
Study outcomes	For each cluster; Development of additional MLTC-M, All-cause mortality, Cause-mortality, Worsening frailty, Health and social care costs	Trajectories of distinct phenotypes in clusters and time-points for prevention. Predictive algorithms for future morbidities, . Predictive algorithm for choice of	Objective 1: morbidity clusters. Objective 2: morbidity clusters from objective 1. Objective 3: morbidity clusters from objective 1 and a priori multimorbidty groups.	1. Patterns and clusters of conditions, medications, tests and clinical contacts preceding adverse events identified by AI in key high-risk groups (older people with frailty; coexisting physical and mental health problems;	Polypharmacy All-cause mortality Premature mortality Healthcare utilisation MLTC-M inflection point trial emulation Interdisciplinary, theory-informed guidance	New knowledge on clusters and trajectories of MLTCs in people with intellectual disabilities	A suite of burdensomeness/complexity indicators. Novel, burdensome MLTC-M clusters Summary of early-life and wider determinants of burdensome clusters Summary of critical time points and targets for intervention

		medication in the context of multimorbidity. Discovery of medications for repurposing.	Objective 4: serious adverse events of various kinds (eg delirium, falls, increasing dependency, hospital admission, care home admission, death).	complex multimorbidity and problematic polypharmacy). Build the patterns into biostatistical causal inference and prediction of (clustered) clinical outcomes 2.Advance visualisation of longitudinal summaries of multi-provider care records overlain combined with key features from AI-learned patterns/structures 3.AI-augmented multimorbidity information built into an existing prescribing audit and feedback system – creating a learning system for medicines optimisation	A defined and operational PPI strategy influencing study outcomes. Refined pipelines of data engineering, AI and ML methodology. Defined MLTC-M-PP clusters, replicated cross-methods and cross-data sets. External validation of MLTC-M-PP clusters. Defined drivers of MLTC-M-PP clusters including influence of intersectional factors. Theoretically informed guidance for future AI-in-health collaborations, based on interdisciplinary understanding of ideas, assumptions and knowledge creation. New interventions to improve long-term treatment of patients, for testing in pragmatic trials. Identification of a partner to develop MLTC-M-PP clinical decision support tools and identification of potential 'tipping' points. Widespread communication and dissemination of results through multiple channels		Models of potential 'preventable moments' of early, burdensome MLTC-M. Co-produced public health implementation recommendations.

					including open-access academic publication and conference presentations. Enhanced training of early career researchers and increased critical mass in interdisciplinary MLTC research.		
Databases to be used for AI modelling	CRPD Aurum and Gold, SAIL, ELSA	CPRD Aurum/GOLD, PIONEER, INSIGHT, Scottish data (Linked Tayside and Fife Database and Greater Glasgow and Clyde)	Objective 1: CPRD Aurum, Lothian DataLoch. Objective 2: UK Biobank, Generation Scotland Objective 3: Scottish Longitudinal Study, Lothian DataLoch Objective 4: Lothian DataLoch	Structured data from care records across Cheshire & Merseyside, Greater Manchester, and Yorkshire & Humber, covering a population of ~11m using the www.cipha.nhs.uk blueprint that links population health management with care workflow, so can provide data for AI test/train and feedback actionable information to prescribers.	UK-Biobank CPRD Aurum/GOLD East London Patient Record Great North Care Record Anonymised dataset Connected Bradford Lothian DataLoch	CPRD Gold SAIL	CPRD Aurum and CPRD Gold SAIL Birth cohorts: National Child Development Study (NCDS), 1970 British Cohort Study (BCS70), Aberdeen Children of the 1950s (ACONF)

Defining MLTC-M

Researchers acknowledged the substantial variations within the literature in defining MLTC-M and in the conditions that might be included under this terminology.[1] [2]Moreover, each consortium may have a slightly different clinical or public health focus meriting inclusion of a diverse range of conditions [ref]. Acknowledging these challenges, the consortia agreed to conform to the definition of MLTC-M set out by Guthrie et al., (paper in submission) and to include a minimum set of conditions within our data requests based on the 59 core conditions summarised in figure 1 below. Not all these conditions will be relevant to every study and indeed researchers will include additional conditions. However, the agreement on the availability of a minimum set of conditions by every consortium will facilitate future external validation across studies.

Figure 1: Core conditions to be included within MLTC-M (Guthrie et al., in submission)

Body System (based on ICD-10 chapters)	Always include	Usually include (unless good reason not to in your context)
Cardiovascular	Stroke Coronary Artery Disease Heart Failure Peripheral Artery Disease	Heart Valve Disorder Arrythmia Venous Thromboembolic Disease Aneurysm Hypertension (treated and untreated)
Metabolic and endocrine	Diabetes Addison's Disease Cystic Fibrosis	Thyroid Disorders

Respiratory	Chronic Obstructive Pulmonary Disease Asthma	Bronchiectasis
Neurological	Parkinson's Epilepsy Multiple Sclerosis Paralysis	Transient Ischaemic Attack Peripheral Neuropathy Chronic Primary Pain
Cancers	Solid Organ Cancer Haematological Cancers Metastatic Cancers	Melanoma Benign Cerebral Tumours that can cause disability
Mental and behavioural	Dementia Schizophrenia	Depression Anxiety Bipolar Disorder Drug/Alcohol misuse Eating disorder Autism Post-traumatic stress disorder

Musculoskeletal	Connective Tissue Disease	Osteoarthritis Long-term musculoskeltal problems due to injury Osteoporosis Gout
Digestive	Chronic Liver Disease Inflammatory Bowel Disease	Chronic Pancreatitis Peptic Ulcers
Urogenital	Chronic Kidney Disease End-stage Kidney Disease	Endometriosis Chronic urinary tract infection
Haematological		Anaemia (including pernicious anaemia and sickle cell disease)
Eye		Vision Impairment that cannot be corrected
Ear		Hearing Impairment that cannot be corrected Meniere's Disease
Infections	HIV/AIDS	Chronic Lyme Disease Tuberculosis Post-acute COVID-19

Congenital		Congenital Disease and Chromosomal abnormalities

Protocols, data requests and coding

Study protocols and analysis plans will be made widely available, and wherever possible the consortia will include within data requests and approvals, opportunities for unspecified replication for other studies within the NIHR AIM programme of research[3]. This is an essential step to achieving the overall objective of cross-collaboration replication and validation. Code-sets and analytical code should also be made available. We constructed clinical code lists using a rigorous process, which involved reviewing existing code lists (e.g., CALIBER and Cambridge CPRD codes), applying comprehensive search terms to identify codes and finally review by clinicians and where necessary a consensus process to produce the final list of codes. The process is documented for transparency and was developed using the DExtER code builder. Initial code list was generated by the University of Birmingham for the 59 agreed conditions and shared freely across the consortia. Individual groups will include additional codes where appropriate according to their projects. Analytical code will be [4,5]available through the HDR-UK phenotype library. Researchers will also consider The FAIR Guiding Principles for scientific data management and stewardship, [6] and utilise appropriate reporting standards such as STROBE, RECORD or TRIPOD. [4,7]This will facilitate study consistency and rigour, improve transparency and reduce ambiguity in both methods and reporting.

Data variables

The consortia agreed to include a core minimum list of variables in their data requests. Individual study aims, objectives and methodologies will vary and thus not all variables will be included in all analytical models. Additional variables could be required accordingly. Depending on study aims, variables might

be used as exposures, covariates or outcomes and thus we have purposely not specified these here. The inclusion of a minimum set of clustering variables means that different research teams will be able to test their algorithms and validate research findings across datasets. We have agreed on the following core groups of variables but acknowledge that they may be measured, reported and defined differently across datasets and by each consortium. Clear reporting and explanations of variables will be included to permit meaningful interpretation and comparison between measures across consortia:

1. Sociodemographic variables
    a. Age
    b. Sex / Gender
    c. Ethnicity
    d. Deprivation
2. Clinical variables (a clear justification will be provided for their inclusion)
3. Health and social care service utilisation
4. Drugs (specify generic name, class of drugs and purpose)
5. Mortality (specify if cause-specific or all-cause)
6. Frailty (specify measure/s used)

AI fairness

The consortia acknowledge ethical concerns around bias and 'unfairness' that occur in AI systems and the impact this can have on the misuse of predictive models in decision making.[6] We acknowledge ethical frameworks that could be applied in the production and deployment of AI methods.

Share learning and conclusions

We agree to have regular engagement as a community of practice aiming to share learning, discuss problems and find collective solutions to the wider challenges of research that uses AI and big data such as those around data governance limitations, data sharing for the purposes of external validation or

methodological limitations around incomplete records. We recognise that further work in this new field is necessary, specifically around how best to effectively foster research collaboration and sharing across groups and individual databases, that encourage external validation of novel findings, maintain transparency and promote cooperation with the purpose of patient benefit.

Competing Interests: CY declares the receipt of remuneration from the Medicines and Healthcare products Regulatory Agency for consultancy work in artificial intelligence. IB is Chief Data Scientist advisor for AstraZeneca.

Funding: All the authors of this report have received funding from the National Institute for Health Research (Artificial Intelligence for Multiple Long-Term Conditions (AIM). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care.

Contributors: All authors contributed equally

Transparency declaration: This manuscript is an honest, accurate, and transparent account; no important aspects have been omitted.

Ethical approval: Not applicable

Data sharing: No other data is available for this consensus statement

References:

1 Johnston MC, Crilly M, Black C, *et al.* Defining and measuring multimorbidity: a systematic review of systematic reviews. *European Journal of Public Health* 2019;**29**:182–9. doi:10.1093/EURPUB/CKY098

2 Payne RA, Mendonca SC, Elliott MN, *et al.* Development and validation of the Cambridge Multimorbidity Score. *CMAJ* 2020;**192**:E107–14. doi:10.1503/CMAJ.190757

3 Artificial Intelligence for Multiple Long-Term Conditions (AIM) - Research Specification | NIHR. https://www.nihr.ac.uk/documents/nihr-artificial-intelligence-for-multiple-long-term-conditions-aim-clusters-call-research-specification-finalised/24646 (accessed 24 May 2022).

4 Benchimol EI, Smeeth L, Guttmann A, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Medicine* 2015;**12**:e1001885. doi:10.1371/JOURNAL.PMED.1001885

5 Collins GS, Reitsma JB, Altman DG, *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *BMC Medicine* 2015;**13**:1–10. doi:10.1186/S12916-014-0241-Z/TABLES/1

6 Wilkinson MD, Dumontier M, Aalbersberg IjJ, *et al.* Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 2016;**3**:1–9. doi:10.1038/sdata.2016.18

7 Elm E von, Altman DG, Egger M, *et al.* Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–8. doi:10.1136/BMJ.39335.541782.AD