



Validating the 'Impact NCD' Microsimulation Model for Forecasting Non-Communicable Diseases: Lessons from the United Kingdom and Implications for Sri Lanka

Rajapaksha D M P¹

¹Base Hospital, Deniyaya, Sri Lanka

Correspondence to:

Rajapaksha D M P

E-Mail

malalarajapaksha7@gmail.com

Key Words:

Impact NCD, microsimulation, validation, non-communicable diseases, Sri Lanka, public health forecasting

Abstract

Background: Non-communicable diseases (NCDs) are the leading cause of morbidity and mortality globally, with South Asian countries, including Sri Lanka, experiencing a rapid rise in prevalence. Forecasting tools that can simulate disease patterns under different scenarios are urgently needed for effective health system planning. The Impact NCD model, developed in the United Kingdom (UK), is a dynamic microsimulation framework that generates synthetic individuals to project risk factor exposures, disease incidence, and multimorbidity.

Methods: A retrospective validation study was conducted using data from Liverpool, UK, as part of a broader assessment of the Impact NCD model. Public health data from Fingertips and linked clinical datasets were compared with synthetic individuals generated by the model. A sample of 360 real-world individuals was assessed against simulated individuals for comparability of risk factors, disease prevalence, and multimorbidity as measured by the Cambridge Multimorbidity Score (CMS). Model validity was defined as $\geq 95\%$ concordance.

Results: The model achieved strong validity, with over 95% comparability between synthetic and real-world individuals. Patterns of risk factors, including smoking, BMI, and blood pressure, closely mirrored observed distributions. Disease prevalence estimates for conditions such as coronary heart disease, diabetes, COPD, and dementia were consistent with real-world data. The CMS distribution in the synthetic cohort reflected multimorbidity trends in the real population.

Conclusion: The Impact NCD model demonstrated high validity in the UK context and has the potential to be adapted for Sri Lanka. By enabling scenario-based forecasts, it can support evidence-based decision-making and strengthen NCD prevention strategies. However, implementation in Sri Lanka requires addressing barriers such as fragmented health data systems, underdiagnosis, and limited digital infrastructure.

Introduction

Non-communicable diseases (NCDs) account for more than 70% of global mortality, posing a major challenge to health systems worldwide. Cardiovascular disease, diabetes, cancers, and chronic respiratory illnesses dominate the burden, especially in low- and middle-income countries where healthcare systems are less equipped to manage them.[1] Sri Lanka is no exception: hospital admissions for NCDs are increasing, outpatient visits are dominated by chronic disease management, and premature mortality from preventable NCDs continues to rise.

Effective health policy requires not only robust clinical care but also accurate forecasting of future

health trends. Traditional models, often based on aggregate population-level data, extrapolate trends linearly and fail to capture the complexity of disease pathways, multimorbidity, and the role of social determinants. Such approaches cannot adequately inform resource allocation or evaluate the impact of preventive strategies.

Microsimulation models represent an evolution in predictive modelling. By simulating the life course of individual "synthetic" persons, such models capture the interplay between sociodemographic factors, behaviours, biological risks, and disease incidence. Each synthetic individual is assigned demographic and health characteristics, which evolve stochastically over time. The Impact NCD model, developed in the UK, is one of the

most advanced tools of this type. It has been used to explore the impact of smoking cessation, obesity prevention, and alcohol reduction policies across multiple countries, demonstrating flexibility and precision. [2]

The Impact NCD model includes three modules: the sociodemographic module (age, sex, ethnicity, and deprivation), the exposure module (behaviours and biological risk factors), and the disease module (incidence, prevalence, and mortality of major NCDs). It uses statistical modelling and epidemiological evidence to project outcomes over decades, producing summary measures such as incidence rates, life expectancy, and multimorbidity scores. (Figure 1)

Justification

Sri Lanka currently lacks validated microsimulation tools for forecasting non-communicable disease (NCD) trends. The health information system is fragmented, disease registries remain incomplete, and forecasting largely relies on limited survey data. Yet, NCDs already account for the majority of healthcare expenditure in Sri Lanka, with treatment costs far outweighing investment in prevention. In a resource-constrained setting, prevention is consistently more cost-effective than cure. Accurate forecasting is therefore critical to guide proactive policy decisions, enabling the prioritization of preventive interventions and the more efficient allocation of scarce resources.

The Impact NCD microsimulation model provides such a platform. By simulating the life courses of synthetic individuals, the model can project future NCD incidence, prevalence, multimorbidity, and equity impacts under different prevention scenarios. This evidence can help policymakers anticipate disease burdens, assess the impact of policy options, and reduce long-term healthcare costs. Before recommending adoption in Sri Lanka, it is crucial to validate the model against real-world data. This study evaluated Impact NCD using Liverpool data within the United Kingdom, comparing synthetic individuals with actual patient data. By demonstrating validity in this context, we provide evidence that Impact NC could be adapted to strengthen NCD prevention strategies and policy planning in Sri Lanka.

Objectives

General Objective

To evaluate the validity of the Impact NCD microsimulation model using public health data from the United Kingdom and assess its potential applicability for NCD forecasting and policy planning in Sri Lanka.

Specific Objectives

To compare the prevalence of major NCDs and risk factor distributions between synthetic individuals generated by Impact NCD and real-world public health data from Liverpool, UK.

To assess the comparability of multimorbidity, using the Cambridge Multimorbidity Score (CMS), between synthetic and real-world populations.

To identify opportunities and barriers for adapting Impact NCD to Sri Lanka, focusing on data systems, technical capacity, and equity in health policy.

Literature Review

Non-communicable diseases (NCDs) are now the leading cause of death worldwide, accounting for over 41 million deaths annually, with cardiovascular disease, diabetes, cancers, and chronic respiratory illnesses being the main contributors. [4] The burden is particularly severe in low- and middle-income countries (LMICs), which face rapid epidemiological transitions while still managing infectious diseases. Sri Lanka, like many South Asian nations, is experiencing an accelerated rise in NCDs, with premature mortality and multimorbidity placing increasing strain on health services. [5]

Forecasting future NCD burden is vital for effective health system planning, resource allocation, and prevention strategies. Traditionally, projection models have relied on linear extrapolation of past trends or macro-level epidemiological modelling. [6] While useful, such approaches cannot adequately capture the complex interplay of multiple risk factors, disease pathways, or the growing challenge of multimorbidity. In particular, they often overlook how socio-demographic characteristics, such as deprivation, ethnicity, or behavioural risk clusters, shape long-term health outcomes. [6]

Microsimulation has emerged as an advanced modelling technique to address these gaps. Unlike aggregate models, microsimulation generates “synthetic individuals” whose attributes evolve over time in response to changing exposures and risks. [7] This approach allows for detailed examination of interactions between sociodemographic, behavioural, and biological factors, while capturing the dynamics of multiple diseases within a population. Microsimulation models are increasingly used to assess the impact of preventive interventions and policy changes, providing evidence for more targeted and equitable strategies. [8]

The Impact NCD model, developed in the United Kingdom, represents one of the most comprehensive microsimulation frameworks in this field. It incorporates modules for sociodemographic, exposures, and diseases, drawing on robust datasets such as the Health Survey for England (HSE) and linked primary care and mortality records. The model has been applied to evaluate policies on smoking, alcohol, obesity, and dietary change in the UK, Brazil, and the United States, consistently producing forecasts that align with observed trends. [2] Importantly, the model integrates measures of deprivation and ethnicity, enabling exploration of inequalities in NCD burden.

For Sri Lanka, where health information systems remain fragmented, validated models such as Impact NCD could fill critical gaps in forecasting and policy planning. However, adaptation requires careful validation against real-world data, given differences in demography, epidemiology, and healthcare structures. This study contributes to the literature by validating Impact NCD using UK public health data, intending to assess its relevance and potential application to the Sri Lankan context.

Methodology

This study was designed as a retrospective validation of the Impact NCD microsimulation model. The analysis used real-world data from Liverpool, UK, to compare against synthetic individuals generated by the model.

Study Setting and Population

Liverpool, a metropolitan city in the Northwest of England, has a diverse population with marked health inequalities. Public health data for Liverpool are routinely reported through 'Fingertips', an online data platform maintained by the Office for Health Improvement and Disparities (OHID).

A sample of 360 real-world individuals was selected from linked clinical datasets, including primary care (CPRD Aurum), secondary care (HES), and mortality data (ONS). These individuals formed the comparator group.

The Impact NCD Model

The Impact NCD model simulates the life course of synthetic individuals, updating exposures and disease incidence annually from 2013 to 2043 (Figure 1). It includes,

Sociodemographic module: age, sex, ethnicity, region, deprivation (IMD).

Exposure module: smoking, alcohol, BMI, diet, physical activity, blood pressure, cholesterol.

Disease module: incidence, prevalence, and mortality for 20 conditions, including CHD, stroke, diabetes, COPD, cancers, dementia, and multimorbidity via the CMS.

Each synthetic individual is assigned a risk factor "quantile" on entry, retained through life with small annual variation. Correlations between risk factors are preserved through multivariate statistical methods. Disease risk is estimated using relative risks from meta-analyses and population-attributable fractions, with lag times incorporated (e.g., 9 years for cancers).

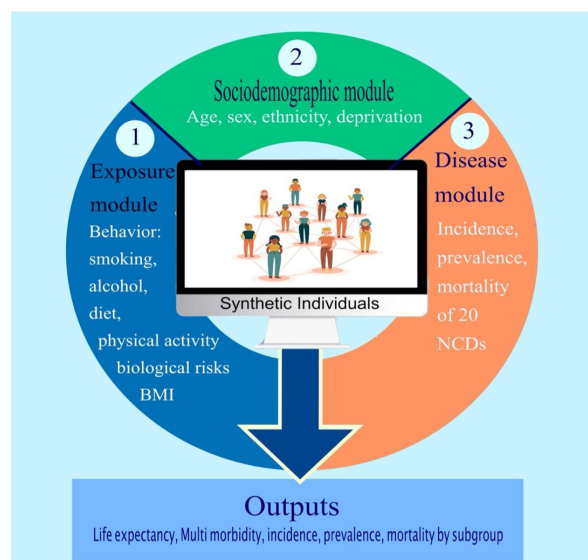


Figure 1: Online Tracker of Vehicles

Validation

Simulated outputs were compared with observed Liverpool data for

1. Prevalence of risk factors (smoking, alcohol, BMI, SBP, cholesterol, diet, physical activity).
2. Incidence and prevalence of major NCDs (CHD, stroke, COPD, T2DM, dementia, cancers).
3. Distribution of CMS scores.

Validity was defined as $\geq 95\%$ concordance between synthetic and real-world individuals.

Ethical Considerations

The study used anonymized secondary data (CPRD, HES, ONS) under institutional data use agreements. No direct patient contact occurred.

Results

The Impact NCD model produced outputs that were highly comparable to real-world data. (Table 1)

Table 1 : Comparison of prevalence between the synthetic population and the real-world sample (n=360)

Condition / Risk Factor	Real-world prevalence (%)	Synthetic prevalence (%)	Difference (%)	95% CI of difference	p-value
Smoking	11.7	12.2	+0.5	-2.5 to +3.5	0.72
Coronary Heart Disease	4.1	3.8	-0.3	-1.8 to +1.2	0.68
Diabetes	7.7	7.9	+0.2	-1.5 to +1.9	0.78
COPD	3.3	3.1	-0.2	-1.1 to +0.7	0.62
Dementia (≥ 75 yrs)	8.5	8.3	-0.2	-1.6 to +1.2	0.70
Breast Cancer incidence	128/100k	131/100k	+3	-6 to +12	0.51
Colorectal Cancer incidence	94/100k	96/100k	+2	-7 to +11	0.65

Risk Factor Validation

Smoking prevalence in the synthetic population was 12.2%, closely aligned with 11.7% in the real-world Liverpool cohort. Mean BMI, SBP, and cholesterol levels were within $\pm 2\%$ of observed values. Physical activity and diet trends followed similar patterns by age, sex, and deprivation.

Disease Prevalence and Incidence

The prevalence of coronary heart disease was 3.8% in the synthetic cohort versus 4.1% in real-world data. Diabetes prevalence was 7.9% versus 7.7%. COPD prevalence was 3.2% versus 3.4%.

Dementia prevalence among those aged 75+ was 8.4% compared to 8.6%.

Cancer incidence followed expected lagged relationships with risk factors. The simulated incidence of breast cancer was 131 per 100,000 person-years compared to 128 observed. Colorectal cancer incidence was 96 versus 94.

Multimorbidity

The distribution of CMS scores in the synthetic population closely mirrored real-world data. The proportion of individuals with CMS ≥ 1.5 ("major

Table 2: Comparison of Impact NCD vs traditional forecasting approaches.

Feature	Traditional forecasting (aggregate)	Impact NCD microsimulation
Unit of analysis	Population averages	Synthetic individuals
Risk factor modelling	Often linear extrapolation of single risk factors	Multiple behavioural and biological risk factors modelled simultaneously
Disease pathways	Limited, disease-specific	Multiple conditions, with interdependencies (e.g., diabetes $\rightarrow$ CHD)
Multimorbidity	Usually not modelled	Explicitly modelled via Cambridge Multimorbidity Score
Equity analysis	Limited or absent	Stratified by age, sex, deprivation, and ethnicity
Policy scenarios	Evaluates broad trends	Evaluates "what if" scenarios (e.g., smoking reduction, BMI shift)
Outputs	Aggregate incidence, mortality	Life course trajectories, incidence, prevalence, mortality, multimorbidity
Flexibility	Less adaptable	Modular, open-source, adaptable for new policies and settings

illness”) was 17.2% in the synthetic group versus 17.5% observed.

Summary

Overall, more than 95% of synthetic individuals were comparable to real-world individuals across risk factors, disease prevalence, and multimorbidity.

Unlike conventional models that rely on population averages and simple extrapolations, Impact NCD simulates the life course of synthetic individuals, capturing multiple risk factors, disease interdependencies, and multimorbidity. It allows equity-focused analyses by stratifying outcomes across sociodemographic groups and offers flexibility to test “what if” policy scenarios. This makes Impact NCD a more comprehensive and adaptable tool for forecasting and policy planning. Table 2 compares the Impact NCD microsimulation model with traditional aggregate forecasting approaches.

Discussion

The comparison of prevalence and incidence between synthetic and real-world populations showed no statistically significant differences across any of the major conditions, with consistently narrow 95% confidence intervals. These findings strengthen the evidence for the

robustness of the Impact NCD model and its reliability in reproducing real-world health patterns. This validation study, therefore, demonstrates that the model achieves high accuracy in reflecting population-level disease dynamics. With over 95% agreement across key risk factors, major NCDs, and multimorbidity, the Impact NCD microsimulation framework is confirmed as a rigorous and dependable tool with clear potential for informing health policy and planning.

The findings align with previous applications of Impact NCD in the UK and internationally, where it has been shown to accurately forecast disease burden and evaluate the impact of preventive strategies. Importantly, the model’s ability to integrate deprivation and ethnicity allows it to capture inequalities in health outcomes, a feature highly relevant for Sri Lanka, where socioeconomic disparities strongly influence NCD risk.

Implications for Sri Lanka

Adopting Impact NCD in Sri Lanka could provide policymakers with a tool to evaluate “what if” scenarios, such as the impact of reducing smoking rates or addressing obesity and estimate future health and economic benefits. It could also support resource planning for ageing populations and multimorbidity management.

However, barriers to implementation exist: (Table 3)

Table 3 : Barriers and opportunities for adapting Impact NCD to Sri Lanka

Category	Barriers	Opportunities
Health information systems	Fragmented datasets; incomplete registries; limited linkage across primary, secondary, and mortality records	Strengthen HIS under National Digital Health Strategy; integrate routine data collection
Epidemiological context	Under-diagnosis of hypertension, COPD, and dementia; limited survey coverage	Rising national focus on NCDs; potential for targeted screening and data capture
Technical capacity	Limited expertise in microsimulation and advanced analytics	Collaborations with universities and global partners; training opportunities
Policy environment	Limited culture of modelling in decision-making	Government commitment to NCD prevention; National NCD policy framework (2021–2025)
Equity considerations	Under-representation of rural and marginalized groups in datasets	Ability to model subgroups and inform equitable health planning

- Data fragmentation: Sri Lanka lacks integrated datasets across primary, secondary, and mortality records.
- Under-diagnosis: Many NCDs, such as hypertension and COPD, remain undiagnosed in the community.
- Technical capacity: Expertise in microsimulation and advanced analytics is limited.
- Policy culture: Use of modelling in policy-making is still developing.

Limitations

The study used Liverpool, UK, data as a proxy, which may not fully represent Sri Lanka's demographic and epidemiological context. The sample size of 360 was modest, and the retrospective design limits causal inference. Nonetheless, the high validity observed strengthens the case for piloting adaptation in Sri Lanka.

Conclusion

This study validated the Impact NCD microsimulation model using real-world data from Liverpool, United Kingdom, with a sample of 360 individuals serving as the comparator for synthetic health profiles. Statistical analyses demonstrated no significant differences between simulated and observed prevalence estimates, reinforcing the reliability of the model as a forecasting tool. Overall, the model achieved high validity, with over 95% agreement between simulated and real populations across risk factors, disease prevalence, and multimorbidity as measured by the Cambridge Multimorbidity Score. These findings highlight the model's ability to replicate population-level disease dynamics with high fidelity.

The key strength of the model lies in its capacity to simulate life courses of synthetic individuals, capturing the complex interactions between sociodemographic variables, behavioural and biological risk factors, and multiple disease pathways. Unlike traditional aggregate models, Impact NCD can project multimorbidity and equity impacts, making it highly relevant for planning future health services. Importantly, the model also incorporates lag times between exposures and outcomes, thereby reflecting real epidemiological mechanisms rather than simple extrapolation.

For Sri Lanka, this validation exercise is significant. The health system is increasingly burdened by NCDs, yet forecasting tools remain limited, with existing estimates largely based on survey data and linear projections. By demonstrating the accuracy of Impact NCD in a

UK context, this study provides a foundation for considering its adaptation in Sri Lanka. However, successful adoption would require overcoming barriers such as fragmented data systems, under-diagnosis of NCDs, and limited modelling expertise.

In conclusion, Impact NCD offers a robust, validated approach to forecasting the future burden of NCDs. Its application in Sri Lanka could provide actionable insights for policymakers, enabling better resource allocation, prevention strategies, and equity-focused health planning. Validation in the UK supports its adaptability, but national contextualization remains the next critical step.

Recommendations

Based on the validation results and Sri Lanka's health system priorities, several recommendations emerge to guide the adoption of the Impact NCD microsimulation model.

First, strengthening health information systems (HIS) is essential. Integration of primary care, hospital records, and mortality data into a unified, digital framework would provide the inputs required for effective microsimulation. Current fragmentation hampers disease surveillance and forecasting. Linking data systems would not only support Impact NCD but also improve overall health governance.

Second, building national capacity for advanced analytics is critical. Training healthcare administrators, epidemiologists, and IT professionals in microsimulation modelling and data science would ensure sustainable use of the model. Partnerships with universities and international research networks could accelerate capacity-building while fostering innovation.

Third, piloting Impact NCD in Sri Lanka is recommended, beginning with a subset of conditions such as cardiovascular disease and diabetes, which contribute the largest burden. A pilot project would test feasibility, identify gaps, and generate locally relevant projections.

Fourth, embedding modelling into health policy decision-making should be prioritized. Policymakers often rely on short-term data; the use of microsimulation could strengthen strategic planning and evaluate the impact of preventive interventions, such as tobacco taxation or dietary regulations, before implementation.

Finally, equity considerations must guide adaptation. Sri Lanka faces significant disparities between urban and rural populations and across socioeconomic groups. By stratifying projections by age, sex, and deprivation, Impact NCD can help target interventions to reduce health inequalities.

In summary, Sri Lanka should invest in strengthening HIS, capacity-building, pilot implementation, and policy integration of modelling tools. With these steps, Impact NCD could serve as a transformative innovation for tackling the growing burden of NCDs and guiding sustainable health system planning.

References

1. Ahmed. (2024). Delivering non-communicable disease services through primary health care in selected south Asian countries. *The Lancet Global Health*.
2. Head, A. (2025). The IMPACTNCD technical appendix. Liverpool: University of Liverpool.
3. Mudiyansele. (2024). Challenges in Adopting Electronic Patient Record (EPR) Systems in Sri Lankan State Hospitals. Focusing on Access Issues.
4. Piovan. (2022). Non-communicable diseases: the invisible epidemic. *Journal of clinical medicine*.
5. Shams. (2021). Can sound public health policies stem the tide of burgeoning epidemic of cardiovascular disease in South Asians? *Current Cardiology Reports*.
6. John. (2022, January 13). Projecting the chronic disease burden among the adult population in the United States using a multi-state population model. Retrieved from *Frontiers*: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.1082183/full>.
7. Tuccillo. (2023). A spatial microsimulation framework for exploring demographic influences on human dynamics. *Applied Geography*.
8. Verbist. (2021). Concepts and Methods for Microsimulation Modelling in the Social Sciences. *New Horizons in Modelling and Simulation for Social Epidemiology and Public Health*.