How do different ways of making a home warmer affect risk of preschool respiratory infections?
Using artificial intelligence to make homes and children healthier.

Abstract

Underheated, damp homes are associated with childhood respiratory infections. Governmental measures such as installing insulation or replacing a boiler try to make homes easier to heat. However, as well as keeping warm air in, some measures also trap in mould, smoke and air pollution, potentially increasing the risk of respiratory infections.

This interdisciplinary PhD will combine artificial intelligence methodology with novel cross-sectoral datasets to investigate how underheated homes and the different approaches to warming them affect preschool acute respiratory infection risk. For the first time, spatio-temporal machine learning methods will identify groups of children with particular characteristics for whom different ways of heating a home may be beneficial or detrimental, with significant policy implications.

Introduction

Respiratory infections are the major cause of illness in preschool children.

Acute respiratory infections (ARIs) are the leading cause of hospital admissions and antibiotic prescriptions in preschool children in the UK,[1,2] with 40/1000 infants admitted with ARIs in Scotland annually.[3] Deprivation and underheated, damp and mouldy housing are associated with preschool ARIs[4–6] which are in turn associated with childhood asthma[7] and premature adult death from respiratory disease.[8] Whilst the association between deprivation and preschool ARIs is clear,[3] we do not know how reducing underheating could affect this.

New Scottish data infrastructure developments have linked electronic health records to individual properties using Unique Property Reference Numbers (UPRNs) for the first time.[9] Using these linked data, we can undertake the first national investigation of the impact of property-level housing conditions on individual health. This will enable data-driven population health policy focused on key areas of societal need.

The health consequences of different interventions to make homes warmer are unclear.

Heating homes generates one third of the UK's carbon footprint.[10] The Scottish Government is investing £1.8 billion retrofitting existing homes to improve home energy efficiency (HEE) and meet Net Zero targets.[11] However, evidence of HEE measures effect on respiratory health is conflicting.[12]

HEE measures work by: A) reducing energy costs without increasing airtightness (e.g., boiler replacement) or B) reducing heat loss by increasing airtightness (e.g., insulation). Group A measures may reduce ARIs by increasing temperature or improving finances. Group B measures retain warm air but may increase ARI risk by trapping indoor air pollutants (e.g., smoke, particulate matter, and viruses[13]) and increasing mould. While Group A HEE measures are associated with fewer asthma exacerbations,[14] Group B showed mixed associations, with some studies suggesting they worsen respiratory health.[15–18] Within these associations, ARI risk may differ in groups of children with particular characteristics (e.g. geographical area, comorbidities). However, the impact of these characteristics are hard to untangle with conventional statistics. Spatio-temporal machine learning models now enable us to explore detailed associations in different groups.

Research Challenge

We do not know how different HEE measures affect preschool ARI risk and which measures have the greatest health co-benefits. Spatio-temporal machine learning models will allow us to explore which HEE measures may be beneficial or detrimental in different groups of children. This information will ensure Net Zero policies do not inadvertently worsen health inequalities.

About the study

This PhD is embedded in the Homes, Heat and Healthy Kids study, a major new interdisciplinary national data linkage study with five years Wellcome Trust funding[19]. The Scottish Government has granted unique access to property-level housing datasets and the study is one of 10 granted Ofgem approval to UK-wide smart meter data.[20]

Data & Methodology

This PhD will apply spatio-temporal machine learning models to a new retrospective birth-cohort of children born in Scotland from 2008-2025, followed until 5th birthday, migration, or death. This will contain ~4.3 million person-years across ~230,000 properties.[21,22]

Electronic health records are now linked to individual properties, permitting new linkages to energy efficiency, energy use and HEE measures.[23] Area-level linkages include aggregated prepayment smart meter data,[24] financial data,[25] air pollution and climate data. This approach will produce a rich, interdisciplinary dataset with information at multiple levels (child, property and area), outlined in **Figure 1**.

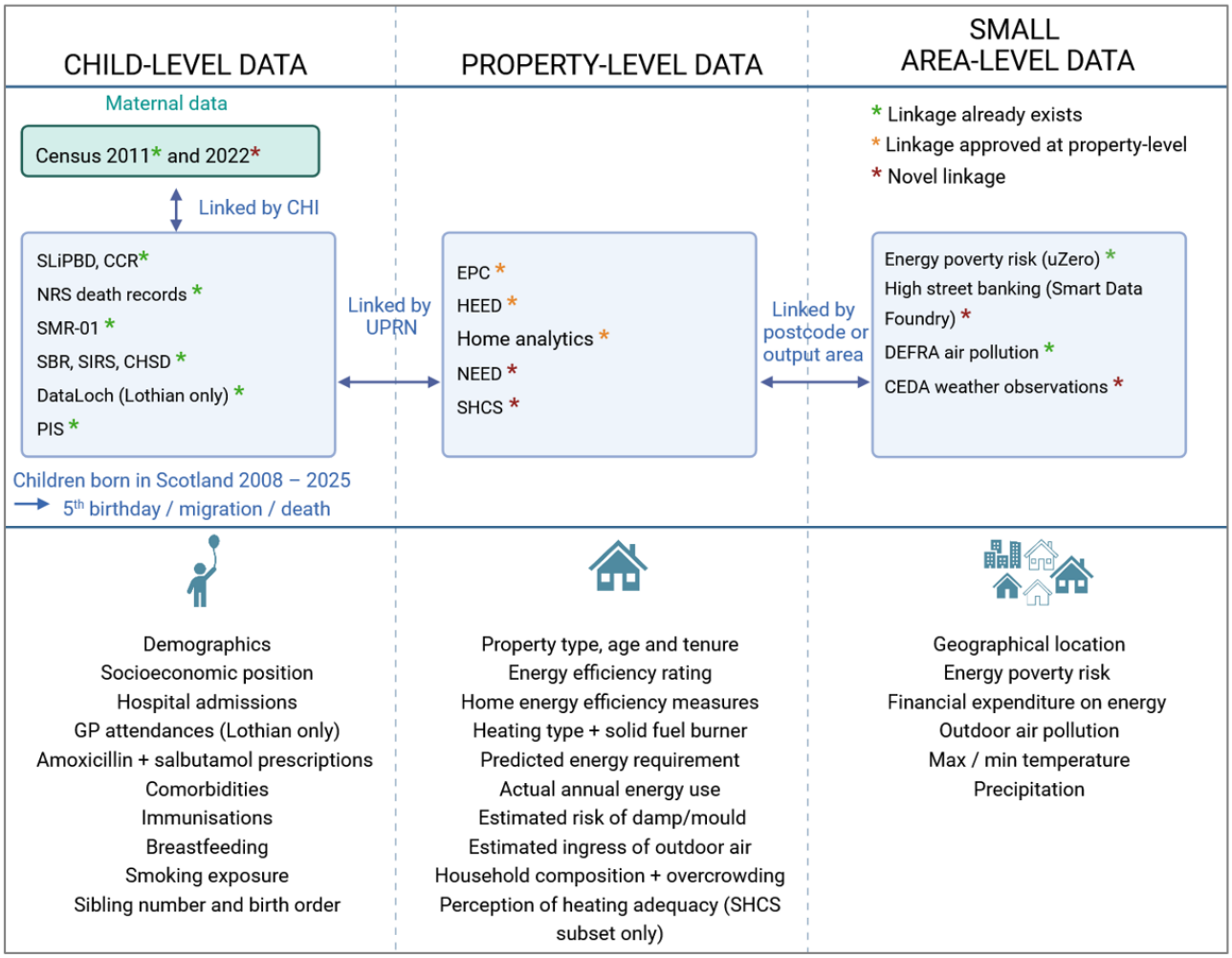

Figure 1. Data linkage diagram and variables for birth cohort. **CHI:** Community Health Index number, **SLiPBD:** Scottish Linked Pregnancy and Birth Dataset, **CCR**: CHI Central Register, **NRS:** National Records of Scotland, **SBR:** Scottish Birth Record, **SMR:** Scottish Morbidity Record, **SIRS:** Scottish Immunisation & Recall System, **CHSD:** Child Health Surveillance Data, **PIS:** Prescribing information System, **URPN:** Unique Property Reference Number, **EPC:** Energy Performance Certificate, **HEED:** Home Energy Efficiency Dataset, **NEED:** National Energy Efficiency Data-Framework, **SHCS:** Scottish Housing Conditions Survey (3000 properties each year), **DEFRA:** Department for Environment, Food & Rural Affairs, **CEDA:** Centre for Environmental Data Analysis. NB: GP ARI attendances available for Lothian only (DataLoch).

Spatio-temporal machine learning methods will be explored to identify patterns of ARI healthcare use over time, in the 12-month periods before and after different types of HEE measures. This will enable the examination of longitudinal healthcare use and identify children with different patterns of ARI healthcare use (trajectories) Machine learning approaches will allow us to group children with similar trajectories and explore whether particular characteristics (e.g., deprivation, climate or air pollution) are associated with each trajectory. This will enable identification of groups of children for whom a specific HEE measure may be beneficial or detrimental, informing policy makers of ways to stratify HEE measures to different populations for maximum benefit.

Opportunities include the use of deep learning models (such as long-short term memory [LSTM] and other neural network architectures) as well as more traditional machine learning approaches (such as decision tress/random forests) or stochastic models (such as dynamic time warping).[26]

Risk Mitigation

National data access is expected by October 2025. In the unlikely event that national data linkage is delayed, a regional dataset has already been linked in DataLoch (population, ~ 700,000) which will provide an environment to rigorously develop methods for the main dataset.

Responsible AI / Ethical considerations

Development of spatio-temporal machine learning models will inform policy makers of ways to stratify HEE measures to different populations. These models will form the basis for investigating HEE impact on other populations and health problems (e.g. cardiovascular disease in the elderly).

All linkage and analyses will be undertaken in the National Safe Haven using linked pseudonymised data. The successful applicant will undertake the Introduction to Administrative Data Research and Analysis course and ONS Safe Researcher Training Course. We are committed to open science and all code will be made publicly available on GitHub.

Expected Outcome and Impact

This project will produce important results which will inform national energy, welfare, climate and building policies. Methodology developed will be applicable to other health conditions, demographics and populations. The interdisciplinary nature of this project will provide the successful applicant with broad skills across disciplines and a wide network for future post-doctoral work.

References

1. NHS Digital. Hospital Admitted Patient Care Activity 2021–22. National Statistics., https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2021-22 (2022).

2. O'Brien K, Bellis TW, Kelson M, et al. Clinical predictors of antibiotic prescribing for acutely ill children in primary care: an observational study. *Br J Gen Pract* 2015; 65: e585–e592.

3. Moore HC, Klerk N de, Blyth CC, et al. Temporal trends and socioeconomic differences in acute respiratory infection hospitalisations in children: an intercountry comparison of birth cohort studies in Western Australia, England and Scotland. *BMJ Open* 2019; 9: e028710.

4. Barnes M, Butt S, Tomaszewski W. The Dynamics of Bad Housing: The Impact of Bad Housing on the Living Standards of Children. National Centre for Social Research.

5. Antova T, Pattenden S, Brunekreef B, et al. Exposure to indoor mould and children's respiratory health in the PATY study. *Journal of Epidemiology & Community Health* 2008; 62: 708–714.

6. Fisk WJ, Eliseeva EA, Mendell MJ. Association of residential dampness and mold with respiratory tract infections and bronchitis: a meta-analysis. *Environmental Health* 2010; 9: 72.

7. Rosas-Salazar C, Chirkova T, Gebretsadik T, et al. Respiratory syncytial virus infection during infancy and asthma during childhood in the USA (INSPIRE): a population-based, prospective birth cohort study. *Lancet* 2023; 401: 1669–1680.

8. Allinson JP, Chaturvedi N, Wong A, et al. Early childhood lower respiratory tract infection and premature adult death from respiratory disease in Great Britain: a national birth cohort study. *The Lancet* 2023; 401: 1183–1193.

9. Clark D, Dibben C. Adding a Residential Dimension to the Scottish Population Spine – CHI-UPRN Residential Linkage (CURL). *International Journal of Population Data Science*; 7. Epub ahead of print 25 August 2022. DOI: 10.23889/ijpds.v7i3.1946.

10. Business, Energy and Industrial Strategy Committee. Decarbonising heat in homes. Seventh Report of Session 2021–22., https://publications.parliament.uk/pa/cm5802/cmselect/cmbeis/1038/report.html (2022).

11. Scottish Government. Heat in Buildings Strategy: Achieving Net Zero Emissions in Scotland's Buildings., https://www.gov.scot/publications/heat-buildings-strategy-achieving-net-zero-emissions-scotlands-buildings/ (2021).

12. Thomson H, Thomas S, Sellstrom E, et al. Housing improvements for health and associated socio-economic outcomes. *Cochrane Database of Systematic Reviews*. Epub ahead of print 2013. DOI: 10.1002/14651858.CD008657.pub2.

13. Niculita-Hirzel H. Latest Trends in Pollutant Accumulations at Threatening Levels in Energy-Efficient Residential Buildings with and without Mechanical Ventilation: A Review. *Int J Environ Res Public Health* 2022; 19: 3538.

14. Howden-Chapman P, Pierse N, Nicholls S, et al. Effects of improved home heating on asthma in community dwelling children: randomised controlled trial. *BMJ* 2008; 337: a1411.

15. Rodgers SE, Bailey R, Johnson R, et al. Emergency hospital admissions associated with a non-randomised housing intervention meeting national housing quality standards: a longitudinal data linkage study. *J Epidemiol Community Health* 2018; 72: 896–903.

16. Sharpe RA, Machray KE, Fleming LE, et al. Household energy efficiency and health: Area-level analysis of hospital admissions in England. *Environment International* 2019; 133: 105164.

17. Hamilton I, Milner J, Chalabi Z, et al. Health effects of home energy efficiency interventions in England: a modelling study. *BMJ Open* 2015; 5: e007298.

18. Tieskens KF, Milando CW, Underhill LJ, et al. The impact of energy retrofits on pediatric asthma exacerbation in a Boston multi-family housing complex: a systems science approach. *Environ Health* 2021; 20: 14.

19. The role of underheated homes in preschool respiratory infections: A national epidemiological and natural experiment study using linked data.- Grants Awarded. *Wellcome*, https://wellcome.org/grant-funding/people-and-projects/grants-awarded/role-underheated-homes-preschool-respiratory (accessed 7 October 2024).

20. Ofgem. Consent granted to DCC under Conditions 9 and 10 of the Smart Meter Communication Licence, and Section M4.3 of the Smart Energy Code- August 2023,

https://www.ofgem.gov.uk/publications/consent-granted-dcc-under-conditions-9-and-10-smart-meter-communication-licence-and-section-m43-smart-energy-code-august-2023 (2023).

21. Office for National Statistics. Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland., https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland (2022).

22. Scottish Government. Scottish Council Area 2011 by dependent children in household by household type. Scotland's Census., https://www.scotlandscensus.gov.uk/webapi/jsf/tableView/tableView.xhtml (2014).

23. Hamilton IG, Steadman PJ, Bruhns H, et al. Energy efficiency in the British housing stock: Energy demand and the Homes Energy Efficiency Database. *Energy Policy* 2013; 60: 462–480.

24. Urban Tide. Summary Report. uZERO:Tackling the growing fuel poverty crisis with AI., https://urbantide.com/uzero (2022).

25. Smart Data Foundry. Smart Data Foundry: Opening Finance for Good, https://smartdatafoundry.com/about.

26. Jayatilaka G, Hassan J, Marikkar U, et al. Use of Artificial Intelligence on spatio-temporal data to generate insights during COVID-19 pandemic: A Review. 2021; 2020.11.22.20232959.