|  |
| --- |
| Host department:Choose an item.Keele |
| Project Title: |
| Does individual level or household data on social determinants of health perform better than neighbourhood level data when looking at health outcomes? |
| Proposed supervisory team: Names and areas of expertise to be included |
| Dr Emma Parry, Lead Supervisor. NIHR Academic Clinical Lecturer in Primary Care, Keele School of Medicine and Lead for Royal Wolverhampton NHS Trust Academic Primary Care Unit. Academic GP with experience in systematic review, quantitative data analysis and interest in health inequalities. Will provide supervision on all aspects of the project and provide opportunity for links with Health Inequalities group at RCGP, NHSE, Health Foundation and local CCGs to ensure results from this PhD impact on national and local policy. Dr Ross Wilkie, Senior Lecturer in Public Health and Epidemiology, Keele School of Medicine. Will provide co-supervision and advice on systematic review, epidemiological analysis and public health expertise.Dr Dahai Yu, Senior Research Fellow, Keele School of Medicine. Will provide expertise on epidemiological analysis including multilevel models. Professor Baldev Singh, Consultant Diabetologist and Clinical Lead for IT, Royal Wolverhampton NHS Trust, Medical Director for School of Medicine and Clinical Practice, Wolverhampton University. Will provide co-supervision and advice on data linkage. |
| Potential for cross consortium networking and educational opportunities: |
| Networking opportunities:The PhD will provide opportunities to link with the Public Health Department and Health Informatics team based at Royal Wolverhampton NHS Trust (RWT) through the Keele School of Medicine link with the RWT Academic Primary Care and Population Health Unit partnership. During the course of the PhD it is expected that the student will link with health inequalities groups at the RCGP, NHSE, Health Foundation and local CCG level leads to ensure findings influence public and national policy**Educational opportunities:**The successful candidate will have completed a relevant Masters degree in health sciences / statistics with evidence of practical skills in quantitative data analysis. They will be encouraged to identify generic and project-specific training needs using the Vitae Framework and to make full use of PGR workshops and training opportunities available from Academic Development at Keele University, as well as informal training opportunities such as journal clubs and seminar programmes in SoM. The student will utilise internal courses and workshops at Keele, for example systematic review and evidence synthesis workshops and statistics and epidemiology courses.Opportunities will be provided for the student to attend external workshops and courses on health inequalities and social determinants of health, for example, courses hosted by Futurelearn and the BMJ. This is in addition to formal training courses on understanding electronic healthcare records and big datasets, and advanced statistical techniques such as multi-level modelling. |
| Project description: |
| BackgroundHealth inequalities are worsening in the UK1 and the Covid-19 pandemic has further exposed disparities in health and health outcomes2. Social determinants of health - the conditions in which people are born, grow, work, live, and age, and the wider set of forces and systems shaping the conditions of daily life3 – critically influence population and individual health outcomes. Despite this, information on individual-level social risks are not collected routinely or in a standardised format in primary care. This may be due to insufficient time during the GP consultation to collect such data or to respond when a risk is identified, lack of services or lack of knowledge on available services to help these people once a need has been identified, and feeling uncomfortable asking these questions4. The Indices of Multiple Deprivation are the routinely used measure in England for deprivation based on geographical area­5 but does not give accurate information at an individual level. Yet there is some evidence that having better information on individual-level social risks and social needs can improve health outcomes through adaptations of care plans, influence on clinical decision making6, being mindful of financial barriers to accessing medication, addressing behavioural health issues, helping with transportation7 and increased referral to community services such as social prescribers8-10. There is increasing scientific enquiry comparing individual level socio-economic factors with area level factors on certain health outcomes with mixed results. There are concerns that using area level variables instead of individual level indicators leads to an increase in measurement error due to inference of individual level data on the basis of grouped data, something called ecologic fallacy11. Area level indicators have also been shown to be independently associated with certain outcomes and therefore should be used cautiously as a proxy measure of individual level factors as they may overestimate associations12-13. There are also concerns that individual and area level indicators do not reflect the same reality, are based on different constructs and contribute independently to outcomes observed13. Furthermore, there are concerns about measurement validity and construct validity when area level variables are used14. In some studies, individual level risk factors showed associations with outcomes that area level factors did not, for example residential instability versus Area Deprivation Index and chronic asthma severity, 365 day ED readmission and hospitalisation rates15. Another study found a significant association with ED attendance in children with asthma for individual level factors such as living in overcrowded accommodation, living in rented accommodation, and access to a car but no significant association with Townsend Deprivation16.Authors, in general support the use of area level indicators where individual level data are not available as the strength and direction of the associations are often found to be comparable17-19. Despite this there is an acknowledgement that area level variables do not capture all of the variability observed and has potential to also underestimate assumptions11. If area level data are to be used, most studies supported the use of smaller areas for example census blocks performed better than census tracts in US based census studies11.Determining whether area or individual level factors are best for determining a relationship with certain outcomes is complex and seems to differ depending on the SDH construct that is being investigated and the health outcome. For example, in a study by Mustard et al who assessed proxies for household income and health status, individual level risk factors performed as well as area level risk factors when looking at household income level and average neighbourhood level income level on the outcome of mortality, nursing home admissions, morbidity related care and fertility. However, individual level risk factors performed better when the outcome was disability and mental health problems14. Individual level indicators also performed better in a study analysing the outcomes of life-expectancy and disability-free life expectancy. Associations for individual level factors were stronger than area level risk factors, for example; no high school diploma, employment, income, being widowed, separated or divorced living alone and being in a one-parent family). However the direction of association was the same for area level variables and the results were statistically reliable and so the authors concluded that area level indicators were a suitable alternative in the absence of individual level indicators13. Area level indicators were found to under-estimate disadvantage-based life expectancy more for women than men. There were some studies where area level factors performed better, particularly with regard to education variables in younger and older populations17-18. There may therefore be instances depending on the SDH domain being investigated or the health outcome, where area or individual level factors are better suited. Area level factors seem to perform less well where race and income are concerned13,19 and when covering non-urban areas due to the increased heterogeneity of people in these areas20. In summary, evidence is conflicting as to whether area or individual level factors on social determinants of health perform better when looking at associations with certain health outcomes. There is a suggestion from the literature that the choice of individual or area level indicators is dependent on particular SDH constructs or the health outcome being investigated. Understanding which indicators are best used will ensure that information gathered going forward on health inequalities accurately describe a population and their risk. Understanding when individual level factors should be collected, particularly within the context of managing patients in primary care will ensure that care and management is targeted to the individual and those that are at risk of certain disease outcomes due to health inequalities are quickly identified. Furthermore, joining up existing datasets where social determinants of health are already collected with the primary care health record will help increase the information available to clinicians on those at risk from health inequalities. Overall aim:This project will aim to understand whether individual level or household data performs better at predicting outcomes linked to social determinants of health (SDH) compared to neighbourhood or area level data. It will also explore the scope for joining up community databases that contain this information to the electronic primary care health record in order to improve the information available to clinicians on social determinants of health. Specific Objectives:1. To design and conduct a review and synthesis of existing evidence that has compared the performance of individual level or household data and area level data on social determinants of health on outcomes for example, unplanned hospital admissions and unscheduled care 2. To conduct a population survey in the underlying musculoskeletal population, and collect information on both area-level and individual-level social determinants, alongside general heath (i.e. EQ-5D) and musculoskeletal health indicators (i.e. MSK-HQ);3. To derive multilevel models incorporating area-level and individual-level social determinants separately, and compare the predicting capacity on population-level mean of general health indicators and musculoskeletal health indicators in a small-area (practice, PCN, CCG, or LSO/MSOA) ;4. To produce data visualisation (map) of the general health indicators and musculoskeletal health indicators with local distribution based on the models with the most predictable social determinant;4. To link the survey data with local primary care electronic health record (EHR) data, collecting routinely recorded patients’ characteristics (ethnicity, comorbidity, et al) with primary care recorded health indicators (i.e. consultation on back pain, osteoarthritis, and other musculoskeletal disorders);5. To compare the neighbourhood-level and individual-level social determinant in the prediction of primary care recorded health indicators, with incorporation of more EHR covariates;3. Explore the potential for linkage of data on SDH collected by other health professionals and services, for example, health visitor and midwifery databases. Create scalable methods to link different data sources using the Wolverhampton population as an example.MethodsStudy 1: *Comparison of area level and individual risk factors of social determinants of health as measures of health outcomes; a systematic review (Objective 1)*A systematic review will be designed and registered on PROSPERO with the purpose of comparing the performance of area level and individual indicators of the social determinants of health and certain health outcomes as described in the literature. The results will also be used to explore if area or individual level indicators perform better for certain predictors for example, to determine if area level information on education is better than individual level data when looking at health outcomes as reported in the literature, for example unplanned Emergency Department attendance. The student will identify further gaps to be addressed within subsequent analysis of this thesis. In parallel they will gain an understanding of multi-level linear modelling21.Study 2: *Comparison of area level and individual level indicators of social determinants in predicting population-level general health and musculoskeletal health in a small-area (Objective 2)*Data: A population-level survey in the underlying musculoskeletal (disorder) population within the MIDAS study Based on findings from the systematic review and discussions with the supervisory team, PPIE group and collaborators the student will need to;* Identify individual level variables of social determinants collected within the MIDAS population survey to measure
* Identify area level variables for the participants from MIDAS population survey
* Derive multilevel linear models incorporating area-level and individual-level social determinants separately, to predict population-level mean of general health and musculoskeletal health, and compare the predictive abilities of area-level and individual-level social determinants
* Produce data visualisation (map) on the distribution of population-level mean of general health and musculoskeletal health in local small-areas (MSOA/LSOA) with most predictive social determinant, and present these for discussion at a PPIE meeting

Study 3: *Explore the capability for linkage of community datasets with the primary care electronic health record* *(Objective 3)*The student will undertake a proof of concept for linking a community dataset that has information on social determinants of health (for, example Health Visitor database, midwifery database or local authority database) with the primary care electronic health record. Royal Wolverhampton NHS Trust has several linked databases housed within the Trust that includes primary care, district nursing, palliative care, and walk in centres. The framework already exists to link the datasets however the student will need to work collaboratively with the Trust data warehouse team, public health and the Institute of Health and Innovation to join one of these datasets with information held in the primary care record. The student will use the results of study one and two, as well as discussions with supervisors and PPIE group to determine which dataset to link and to identify key indicators.**References**1. ONS. Health state life expectancies by national deprivation deciles, England: 2016-2018. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/bulletins/healthstatelifeexpectanciesbyindexofmultipledeprivationimd/2016to2018. Accessed 08.09.2021.
2. Build back fairer: the Covid-19 Marmot review. URL:https://www.health.org.uk/publications/build-back-fairer-the-covid-19-marmot-review. Accessed 08.09.21.
3. WHO. Social determinants of health. URL: https://www.who.int/health-topics/social-determinants-of-health#tab=tab\_1. Accessed: 08.09.21
4. Kostelanetz S, Pettapiece-Phillips M, Weems J, Spalding T, Roumie C., Wilkiins C.H, Kripalani S. Health Care Professionals’ perspectives on Universal Screening of Social Determinants of Health: A mixed-Methods Study. *Population Health Management*, 2021. Doi: <https://doi.org/10.1089/pop.2021.0176>
5. Ministry of housing communities and local government. The English Indices of Deprivation 2019. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/835115/IoD2019\_Statistical\_Release.pdf. Accessed: 08.09.21
6. Pinto AD, Bondy M, Ruccetto A, Ihnat j, Kaufman A. screening for poverty and intervening in a primary care setting: an acceptability and feasibility study. Family Practice. 2019; 36(5):634–638. Doi: https://doi.org/10.1093/fampra/cmy129
7. Chhabra M, Sorrentino AE, Cusack M. *et al.* Screening for Housing Instability: Providers’ Reflections on Addressing a Social Determinant of Health. *J GEN INTERN MED. 2019;*34, 1213–1219.
8. Kreuter MW, Thompson T, McQueen A, Garg R. Addressing Social Needs in Health Care Settings: Evidence, Challenges, and Opportunities for Public Health. *Annual Review of Public Health.* 2021 42:1, 329-344. Doi: https://doi.org/10.1146/annurev-publhealth-090419-102204.
9. Chen M, Tan X, Padman R, Social determinants of health in electronic health records and their impact on analysis and risk prediction: A systematic review. *Journal of the American Medical Informatics Association*, 2020; 27(11), 1764–1773. Doi:  https://doi.org/10.1093/jamia/ocaa143
10. Tong ST, Liaw WR, Kashiri PL, Pecsok J, Rozman J, Bazemore AW, Krist AH. Clinical experiences of screening for social needs in primary care. *Journal of the American Board of Family Medicine,* 2018; 31(3): 351-363. Doi:10.3122/jabfm.2018.03.170419
11. Krieger N. Overcoming the absence of socio-economic data in medical records: validation and application of a census- based methodology. *Am J Public Health,* 1992;82:703–10.
12. Robert SA, Strombom I, Trentham-Dietz A, Hampton JM, McElroy JA, Newcomb PA, Remington PL. Socioeconomic risk factors for breast cancer: distinguishing individual- and community-level effects. *Epidemiology,* 2004;15(4):442-50. Doi: 10.1097/01.ede.0000129512.61698.03.
13. Pampalon R, Hame D, Gamache P. A comparison of individual and area-based socio-economic data for monitoring social inequalities in health. *Component of Statistics Canada Catalogue Health Reports,* 2009; 20(3): 85-94.
14. Mustard CA, Derksen S, Berthelot JM, Wolfson M. Assessing ecologic proxies for household income: a comparison of household and neighbourhood level income measures in the study of population health status. *Health and Place,* 1999; 5(2): 157-71.
15. Molina AL, Molina Y, Walley SC, Wu CL, Zhu A, Oates G. Residential Instability, neighbourhood deprivation, and paediatric asthma outcomes. *Pediatric Pulmonology,* 2020; 55(6): 1340-1348.
16. Forbes L, Harvey S, Newson R, et al. Risk factors for accident and emergency (A&E) attendance for asthma in inner city children. *Thorax*. 2007;62(10):855-860. doi:10.1136/thx.2006.058362
17. Subramanian SV, Chen JT, Rehkopf DH, Waterman PD, Krieger N. Comparing Individual- and Area-based Socioeconomic Measures for the Surveillance of Health Disparities: A Multilevel Analysis of Massachusetts Births, 1989-1991. *Am J Epidemiol,* 2006; 164:823-834
18. Rehkopf DH, Haughton LT, Chen JT, Waterman PD, Subramanian SV, Krieger N.

Monitoring Socioeconomic Disparities in Death: Comparing Individual-Level Education and Area-Based Socioeconomic Measures. *American Journal of Public Health,* 2006;96, 2135-2138. Doi: https://doi.org/10.2105/AJPH.2005.0754081. Diez-Roux AV, Kiefe CI, Jacobs DR Jr, Haan M, Jackson SA, Nieto FJ, Paton CC, Schulz R. Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Ann Epidemiol,* 2001;11(6):395-405. Doi: 10.1016/s1047-2797(01)00221-6.
2. Soobader M, LeClere FB, Hadden W, Maury B. Using aggregate geographic data to proxy individual socioeconomic status: does size matter? *Am J Public Health*. 2001;91(4):632-636. Doi:10.2105/ajph.91.4.632

Subramanian SV,The relevance of multilevel statistical methods for identifying causal neighborhood effects. *Social Science & Medicine* 2004; 58(10): 1961-1967. Doi:https://doi.org/10.1016/S0277-9536(03)00415-5. |

|  |
| --- |
| Training and development provision by host: |
| *Formal training:*It is expected that the student will participate in:* Systematic literature reviewing and evidence synthesis workshops hosted by Keele SoM
* Understanding primary care EHR data and practical handling of large EHR datasets – by online training, self-directed reading, supervised analysis, discussion with supervisors and external course attendance, e.g. Manchester Digital Epidemiology Summer School (£600 + accommodation/travel)
* Multilevel linear regression concepts and methods by external course attendance (e.g. RSS Multi-level Modelling (£623 plus VAT)

  |
| *Informal training:* It is expected that the student will participate in:* Keele SOM internal and external seminar programme
* Keele SoM Postgraduate student group
* Keele SoM post-graduate research symposium
 |
| *PPIE:* The student will engage with PPIE at a number of stages during this project. This will include but is not limited to advice on understanding results of study 1-3 above, advice on variables to select concerning the individual level social determinants of health from the MIDAS population survey for analysis of area versus individual level data, advice on key indicators to select for the data linkage study (study 3), advice on dissemination strategies. |