An application of matching algorithms to generalize small-area estimates of chronic pain prevalence to neighbourhoods across England

QUINN, L <http://orcid.org/0009-0004-4076-3827>, YU, D <http://orcid.org/0000-0002-8449-7725>, LYNCH, M <http://orcid.org/0000-0003-4882-6875>, JORDAN, KP <http://orcid.org/0000-0003-4748-5335>, WILKIE, R <http://orcid.org/0000-0003-4825-714X> and PEAT, George <http://orcid.org/0000-0002-9008-0184>

Journal of Public Health | pp. 1–9 | https://doi.org/10.1093/pubmed/fdaf002

An application of matching algorithms to generalize small-area estimates of chronic pain prevalence to neighbourhoods across England

L. Quinn[1,2], **D. Yu**[3], **M. Lynch**[4], **K.P. Jordan**[3], **R. Wilkie**[3] and **G. Peat**[2,5,*]

[1]Adult Care, Housing and Public Health, Rotherham Metropolitan Borough Council, Riverside House, Main Street, Rotherham S60 1AE, UK
[2]Centre for Applied Health & Social Care Research (CARe), Robert Winston Building, Broomhall Road, Sheffield Hallam University, Sheffield S10 2BP, UK
[3]Primary Care Centre Versus Arthritis, School of Medicine, Keele University, Keele, Staffordshire, ST5 5BG, UK
[4]School of Medicine, Keele University, Keele, Staffordshire ST5 5BG, UK
[5]Honorary Professor, Primary Care Centre Versus Arthritis, School of Medicine, Keele University, Keele, Staffordshire ST5 2HE, UK

*Corresponding author. Professor George Peat, Centre for Applied Health & Social Care Research (CARe), Robert Winston Building, Broomhall Road, Sheffield Hallam University, Sheffield S10 2BP, UK E-mail: g.peat@shu.ac.uk

ABSTRACT

Background: Local decision-makers lack granular data on the prevalence of chronic pain in their populations. We applied matching methods to generalize estimates from one local survey in England to other neighborhoods across the country with a similar sociodemographic composition.

Methods: We used propensity score matching to match lower-layer super output areas (LSOA) across England with 230 surveyed LSOAs in North Staffordshire by age, sex, ethnicity, deprivation, and rurality. LSOA-specific crude prevalence of chronic pain and high-impact chronic pain in adults aged 35+ years were estimated and mapped for matched LSOAs.

Results: Satisfactory matching was achieved for 24 871 of 31 580 LSOAs (79%). The 6709 LSOAs identified as either "off common support" or unmatched were principally inner-city neighborhoods with younger, more ethnically diverse populations. LSOA-specific estimates of chronic pain and high impact chronic pain ranged from 14% to 52% and from 4% to 31% respectively. Integrated Care Board estimates ranged from 27% to 38% and from 10% to 18%, respectively. Estimates for England were 31.9% and 12.6%, respectively.

Conclusions: Using matching methods we have produced the first detailed map of the distribution of chronic pain in England but with several strong assumptions. Our estimates highlight substantial variation in prevalence within ICBs.

Keywords: epidemiology; health intelligence; public health

Introduction

The need for better, more granular data on health, care and their determinants to support local decision-making has been repeatedly highlighted in recent national reports,[1,2] with one expert roundtable concluding that "data at local and hyper-local levels must be addressed if we are to have a clearer picture of health inequalities, those affected, and to design appropriate policies".[3] Realizing this goal is particularly challenging for health indicators that are not prioritized within national outcomes frameworks. Chronic pain is one such case: it affects an estimated 19–28 million adults in the UK,[4] healthcare costs attributable to back pain alone exceed £3 billion,[5] and the costs of informal care and productivity losses are likely to be several times greater still.[6] However, chronic pain is not specifically prioritized within either the Public Health Outcomes Framework or the Quality and Outcomes Framework (QOF) in general practice, and there is "a near absence of research on the geographic distribution of pain at subnational levels".[7] Analysis of 2011 Health Survey for England data produced prevalence estimates of chronic pain but only down to the level of local authorities.[8]

L. Quinn, Public Health Intelligence Principal & NIHR Predoctoral Local Authority Fellow
D. Yu, Senior Research Fellow
M. Lynch, Lecturer in Public Health
KP. Jordan, Professor of Biostatistics
R. Wilkie, Professor of Public Health and Epidemiology
G. Peat, Professor of Clinical Epidemiology

Substantial variation in the prevalence of disabling chronic pain between neighborhoods [lower-level super output areas (LSOA)] within a single Integrated Care Board (ICB) was recently revealed by a local survey[9] suggesting that LSOA-level estimates may be important for targeting actions and services. In the absence of expensive, large surveys of chronic pain using national probability sampling that can produce such granular data across all local authorities, an alternative is to critically consider whether, how, and how far existing modeled estimates from local surveys (i.e. using non-probabilistic sampling of neighborhoods) can be generalized to other parts of the country.

In this study we drew on Rosenbaum and Rubin's seminal work on propensity scores[10] as a framework for considering generalizability of findings from samples to populations by balancing groups on a set of observed covariates.[11–14] Specifically, we took modeled LSOA-specific prevalence estimates of chronic pain from a local survey and used matching on a balancing score to identify LSOAs in England with similar characteristics as a means of highlighting where prevalence estimates might be generalized with greater or lesser degrees of confidence.

Methods

Study design, setting, and population

Descriptive study with propensity score matching, in the general population in England, adults aged 35 years and over.

Data sources

We used three main data sources for this study. Firstly, we used regression model coefficients that had been previously derived to produce modeled estimates of the prevalence of chronic pain and high-impact chronic pain in adults aged over 35 years at the level of Lower-level Super Output Area. These had been produced from an analysis of the PRELIM survey of adults aged 35 years and over registered with 11 general practices in North Staffordshire & Stoke-on-Trent (n = 4389 respondents, adjusted response rate 48%).[9,15] The PRELIM survey mailed a single questionnaire, with an invitation letter from their general practice, an information sheet, and prepaid return envelope, to a random sample of all adults aged 35 years and over in June 2017. A repeat pack was sent to non-respondents at 2 weeks, with the offer of online questionnaire completion, and after a further 2 weeks, non-respondents were mailed a shortened questionnaire containing selected outcome measures and descriptive fields. We used previously validated self-report measures of chronic pain. Chronic pain was defined as pain on most days or more in the past 6 months; high-impact chronic pain as pain on most days or more in the past 6 months which limited activities on most days or more over the same period.[16] Covariates in the regression models were LSOA age-sex composition, deprivation (Index of Multiple Deprivation (2015) decile based on national rank), rurality (classified as urban or rural based on the 2011 Rural–Urban Classification), and a correction factor for ethnicity (due to limited numbers of survey respondents from, Black, Asian, Mixed, and multiple ethnic backgrounds). Further details of the survey methodology and modeling procedures is provided in Lynch *et al*.[9]

Secondly, we used information on age, sex, sex by age, rurality, and ethnicity of all LSOAs in England extracted from Office for National Statistics 2021 Census (datasets TS007A, TS021, RM121, TS008), English indices of deprivation 2019 decile scores, and the 2011 rural/urban classification for LSOA boundaries. Thirdly, we used shapefiles for 2021 LSOA boundaries and ICB boundaries in England obtained from the Open Geography Portal and transformed using ArcGIS, and displayed in Power BI.

Statistical analysis

Matching analyses using propensity scores were performed to match "control" LSOAs (all LSOAs in England, minus focal LSOAs) to "focal" LSOAs (LSOAs in North Staffordshire & Stoke-on-Trent that had at least one PRELIM survey respondent—the minimal level of observed data per LSOA). Matching was conducted in two steps. After comparing the age, sex, deprivation, rurality and ethnicity distributions of focal and control LSOAs, we excluded control LSOAs that were outside the region of common support (i.e. propensity score distances beyond the range of distance measures of focal LSOAs). We then performed full optimal matching of the remaining control LSOAs to the focal LSOAs. Satisfactory matching was defined as absolute standardized mean differences of <0.1 on all covariates.[17] Matching analyses were performed using the *MatchIt* package,[18] with covariate distributions at each stage assessed using "cobalt",[19] both in R 4.3.0.

Upon completion of the matching process, we applied the regression model coefficients to estimate the prevalence of chronic pain and high-impact chronic pain in all matched control LSOAs in England. We produced choropleth maps for England and for each of the 42 ICBs in England using Microsoft PowerBI Version 2.107.683.0 64-bit (July 2022).

To explore the potential for residual bias after matching we compared focal and control LSOAs on their prevalence of poor general health and a long-term disabling condition. These two measures were available for all LSOAs in England from the national Census 2021 (datasets TS037 and TS038, extracted from NOMIS) and showed moderate-to-strong correlations with modeled estimates of chronic pain and

high-impact chronic pain (r = 0.55–0.76) in the focal LSOAs where both sets of estimates were available. Mean differences between focal and control LSOAs in the prevalence of poor general health and in the prevalence of a long-term disabling condition were estimated using the "marginal effects" package[20] after full optimal matching on the propensity score.

In the absence of direct comparative estimates of chronic pain, we also explored the relationship between our estimates and available prevalence estimates of two musculoskeletal conditions—rheumatoid arthritis and osteoporosis—which, although uncommon, are associated with chronic pain.[21,22] Estimates for these two conditions were obtained from QOF 2022/2023 data from general practice.[23]

Patient and public involvement

The MIDAS program has a dedicated Public Advisory Group (PAG) comprising seven people with lived experience of musculoskeletal conditions drawn from Keele University's Research User Group. The PAG met with the MIDAS Program Lead, Chief Investigators, Trial Manager, and other members of the research team on a monthly basis via MS Teams. PAG members advised on the design of the study and interpretation of the findings.

Results

Of 33,755 LSOAs in Census 2021, 1945 had no deprivation score at the time of analysis due to boundary changes, leaving a total of 31 810 LSOAs with complete data on age, sex, ethnicity, deprivation, and rurality. PRELIM survey respondents came from 230 of these LSOAs across North Staffordshire, Stoke-on-Trent and bordering areas (n = 230 "focal" LSOAs). Compared with the 31 580 "control" LSOAs elsewhere in England, focal LSOAs had, on average, lower ethnic diversity, older age distribution, and higher deprivation. Satisfactory matching was achieved for 24 871 of 31 580 (79%) LSOAs elsewhere in England (Fig. 1 and Table 1). Based on non-overlapping propensity score distributions, a total of 6204 "control" LSOAs were found to be "off common support," with a further 505 discarded or unmatched during full optimal matching. Detailed tables and figures comparing covariate distributions and balance between focal and control LSOAs are provided in the Supplementary Data (Table S1).

LSOA-specific estimates of chronic pain prevalence in the population aged 35+, among matched LSOAs ranged from 14.1% to 50.6% (Fig. 2 and Table 2). ICB estimates ranged from 26.8% in NHS Buckinghamshire, Oxfordshire and Berkshire West ICB to 37.2% in NHS Black Country ICB (Table 2 and Fig. S1). The aggregated national prevalence estimate for England was 31.8%.

LSOA-specific estimates of the prevalence of high impact chronic pain among matched LSOAs in adults aged 35+ years ranged from 4.3% to 30.4%. At ICB-level, estimates ranged from 9.9% in NHS Buckinghamshire, Oxfordshire and Berkshire West ICB to 16.5% in NHS Birmingham and Solihull ICB. The aggregated national prevalence estimate for England was 12.6%. All LSOA- and ICB-level estimates, including for unmatched LSOAs are provided in Supplementary Data (Table S2). All choropleth maps in PowerBI are freely available at the website listed under Data Availability at the end of this manuscript.

Exploring potential residual bias

With full optimal matching on age, sex, ethnicity, deprivation, and rurality, focal LSOAs in North Staffordshire and Stoke-on-Trent had a marginally higher average prevalence of people reporting bad or very bad general health [6.5% (95%CI: 6.4, 6.7) vs. 6.1% (6.1, 6.2)] and those living with a long-term disabling condition [20.5% (20.3, 20.7) vs. 19.9 (19.8, 20.0)] compared to LSOAs elsewhere in England. This difference persisted when we re-ran full optimal matching with exact matching on IMD decile (data not shown).

We found little or no correlation at local authority level between our chronic pain and high impact chronic pain estimates and pooling of available GP-level estimates of osteoporosis and rheumatoid arthritis from QOF data (Supplementary Data).

Discussion

Main finding of this study

Our study applied matching methods to explore the generalizability of LSOA-specific modeled prevalence estimates of chronic pain and high-impact chronic pain from one geographical area to neighborhoods across England. Based on matching LSOAs on five covariates—age, sex, deprivation, rurality, and ethnicity—we found satisfactory matching to 79% of LSOAs with available data. Conversely, roughly one in five LSOAs could not be matched: these were principally neighborhoods with younger, more ethnically diverse populations. LSOA-specific prevalence estimates for chronic pain and high-impact chronic pain for matched neighborhoods showed substantial variation within ICBs.

What is already known on this topic

Previous estimates for the prevalence of chronic pain and moderate-severely disabling chronic pain among adults in the UK range from 35% to 51% and 10% to 14%, respectively.[4] Based on 2011 Health Survey for England data, Todd et al.[8] reported local authority estimates of the prevalence of

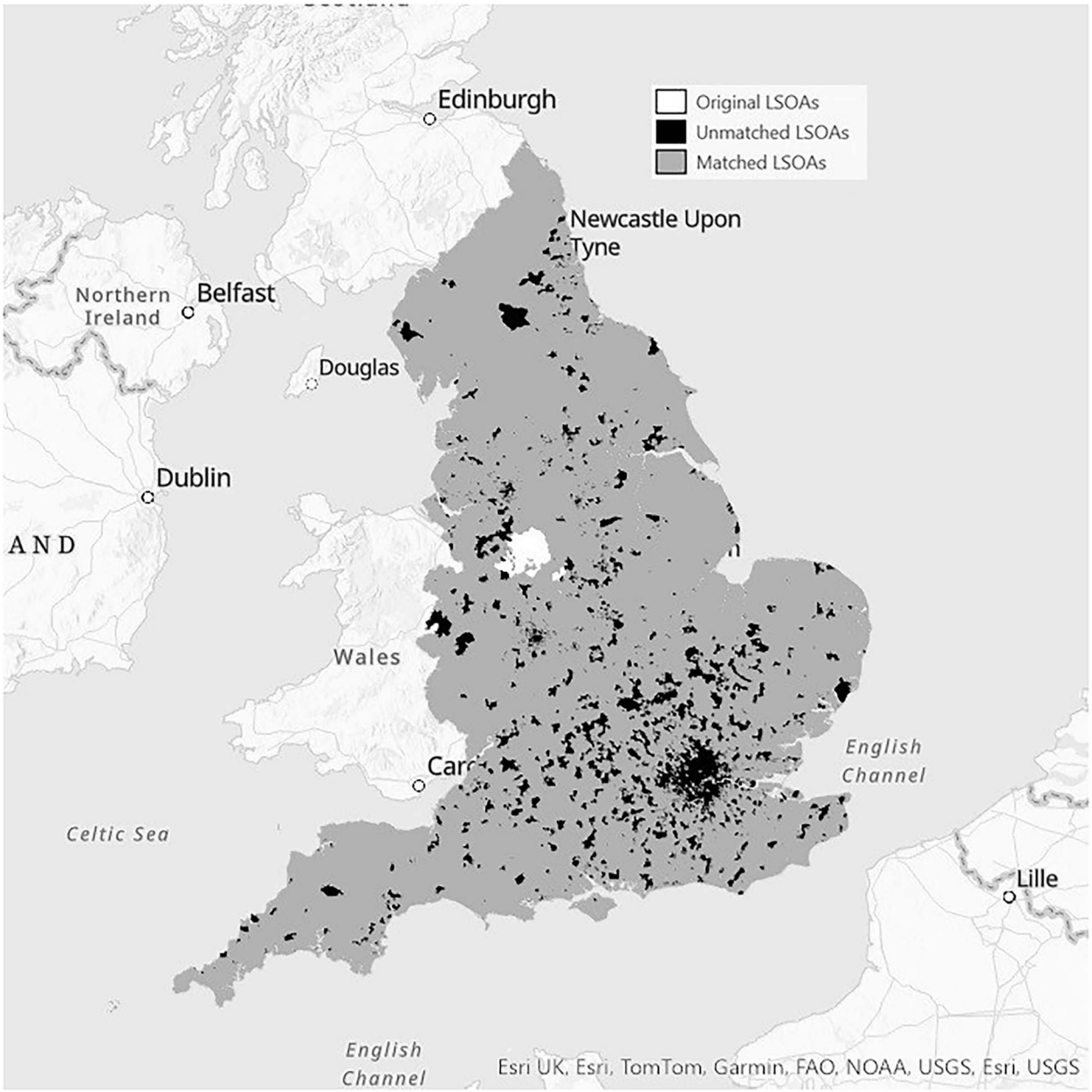

Figure 1. Map of matched and unmatched LSOAs, by lower-level super output area (LSOA), England.

chronic pain ranging from <21% to over 60%, with evidence of a North–South divide. The report on the Health Survey for England 2017 chronic pain survey confirmed the association between chronic pain and age, sex, socioeconomic position, area-level deprivation, and ethnicity but did not produce modeled estimates at lower geographies.[24]

What this study adds

To our knowledge, the current analyses have produced the first nationwide LSOA-level maps, and ICB-level estimates of chronic pain and high-impact chronic pain prevalence in England.

Limitations of this study

Our study explored the external validity or generalizability of modeled prevalence estimates from a local survey, assuming that the modeled estimates were internally valid, i.e. that multilevel regression and poststratification of a relatively large local survey with 48% response yielded unbiased estimates of the prevalence of chronic pain in adults aged 35 years and over

Table 1. Characteristics of neighborhoods before and after matching

	Focal LSOAs in North Staffordshire and Stoke-on-Trent and surrounds[a] *N* = 230		*All other LSOAs in England*					
			Before excluding those off common support *N* = 31 580		*After excluding those "off common support"* *N* = 25 376		*After further excluding those "off common support" and unmatched*[b] *N* = 24 871	
	n	%	*n*	%	*n*	%	*n*	%
Deprivation[c]								
d1 (most)	40	*17.4*	3147	*10.0*	2619	*10.3*	2564	*10.3*
d2	19	*8.3*	3169	*10.0*	2375	*9.4*	2339	*9.4*
d3	26	*11.3*	3138	*9.9*	2311	*9.1*	2277	*9.2*
d4	23	*10.0*	3152	*10.0*	2425	*9.6*	2385	*9.6*
d5	15	*6.5*	3166	*10.0*	2540	*10.0*	2500	*10.1*
d6	24	*10.4*	3128	*10.0*	2482	*9.8*	2440	*9.8*
d7	24	*10.4*	3160	*10.0*	2611	*10.3*	2566	*10.3*
d8	20	*8.7*	3149	*10.0*	2618	*10.3*	2571	*10.3*
d9	29	*12.6*	3162	*10.0*	2620	*10.3*	2533	*10.2*
d10 (least)	10	*4.4*	3209	*10.2*	2775	*10.9*	2696	*10.8*
Rural[d]	24	*10.4*	5374	*17.0*	4852	*19.1*	4464	*18.8*
% male: median (IQR)	49.2 (48.4, 50.1)		48.8 (47.8, 49.7)		48.9 (48.0, 49.8)		48.9 (48.0, 49.8)	
% age group: median (IQR)								
35–44	11.5 (10.1, 13.2)		12.7 (10.9, 14.6)		12.1 (10.5, 13.7)		12.1 (10.5, 13.7)	
45–54	13.5 (12.0, 14.7)		13.5 (12.3, 14.7)		13.5 (12.4, 14.7)		13.5 (12.4, 14.7)	
55–64	13.8 (11.6, 15.6)		13.0 (11.0, 15.0)		13.6 (11.7, 15.4)		13.6 (11.7, 15.4)	
65–74	11.6 (8.3, 14.7)		10.1 (7.3, 13.1)		11.0 (8.3, 13.8)		11.0 (8.4, 13.8)	
75–84	7.0 (5.0, 9.9)		6.0 (4.0, 8.5)		6.7 (4.7, 9.1)		6.7 (4.7, 9.1)	
85+	2.5 (1.8, 3.5)		2.2 (1.4, 3.4)		2.4 (1.6, 3.5)		2.4 (1.6, 3.5)	
% ethnic background[e]: median (IQR)								
White	94.8 (89.9, 97.5)		91.8 (75.8, 96.5)		94.2 (86.4, 97.0)		94.2 (86.4, 97.0)	
Black	0.7 (0.2, 2.1)		1.1 (0.3, 4.1)		0.7 (0.2, 2.1)		0.7 (0.3, 2.1)	
Asian	2.0 (0.8, 4.6)		3.3 (1.2, 9.8)		2.2 (1.0, 6.0)		2.2 (1.0, 6.0)	
Mixed/Multiple	2.0 (1.2, 3.2)		3.2 (1.8, 6.7)		2.5 (1.5, 4.5)		2.5 (1.5, 4.5)	
% bad/very bad general health	6.2 (4.7, 8.2)		4.9 (3.6, 6.5)		5.1 (3.8, 6.8)		5.1 (3.8, 6.8)	
% long-term disabling condition	20.0 (17.6, 23.3)		17.2 (14.2, 20.7)		18.0 (15.1, 21.4)		18.0 (15.1, 21.4)	

d1 . . . d10 decile (national rank) IQR Interquartile range; LSOA Lower-level Super Output Area.

[a]Defined as LSOAs with one or more respondent to PRELIM survey.

[b]Defined by propensity score of all covariates listed in table.

[c]Index of multiple deprivation (2019).

[d]Rurality index (2011) dichotomized at urban (A1-C2) and rural (D1-F2).

[e]Self-reported, and grouped.

in North Staffordshire and Stoke-on-Trent. We were limited in the covariates we used, both in our original models and in the current study for matching, to five that were available at LSOA level and associated with chronic pain. The availability of more LSOA-level covariates strongly associated with the occurrence of chronic pain would be expected to improve both the models and the quality of matches. Our finding that neighborhoods elsewhere in England still had slightly lower rates of poor general health and long-term disabling conditions than focal LSOAs in North Staffordshire and Stoke-on-Trent even after optimal full matching may reflect this limitation. However, we cannot know if such differences would reduce if we could additionally include the prevalence of anxiety, depression, or obesity, for example, as covariates

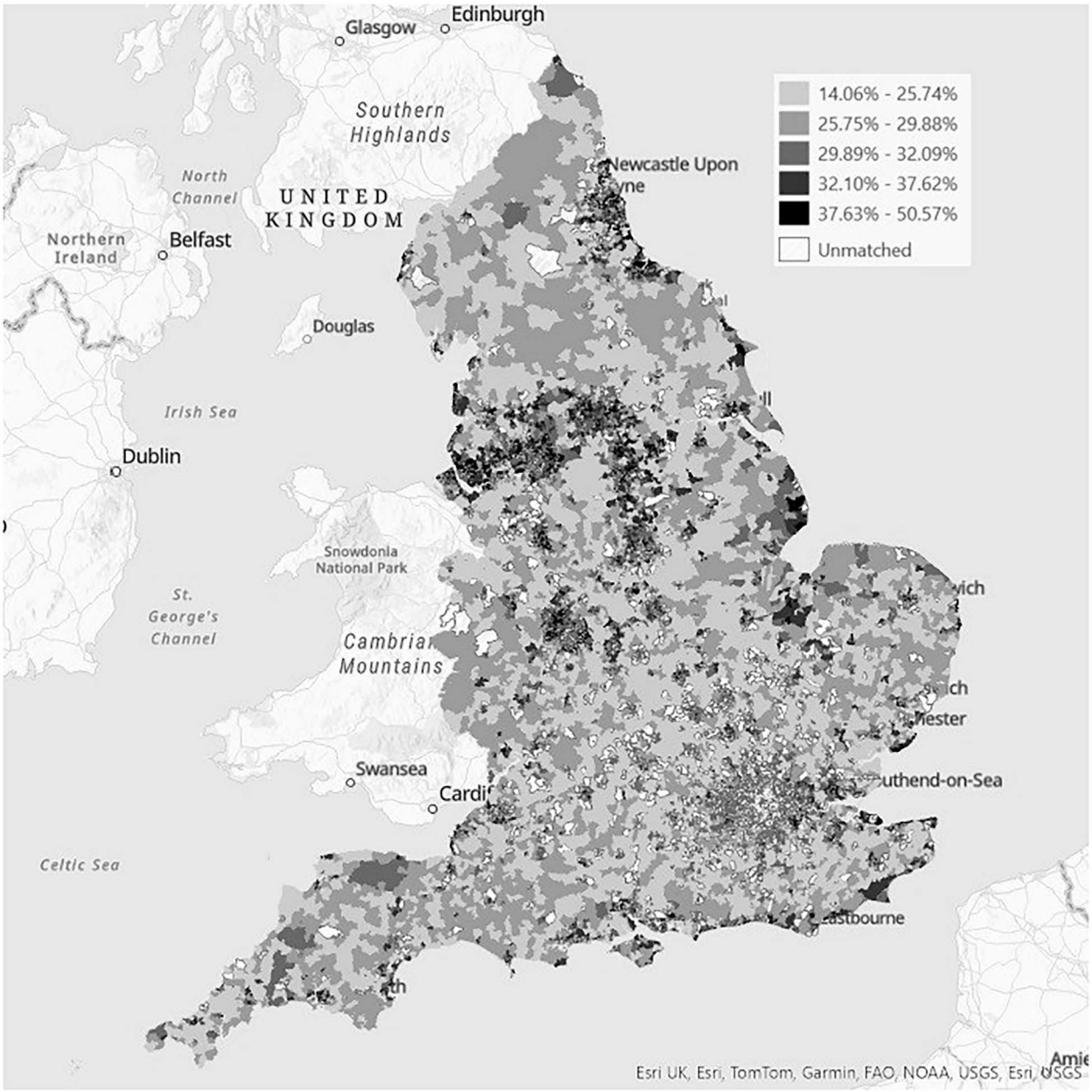

Figure 2. Map of chronic pain prevalence in matched LSOAs, in adults aged 35 years and older, by lowerlevel super output area (LSOA).

in our models and in our matching algorithms. Nevertheless, the magnitude of any residual bias, in the direction of over-estimating the prevalence of chronic pain and high impact chronic pain elsewhere in England is, on average, likely to be small (<1% difference in prevalence). Lack of coverage of younger, more ethnically diverse neighborhoods in the original PRELIM survey is made transparent in the matching process but other limitations are not, e.g. lack of coastal neighborhoods in North Staffordshire survey. One solution to this would be to apply a correction factor to the modeled estimates. The estimates were obtained from a survey conducted in 2017 before the COVID-19 pandemic. While there is concern that the prevalence of chronic pain may have increased after COVID-19, previous repeated cross-sectional surveys suggest that the prevalence of chronic pain is relatively stable over intervals of several years.[24] The lack of correlation with available GP estimates of osteoporosis and rheumatoid arthritis provided limited insight into the validity of our estimates. Both conditions are relatively uncommon causes of chronic pain in the adult population and there was

Table 2. ICB-specific estimates (in ascending order of estimated prevalence of chronic pain) of matched LSOAs

	LSOAs		*Chronic pain*			*High impact chronic pain*		
NHS ICB	*Total N*	*Matched N (%)*	*%*	*Min*[a]	*Max*	*%*	*Min*	*Max*
Buckinghamshire, Oxfordshire and Berkshire West	964	711 (74)	26.8	17.5	46.9	9.9	6.2	25.1
Frimley	426	327 (77)	27.7	19.1	43.7	10.3	6.6	21.5
Surrey Heartlands	601	466 (78)	27.8	18.6	42.2	10.2	6.9	20.1
Hertfordshire and West Essex	829	591 (71)	28.5	18.4	46.2	10.4	6.5	24.7
South West London	842	111 (13)	28.7	23.1	47.0	10.6	7.1	25.3
Bath and North East Somerset, Swindon and Wiltshire	511	486 (95)	28.8	16.3	47.7	10.7	4.3	26.5
Gloucestershire	354	339 (96)	29.1	19.1	47.8	10.9	6.7	26.4
Somerset	313	309 (99)	29.6	19.3	47.5	11.2	6.0	26.0
Lincolnshire	406	395 (97)	29.7	15.2	50.0	11.4	5.3	29.0
Cambridgeshire and Peterborough	466	369 (79)	29.8	18.4	48.1	11.4	6.7	27.1
Herefordshire and Worcestershire	470	463 (99)	30.0	18.6	47.3	11.4	6.7	25.7
Cornwall and the Isles of Scilly	315	303 (96)	30.1	22.5	48.9	11.6	7.2	27.6
Leicester, Leicestershire and Rutland	587	509 (87)	30.1	17.6	50.0	11.5	6.1	28.8
South East London	974	213 (22)	30.2	23.5	43.9	11.2	8.2	22.1
Shropshire, Telford and Wrekin	286	277 (97)	30.3	19.9	48.3	11.7	6.3	27.2
Northamptonshire	403	376 (93)	30.3	18.0	48.2	11.4	6.1	27.2
Hampshire and Isle of Wight	1069	1000 (94)	30.4	15.2	48.1	11.7	5.5	26.7
Mid and South Essex	690	639 (93)	30.5	18.4	47.9	11.5	6.8	26.4
Humber and North Yorkshire	999	963 (96)	30.5	18.0	50.0	12.1	6.5	28.7
Suffolk and North East Essex	545	500 (92)	30.6	17.5	49.6	11.8	6.3	28.5
Norfolk and Waveney	590	567 (96)	30.6	14.1	47.9	12.0	4.8	27.0
Coventry and Warwickshire	515	479 (93)	30.6	19.1	48.4	11.7	6.1	27.2
Dorset	438	416 (95)	30.7	19.9	48.5	11.7	6.3	27.2
Bedfordshire, Luton and Milton Keynes	504	379 (75)	30.8	18.4	49.3	11.6	6.8	28.4
Sussex	970	880 (91)	30.9	17.7	50.6	11.8	5.5	30.4
Devon	689	669 (97)	31.0	18.9	49.5	11.9	6.4	28.3
Bristol, North Somerset and South Gloucestershire	549	484 (88)	31.1	18.2	49.0	12.1	7.0	28.6
Kent and Medway	1028	955 (93)	31.3	17.7	48.7	12.1	6.4	27.8
Derby and Derbyshire	633	589 (93)	31.5	18.9	48.5	12.2	6.7	27.3
Staffordshire and Stoke-on-Trent	674	663 (98)	32.0	17.8	49.0	12.5	6.8	27.9
Nottingham and Nottinghamshire	658	551 (84)	32.3	18.2	49.0	13.0	7.2	28.4
North West London	1105	94 (9)	32.4	25.5	45.2	11.9	8.4	23.4
Lancashire and South Cumbria	1034	1011 (98)	33.6	19.0	50.2	13.9	6.6	29.5
North East and North Cumbria	1800	1730 (96)	33.8	18.8	49.9	14.0	6.0	28.9
North East London	966	336 (35)	33.8	24.3	46.6	12.9	8.1	25.0
Cheshire and Merseyside	1518	1438 (95)	34.6	18.5	49.6	14.5	6.6	28.4
West Yorkshire	1395	1249 (90)	34.6	18.7	49.5	14.4	6.1	28.9
North Central London	772	5 (1)	34.6	30.2	43.4	14.3	10.9	21.1
South Yorkshire	835	756 (91)	35.4	19.3	49.9	14.9	6.7	29.1
Greater Manchester	1636	1427 (87)	35.7	21.5	49.7	15.1	6.6	29.4
Birmingham and Solihull	750	505 (67)	36.9	18.8	49.8	16.5	7.0	28.7
Black Country	701	571 (81)	37.2	22.9	50.0	16.2	8.7	29.0

[a]Minimum and maximum estimates for matched LSOAs within the ICB.

wide variation in osteoporosis and rheumatoid arthritis estimates between practices within the same local authority. This underscores the dearth of chronic pain estimates currently available at ICB, local authority, and neighborhood levels. We have presented only point estimates. In our previous study,[7] obtaining confidence intervals for just 300 LSOAs was very computationally intensive and was judged prohibitive in the current analysis. The LSOA-level point estimates presented in this work must be understood to have a high degree of uncertainty. However, for many applications, identifying LSOAs that are in the highest decile or quintile of chronic pain prevalence may be sufficient. Bearing in mind the above limitations, and recognizing the several strong assumptions that these estimates still require, we argue that the modeled point estimates nevertheless provide potentially new insights into variations in chronic pain within local health systems in England that may be of use for future research (e.g. targeting areas of presumed high need for recruitment of participants to interventional studies) and service planning.

Acknowledgements

We are extremely grateful for the input of Ben Anderson, Director of Public Health, Rotherham Metropolitan Borough Council, and for our public contributors to the study. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

Author contributors

Conception and design: G.P., L.Q. Data acquisition and analysis: L.Q., G.P. Interpretation of data: L.Q., G.P. Drafting and critical review of content: G.P., L.Q. Final approval of submitted manuscript: L.Q, G.P.

Supplementary data

Supplementary data are available at the *Journal of Public Health* online.

Conflict of interest

The authors have no conflicts to declare.

Funding

Lorna Quinn, Pre-doctoral Local Authority Fellow [NIHR303505], is funded by the NIHR for this research project. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, NHS or the UK Department of Health and Social Care. There was project funding from Versus Arthritis for the PRELIM Survey [21403] and from Nuffield Foundation and Versus Arthritis for the MIDAS Project [OBF/43390]. Visit http://www.nuffieldfoundation.org/. The views expressed are those of the authors and not necessarily the funders.

Data availability

The data underlying this article are available in [MIDAS OSF], at https://osf.io/e542w/. Interactive maps displaying the prevalence estimates derived in this study are publicly available at https://app.powerbi.com/view?r=eyJrIjoiYjE5MTAzMTgtZjlkZC00ZWZiLWJkYmYtYWYxYWRjMzgxZGY4IiwidCI6IjQ2ZmJlNmZkLTc4YWUtNDc2OS05YzFkLWJjZWE5NzM3OGFmNiJ9. The datasets were derived from sources in the public domain: [PRELIM regression model coefficients and LSOA-specific estimates for North Staffordshire & Stoke-on-Trent available at https://onlinelibrary.wiley.com/doi/10.1002/ejp.2148; Office for National Statistics 2021 Census datasets TS007A https://www.ons.gov.uk/datasets/TS007/editions/2021/versions/1, TS021 https://www.ons.gov.uk/datasets/TS021/editions/2021/versions/1, RM121 https://www.ons.gov.uk/datasets/RM121/editions/2021/versions/1, TS008 https://www.ons.gov.uk/datasets/TS008/editions/2021/versions/4;

English indices of deprivation 2019 English indices of deprivation 2019 - GOV.UK (www.gov.uk).

Ethical approval

Ethical approval for the PRELIM survey was obtained from the North West—Greater Manchester East Research Ethics Committee (Reference: 15/NW/0735).

References

1. Goldacre B, Morley J. *Better, Broader, Safer: Using Health Data for Research and Analysis. A Review Commissioned by the Secretary of State for Health and Social Care*. Department of Health and Social Care, London, 2022. https://www.goldacrereview.org/.
2. Department of Health & Social Care. Data saves lives: reshaping health and social care with data. 2022. https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data.
3. The British Academy. *Historic and Geographic Patterns of Health Inequalities. Report of a Roundtable*. The British Academy & The Academy of Medical Sciences, London, 2022. https://www.thebritishacademy.ac.uk/publications/historic-and-geographic-patterns-of-health-inequalities/.
4. Fayaz A, Croft P, Langford RM. *et al.* Prevalence of chronic pain in the UK: A systematic review and meta-analysis of population studies. *BMJ Open* 2016;**6**:e010364–4.

5. Zemedikun DT, Kigozi J, Wynne-Jones G. *et al.* Healthcare resource utilisation and economic burden attributable to back pain in primary care: A matched case-control study in the United Kingdom. *Br J Pain* 2024;**18**:137–47.
6. Maniadakis N, Gray A. The economic burden of back pain in the UK. *Pain* 2000;**84**:95–103.
7. Zajacova A, Lee J, Grol-Prokopczyk H. The geography of pain in the United States and Canada. *J Pain* 2022;**23**:2155–66.
8. Todd A, Akhter N, Cairns J. *et al.* The pain divide: A cross-sectional analysis of chronic pain prevalence, pain intensity and opioid utilisation in England. *BMJ Open* 2018;**8**:e023391–1.
9. Lynch M, Peat G, Jordan K. *et al.* Where does it hurt? Small area estimates and inequality in the prevalence of chronic pain. *Eur J Pain* 2023;**27**:1177–86.
10. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;**70**: 41–55.
11. Stuart EA, Cole SR, Bradshaw CP. *et al.* The use of propensity scores to assess the generalizability of results from randomized trials. *JR Stat Soc Ser A* 2011;**174**:369–86.
12. Stuart EA. What is a propensity score? Applications and extensions of balancing score methods. *Observational Studies* 2023;**9**:113–7.
13. Mercer AW, Kreuter F, Keeter S. *et al.* Theory and practice in nonprobability surveys: Parallels between causal inference and survey inference. *Public Opin Q* 2017;**81**:250–71.
14. Tipton E, Hartman E. Generalization and transportability. In: Zubizaretta JR, Stuart EA, Small DS, Rosenbaum PR (eds.), *Handbook of Matching and Weighting Adjustments for Causal Inference*, 1st edn. Boca Raton: Chapman & Hall/CRC, 2023, 39–59.
15. Wilkie R, Yu D, Jordan K. *et al.* PRELIM survey (2017). *Keele Data Repository* 2020. https://doi.org/10.21252/5ag3-ta31.
16. Von Korff M, Scher AI, Helmick C. *et al.* United States National Pain Strategy for population research: Concepts, definitions, and pilot data. *J Pain* 2016;**17**:1068–80.
17. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;**46**:399–424.
18. Ho D, Imai K, King G. *et al.* MatchIt: Nonparametric Preprocessing for parametric causal inference. *J Stat Softw* 2011;**42**:1–28.
19. Greifer N. *_Cobalt: Covariate Balance Tables and Plots_. R Package Version 4.5.5.9000*. 2024. https://ngreifer.github.io/cobalt/articles/cobalt.html.
20. Arel-Bundock V, Griefer N, Heiss A. How to interpret statistical models using marginaleffects in R and Python. *J Stat Software* 2024;**11**:1–32.
21. Paolucci T, Saraceni VM, Piccinini G. Management of chronic pain in osteoporosis: Challenges and solutions. *J Pain Res* 2016;**9**:177–86.
22. Mathias K, Amarnani A, Pal N. *et al.* Chronic pain in patients with rheumatoid arthritis. *Curr Pain Headache Rep* 2021;**25**:1–11.
23. NHS Digital. Quality and Outcomes Framework, 2022–2023. 2023. Quality and Outcomes Framework, 2022-23 - NHS England Digital.
24. Public Health England. *Chronic Pain in Adults in 2017*. Health Survey for England, London, 2020. https://www.gov.uk/government/publications/chronic-pain-in-adults-2017.