International Journal of Population Data Science

Journal Website: www.ijpds.org

Inequalities in overcrowding in households with children in an ethnically diverse urban population: a cross-sectional study using linked health and housing records

Marta Wilk[1,*], Gill Harper[1], Silvia Liverani[2], Nicola Firman[1], Paul Simon[3], and Carol Dezateux[1]

Submission History

Submitted:	22/03/2024
Accepted:	21/11/2024
Published:	23/01/2025

[1]Centre for Primary Care, Wolfson Institute of Population Health, Faculty of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK
[2]School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK
[3]Endeavour Health Charity, Malton YO17 9RW, UK

Abstract

Introduction
Household overcrowding is an important determinant of health and is associated with adverse child health, educational and social outcomes.

Objectives
We aimed to determine whether households with dependent children were more likely to be overcrowded after taking into account household ethnicity and housing tenure in an urban, ethnically diverse, and disadvantaged London population by pseudonymously linking health and property data.

Methods
We used pseudonymised Unique Property Reference Numbers to link electronic health records to Energy Performance Certificate property data in north-east London and identified 332,473 households comprising 1,093,047 people. Our primary outcomes were overcrowding measures based on a bedroom standard and a space standard (space per person; m^2). We examined household level associations of overcrowding with presence of children in the household before and after adjusting for household ethnicity and tenure. We used multivariable logistic regression to estimate the adjusted odds (aOR) and 95% Confidence Intervals (CI) of bedroom standard overcrowding and linear regression to estimate effects (95% CI) on space per person.

Results
Overall, 42.8% (142,401/332,473) of households included children, 54.5% were of White household ethnicity, and 58.4% in private or social rented accommodation. 22.5% (32,075/142,401) and 45.9% (65,388/142,401) of households with children were overcrowded by the bedroom and space standards respectively compared with 4.7% (8,953/190,072) and 9.6% (18,229/190,072) without children. After adjusting for household ethnicity and housing tenure, households with children were more likely to be overcrowded (aOR [95% CI] 5.54 [5.40-5.68] and had 22.61m^2 (95%CI: -22.75,-22.46) less space per person than those without children.

Conclusions
Up-to-date estimates of household overcrowding measured by bedroom and space standards can be derived from linked housing and health records. Our findings highlight the inequalities in overcrowding experienced by households with children and enable future work using linked data to evaluate impacts of overcrowding on children's health.

Keywords
overcrowding; household; children; data linkage; social determinants of health; inequalities

*Corresponding Author:
Email Address: marta.wilk@qmul.ac.uk (Marta Wilk)

https://doi.org/10.23889/ijpds.v10i1.2408

Introduction

Household overcrowding is an important wider determinant of health. It has been implicated in the household transmission of infections, particularly in urban areas [1–5]. It is also a significant risk factor for condensation which can contribute to respiratory problems [6]. In their systematic review and meta-analysis, Baker et al. found strong evidence that household overcrowding was associated with increased risk of transmission of close contact infectious diseases, including gastroenteritis, pneumonia, lower respiratory tract infections, and tuberculosis [2]. Most of the literature included in this review focussed on the risks of overcrowding in preschool-aged children. They estimated that, in New Zealand, overcrowding accounted for around 10% of hospital admissions for one of the nine infection categories included in the meta-analyses. Recently, the 2019 Coronavirus (COVID-19) pandemic has highlighted the importance of household overcrowding as a factor in transmission [7–10]. However, the potential impact of overcrowding goes beyond transmission of infections, with evidence of associations with poorer self-rated health [11], and, for children, poorer educational attainment, behavioural problems [12], reduced well-being [13], increased risks of injury, including burns [14, 15], and health risks in later life [16]. Hence, timely estimates of household overcrowding are key to evaluating the wider determinants of health and informing local and central government policies.

Household overcrowding estimates in the UK are provided by the Office of National Statistics (ONS) and are based on data collected as part of the national Census [17]. Conducted once every 10 years, the most recent estimates of overcrowding from the Census relate to the population as it was in 2021, with data available at Lower layer Super Output Area (LSOA) geographies, comprising between 400 and 1,200 households [18]. Alternatively, conducted annually, the English Housing Survey [19] usually includes fewer than 20,000 households and lacks the coverage, geographic, and demographic categorisation provided by the Census. Neither provide an up-to-date picture of the prevalence of, and inequalities in, household overcrowding at the individual household level in England and Wales given their infrequency and scale, and furthermore, neither can be linked to health outcomes at the individual level.

Measurements of overcrowding in these surveys have focussed on estimating the number of bedrooms required in relation to the number of occupants. Findings from the 2022–23 English Housing Survey reveal that households with dependent children are more likely to experience overcrowding, especially when living in private and social rented accommodation [20]. This is consistent with findings from the 2021 Census [21]. There is evidence that young children living in overcrowded households experience adverse effects on their physical and mental health, education and family cohesion [13, 22].

It has been suggested that space per person provides a more sensitive measure of overcrowding as it takes account of dwelling size and number of occupants and allows household members to contribute equally to its calculation irrespective of age, gender, and relationships. It also includes all space shared in the household. Dwelling size is smaller, and numbers of occupants and costs per square metre higher, in the private and social rented sectors relative to owner occupiers [23–25]. However, agreed standards for assessing how much space a household needs are lacking, and little is known about inequalities in overcrowding using measures of space.

In this study, we propose a novel method for estimating overcrowding at a household level, using routinely collected primary care electronic health records (EHR) linked pseudonymously to publicly available housing information for an ethnically diverse and disadvantaged region of north-east London.

Specifically, we set out to answer the following research questions:

- What is the prevalence of household overcrowding using space per person and how does this compare to the bedroom standard?
- Are households with dependent children more likely to be overcrowded by the space standard and the bedroom standard than those without children?
- How does the overcrowding experienced by households with dependent children vary by household ethnicity and housing tenure?

We hypothesised that households with children would be more overcrowded, and this would be more likely in those living in the private and social rented sector and in households headed by a person from a non-white ethnic group.

Methods

Data sources

Pseudonymised primary care EHR were obtained from the Discovery Data Service (DDS) which receives primary care EHR data from all general practitioners (GPs) in north-east London (NEL), covering eight local authorities (Barking and Dagenham, City of London, Hackney, Havering, Newham, Redbridge, Tower Hamlets, and Waltham Forest). These practices publish primary care EHR data daily, directly from their electronic patient record enterprise systems, into the NEL DDS [26]. This data is pseudonymised and provided as a deidentified dataset for research. We downloaded deidentified DDS data on 7$^{\text{th}}$ June 2023.

Energy Performance Certificate (EPC) records [27] are publicly available datasets which provide information about the energy rating of properties. EPC records also include the postal address, number of rooms, and total dwelling area of each property. They have been required by law for every property that has been newly built, sold, or let since 2008 in England and Wales. Issued certificates are valid for 10 years. We downloaded EPC data for NEL in January 2023 [28].

Data linkage

Every addressable location in Great Britain is assigned a Unique Property Reference Number (UPRN) [29]. UPRNs identify a place of residence at a granular level, identifying individual properties, including houses or flats within a block [30]. Within DDS, UPRNs are allocated to GP-recorded addresses in real-time using the validated Address Matching

to Unique Property Reference Numbers (ASSIGN) algorithm [31]. They are pseudonymised into Residential Anonymised Linkage Fields (RALFs) using a study-specific SALT or encryption key [32]. The same study-specific encryption key was used to pseudonymise UPRNs in the EPC records to enable the linkage of the two data sources using RALFs.

Study design

Cross-sectional study.

Study population

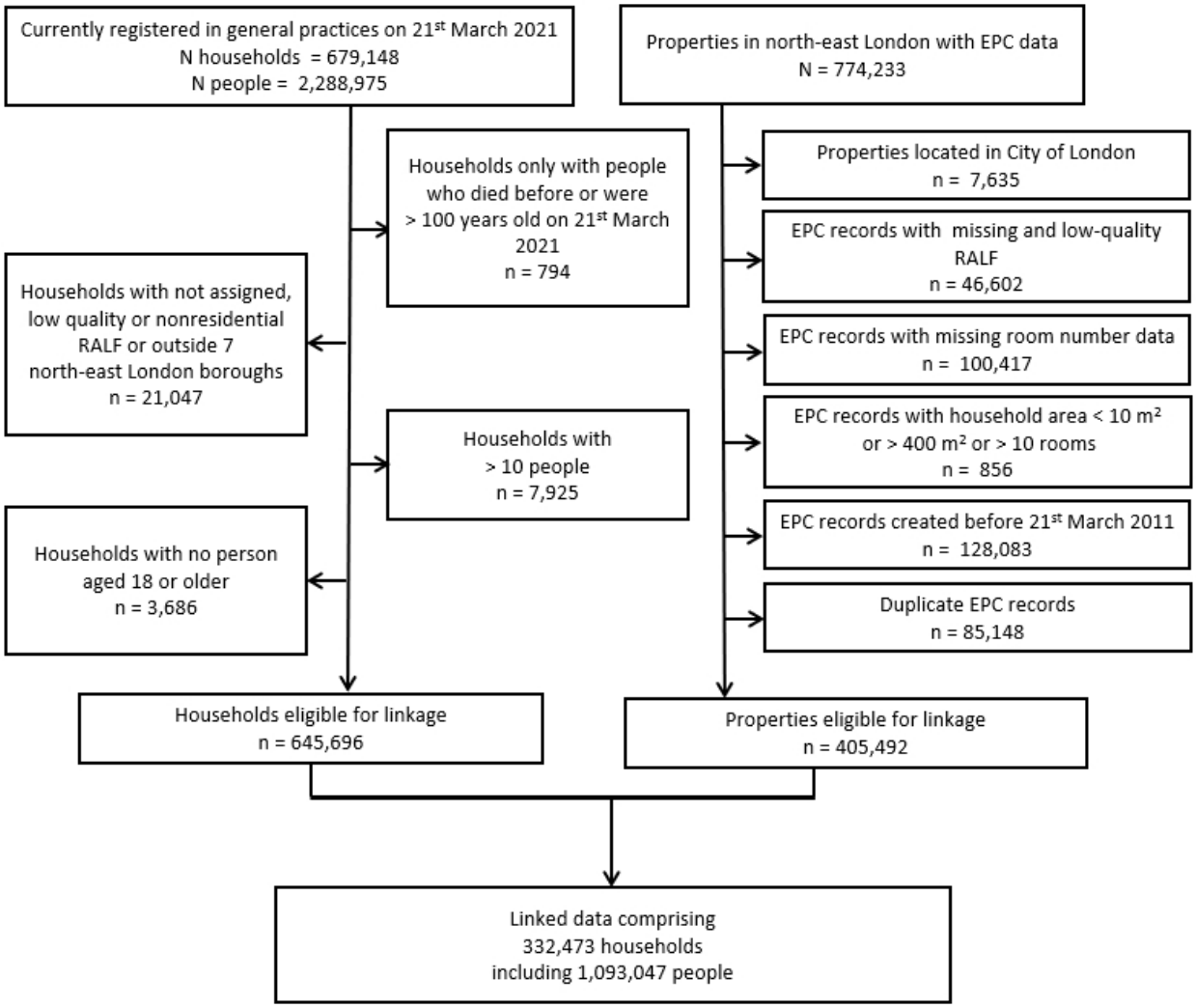
We identified 679,148 households containing 2,288,975 patients registered with a NEL GP, on 21st March 2021, the date of the 2021 Decennial Census. We assumed that people who shared a RALF on this date were members of the same household and that every household had a unique RALF. We excluded households with no RALF, with a low-quality RALF match, those located in the City of London - due to very small number of households - and those located outside the seven remaining NEL boroughs. We also excluded households with: only people aged >100 years, people who had died before 21st March 2021; >10 people; no adults aged ≥18 years. This left 645,696 households eligible for linkage to EPC records (Figure 1).

EPC data was downloaded for 774,233 properties located in NEL. We excluded records of properties with missing or low-quality RALF matches, those located in the City of London, those without data on number of rooms, with implausible areas (<10m^2 or >400m^2) or implausible number of rooms (>10). We excluded expired EPC certificates with dates earlier than 21st March 2011. We selected the most recent certificate for properties with more than one certificate within 10 years of 21st March 2021. This left 405,492 properties eligible for linkage. We linked eligible EPC to eligible EHR records using RALFs. The final study population comprised 332,473 households including 1,093,047 people (Figure 1).

Primary outcome

The primary outcome of interest was household overcrowding. We created two household overcrowding measures: a bedroom standard [33] and a space standard. Although both estimate

Figure 1: Flow diagram illustrating derivation of linked sample of households in primary care and EPC records



EPC – Energy Performance Certificates. LSOA – Lower layer Super Output Area. RALF – Residential Anonymous Linking Field.

Box 1: Description of selected overcrowding measures

Measure	Description
Bedroom standard	Based on the number of bedrooms in a dwelling, the number of people that share them, and their sex, age, and relationships. A household is overcrowded if the number of bedrooms is considered insufficient to provide separate bedrooms for every: • Married/cohabiting couple, • Adult aged $\geq$21 years, • Two adolescents of the same sex aged 10–20 years, • Two children aged $<$10 years regardless of sex [33].
Space standard	A continuous measure based on the total floor area of a dwelling and the number of people that share space in a dwelling: total floor area in m^2/ total number of people registered as residents in the dwelling.

overcrowding, their calculations and implications differ (Box 1).

We used modified ONS criteria to create a bedroom standard measure of overcrowding [17] based on the sex and age of household members and relationships between them. This enabled us to estimate the number of bedrooms required compared to the number available. The ONS bedroom standard rules require information on biological or marital/partner relationships between household members. As these are not available in EHRs, we used the ages of household members to infer these relationships, using as guidance the ONS Household Composition measure derived using administrative data [34]. See Supplementary Appendix 1 for further details.

We derived a space standard from the space per person, calculated by dividing the total floor area of the dwelling by the number of household members regardless of their age or relationships (Supplementary Appendix 1). As there is no agreed threshold defining sufficient or minimum space per person in a household, we estimated quartiles of the distribution of continuous space per person and defined households in the lowest quartile as overcrowded. We used this definition to estimate the prevalence of overcrowding using the space standard and to describe characteristics of households overcrowded by this measure.

Explanatory variables

We created a binary variable to identify households with one or more children under eighteen years of age on 21st March 2021.

We derived a household ethnicity measure by modifying the Office for National Statistics method which uses the ethnic group of the address reference person[35]. We defined household ethnicity based on the GP-recorded self-reported ethnicity of the oldest household member grouped using NHS categories [36]. We reported household ethnicity in five mutually exclusive groups: White, Black (any self-reported Black ethnicity excluding any Mixed Black), South Asian (any self-reported South Asian ethnicity excluding any Mixed Asian), Mixed and Other (Mixed: Mixed Black-White or South Asian-White ethnicity, Other: any self-reported category not included White, Black, South Asian or Mixed), Not known (Not Stated or Missing).

We identified housing tenure from the EPC record (owner occupied, private rented, social rented or unknown).

We calculated the average age of household members by dividing the sum of ages of all household members, including children, by the number of household members.

We assigned an area-level measure of relative deprivation – the 2019 Index of Multiple Deprivation (IMD) decile [37] – to each household based on the Lower layer Super Output Area (LSOA) of the property and concatenated this into five quintiles ranging from most to least deprived.

We assigned households to one of seven local authorities, an administrative unit for local government, with the following mutually exclusive categories: Barking and Dagenham, Hackney, Havering, Newham, Redbridge, Tower Hamlets, and Waltham Forest.

Statistical analyses

Descriptive analyses

We assessed potential biases in linkage between primary care and EPC records by comparing differences between linked and unlinked households in relation to presence of children, household ethnicity, average age of household members, local authority, and IMD quintile. We explored bivariate associations between household ethnicity, household presence of children and housing tenure. We compared the prevalence of overcrowding estimated using the bedroom standard and the space standard by region and local authority.

We created inverse probability weights using the following covariates: household ethnicity, household presence of children, average age of household members, local authority and IMD quintile of household, and applied these to estimate overcrowding prevalence in a sensitivity analysis to adjust for linkage bias. We calculated the proportion and 95% Confidence Intervals (CI) of overcrowded and not overcrowded households by bedroom standard and space standards by household presence of children, household ethnicity, and housing tenure. We illustrated the percentage of overcrowded households by bedroom standard and space per person in choropleth maps at the Middle layer Super Output Area (MSOA) level.

Multivariable analyses

We fitted logistic regression models to estimate the mutually adjusted odds ratios (aOR), and 95% CI, of household overcrowding defined by the bedroom standard for households with, compared to those without, children. We fitted linear regression models to estimate the mutually adjusted effects, and 95% CI, of household overcrowding defined by space per person in households with, compared to those without, children. We reported the adjusted R-squared, which summarises the variance explained by the final models. Both models were adjusted stepwise for household ethnicity (reference category: White ethnicity), and then housing tenure (reference category: owner occupied).

We derived predicted space per person from the linear model for households with selected characteristics, to compare differences in space per person by household presence of children, household ethnicity, and housing tenure.

All analyses were performed using R Studio [38].

Results

Descriptive statistics

The final study sample comprised 332,473 households, a linkage rate of 51.5%. The most common reason for no linkage was the lack of an EPC record (Figure 1). Households with children were overrepresented in the linked sample as were those of South Asian and Mixed, or Other household ethnicity, or with lower average age of the household members. They were more likely to be located in Barking and Dagenham, Newham, and Waltham Forest, and in areas included in the second most deprived IMD quintile (Table 1).

Overall, 42.8% of households in the linked sample included children, 54.5% were of White household ethnicity, and 35.2%

Table 1: Comparison of characteristics of 645,696 households according to whether their primary care electronic health record (EHR) was linked to an Energy Performance Certificate (EPC) record

	Linked	Not linked	All
N (%)	332,473 (51.5%)	313,223 (48.5%)	645,696
Children in household			
Yes	142,401 (42.8%)	93,785 (29.9%)	236,186 (36.6%)
No	190,072 (57.2%)	219,438 (70.1%)	409,510 (63.4%)
Household ethnicity			
White	181,096 (54.5%)	178,077 (56.9%)	359,173 (55.6%)
South Asian	72,348 (21.8%)	59,075 (18.9%)	131,423 (20.4%)
Black	45,413 (13.7%)	46,443 (14.8%)	91,856 (14.2%)
Mixed, Other	27,295 (8.2%)	23,596 (7.5%)	50,891 (7.9%)
Not known	6,321 (1.9%)	6,032 (1.9%)	12,353 (1.9%)
Tenure*			
Owner occupied	134,002 (40.3%)	–	–
Private rented	117,005 (35.2%)	–	–
Social rented	77,133 (23.2%)	–	–
Unknown	4,333 (1.3%)	–	–
Average age of household members (years)			
Mean (SD)	36.63 (15.95)	45.19 (18.61)	40.78 (17.81)
Local authority			
Barking and Dagenham	36,653 (11.0%)	31,345 (10.0%)	67,998 (10.5%)
Hackney	46,559 (14.0%)	46,581 (14.9%)	93,140 (14.4%)
Havering	42,991 (12.9%)	51,213 (16.4%)	94,204 (14.6%)
Newham	53,828 (16.2%)	47,859 (15.3%)	101,687 (15.7%)
Redbridge	47,005 (14.1%)	44,042 (14.1%)	91,047 (14.1%)
Tower Hamlets	54,505 (16.4%)	52,007 (16.6%)	106,512 (16.5%)
Waltham Forest	50,932 (15.3%)	40,176 (12.8%)	91,108 (14.1%)
IMD quintile			
Most deprived 1	84,372 (25.4%)	79,643 (25.4%)	164,015 (25.4%)
2	140,001 (42.1%)	123,465 (39.4%)	263,466 (40.8%)
3	62,322 (18.7%)	60,476 (19.3%)	122,798 (19.0%)
4	30,196 (9.1%)	30,463 (9.7%)	60,659 (9.4%)
Least deprived 5	15,582 (4.7%)	19,176 (6.1%)	34,758 (5.4%)

*Information on tenure is obtained from the Energy Performance Certificates records and is therefore not available for the unlinked sample.

Figure 2a: Presence of children in household by household ethnicity

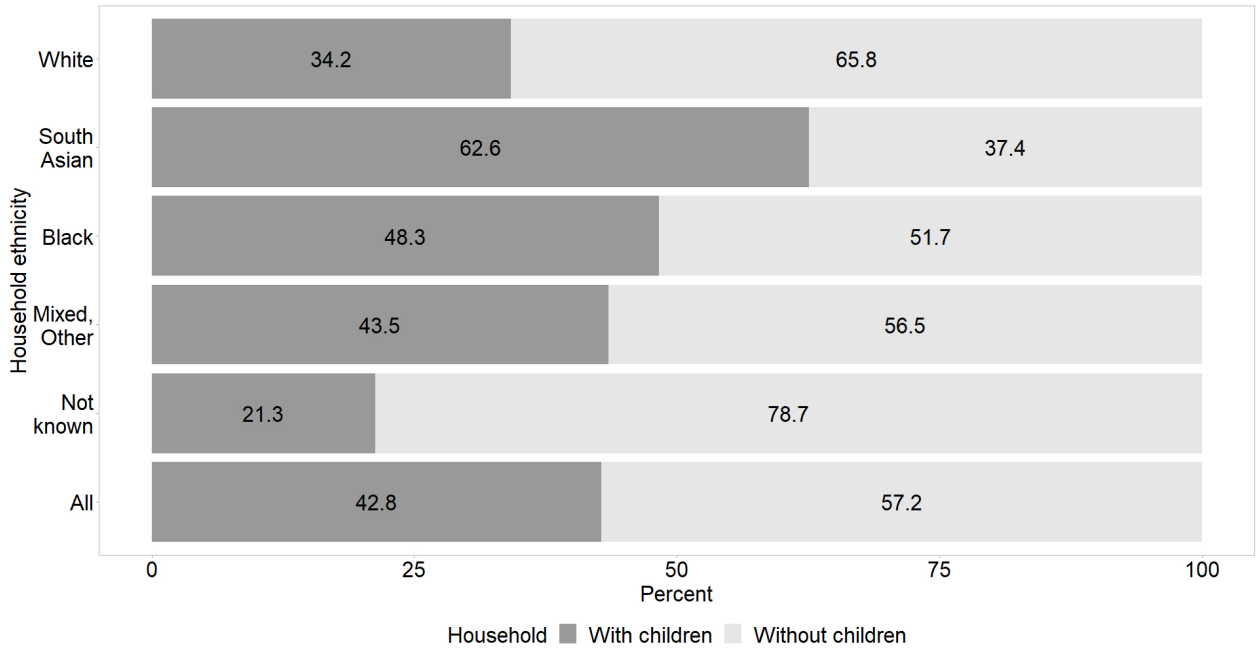
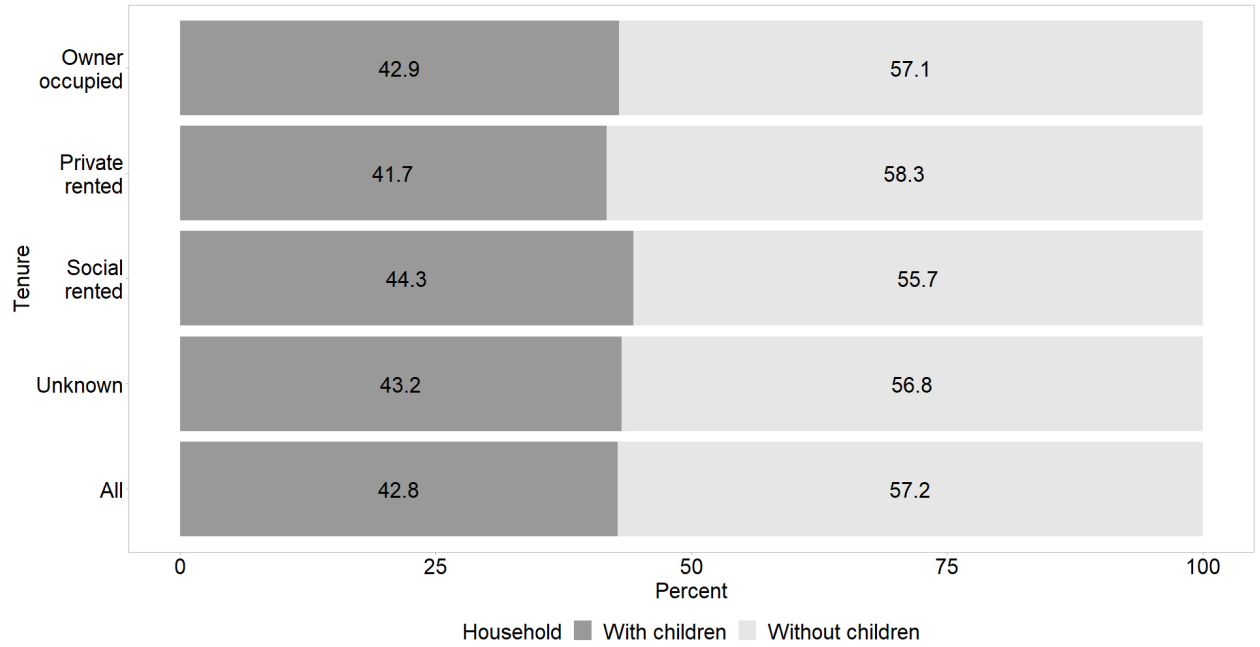


Figure 2b: Presence of children in household by housing tenure



and 23.2% were households in the private or social rented sectors respectively (Table 1). In the linked sample, there was a mean of 3.3 people per household (95% CI: 3.3, 3.3), and a mean property size of 77.3m^2 (77.2, 77.4). Households of South Asian ethnicity were overrepresented in households with children compared to those without (Figure 2a). Households in the social rented sector were more likely, and those in the private rented sector, least likely to include children, but these differences were very small (Figure 2b). Households of White and South Asian ethnicity were more likely to be in the owner occupied, while those of Black ethnicity were more likely to be in the social rented sector (Figure 2c).

Households with children, of South Asian ethnicity, or in private rented properties were more likely to be overcrowded by the bedroom standard (Supplementary Table 2a) as were households overcrowded by the space standard (Supplementary Table 2b).

Out of 332,473 households, 41,028 (12.3%; 95% CI: 12.2%,12.5%) were overcrowded by the bedroom standard. The minimum space per person was 5m^2 and maximum 396m^2. The lowest quartile of households had less than 16.6 m^2 per person (16.6, 16.7), and the second and third quartiles less than 25.3 m^2 (25.3, 25.5) and 43.0 m^2 (43.0, 43.5) per person, respectively.

Figure 2c: Housing tenure by household ethnicity

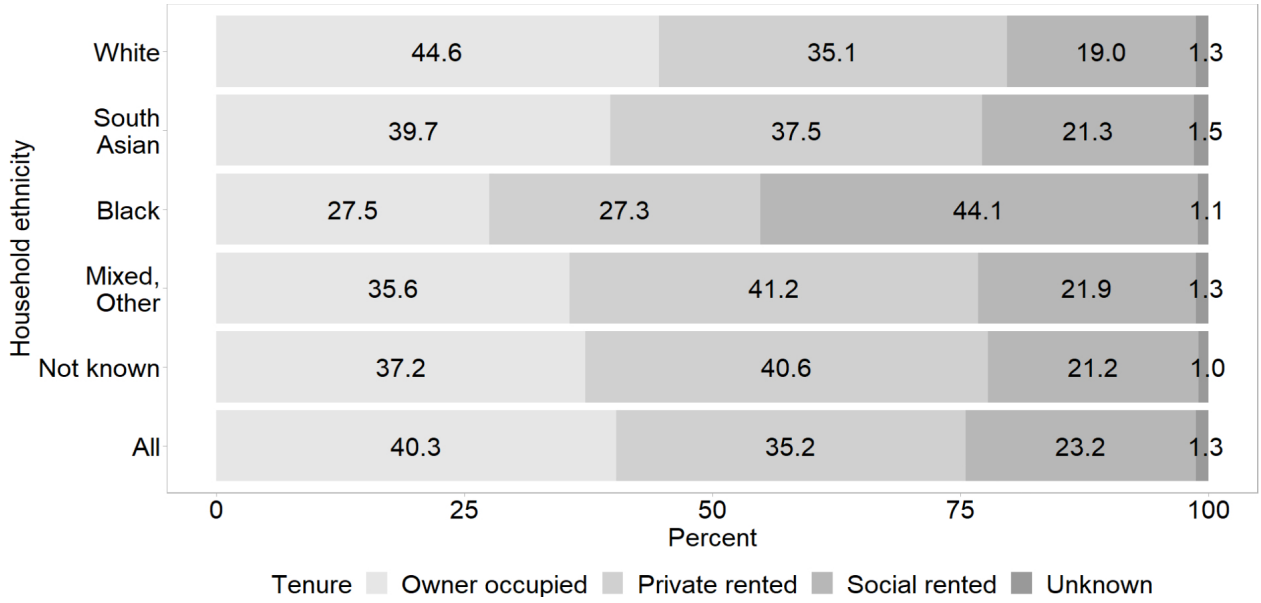


Table 2a: The number and percentage of overcrowded households in north-east London estimated using bedroom standard and space per person in overcrowded households by local authority

Local authority	Total number of households	Households overcrowded by bedroom standard* n; % (95% CI)		Space per person m^2 (95% CI) in in households overcrowded by space standard**
		n	% (95% CI)	m^2 (95% CI)
Newham	53,828	9,858	18.3 (18.0, 18.6)	14.4 (14.3, 14.5)
Barking andDagenham	36,653	5,200	14.2 (13.8, 14.5)	15.0 (15.0, 15.2)
Tower Hamlets	54,505	7,665	14.1 (13.8, 14.4)	16.3 (16.2, 16.5)
Redbridge	47,005	6,146	13.1 (12.8, 13.4)	16.4 (16.3, 16.6)
Waltham Forest	50,932	5,534	10.9 (10.6, 11.1)	17.2 (17.2, 17.4)
Hackney	46,559	4,732	10.2 (9.9, 10.4)	17.8 (17.7, 18.0)
Havering	42,991	1,893	4.4 (4.2, 4.6)	20.8 (20.8, 21.0)
Total	**332,473**	**41,028**	**12.3 (12.2, 12.5)**	**16.6 (16.6, 16.7)**

*The bedroom standard is calculated by determining the number of bedrooms required based on the age and sex and relationships between household members. A separate bedroom is allocated to each couple in a relationship, every person aged 21 or over, each pair of people aged 10-20 of the same sex, and each pair of children under 10. Any remaining person aged 10-20 and a child under 10 of the same sex are allocated the shared bedroom. The remaining people are given separate bedrooms. If the required number of bedrooms is less than the observed number of bedrooms, a household is considered to be overcrowded.
**Household overcrowded by space standard are those in the first quartile of space per person.

Table 2b: Comparison of the number (percentage) of overcrowded households, by the bedroom and space standards

		Overcrowded by space standard		Total
		Yes n (%)	No n (%)	
Overcrowded by bedroom standard	Yes n (%)	38,016 (11.4)	3,012 (0.9)	**41,028 (12.3)**
	No n (%)	45,601 (13.7)	245,844 (74.0)	**291,445 (87.7)**
	Total	**83,617 (25.1)**	**248,856 (74.09)**	**332,473 (100.0)**

The prevalence of overcrowding estimated by the bedroom standard was highest in Newham, and lowest in Havering, with one in five and one in twenty-three households overcrowded respectively. Findings were similar by the space standard with households in the lowest quartile having up to 14.4 m^2 (14.3, 14.5) and 20.8 m^2 (20.8, 21.0) per person in Newham and Havering respectively (Table 2a).

The majority (38,016/41,028; 92.7%) of households overcrowded by the bedroom standard were overcrowded by the space standard. However more than half (45,601/83,617;

Table 3a: Sensitivity analysis: The number and percentage of overcrowded households in north-east London estimated using bedroom standard and space per person in overcrowded households by local authority using the weighted sample

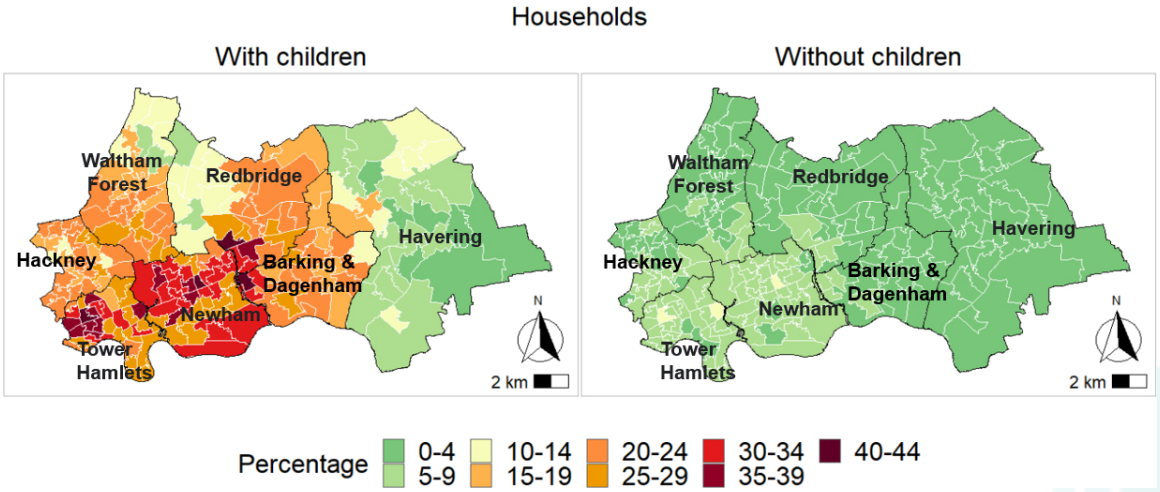
Local authority	Total number of households	Households overcrowded by bedroom standard* Weighted sample***		Households overcrowded by space standard** Weighted sample***
		N	% (95% CI)	m^2 (95% CI)
Newham	53,828	8,935	16.6 (16.3, 16.9)	15.0 (15.0, 15.2)
Tower Hamlets	54,505	7,086	13.0 (12.7, 13.2)	17.0 (17.0, 17.2)
Barking and Dagenham	36,653	4,582	12.5 (12.2, 12.8)	16.0 (16.0, 16.2)
Redbridge	47,005	5,312	11.3 (11, 11.5)	17.8 (17.6, 18.0)
Waltham Forest	50,932	4,940	9.7 (9.4, 9.9)	18.3 (18.2, 18.5)
Hackney	46,559	4,190	9.0 (8.8, 9.3)	19.0 (19.0, 19.3)
Havering	42,991	1,505	3.5 (3.4, 3.7)	23.0 (23.0, 23.3)
Total	**332,473**	**35,907**	**10.8 (10.7, 10.9)**	**17.8 (17.8, 17.9)**

*The bedroom standard is calculated by determining the number of bedrooms required based on the age and sex and relationships between household members. A separate bedroom is allocated to each couple in a relationship, every person aged 21 or over, each pair of people aged 10-20 of the same sex, and each pair of children under 10. Any remaining person aged 10-20 and a child under 10 of the same sex are allocated the shared bedroom. The remaining people are given separate bedrooms. If the required number of bedrooms is less than the observed number of bedrooms, a household is considered to be overcrowded.
**Household overcrowded by space standard are those in the lowest quartile of space per person.
***Weighted estimates were obtained using the sample of households whose Electronic Health Records (EHR) were linked to Energy Performance Certificates (EPC) adjusted with weights for each household, calculated as the inverse of the probability of having their EHR linked to EPC.

Table 3b: Comparison of the number (percentage) of overcrowded households, by the bedroom standard and space quartiles

		Overcrowded by space standard		Total
		Yes n (%)	No n (%)	
Overcrowded by bedroom standard	Yes n (%)	34,258 (10.3)	1,649 (0.5)	**35,907 (10.8)**
	No n (%)	49,407 (14.9)	247,159 (74.3)	**296,566 (89.2)**
	Total	**83,665 (25.2)**	**248,808 (74.8)**	**332,473 (100.0)**

Figure 3a: Percentage of overcrowded households with and without children by the bedroom standard

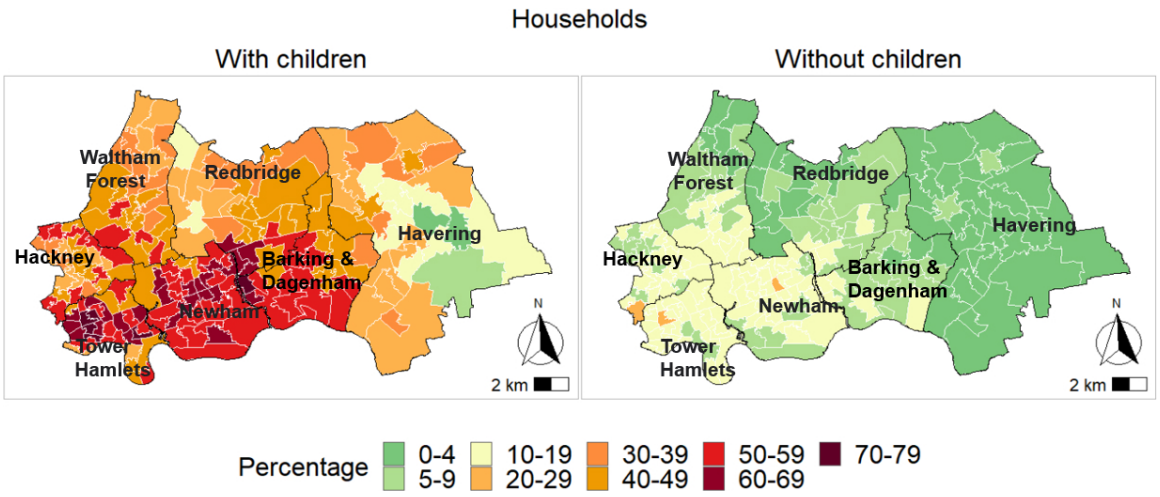


MSOA (Middle layer Super Output Area) boundaries are marked in white and local authority boundaries are marked in black. Households with children shown on the left and household without children on the right.

54.5%) of households considered overcrowded by the space standard were not overcrowded by the bedroom standard (Table 2b).

In the sensitivity analysis using weighted data, the prevalence of overcrowding by the bedroom standard fell from 12.3% to 10.8% (Table 3a). The maximum space per person in the lowest quartile increased from 16.6 m^2 to 17.8 m^2. The proportion of households overcrowded by the bedroom standard, which were also overcrowded by the space standard, was slightly higher in the weighted sample (Table 3b).

Figure 3b: Percentage of households with and without children overcrowded by the space standard i.e. <16.6m^2 space per person – by MSOA (Middle layer Super Output Area)



MSOA boundaries are marked in white and local authority boundaries are marked in black. Households with children shown on the left and household without children on the right.

Table 4: Odds of being overcrowded by the bedroom standard

	Unadjusted (OR 95% CI)	**Adjusted by household ethnicity (OR 95% CI)**	**Adjusted by) household ethnicity and tenure (OR 95% CI)**
Child in household			
No	**1**	**1**	**1**
Yes	**5.88 (5.74–6.03)**	**5.29 (5.15–5.42)**	**5.54 (5.40–5.68)**
Household ethnicity			
White	1	1	1
South Asian	2.46 (2.40–2.52)	1.66 (1.62–1.70)	1.60 (1.56–1.65)
Black	1.71 (1.66–1.76)	1.39 (1.35–1.43)	1.34 (1.30–1.39)
Mixed, Other	1.55 (1.49–1.61)	1.35 (1.30–1.41)	1.26 (1.21–1.32)
Not known	0.14 (0.11– 0.17)	0.17 (0.14–0.21)	0.16 (0.13–0.19)
Tenure			
Owner occupied	1		1
Private rented	2.48 (2.42–2.54)	–	2.68 (2.61–2.75)
Social rented	1.62 (1.57–1.67)	–	1.57 (1.52–1.62)
Unknown	1.54 (1.40–1.69)	–	1.54 (1.39–1.70)

The bedroom standard is calculated by determining the number of bedrooms required based on the age and sex composition of household members. A separate bedroom is allocated to each couple in a relationship, every person aged 21 or over, each pair of people aged 10-20 of the same sex, and each pair of children under 10. Any remaining person aged 10-20 and a child under 10 of the same sex are allocated the shared bedroom. The remaining people are given separate bedrooms. If the required number of bedrooms is less than the observed number of bedrooms, a household is considered to be overcrowded.
Odds ratios for households with children being overcrowded compared to the adults only ones were mutually adjusted for household ethnicity, and property tenure.

Households with children were more likely to be overcrowded by both standards: 22.5% and 45.9% of households with children were overcrowded by the bedroom and space standards respectively compared with 4.7% and 9.6% without children (Supplementary Table 2a, 2b). There are geographic inequalities in the prevalence of overcrowding for households with and without children by both standards (Figures 3a,3b). Almost all MSOAs have significantly higher levels of household overcrowding for households with children compared to those without children. The MSOAs with the highest prevalence of overcrowding by the bedroom standard also have the highest prevalence by the space standard.

Multivariable analyses

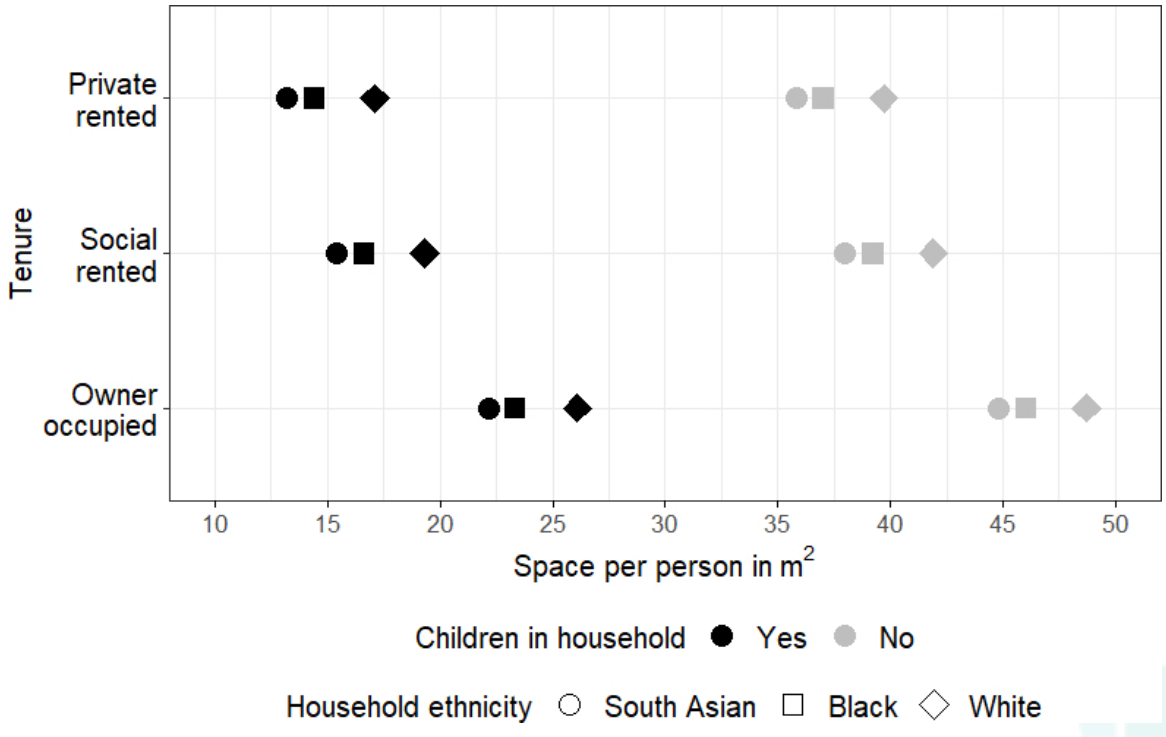
In the unadjusted logistic regression model, households with children were nearly six times more likely to be overcrowded by the bedroom standard than households without children (OR 5.88; 95%CI 5.74,6.03). These odds attenuated slightly

Table 5: Space per person in households with children before and after adjustment for household ethnicity and housing tenure

	Unadjusted space per person m^2 (95%CI)	**Adjusted by household ethnicity Space per person m^2 (95%CI)**	**Adjusted by household ethnicity and tenure Space per person m^2 (95%CI)**
Child in household			
No	**1**	**1**	**1**
Yes	**−23.57 (−23.70,−23.44)**	**−22.47 (−22.61,−22.32)**	**−22.61 (−22.75,−22.46)**
Household ethnicity			
White	1	1	1
South Asian	−10.70 (−10.89,−10.51)	−4.32 (−4.51,−4.14)	−3.90 (−4.08,−3.72)
Black	−6.91 (−7.13,−6.68)	−3.73 (−3.95,−3.52)	−2.7 (−2.93,−2.50)
Mixed, Other	−5.08 (−5.38,−4.78)	−2.98 (−3.25,−2.72)	−2.23 (−2.49,−1.97)
Not known	16.43 (15.69,−17.17)	13.52 (13.00, −14.05)	14.13 (13.62, −14.65)
Tenure			
Owner occupied	1	–	1
Private rented	−8.77 (−8.96,−8.58)	–	−8.94 (−9.10,−8.78)
Social rented	−7.50 (−7.70,−7.30)	–	−6.79 (−6.98,−6.61)
Unknown	−3.88 (−4.66,−3.10)	–	−3.62 (−4.23,−3.00)

The space standard is calculated by dividing a total floor area of a dwelling by the number of household members. Effects for change of space per person for households with children compared to the adults only ones were mutually adjusted for household ethnicity, and property tenure. CI: Confidence Interval.

Figure 4: Predicted average space per person for households with and without children by household ethnicity and tenure



on adjustment for household ethnicity (aOR 5.29; 5.15,5.42) and housing tenure (aOR 5.54; 5.40,5.68) (Table 4).

In the unadjusted linear regression model, households with children had less space per person (−23.57m^2; 95%CI −23.70,−23.44) than households without children. This attenuated slightly after adjustment for household ethnicity (−22.47 m^2; −22.61,−22.32) and housing tenure (−22.61 m^2; 95% CI −22.75,−22.46) (Table 5). The R-squared value for the fully adjusted model was 0.28.

Based on the predictions from the linear model households of White ethnicity, without children in owner occupied properties had the most, and those of South Asian ethnicity, with children and in private rentals, the least space per person (Figure 4, Supplementary Table 3).

Discussion

Key findings

This is to our knowledge the first study to estimate household overcrowding in households with children using linked primary care and housing records. Overall, in NEL the prevalence of household overcrowding based on the space standard (25.1%) was double that based on the bedroom standard (12.3%). More than 90% of households overcrowded by the bedroom standard were also overcrowded by the space standard. Households with children were almost six times more likely to be overcrowded by the bedroom standard and had on average 23 m^2 less space per person than households without children after adjusting for household ethnicity and housing tenure. We have demonstrated marked geographic inequalities in the distribution of household overcrowding in households with and without children by both measures. We have shown that within each ethnic and tenure category, households with children have one third to one half of the space per person less than the equivalent households without children.

Strengths and limitations

We applied a novel method for estimating up-to-date overcrowding estimates using linked health and housing data, which can be replicated in other datasets. The study is based on timely real-world data for over 300,000 households, which provides statistical power to examine ethnic and geographic inequalities in household overcrowding for households with children. This was enabled by access to high quality information on residents obtained using primary care records from NEL – an ethnically diverse, urban population in a large region of England. This is important as evidence to date shows that ethnic minority households in urban environments are more likely to be overcrowded [21, 39].

We developed a novel method of assigning individual people in households using RALFs. RALFs were also utilised to link EHRs and publicly available EPC data. This facilitated the derivation of overcrowding indicators at a household level. This is an important advantage of the method, as it has been shown that using area-level estimates to make inferences about individuals can result in ecological fallacies, with misclassification of individuals based on the demographic or geographic characteristics of the area rather than its residents [40, 41].

We assumed that EHR-recorded patient addresses are up-to-date and accurate, however there is a possibility of incorrect assignment of people in some households. We used a conservative approach to identify household members, but it is possible we included people who no longer live at their GP-registered address. This can happen if people do not deregister when leaving the area or are registered with a GP without being permanent residents in the area. We excluded temporary residents, as well as residents who were registered with practices but lived outside of the region and included only currently registered patients.

EPC data were not available for nearly half of the households in this region. The estimates of overcrowding prevalence and space per person attenuated slightly after weighting for the characteristics of households over-represented in the linked dataset. Households with children were on average larger and more likely to be included in the linked sample. This may explain our finding of a slightly higher average household size than reported from the 2021 Census for this region (3.3 vs 2.7 people per household respectively) [42].

We derived a measure of overcrowding using the bottom quartile of space per person as agreed standards on space per person are lacking. In our sample 25% of households had less than 16.6 m^2 per person, around half of the 32.5 m^2, the estimated average space per person for the whole London area [23]. We made assumptions about relationships based on age differences to create the bedroom standard and restricted the number of couples to one per generation per household, which may not always be the case. We defined household ethnicity by the ethnicity of the oldest resident, which does not reflect the ethnic diversity within households.

While including variables unmeasured for this study, including household income, which is known to be an important factor in household overcrowding, could explain at least partially the remaining variance of outcomes in our models, we found large, significant relative effects between households with and without children and the likelihood of household overcrowding in the fully adjusted models.

Comparison with existing evidence

Our finding that households with children are more likely to be overcrowded than those without children is consistent with findings from the Census 2021 and the English Housing Survey, both of which use the bedroom standard to estimate overcrowding prevalence [21]. The overcrowding estimates using the bedroom standard in this study were slightly lower than those estimated using the bedroom standard from the 2021 Census [17] at the local authority level. This may reflect differences in methods used to define households and count rooms as well as some differences in populations included. We plan to explore these differences in more detail in future work.

Ethnic and geographic inequalities in household overcrowding using the bedroom standard for households with and without children have been reported from the Census 2021 and are consistent with our findings. Census estimates based on space per person are not available for comparison. The differences between overcrowding prevalence estimates based on bedrooms and those based on space reflect different assumptions in their calculation. The former depends on the age, sex, and relationship of household members, not all of whom qualify for the same bedroom entitlement. In contrast, a measure based on space per person allows all members to have an equal entitlement to space irrespective of age or sex and takes account of dwelling size as well as household size and all the space available to a household. Hence, space per person provides a more direct and sensitive estimate of overcrowding [43] in terms of the living, working, and sleeping space household members share. This may be more relevant for health including, for example, in transmission of infections and effects of passive smoking, and may impact on physical and mental health.

Implications for research, policy and practice

A household should provide safe shelter, privacy, a place to sleep, eat, study, interact with other household members, and to enjoy privacy and personal space. Household overcrowding has been linked to mental health and sleep problems [44–46] as well as to increased risks of transmitting infectious diseases between household members due to shared workspaces, spending time and sleeping in very close proximity [1, 47]. The recent COVID-19 pandemic highlighted the impact of poor housing on people's wellbeing with 15.9 million (31%) adults in the UK suffering from mental or physical health problems because of housing conditions or insufficient space during periods of lockdown [48].

It is important to consider household overcrowding in urban populations in the context of the wider environment and its implications for national and local policies. NEL includes multiple MSOAs with a population density greater than the whole London area average of 5,600 people per km^2 [49] with some MSOAs in Tower Hamlets, Newham, and Hackney exceeding 20,000 people per km^2 (Supplementary Figure 1). Between the 2011 and 2021 Censuses, the populations of Tower Hamlets, Barking and Dagenham, and Newham, local authorities with the highest levels of household overcrowding and the smallest space per person [23], increased by 22%, 18% and 14% respectively [50]. Demographic change in these overpopulated areas is likely to create additional pressures on the well-being of children living in overcrowded households. It is important that local authorities have access to timely estimates of overcrowding in their areas to support planning and policies. Neither the decennial Census nor the English Housing Survey provide sufficiently large, detailed or timely estimates of overcrowding to enable that. Our method of calculating household overcrowding enables its estimation in local populations and provides information about the characteristics of households at most risk of being overcrowded. Using two indicators provides a comprehensive view of overcrowding in terms of space and bedrooms available for household members. This method can be replicated in other areas using the same information to investigate overcrowding in local populations.

Increased housing provision for households with children is a strategic goal for the current UK government but requires policies and investment in building new housing with sufficient space and access to publicly accessible infrastructure to prevent further overcrowding and to protect the mental and physical health of those families who are overcrowded. Assessments of need based on the bedroom standard are being challenged as no longer in line with public expectations with qualitative research suggesting that space per person for children as well as for adults is a more relevant measure [51].

Conclusions and future work

Housing and household overcrowding are important determinants of childhood health with impacts experienced across the life course for those affected. Methods to estimate overcrowding for local populations in real time and from all available data is essential for policy and research. We have demonstrated a novel and replicable method of deriving overcrowding estimates by linking health and property records and shown that it is feasible to create overcrowding measures based on space per person. Our study has added to the evidence on inequalities in overcrowding experienced by households with children and the intersection with household ethnicity and housing tenure. Further work is underway as part of the ADR UK Healthy Households project [52] to compare household estimates with Census data at the person level, to access Valuation Office Agency information obtained for Council tax purposes [53] to estimate overcrowding for properties without an EPC record, and to understand the association of overcrowding with poor health outcomes.

Authors contribution

Carol Dezateux conceptualised and obtained funding for the study.

Carol Dezateux, Marta Wilk, and Paul Simon contributed to data specification.

Paul Simon implemented the ASSIGN algorithm, and performed the data extraction from the Discovery Data Service.

Carol Dezateux, Gill Harper, Silvia Liverani, Nicola Firman and Marta Wilk conceptualised and designed the analyses.

Marta Wilk carried out the literature search, conducted the analyses, generated tables and figures.

Marta Wilk and Carol Dezateux wrote and revised the initial manuscript.

Carol Dezateux, Gill Harper, Silvia Liverani, Nicola Firman and Marta Wilk contributed to the development of the methodology, and interpretation of analyses and reviewed and revised the final manuscript.

All authors were involved in writing the paper and had final approval of the submitted and published manuscript.

The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Acknowledgments

This work was funded by Barts Charity ref: MGU0419 and MGU0504. It was also supported by ADR UK (Administrative Data Research UK), an Economic and Social Research Council investment (part of UK Research and Innovation) [Grant number: ES/X00046X/1]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Statement of conflict of interests

None declared.

Ethics statement

Data access was approved by the NEL Discovery Programme Board on behalf of the data controllers of primary care EHRs. As only routinely acquired de-identified data were analysed no research ethics committee approval was required by the Health

Research Authority. Access to general practice data is enabled by data sharing agreements between the DDS and GP data controllers. EPC data are publicly available.

Only aggregated patient data are reported in this study.

The manuscript was prepared according to the Equator Record standard (Benchimol et al., 2015). See Supplementary Appendix.

Data availability statement

The data controllers have not permitted onward sharing by the study authors of data used for this research.

References

1. World Health Organisation, 2018. WHO Housing and Health Guidelines. https://www.ncbi.nlm.nih.gov/books/NBK535293/

2. Baker MG, McDonald A, Zhang J, Howden-Chapman P. Infectious Diseases Attributable to Household Crowding in New Zealand: A Systematic Review and Burden of Disease Estimate. 2013. https://www.health.govt.nz/publication/infectious-diseases-attributable-household-crowding-new-zealand-systematic-review-and-burden-disease

3. Reynolds L. Crowded House: Cramped Living in England's Housing. Shelter England. 2004. https://england.shelter.org.uk/professional_resources/policy_and_research/policy_library/crowded_house_cramped_living_in_englands_housing

4. Nguipdop-Djomo P, Rodrigues LC, Smith PG, Abubakar I, Mangtani P, 2020. Drug misuse, tobacco smoking, alcohol and other social determinants of tuberculosis in UK-born adults in England: a community-based case-control study. Scientific Reports.https://doi.org/10.1038/s41598-020-62667-8

5. Office for National Statistics. General Health in Overcrowded and Under-occupied Households in England and Wales. Office for National Statistics; 2011 [August 2021]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/articles/generalhealthinovercrowdedandunderoccupiedhouseholdsinenglandandwales/2015-02-19.

6. London Assembly Environment Committee. Keeping out the chill: fixing London's cold, damp and mouldy homes. 2019. https://www.london.gov.uk/sites/default/files/keeping_out_the_chill_-_final.pdf

7. Aldridge R, Pineo H, Fragaszy E, Eyre M, Kovar J, Nguyen V, et al., 2021. Household overcrowding and risk of SARS-CoV-2: analysis of the Virus Watch prospective community cohort study in England and Wales. Wellcome Open Research. https://doi.org/10.12688/wellcomeopenres.17308.1

8. Boukari Y, Beale S, Nguyen VG, Fong WLE, Burns R, Yavlinsky A, et al., 2022. SARS-CoV-2 infections in migrants and the role of household overcrowding: A causal mediation analysis of Virus Watch data. https://doi.org/10.1136/jech-2022-220251

9. Dioun S, Chen L, Hillyer G, Tatonetti NP, May BL, Melamed A, et al., 2023. Association between neighborhood socioeconomic status, built environment and SARS-CoV-2 infection among cancer patients treated at a Tertiary Cancer Center in New York City. Cancer Reports. https://doi.org/10.1002/cnr2.1714

10. Leite A, Leão T, Soares P, Severo M, Moniz M, Lucas R, et al., 2021. A Case-Control Study of Contextual Factors for SARS-CoV-2 Transmission. Frontiers in Public Health. https://doi.org/10.3389/fpubh.2021.772782

11. Hansen CB, Larsen CVL, Bjerregaard P, Riva M, 2021. The effect of household crowding and composition on health in an Inuit cohort in Greenland. Scandinavian Journal of Public Health, 921-30. https://doi.org/10.1177/1403494820929496

12. Marsh R, Salika T, Crozier S, Robinson S, Cooper C, Godfrey K, et al., 2019. The association between crowding within households and behavioural problems in children: Longitudinal data from the Southampton Women's Survey. Paediatric and Perinatal Epidemiology, 195-203. https://doi.org/10.1111/ppe.12550

13. Solari CD, Mare RD, 2012. Housing crowding effects on children's wellbeing. Social Science Research, 464-76. https://doi.org/10.1016/j.ssresearch.2011.09.012

14. Delgado J, Ramirez-Cardich ME, Gilman RH, Lavarello R, Dahodwala N, Bazán A, et al., 2002. Risk factors for burns in children: crowding, poverty, and poor maternal education. Injury Prevention, 38-41. https://doi.org/10.1136/ip.8.1.38

15. Martin T, Nanjebe D, Atwine D, 2023. Assessment of the proportion of households with burn victims, associated risk factors and knowledge of burn injury prevention strategies in South Western Uganda. A population based cross sectional survey. Burns.https://doi.org/10.1016/j.burns.2023.03.016

16. Korhonen K, Leinonen T, Tarkiainen L, Einiö E, Martikainen P, 2023. Childhood socio-economic circumstances and dementia: prospective register-based cohort study of adulthood socio-economic and cardiovascular health mediators. International Journal of Epidemiology, 523-35. https://doi.org/10.1093/ije/dyac205

17. Office for National Statistics. Occupancy rating for bedrooms variable: Census 2021. 2023 [Available from: https://www.ons.gov.uk/census/census2021dictionary/variablesbytopic/housingvariablescensus2021/occupancyratingforbedrooms].

18. Office for National Statistics. Census 2021 geographies. 2022 [Available from: https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies].

19. Department for Levelling Up Housing and Communities. English Housing Survey. 2023 [Available from: https://www.gov.uk/government/collections/english-housing-survey].

20. Ministry of Housing CLG. English Housing Survey 2022 to 2023: rented sectors. https://www.gov.uk/government/statistics/english-housing-survey-2022-to-2023-rented-sectors/english-housing-survey-2022-to-2023-rented-sectors#introduction-and-main-findings:gov.uk; 2024 Available from: https://www.gov.uk/government/statistics/english-housing-survey-2022-to-2023-rented-sectors/english-housing-survey-2022-to-2023-rented-sectors#introduction-and-main-findings.

21. Office for National Statistics. Overcrowding and under-occupancy by household characteristics, England and Wales: Census 2021. 2023 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/overcrowdingandunderoccupancybyhouseholdcharacteristicsenglandandwales/census2021].

22. National Housing Federation. Overcrowding in England. 2023. https://www.housing.org.uk/resources/overcrowding-in-england-2023/

23. Gleeson J. An analysis of housing floorspace per person, Housing Research Note 6. https://data.london.gov.uk/dataset/housing-research-notes: Greater London Authority; 2021 February 2021.

24. The Resolution Foundation. Home truths. Putting housing policy in context for the 2024 general election. https://www.resolutionfoundation.org/publications/home-truths/; 2024. https://www.resolutionfoundation.org/publications/home-truths/

25. Willson W, Barton C, Carthew H. Overcrowded housing (England). House of Commons Library. commonslibrary.parliament.uk; 2023. https://commonslibrary.parliament.uk/research-briefings/sn01013/

26. Discovery Health and Care Data Service. Discovery East London. 2022 [cited 2022. Available from: https://www.discoverydataservice.org/Content/Regional_programmes/Discovery_East_London/Discovery_East_London.htm].

27. Department for Levelling Up Housing & Communities. A guide to energy performance certificates for the marketing, sale and let of dwellings. Department for Communities and Local Government; 2017 [March 2022]. Available from: https://www.gov.uk/government/publications/energy-performance-certificates-for-the-construction-sale-and-let-of-dwellings].

28. Department for Levelling Up Housing & Communities. Access to Energy Performance Certificates and Display Energy Certificates data for buildings in England and Wales. 2023 [Available from: https://epc.opendatacommunities.org/].

29. Ordnance Survey. AddressBase. Postal Addresses Matched to Unique Property Reference Numbers. Vector Map Data. 2021 [Available from: https://www.ordnancesurvey.co.uk/business-government/products/addressbase].

30. Ideal Postcodes. UPRN: Unique Property Reference Number. https://ideal-postcodes.co.uk/guides/uprn; 2021 [Available from: https://ideal-postcodes.co.uk/guides/uprn].

31. Harper G, Stables D, Simon P, Ahmed Z, Smith K, Robson J, et al., 2021. Evaluation of the ASSIGN open-source deterministic address-matching algorithm for allocating Unique Property Reference Numbers to general practitioner-recorded patient addresses. International Journal of Population Data Science. https://doi.org/10.23889/ijpds.v6i1.1674

32. OpenPseudonymiser. OpenPseudonymiser, About. Julia Hippisley-Cox, University of Nottingham 2011; 2011 [Available from: https://www.openpseudonymiser.org/About.aspx].

33. Department for Levelling Up Housing & Communities. Allocation of accommodation: guidance for local housing authorities in England. 2023 October 2023. https://www.gov.uk/guidance/allocation-of-accommodation-guidance-for-local-authorities/download-this-guidance

34. Office for National Statistics. Household Composition using administrative data: initial research. ons.gov.uk; 2018 [updated 2018. Available from: https://www.slideshare.net/secret/iK9pjZsOpHxbWx].

35. Statistics OfN, 2023. Developing subnational multivariate housing by ethnicity statistics from administrative data, England and Wales: 2020. https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/developingsubnationalmultivariatehousingbyethnicitystatisticsfromadministrativedataenglandandwales/2020

36. National Health Service England. Ethnic category. 2023 [Available from: https://www.datadictionary.nhs.uk/data_elements/ethnic_category.html#:~:text=CATEGORY%20CODE%202001.-,The%2016%2B1%20ethnic%20data%20categories%20defined%20in%20the%202001,is%20optional%20for%20use%20locally.].

37. Ministry of Housing Communities & Local Government. The English Indices of Deprivation 2019, Frequently Asked Questions. 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853811/IoD2019_FAQ_v4.pdf

38. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. In: Foundation TR, editor. R version 4.2.1 ed: The R Foundation; 2020 Available from: https://www.R-project.org/.

39. GOV.UK English Housing Survey. Overcrowded households , The English Housing Survey. Ministry of Housing, Communities and Local Government; 2020 [Available from: https://www.ethnicity-facts-figures.service.gov.uk/housing/housing-conditions/overcrowded-households/latest].

40. McCartney G, Hoggett R, 2023. How well does the Scottish Index of Multiple Deprivation identify income and employment deprived individuals across the urban-rural spectrum and between local authorities? Public Health, 26-32.https://doi.org/10.1016/j.puhe.2023.01.009

41. McCartney G, Hoggett R, Walsh D, Lee D, 2023. How well do area-based deprivation indices identify income- and employment-deprived individuals across Great Britain today? Public Health, 22-5. https://doi.org/10.1016/j.puhe.2023.01.020

42. Office for National Statistics. Household size, UK Statistics. Census 2021. Dataset ID: TS017. Census 2021; 2023 [Available from: https://www.ons.gov.uk/datasets/TS017/editions/2021/versions/3].

43. Cable N, Sacker A, 2019. Validating overcrowding measures using the UK Household Longitudinal Study. SSM - Population Health, 3. https://doi.org/10.1016/j.ssmph.2019.100439

44. National Housing Federation. Overcrowding in England. 2023 [Available from: https://www.housing.org.uk/resources/overcrowding-in-england-2023/].

45. Finch D, Farrington-Douglas J, Johnes C. Moving to healthy homes. 2023 [Available from: https://www.health.org.uk/publications/long-reads/moving-to-healthy-homes#: :text=A%20healthy%20home%20should%20be,would%20need%20to%20be%20ambitious.].

46. Chambers EC, Pichardo MS, Rosenbaum E, 2014. Sleep and the Housing and Neighborhood Environment of Urban Latino Adults Living in Low-Income Housing: The AHOME Study. Behavioral Sleep Medicine. https://doi.org/10.1080/15402002.2014.974180

47. Soltan M, Crowley L, Melville C, Varney J, Cassidy S, Mahida R, et al. To what extent are social determinants of health, including household overcrowding, air pollution and housing quality deprivation, modulators of presentation, ITU admission and outcomes among patients with SARS-COV-2 infection in an urban catchment area.: https://www.researchsquare.com/article/rs-35617/v1; 2020 2020-07-08. https://www.researchsquare.com/article/rs-35617/v1

48. National Housing Federation. Housing issues during lockdown: health, space and overcrowding. 2020 [Available from: https://www.housing.org.uk/news-and-blogs/news/poor-housing-causing-health-problems-for-nearly-a-third-of-brits-during-lockdown/].

49. Office for National Statistics. Population density. 2021 [Available from: https://www.nomisweb.co.uk/datasets/c2021ts006].

50. Office for National Statistics. How the population changed in Tower Hamlets: Census 2021. 2022 [Available from: https://www.ons.gov.uk/visualisations/censuspopulationchange/E09000030/].

51. The Resolution Foundation. Housing Outlook Q2 2024. https://www.resolutionfoundation.org/publications/housing-outlook-q2-2024/; 2024. https://www.resolutionfoundation.org/publications/housing-outlook-q2-2024/

52. Administrative Data Research UK. Data linkage funding will drive new insights into the health of households. 2023 [Available from: https://www.adruk.org/news-publications/news-blogs/data-linkage-funding-will-drive-new-insights-into-the-health-of-households-821/#: :text=ADR%20UK%20is%20funding%20a,for%20England%2C%20Scotland%20and%20Wales.].

53. Valuation Office Agency. Valuation Office Agency. 2023 [October 2023]. Available from: https://www.gov.uk/government/organisations/valuation-office-agency].

54. Office for National Statistics. Live Births in England and Wales by Characteristics of Mother. Office for National Statistics; 2011 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/livebirthsinenglandandwalesbycharacteristicsofmother1/2015-01-24].

55. Office for National Statistics. Birth characteristics in England and Wales: 2021. Office for National Statistics; 2023 [October 2023]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthcharacteristicsinenglandandwales/2021#: :text=In%202021%2C%20the%20standardised%20mean,33.7%20years%20(Figure%201).].

56. Office for National Statistics. Population estimates by marital status and living arrangements, England and Wales – Office for National Statistics. Office for National Statistics – 2011 Census and Census 2021; 2023 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/livingarrangementsofpeopleinenglandandwales/census2021#couples-age-difference].

57. Office for National Statistics. Household Composition using administrative data: initial research, slide

19. 2018 [Available from: https://www.slideshare.net/secret/iK9pjZsOpHxbWx].

Abbreviations

OR:	adjusted odds ratios
ASSIGN:	AddreSS matchInG to unique property reference Numbers
EHR:	electronic health records
EPC:	Energy Performance Certificates
GP:	general practice/practicioner
LSOA:	Lower Layer Super Output Area
MSOA:	Middle Layer Super Output Area
NEL:	north-east London
RALF:	Residential Anonymised Linkage Field
UPRN:	Unique Property Reference Number

Supplementary Appendix

Bedroom and space standard calculation methods

Bedroom standard calculation method

We estimated the bedroom standard using primary care EHRs using rules set out in the bedroom standard calculation [17, 33].

The bedroom standard states that a separate bedroom must be allocated to:

- an adult couple
- any remaining adult (aged 21 years or over)
- two adolescents (aged 10 to 20 years) of the same sex
- one adolescent (aged 10 to 20 years) and one child (aged 9 years or under) of the same sex
- two children (aged 9 years or under) regardless of sex
- any remaining child (aged 9 years or under)"

If the conditions above cannot be fulfilled with the number of bedrooms available in a dwelling, the dwelling is classified as overcrowded. We used these rules by rounding down ages to whole years.

As we did not have information regarding the relationships between people, we assigned people to first, second and third generations based on the following assumptions about the ages of and between household members.

1. **We assigned people aged $\leq$19 years to the child generation.**

   This was based on age at the end of the calendar year in which secondary school education is usually completed. As only 5% and 1.8% of mothers and fathers respectively are aged $<$20 years at the birth of their child [54], we assumed that people aged $\leq$19 years were unlikely to be parents (Supplementary Table 1). We based this on data from the 2011 Census, as data published from the 2021 Census has to date only reported the standardised mean age of mothers (30.9 years) and fathers (33.7 years) [55].

2. **We assigned people to a second generation**, if they were the youngest of all the non-child members and 20-50 years older than the oldest child. If no one was assigned to this category, we reduced this minimum age difference to 19 years.

3. **We assigned partners in the second generation** based on the distribution of age differences between cohabiting or married partners [56, 57] and assuming that the age of the partner was between the age of the next youngest and $\leq$12 years older than the partner, irrespective of sex.

4. **We assigned people to a third generation**, if they were $\geq$50 years than the oldest person in the first generation. This was based on the distribution in Table 1, which shows that a very small number of parents (mainly men) are more than 45 years older than their child at birth.

5. **We assigned partners in the third generation**, as in step three, by assuming that the age of the partner was between the age of the next youngest and $\leq$12 years older than the partner, irrespective of sex.

6. **All other household members were classified as other.**

In adult-only households, the first stage was omitted and stages 2 and 5 are calculated based on a dummy first generation variable, age 1 year. All other stages remained the same.

The set of rules presented above is relatively simple and achieves the aim of imposing plausible household relationships between members, which in consequence allowed us to estimate which household members should be allocated separate rooms/bedrooms and which members could share bedrooms with assigned partners. It is important to note that observational data does not provide enough information in houses with multiple adults to differentiate between the house shares of single households, adult children living with spouses and parents, and other family set-ups. Here, we assumed that there was only one married/cohabiting couple in household per second or third generation.

Space per person calculation method

We calculated space per person by dividing the total floor area of a dwelling in the EPC record by the number of persons living in that dwelling according to the primary care EHR. We divided the distribution of space per person into four quartiles for descriptive statistics and used the continuous space per person as an outcome in multivariate models.

Supplementary Table 1: Live births by age group of mother and father, 2011

Age (years)	Mothers	Fathers
Under 20	5%	2%
20-24	19%	11%
25-29	28%	22%
30-34	29%	29%
35-39	16%	21%
40-44	4%	10%
45 and over	0%	5%

Source [54].

Supplementary Table 2a: Characteristics of overcrowded and not overcrowded households by the bedroom standard

	Not overcrowded (n)	Not overcrowded (95% CI)	Overcrowded (n)	Overcrowded (95% CI)
Child in household				
No	181,119	95.3 (95.2, 95.4)	8,953	4.7 (4.6, 4.8)
Yes	110,326	77.5 (77.3, 77.7)	32,075	22.5 (22.3, 22.7)
Household ethnicity				
White	164,663	90.9 (90.8, 91.1)	16,433	9.1 (8.9, 9.2)
South Asian	58,103	80.3 (80.0, 80.6)	14,245	19.7 (19.4, 20.0)
Black	38,804	85.4 (85.1, 85.8)	6,609	14.6 (14.2, 14.9)
Mixed, Other	23,640	86.6 (86.2, 87.0)	3,655	13.4 (13.0, 13.8)
Not known	6,235	98.6 (98.3, 98.9)	86	1.4 (1.1, 1.7)
Tenure				
Owner occupied	123,411	92.1 (92.0, 92.2)	10,591	7.9 (7.8, 8.0)
Private rented	96,490	82.5 (82.2, 82.7)	20,515	17.5 (17.3, 17.8)
Social rented	67,716	87.8 (87.6, 88.0)	9,417	12.2 (12.0, 12.4)
Unknown	3,828	88.3 (87.3, 89.3)	505	11.7 (10.7, 12.7)

Footnote: CI: Confidence Intervals.

Supplementary Table 2b: Characteristics of overcrowded and not overcrowded households by the space standard

Child in household				
No	171,843	90.4 (90.3, 90.5)	18,229	9.6 (9.5, 9.7)
Yes	77,013	54.1 (53.8, 54.3)	65,388	45.9 (45.7, 46.2)
Household ethnicity				
White	147,792	81.6 (81.4, 81.8)	33,304	18.4 (18.2, 18.6)
South Asian	43,417	60.0 (59.7, 60.4)	28,931	40.0 (39.6, 40.3)
Black	31,773	70.0 (69.5, 70.4)	13,640	30.0 (29.6, 30.5)
Mixed, Other	19,912	73.0 (72.4, 73.5)	7,383	27.0 (26.5, 27.6)
Not known	5,962	94.3 (93.7, 94.9)	359	5.7 (5.1, 6.3)
Tenure				
Owner occupied	111,279	83.0 (82.8, 83.2)	22,723	17.0 (16.8, 17.2)
Private rented	77,903	66.6 (66.3, 66.9)	39,102	33.4 (33.1, 33.7)
Social rented	56,411	73.1 (72.8, 73.4)	20,722	26.9 (26.6, 27.2)
Unknown	3,263	75.3 (74.0, 76.6)	1,070	24.7 (23.4, 26.0)

Footnote: CI: Confidence Intervals.

Supplementary Table 3: Predicted average space per person for households with and without children by household ethnicity and tenure

Household ethnicity	Tenure	Space per person (m^2) Child in household	
		No	Yes
White	Owner occupied	48.7	26.1
	Social rented	41.9	19.3
	Private rented	39.7	17.1
South Asian	Owner occupied	44.8	22.2
	Social rented	38.0	15.4
	Private rented	35.8	13.2
Black	Owner occupied	46.0	23.3
	Social rented	39.2	16.6
	Private rented	37.0	14.4

Supplementary Figure 1: Population density [49]

*1000 people per square kilometre. MSOA (Middle Layer Super Output Area) boundaries are marked in white and local authority boundaries are marked in black.