

Estimating the scale of hospital admissions for people experiencing homelessness in England: a population-based multiple systems estimation study using national Hospital Episode Statistics

Serena April Luchenski ¹, Dankmar Böhning ², Robert Aldridge ^{3,4},
Fiona Stevenson ¹, Shema Tariq ⁵, Andrew C Hayward ^{1,6}

To cite: Luchenski SA, Böhning D, Aldridge R, *et al.* Estimating the scale of hospital admissions for people experiencing homelessness in England: a population-based multiple systems estimation study using national Hospital Episode Statistics. *BMJ Public Health* 2025;**3**:e002978. doi:10.1136/bmjph-2025-002978

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjph-2025-002978>).

Received 16 April 2025
Accepted 24 September 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to
Dr Serena April Luchenski;
s.luchenski@ucl.ac.uk

ABSTRACT

Background People experiencing homelessness have substantial health needs and poor access to primary healthcare, resulting in high rates of hospital care. Housing status is not routinely recorded in English electronic health records, undermining service planning. We developed methods to estimate the scale of hospital admissions for people experiencing homelessness in England.

Methods We analysed admissions for people experiencing homelessness using Hospital Episode Statistics for 2013/2014, 2015/2016 and 2017/2018. We applied multiple systems estimation Poisson regression methods to estimate total admissions and an inflation factor to correct for under-reporting. We calculated unadjusted admission rates per 1000 population per year and admission rate ratios compared with the housed population.

Results We observed 34 790 admissions in 2017/2018, with total homeless admissions estimated at 176 342 (95% CI 164 031 to 188 654) (inflation factor=5.07 (95% CI 4.71 to 5.42)). The unadjusted admission rate for the 2017/2018 homeless population was 879.0 admissions per 1000 population per year (95% CI 817.7 to 940.4), 2.5 (95% CI 2.3 to 2.7) times higher than the housed population. Restricted to rough sleepers and hostel residents, the unadjusted rate was 3516.7 per 1000 (95% CI 3271.2 to 3762.2), with a rate ratio of 10.0 (95% CI 9.3 to 10.7) compared with the housed population.

Conclusions We estimated five times as many hospital admissions for people experiencing homelessness than we observed directly. We advise caution when applying these inflation factors to other datasets because of methodological limitations in this study and sensitivities to local coding practices. In the absence of routine housing status recording, multiple systems estimation could facilitate improved service planning.

INTRODUCTION

Homelessness is not solely defined by rooflessness or literal homelessness, but also includes

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ People experiencing homelessness have some of the worst health outcomes in society, but the scale of healthcare needs is poorly understood because of the lack of routine recording of housing status in electronic health records.

WHAT THIS STUDY ADDS

⇒ We applied a new set of code lists to identify hospital admissions for people experiencing homelessness and used multiple systems estimation (also known as capture–recapture estimation) to quantify and adjust for underreporting within English hospital data. We showed that there were approximately five times more admissions for people experiencing homelessness than we could observe directly in national data, but there were important methodological limitations which necessitate a tempered interpretation of the findings.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study provides valuable insights into the potential scale of hospitalisation for people experiencing homelessness in England and a recommended set of methods for further investigating this problem. While the magnitude of underestimation should be interpreted cautiously, this study contributes important evidence for developing effective policies and services that can meet the needs of this underserved population.

individuals living in insecure, unstable or inadequate housing situations.¹ The transient and hidden nature of this population makes it notoriously difficult to measure.^{2–4} The Crisis Homelessness Monitor⁴ models annual estimates of ‘core homelessness’ in

England, which includes rough sleepers, people living in unconventional accommodation (eg, squatting), hostels, unsuitable temporary accommodation (eg, bed and breakfasts) and sofa surfing (eg, staying with non-family, on a short-term basis, in overcrowded conditions). There were an estimated 242 000 core homeless households in England in 2022, up from 221 000 in 2020, 224 000 in 2018 and 206 000 in 2012.⁴

People experiencing homelessness often have complex health needs, including intersecting physical illness, mental illness and alcohol and drug use disorders.⁵ It has been estimated that nearly a third of deaths among people experiencing homelessness in England are amenable to timely and effective healthcare, compared with about a quarter in the most deprived areas of the general population.⁶ The all-cause standardised mortality ratio has been estimated to be 7.88 (95% CI 7.03 to 8.74) higher than the general population for men and 11.86 (95% CI 10.42 to 13.30) higher for women.⁷

This population encounters multiple barriers to accessing primary care and preventative services in the community and often only use healthcare in a crisis.^{3 8–12} This results in higher levels of emergency department attendances, and emergency admissions and readmissions to hospital,^{3 13 14} than the general population. In turn, this contributes to poor health outcomes and high healthcare costs.^{13 15–17}

Estimating the magnitude of homelessness-related inpatient activity is fundamental to assessing the need for and impact of specialist and mainstream health and social care services for this underserved population. However, housing status is not systematically collected in English NHS health services.^{13 18} Moreover, despite there being an ICD-10 code for homelessness, coding of homelessness within routine hospital data is inconsistent.

Prior to this study, assessment of hospital care utilisation of people experiencing homelessness in England has often relied on the 'No Fixed Abode' (NFA) code^{13 17–19} as a proxy for single people sleeping rough or in a hostel. This approach has notable limitations, including underascertainment (some may give a temporary address), misclassification of people who are not experiencing homelessness but decline to give an address (eg, fearing disclosure of treatment such as termination of pregnancy), and misclassification due to poor data quality or errors. Other research has examined the hospital records of people identified through specialist homeless healthcare services.^{5 14 20–22} However, sampling through specialist services introduces selection biases in terms of who is referred or attends services. There are also surveys of hospital care utilisation among people experiencing homelessness.²³ Such surveys are also prone to selection biases in who responds or not as well as measurement bias because they rely on self-report of healthcare utilisation.

We aimed to develop a novel and robust methodology using multiple systems estimation (MSE) to overcome these limitations and estimate the scale of hospital

admissions among people experiencing homelessness in England. The specific objectives were:

1. Accurately estimate the number of hospital admissions for people experiencing homelessness in England, addressing underreporting in routine hospital data.
2. Quantify the healthcare utilisation of the homeless population by calculating the rate of hospital admissions per 1000 people and compare it to the housed population.

METHODS

Study design

This research employed a population-based, repeated cross-sectional study design, using anonymised hospital records from England's Hospital Episodes Statistics (HES) for the fiscal years 2013/2014, 2015/2016 and 2017/2018. The analysis focused on 'continuous inpatient spells', which represent complete hospital admissions, accounting for transfers between different hospital providers within a single admission. These spells were identified using unique patient identifiers (HESID).

Participants and setting

The study population encompassed all publicly funded National Health Service (NHS) inpatient hospitals in England, including acute mental health trusts. This broad inclusion ensured that the findings were representative of the national healthcare system. Individuals aged 18 years and older who were experiencing homelessness during their hospital admission constituted the target population.

Housing status is not routinely recorded within English hospital data, so we developed a homelessness phenotype to more accurately identify people experiencing homelessness within the HES data.²⁴ The phenotype is described in full in the Health Data Research UK Phenotype Library²⁴ and includes:

1. NFA: address recorded as NFA, with certain exclusions as per previous research¹³
2. HGP (homeless general practitioner): registered at a known HGP practice that exclusively serves those experiencing homelessness (as mapped in a previous study²⁵). We used this existing list from this study to produce a corresponding list of GP practice codes using the NHS Digital ODS Portal²⁶ to identify registered people within HES.
3. Z59.0: a diagnosis that includes the International Classification of Diseases version 10 (ICD-10) code for homelessness (Z59.0), usually as a secondary diagnostic code.

Our phenotype builds on previous studies which have relied on NFA alone^{13 27} to identify people experiencing homelessness. Although we have not done a formal validation study, it is highly likely, however, that the phenotype still underestimates homelessness because there are no national policies on homelessness coding and practices will vary from hospital to hospital. For example, a

service evaluation showed that 58%–75% of a known hospitalised homeless population were identifiable using a similar phenotype to ours.²⁸ It is this unknown level of underestimation at a national level which is the rationale for conducting this study using MSE.

Data access and ethical considerations

Data access was facilitated through the UCL Institute of Health Informatics' secure Data Safe Haven. This environment provided a secure and controlled setting for the analysis of sensitive patient data, ensuring compliance with data protection regulations. The UCL Institute of Health Informatics held a full copy of HES for the duration of this study. Data access was granted conditional on a data sharing agreement, consistent with prior agreements with NHS Digital. Ethical approval for this study was granted by the UCL Research Ethics Committee (Ref: 11607/001). The study used anonymised data, ensuring that individual patient identities were protected throughout the research process.

Outcomes

The primary outcome of this study was the observed and estimated frequencies of hospital admissions per year for people experiencing homelessness in England, addressing the known underreporting issue. Secondary outcomes included the estimated admission rate per 1000 population per year, providing a standardised measure of healthcare utilisation, and the overall admission rate ratio comparing the homeless population to the housed population, highlighting potential disparities in hospital use.

Analysis

The analytical approach aimed to estimate the true prevalence of hospital admissions among people experiencing homelessness in England, accounting for underreporting in routine hospital data, and to quantify their healthcare utilisation. Analyses were conducted using STATA V.17.

Data preparation and cleaning

Continuous inpatient spells, representing complete hospital admissions, were constructed using established data cleaning rules from the University of York Centre for Health Economics and the Department of Health. Unique patient identifiers (HESID) were used to link episodes of care within a single admission, ensuring accurate representation of admission events. The predefined homelessness phenotype was applied to the cleaned data.

Descriptive analysis

Descriptive statistics were generated to characterise the observed hospital inpatient activity. Frequencies of admissions were calculated overall and stratified by each homelessness code (NFA, HGP, Z59.0). The overlap between admissions identified by different codes was visualised to illustrate the extent of co-occurrence. Basic demographic data (age, sex and ethnicity) were calculated for both the housed and homeless populations within the dataset,

providing context for potential demographic influences on hospital utilisation.

Multiple systems estimation

To estimate the true number of hospital admissions among people experiencing homelessness, MSE was employed. Each homelessness code (NFA, HGP, Z59.0) was treated as a separate 'list'. Log-linear (Poisson) regression was used to model the dependence structure between these lists, accounting for potential correlations. Eight possible models were fitted, varying in the inclusion of two-way interaction terms between the lists. The Bayesian information criterion (BIC) was used to assess model fit. Due to the wide range of estimates produced for the unobserved population, Bayesian model averaging was conducted. This involved calculating weighted averages of the three best-fitting models (those with the lowest BIC) to produce more stable estimates of the total number of admissions and corresponding SEs. A final 95% CI was computed based on the normal distribution. An inflation factor was calculated by dividing the MSE-estimated total admissions by the observed total admissions, providing an estimate of the true scale of homeless in-patient activity.

To ensure the validity of our MSE, we addressed the four core assumptions: population closure within a single admission, accurate matching between lists using HESID, independence of lists and homogeneous capture probabilities. While the first three assumptions were adequately met, the fourth, concerning homogeneous capture probabilities, presented a challenge. Ideally, this assumption would be addressed by incorporating covariates that might influence the likelihood of an individual being identified by each homelessness code. However, the low rate of duplicate coding across lists (specifically, 11.6% of admissions with two codes and 0.5% with all three) precluded the inclusion of covariates. This limitation arose from the potential for numerical instability that would accompany stratification of the data with such sparse overlap. Consequently, we proceeded with the MSE without covariate adjustment, acknowledging this constraint in our discussion of study limitations. For a detailed discussion on the four assumptions and how they were met (or not), please consult the analysis section of the online supplemental appendix.

Estimation of admission rates and rate ratios

To quantify healthcare utilisation, admission rates for people experiencing homelessness were calculated for the most recent year available (2017/2018) using both observed and MSE-estimated admission numbers. Population denominator estimates for people experiencing homelessness were obtained from 'The Homelessness Monitor'.⁴ We used the denominator definition for 'core' homelessness (rough sleepers, people living in unconventional accommodation, hostels, unsuitable temporary accommodation and sofa surfing). We also conducted a sensitivity analysis restricting the denominator to people

sleeping rough or staying in hostels only, which we have named 'visible' homelessness to calculate rates. Admission rates were calculated per 1000 population per year. Admission rate ratios were then calculated by comparing the homeless admission rates to the housed population's admission rate, using data from the HES-APC dataset and published population denominators.²⁹ 95% CIs were calculated for the rate ratios.

Patient and public involvement

This study forms part of SL's mixed methods PhD project on the preventative role of hospitals for people experiencing homelessness, where Stan BurrIDGE from Expert Focus was the lead patient and public involvement (PPI) advisor. Over 25 people with lived experience of homelessness were involved through workshops and consultations to develop the research priorities, design the studies and review ethical issues. Stan BurrIDGE had a leading role in the qualitative research. Due to the technical nature of this particular study, PPI participants were only involved in the development of the research question and in discussions about the validity of the homelessness phenotype.

RESULTS

Observed number of hospital admissions

The total number of admissions where at least one episode of care was coded as homelessness (ie, with one of the three homelessness codes) was 27 124 in 2013/2014, increasing to 31 933 in 2015/16 and 34 790 in 2017/2018. The Venn diagram (figure 1) shows the coding structure of 2017–2018 admission data. Admissions which were singly coded were the most frequent, with NFA having the highest number of coded admissions (n=11 527). Admissions with overlapping homeless codes were relatively infrequent, particularly those coded with all three homeless codes (n=162). Results

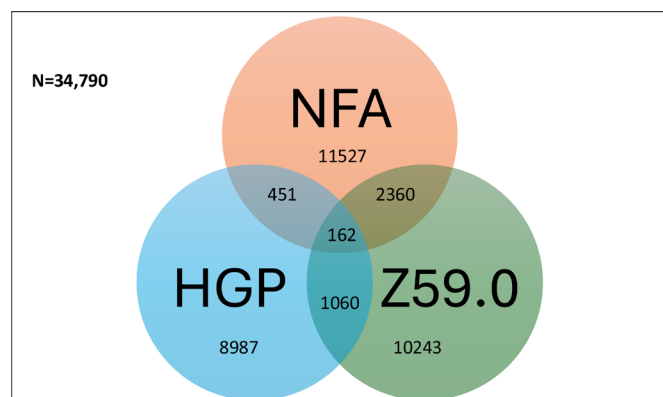


Figure 1 Venn diagram illustrating the observed number of admissions attributed to people experiencing homelessness in England, 2017–2018. The size of the Venn segment is not proportional to the frequency of admissions in that segment. NFA, no fixed abode; HGP, homeless GP; Z59.0, International Classification of Diseases version 10 (ICD-10) code for homelessness.

Table 1 Demographic characteristics of National Health Service (NHS) hospital admissions in England by housing status, 2017/2018

	Homeless N=34 790		Housed N=15 514 367	
	n	%	n	%
Age				
18–25	4349	12.5	985 804	6.4
26–45	16 317	46.9	3 323 951	21.4
46–65	11 376	32.7	4 571 135	29.5
Over 65	2748	7.9	6 633 477	42.8
Sex				
Male	25 362	72.9	6 819 157	44.0
Female	9428	27.1	8 695 210	56.0
Ethnicity				
White (base)	24 944	71.7	12 067 184	77.8
Black/Black British	1983	5.7	444 780	2.9
Asian/Asian British	1600	4.6	785 923	5.1
Mixed	522	1.5	112 953	0.7
Other	1600	4.6	297 953	1.9
Unknown	4140	11.9	1 805 574	11.6

for other years followed the same overall pattern (see online supplemental table 1).

Demographics

The age, sex and ethnicity distribution of admissions for people experiencing homelessness compared with the housed population are presented in table 1. People from younger age groups made up the majority of admissions among those experiencing homelessness (46.9% of admissions were for people aged 26–45 years compared with 7.9% for those over 65). In contrast, among the housed population, those from the middle and older age categories were the largest proportion of hospital admissions (admissions for the two oldest categories combined were 72.3%). Nearly three-quarters of admissions for people experiencing homelessness were in men, compared with 44.0% for the housed population. People experiencing homelessness admitted to hospital had a slightly greater proportion of racially minoritised people than those who were housed (white participants comprised 71.7% for people experiencing homelessness and 77.8% for housed people).

Multiple systems estimation

The estimated number of unobserved admissions and corresponding BICs based on the eight Poisson models are presented in online supplemental table 2. Model 8 (the full model) had the lowest BIC (80), indicating a superior fit to the data. However, it was not considered for further analysis as it perfectly fitted the data, leaving no df for generalisation and rendering it susceptible to overfitting. Models 5, 6 and 7 also demonstrated low

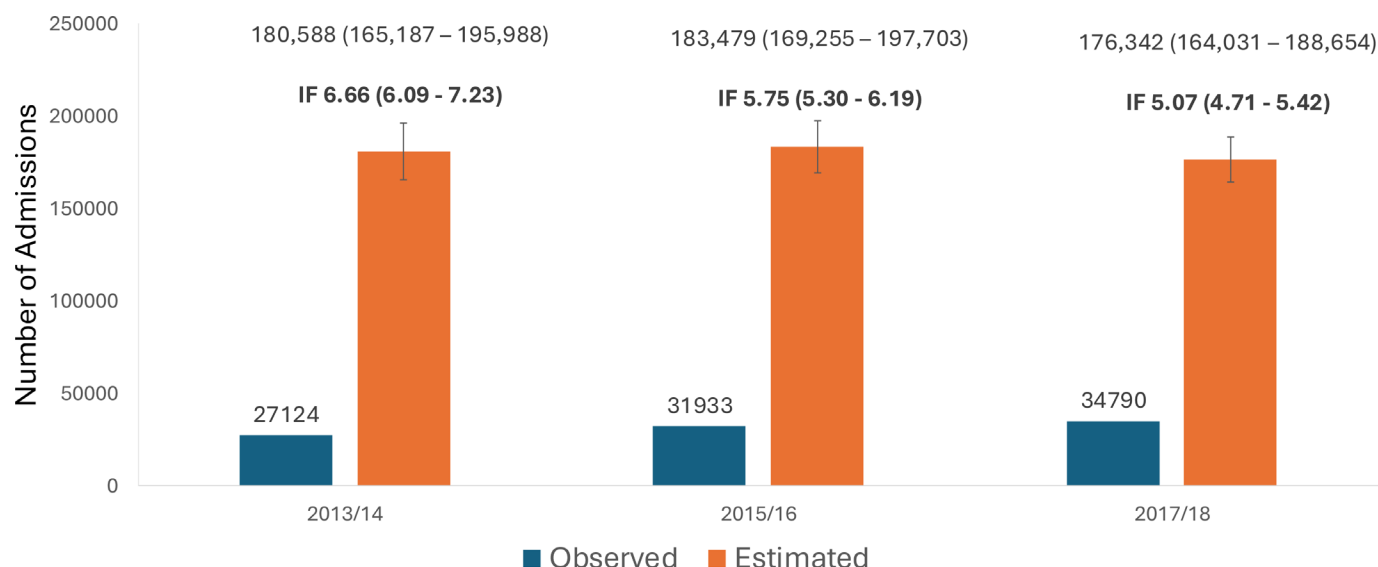


Figure 2 Observed and estimated total number of admissions and inflation factors (IF) by year. Error bars are the 95% CIs around the estimated total. Numerical data are shown within the figure with 95% CIs in brackets. IF, ratio of total estimated admissions to observed admissions.

BICs (111, 190 and 102, respectively), but the range of estimates of the unobserved admissions was large and unstable (model 5–86 843 unobserved admissions, model 6–50 030 and model 7–229 697). To mitigate the risk of relying on a single, potentially unstable model, we adopted a model averaging approach. This involved averaging the three best-fitting, non-overfit models (ie, 5,6,7, excluding 8), thereby producing a more robust and reliable overall estimate. The resulting combined total estimate was 176 342 admissions (95% CI 164 031 to 188 654) for 2017/2018.

The average estimated total number of admissions attributed to people experiencing homelessness decreased slightly over time, but the CIs were overlapping (figure 2). The number of observed admissions increased over time and the corresponding inflation factor, the ratio between the estimated total and the observed total, decreased over time. CIs around the inflation factors overlapped between the first and second years of data and the second and third years, but not the first and third. The observed and estimated numbers of admissions, corresponding inflation factors, and 95% CIs for all time periods are presented in figure 2.

Admission rates and ratios

Based on the observed number of admissions for people experiencing homelessness in 2017/2018 (ie, 34 790 admissions recorded as NFA, HGP or ICD Z59.0), the unadjusted admission rate was 173.4 admissions per 1000 population using the ‘core’ homeless population denominator defined in the Crisis Homeless Monitor⁴ (rough sleepers, people living in unconventional accommodation, hostels, unsuitable temporary accommodation and sofa surfing). As a sensitivity analysis, we restricted the denominator to the ‘visible’ homeless population (rough sleeping and hostel population only) and the unadjusted

admission rate was 693.8 admissions per 1000 population. In the housed population, there were 15 514 367 admissions for an estimated 44 168 935 adults in England in 2018,³⁰ translating to an unadjusted admission rate of 351.3 admissions per 1000 population per year. compared with the housed population, the admission rate ratios for those recorded as homeless were 0.5 and 2.0 for the core and visible homeless population denominators, respectively (table 2).

We estimated that there were 176 342 (95% CI 164 031 to 188 654) admissions in 2017–2018 using MSE, translating to an unadjusted admission rate of 879.0 (95% CI 817.7 to 940.4) admissions per 1000 population per year using the core homeless population denominator. When we restricted to the visible homeless denominator, this resulted in an estimated unadjusted admission rate of 3516.7 (95% CI 3271.2 to 3762.2) admissions per 1000 population per year. We estimated the unadjusted admission rate ratio as 2.5 (95% CI 2.3 to 2.7) for people experiencing homelessness compared with the housed population in England using the core homeless denominator. For the visible homeless population denominator, the unadjusted admission rate ratio was estimated as 10.0 (95% CI 9.3 to 10.7) compared with the housed population (table 2).

DISCUSSION

This population-based repeated cross-sectional study used national hospital records and MSE to investigate the scale of inpatient activity for people experiencing homelessness in England. We provide a plausible range of admission numbers, rates and ratios to support service planning as well as advocacy for improved data collection on housing status in the English national health service. From an observed 34 790 admissions in 2017/2018, we

Table 2 Hospital admission rates for people experiencing homelessness compared with the housed population in England

Numerator	Number of admissions in HES† 2017/2018	Denominator*	Denominator definition and reference	Admission rate per 1000 population per year	Admission rate ratio
Observed admissions for people experiencing homelessness	34 790	200 609	Core homelessness ³⁷	173.4	0.5
Observed admissions for people experiencing homelessness	34 790	50 144	Visible homelessness ³⁷	693.8	2.0
Estimated admissions for people experiencing homelessness	176 342 (95% CI 164 031 to 188 654)	200 609	Core homelessness ³⁷	879.0 (95% CI 817.7 to 940.4)	2.5 (95% CI 2.3 to 2.7)
Estimated admissions for people experiencing homelessness	176 342 (95% CI 164 031 to 188 654)	50 144	Visible homelessness ³⁷	3516.7 (95% CI 3271.2 to 3762.2)	10.0 (95% CI 9.3 to 10.7)
Housed admissions	15 514 367	44 168 935	General population ³⁰	351.3	1.0

*There were no CIs provided in estimates of homeless population denominators.
†HES, Hospital Episode Statistics

estimated the total number of homeless admissions to be five times greater at 176,342 (95% CI 164 031 to 188 654). The overall unadjusted admission rate for 2017/2018 using MSE was 879.0 (95% CI 817.7 to 940.4) admissions per 1000 population per year using the ‘core’ homeless population denominator and 3516.7 (95% CI 3271.2 to 3762.2) admissions per 1000 population per year when restricting to the ‘visible’ homeless population. The unadjusted admission rate ratios were estimated as 2.5 (95% CI 2.3 to 2.7) and 10.0 (95% CI 9.3 to 10.7), respectively, compared with the housed adult population of England.

Findings in the context of the wider literature

Previous research has highlighted the issue of underestimation of homelessness in routine data sources. In England during the COVID-19 pandemic, the ‘Everyone In’ campaign housed 37 000 homeless individuals,³¹ vastly outnumbering the 4266 rough sleepers estimated in 2019 statistics²—a clear example of underestimation in official data. Similarly, the Office for National Statistics (ONS) uses MSE methodology to estimate deaths among people experiencing homelessness. In 2018, ONS observed 541 deaths but estimated 726, yielding an inflation factor of 1.34.³² Underestimation of homelessness in electronic health data has also been described in other high-income countries, such as the USA^{33–35} and Canada.³⁶

We compared our findings with two English studies which relied on the NFA code alone to demonstrate the added value of our extended homelessness phenotype for estimating observable (without MSE) inpatient activity. The UK Department of Health’s 2007/2008 assessment of hospital inpatient care for people experiencing homelessness, which identified 17 400 inpatient episodes for 15 800 individuals using only the NFA code. This study did not construct admissions as we have done here, so we used our phenotype to count the observed number of

episodes and individuals for comparison. We identified 46 631 episodes for 23 956 individuals in 2017/2018 using the observed data, which is substantially higher than using NFA alone. Similarly, a recent study of emergency admissions of people experiencing homelessness coded with NFA²⁷ identified far fewer admissions than our study (14 858 emergency inpatient admissions in 2018/2019). In our study, 82% of our admissions were for an emergency (28 528 admissions in 2017/2018).

The Department of Health study also estimated hospital admission rate ratios for homeless versus housed populations using data from four specialist homeless services in London, Leicester and Cambridge.¹³ Admissions were on average 3.2 times higher for people experiencing homelessness than for the housed population, compared with an admission rate ratio of 2.0 in our study (based on observed admissions relative to ‘visible’ homeless populations (people rough sleeping and living in hostels). However, when we used the CRC estimate as the numerator, we calculated an admission rate ratio of 10.0 (95% CI 9.3 to 10.7) for this population. Our estimates are likely to include a small, but unknown, number of individuals from the broader core homeless population. Consequently, our admission rates for the visibly homeless population may be an overestimate, with rates for the core homeless potentially being an underestimate. Despite this uncertainty, our study provides potential bounded estimates for the admission rate ratio for people experiencing homelessness compared with the general population.

We observed a slight decrease in inflation factors over time, likely due to changes in hospital coding practice which are largely driven by a shifting policy and funding landscape.^{37 38} Online supplemental figure 1 and online supplemental table 1 support this hypothesis,

demonstrating an increase in the use of the ICD-10 code for homelessness over time. Although the observed number of homeless admissions rose over time, the total estimated admissions remained stable, leading to a decreasing inflation factor. The stability and relevance of these inflation factors should be confirmed through analyses of additional recent years.

Strengths

This is the first population-based MSE analysis of hospital admissions for people experiencing homelessness in England and the largest study of this population in the UK to date. For the first time, it quantifies the extent of the significant underestimation of the number of hospital admissions for people experiencing homelessness in routine data. The precise CIs enhance the reliability of the estimates. Our modelling approach accounted for the interdependence of various homeless codes, addressing a common issue in MSE studies. Additionally, the model averaging strategy provided a stable estimate of total admissions and inflation factors. Since the data were collected before the COVID-19 pandemic, they better reflect long-term hospital care utilisation patterns.

Limitations

MSE can be used to estimate the total number of admissions when accurate data are unavailable, but it cannot attribute outcomes to specific individuals. The validity of the homelessness phenotype used in this study has not been subject to a formal validation study. A recent Canadian validation study demonstrated that the use of the ICD-10 code, Z59.0, which is mandatory in Canada, is highly specific (99.5%), but has lower sensitivity (52.9%). This was likely because people were coded if they had a history of homelessness, rather than current homelessness. It is possible that the codes used in this study have similar issues and that validation and improved coding of housing status is needed.

The original MSE method relies on four key assumptions, as outlined earlier in this paper. Some assumptions were likely violated, yet we consider the overall impact on estimates to be minimal. We assumed our population was sufficiently closed as most individuals do not fully transition in or out of homelessness during a single admission.³⁹ The unique identifier, HESID, enabled largely accurate matching between homeless codes when there were multiple episodes of care, although this does not address coding errors.⁴⁰ It is possible that some residual bias was unaccounted for when we adjusted for the lack of independence between homeless codes, which would result in underestimation of the total number of admissions. The greater the overlap between codes, the closer the estimate is to the observed number of admissions.

A key limitation was our inability to adjust for confounders such as age, sex and ethnicity. This arose because we used three different homeless code lists (due to the incomplete nature of homelessness data) which identified very few admissions coded with more than one

homelessness code. The low degree of overlap between codes meant our data were too fragmented to create a clear picture of the population size and led to model instability. We used model averaging to address model instability, but because our data were so fragmented, we could only do this for the entire homeless population as a whole. We were, therefore, unable to perform a more detailed analysis that adjusted for important confounders that might influence the likelihood of an admission being coded with one of the homelessness codes and the resulting inflation factor which we calculated. In other words, we may have underestimated or overestimated hospital activity for different subgroups, particularly for more hidden homeless populations such as women.

Recommendations

In the absence of routine identification and coding of homelessness in healthcare records in England, policy-makers and service providers should consider applying MSE methods to improve estimates of hospital care utilisation for people experiencing homelessness in their specific context. However, we do not advise applying this study's inflation factors directly as these factors are contingent on context-specific coding practices. For example, areas where there is greater overlap between the homeless codes will likely have better model stability and may, therefore, be able to include covariates, further strengthening the analysis.

Further research should aim to adjust for covariates, particularly age, sex/gender and ethnicity, as these factors vary considerably between homeless and housed populations, as shown in this study. This study's data predates the pandemic, raising questions about the generalisability of the estimates and inflation factors post-2020. Research is needed to compare these factors during the pandemic, when homelessness decreased due to initiatives like the Everyone In campaign to house all people, and after the pandemic, when rates rose again due to the cost-of-living crisis.³⁷

However, a critical step in addressing health inequities in people experiencing homelessness is system-wide uptake and implementation of improved routine coding practices for homelessness. Quite simply, if people are not counted, then they will not count. Researchers have attempted to address this issue through using multiple data sources to identify people experiencing homelessness^{33–35 41} (although none to our knowledge have applied MSE as we have), using AI and machine learning approaches to identify people,⁴² and through developing bespoke data capture systems.⁴³ Making data collection on housing status a routine practice, just like collecting a person's birthdate or sex/gender, would allow us to more accurately understand how homelessness affects people's health. While current methods help, they only go so far, and a more standardised approach is needed to capture the full extent of hospital admissions among people experiencing homelessness.

Clinicians should routinely ask about housing status and accurately record it; this would be supported by inclusion of housing status in routine proformas and emphasis of its importance among healthcare providers. Lessons on specific clinical workflows and optimisation of coding practices can be learnt from other contexts.⁴³ Similarly, professional coders need specific guidance and training to code homelessness in administrative datasets. For example, Canada's legal mandate for coding homelessness using the ICD-10 Z59.0 code dramatically increased coding frequency and sensitivity, though specificity needed further improvements.³⁶ Although the ICD-10 code is the most specific method for recording homelessness currently available in English hospital records, it fails to capture the diversity of homelessness types, which is linked to different health needs. Moreover, it prevents the calculation of accurate admission rates for a defined population denominator (eg, 'core' homeless vs 'visible' homeless). Primary care records in England use a wider variety of descriptive codes (eg, for a list of homelessness READ codes and SNOMED codes see⁴¹), but systematic categorisations of homelessness linked to health risk (such as the EHTOS typology¹) are also lacking in primary care. National implementation of system-wide improved coding practices is urgently needed to enhance service planning and delivery for people experiencing homelessness.

CONCLUSIONS

We have developed new methods using MSE to identify the scale of hospital admissions for people experiencing homelessness in England. The observed number of homeless admissions increased over the study period, and significant demographic differences were noted between homeless and housed populations. MSE estimates were substantially higher than observed counts, indicating significant underreporting. However, given the methodological concerns, particularly the limited covariate adjustment and potential model instability, the magnitude of this underreporting should be interpreted cautiously. While this study provides valuable insights into the potential scale of hospital admissions among people experiencing homelessness, the methodological limitations necessitate a tempered interpretation of the findings. Future research should prioritise improved identification strategies that allow for covariate adjustment in MSE and ensure model stability. Additionally, further validation of the homelessness phenotype is needed.

Author affiliations

¹Institute of Epidemiology and Healthcare, University College London, London, UK

²S3RI & Mathematical Sciences, University of Southampton, Southampton, UK

³Institute of Health Informatics, University College London, London, UK

⁴Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

⁵Institute of Global Health, University College London, London, UK

⁶UK Health Security Agency, London, UK

Acknowledgements We would like to acknowledge Dr Nigel Hewett, retired medical director of Pathway, for sharing his experience of Pathway teams

identifying people experiencing homelessness within hospitals in England. Early conversations with Dr Hewett inspired the need for this research. We would also like to thank Mr Stan Burridge, Director of Expert Focus and lead PPI advisor for this study, for sharing his lived experience and expertise for the benefit of this research.

Contributors SAL and ACH conceived and designed the study and ACH is the guarantor. DB provided methodological advice and statistical support on the capture-recapture methodology. SAL conducted the analyses and drafted the initial manuscript. All authors supported the interpretation of the findings and contributed to editing the manuscript. All authors had final responsibility for the decision to submit the manuscript for publication.

Funding SAL was funded by an NIHR/HEE Clinical Doctoral Research Fellowship (ICA-CDRF-2016-02-042).

Competing interests SAL is a Fellow of the Pathway charity and has received funding from NIHR. RA has received funding from the Wellcome Trust, NIHR and the European Union. ACH is UK Health Security Agency national lead for Inclusion Health—this includes remit to advise on health protection for people experiencing homelessness. DB, FS and ST have declared no interests.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by UCL Research Ethics Committee (Ref: 11607/001). These were fully anonymised routine electronic health data (Hospital Episodes Statistics).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Hospital Episodes Statistics data are controlled by NHS England and researchers will need to apply for access directly. Further information about obtaining access is available here: <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/users-uses-and-access-to-hospital-episode-statistics>. We have created an online supplementary appendix including the detailed methods and statistical code used in this study to support reproducibility of this research.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Serena April Luchenski <https://orcid.org/0000-0002-6260-4606>

Dankmar Böhning <https://orcid.org/0000-0003-0638-7106>

Robert Aldridge <https://orcid.org/0000-0003-0542-0816>

Fiona Stevenson <https://orcid.org/0000-0003-2139-2309>

Shema Tariq <https://orcid.org/0000-0001-9802-7727>

Andrew C Hayward <https://orcid.org/0000-0002-3549-6232>

REFERENCES

- 1 ETHOS typology on homelessness and housing exclusion. Available: <https://www.feantsa.org/en/toolkit/2005/04/01/ethos-typology-on-homelessness-and-housing-exclusion> [Accessed 13 Jun 2023].
- 2 UK Department for Communities and Local Government. Homelessness statistics. 2015. Available: <https://www.gov.uk/government/collections/homelessness-statistics> [Accessed 25 Nov 2015].

- 3 Unhealthy state of homelessness 2022: findings from the homeless health needs audit. Homeless Link. Available: <https://homeless.org.uk/knowledge-hub/unhealthy-state-of-homelessness-2022-findings-from-the-homeless-health-needs-audit/> [Accessed 21 Dec 2023].
- 4 Fitzpatrick S, Bramley G, McMordie L, *et al.* The homelessness monitor: England 2023. London, 2023.
- 5 Himsworth C, Paudyal P, Sargeant C. Risk factors for unplanned hospital admission in a specialist homeless general practice population: case-control study to investigate the relationship with tri-morbidity. *Br J Gen Pract* 2020;70:e406–11.
- 6 Aldridge RW, Menezes D, Lewer D, *et al.* Causes of death among homeless people: a population-based cross-sectional study of linked hospitalisation and mortality data in England. *Wellcome Open Res* 2019;4:49.
- 7 Aldridge RW, Story A, Hwang SW, *et al.* Morbidity and mortality in homeless individuals, prisoners, sex workers, and individuals with substance use disorders in high-income countries: a systematic review and meta-analysis. *Lancet* 2018;391:241–50.
- 8 Elwell-Sutton T, Fok J, Albanese F, *et al.* Factors associated with access to care and healthcare utilization in the homeless population of England. *J Public Health* 2016;fdw008.
- 9 Rathod SD, Guise A, Annand PJ, *et al.* Peer advocacy and access to primary healthcare for people who are homeless in London, UK: a mixed method impact, economic and process evaluation protocol. *BMJ Open* 2021;11:e050717.
- 10 Perkin S, Visram S, Lindsey L. 'What does good look like'—exploring access to healthcare for the homeless population in Gateshead, England. *J Public Health (Bangkok)* 2023;45:e486–93.
- 11 Gunner E, Chandan SK, Marwick S, *et al.* Provision and accessibility of primary healthcare services for people who are homeless: a qualitative study of patient perspectives in the UK. *Br J Gen Pract* 2019;69:e526–36.
- 12 Hudson BF, Flemming K, Shulman C, *et al.* Challenges to access and provision of palliative care for people who are homeless: a systematic review of qualitative research. *BMC Palliat Care* 2016;15:96.
- 13 Department of Health. *Healthcare for single homeless people*. 2010.
- 14 Lewer D, Menezes D, Cornes M, *et al.* Hospital readmission among people experiencing homelessness in England: a cohort study of 2772 matched homeless and housed inpatients. *J Epidemiol Community Health* 2021;75:681–8.
- 15 Khan Z, Koehne S, Haine P, *et al.* Improving outcomes for homeless inpatients in mental health. *HCS* 2018;22:77–90.
- 16 Wadhwa RK, Choi E, Shen C, *et al.* Trends, Causes, and Outcomes of Hospitalizations for Homeless Individuals: A Retrospective Cohort Study. *Med Care* 2019;57:21–7.
- 17 McCormick B, White J. Hospital care and costs for homeless people. *Clin Med (Northfield)* 2016;16:506–10.
- 18 Moss C, Sutton M, Cheraghi-Sohi S, *et al.* Comparative 4-year risk and type of hospital admission among homeless and housed emergency department attendees: longitudinal study of hospital records in England 2013–2018. *BMJ Open* 2021;11:e049811.
- 19 Hajat S, Sarran CE, Bezgrebelna M, *et al.* Ambient Temperature and Emergency Hospital Admissions in People Experiencing Homelessness: London, United Kingdom, 2011–2019. *Am J Public Health* 2023;113:981–4.
- 20 Blackburn RM, Hayward A, Cornes M, *et al.* Outcomes of specialist discharge coordination and intermediate care schemes for patients who are homeless: analysis protocol for a population-based historical cohort. *BMJ Open* 2017;7:e019282.
- 21 Lewer D, Menezes D, Cornes M, *et al.* Hospital Readmissions Among People Experiencing Homelessness: A Cohort Study of Linked Hospitalisation and Mortality Data in England for 3,222 Homeless Inpatients. *SSRN Journal* 2020.
- 22 Nadicksbernd J, Nguyen T, Jackson T, *et al.* Health and care needs of hospitalised people experiencing homelessness: an inpatient audit. *Clin Med (Northfield)* 2023;23:395–402.
- 23 Homeless Link. The unhealthy state of homelessness: health audit results. 2014.
- 24 Luchenski S. Homelessness phenotype. HDRUK P; 2021. Available: <https://phenotypes.healthdatagateway.org/phenotypes/PH344/version/688/detail/> [Accessed 19 Dec 2023].
- 25 Crane M, Cetrano G, Joly L, *et al.* Mapping of specialist primary health care services. 2018.
- 26 NHS Digital. NHS digital ODS portal. 2024. Available: <https://odsportal.digital.nhs.uk/> [Accessed 10 Oct 2024].
- 27 Paudyal V, Vohra N, Price M, *et al.* Key causes and long-term trends related to emergency department and inpatient hospital admissions of homeless persons in England. *Int J Emerg Med* 2023;16:48.
- 28 King's Health Partners Pathway Team. KHP pathway homeless team - first year report. 2014.
- 29 ONS. Population estimates for the UK, England and Wales, Scotland and northern Ireland: mid- 2018 national. Office for National Statistics; 2019.
- 30 Office for National Statistics. Estimates of the population of for England and Wales: dataset 1838 - 2023. 2024.
- 31 Garvie D, Rich H, Berry C, *et al.* Everyone in: where are they now? London; 2021.
- 32 Deaths of homeless people in England and Wales. Office for National Statistics. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsofhomelesspeopleinenglandandwales/2019registrations> [Accessed 18 Jul 2022].
- 33 Stella SA, Hanratty R, Davidson AJ, *et al.* Improving Identification of Patients Experiencing Homelessness in the Electronic Health Record: A Curated Registry Approach. *J Gen Intern Med* 2024;39:3113–9.
- 34 Zech J, Husk G, Moore T, *et al.* Identifying homelessness using health information exchange data. *J Am Med Inform Assoc* 2015;22:682–7.
- 35 Biederman DJ, Modarai F, Gamble J, *et al.* Identifying Patients Experiencing Homelessness in an Electronic Health Record and Assessing Qualification for Medical Respite: A Five-Year Retrospective Review. *J Health Care Poor Underserved* 2019;30:297–309.
- 36 Richard L, Carter B, Nisenbaum R, *et al.* Identification of homelessness using health administrative data in Ontario, Canada following a national coding mandate: a validation study. *J Clin Epidemiol* 2024;172:111430.
- 37 Fitzpatrick S, Pawson H, Bramley G, *et al.* The homelessness monitor: England. London, 2021.
- 38 Stuckler D, Reeves A, Loopstra R, *et al.* Austerity and health: the impact in the UK and Europe. *Eur J Public Health* 2017;27:18–21.
- 39 Cornes M, Whiteford M, Manthorpe J, *et al.* Improving hospital discharge arrangements for people who are homeless: A realist synthesis of the intermediate care literature. *Health Soc Care Community* 2018;26:e345–59.
- 40 Hardy F, Heyl J, Tucker K, *et al.* Data consistency in the English Hospital Episodes Statistics database. *BMJ Health Care Inform* 2022;29:e100633.
- 41 Nanjo A, Evans H, Direk K, *et al.* Prevalence, incidence, and outcomes across cardiovascular diseases in homeless individuals using national linked electronic health records. *Eur Heart J* 2020;41:4011–20.
- 42 Xie F, Wang S, Viveros L, *et al.* Using natural language processing to identify the status of homelessness and housing instability among serious illness patients from clinical notes in an integrated healthcare system. *JAMIA Open* 2023;6:ooad082.
- 43 Angoff GH, O'Connell JJ, Gaeta JM, *et al.* Electronic medical record implementation for a healthcare system caring for homeless people. *JAMIA Open* 2019;2:89–98.