A population level study of SARS-CoV-2 prevalence amongst people experiencing homelessness in Wales, UK

Ian Thomas[1,*] and Peter Mackie[2]

Abstract

Introduction

Prior research into the prevalence of SARS-CoV-2 infection amongst people experiencing homelessness (PEH) largely relates to people in communal forms of temporary accommodation in contexts where this type of accommodation remained a major part of the response to homelessness during the COVID-19 pandemic. Little is known about the prevalence of SARS-CoV-2 amongst PEH more broadly, and in a policy and practice context that favoured self-contained accommodation, such as Wales, UK.

Objective

Describe the prevalence of SARS-CoV-2 amongst PEH in Wales, UK, using routinely collected administrative data from the Secure Anonymised Information Linkage Databank.

Methods

Routinely collected data were used to identify PEH in Wales between 1st March 2020 and 1st March 2021. Using SARS-CoV-2 pathology testing data, prevalence rates were generated for PEH and three comparator groups: (1) the not-homeless population; (2) a cohort 'exact matched' for age, sex, local authority and area deprivation; and (3) a matched comparison group created using these same variables and Propensity Score Matching (PSM). Three logistic regressions were run on samples containing each of the comparator groups to explore the effect of experiencing homelessness on testing positive for SARS-CoV-2.

Results

The prevalence of SARS-CoV-2 infection amongst PEH was 5.0%, compared to the not-homeless population at 5.6%. For the exact matched and PSM match comparator groups, prevalence was 6.9% and 6.7%, respectively. Logistic regression found that SARS-CoV-2 infection was 0.9 times less likely amongst PEH compared to people not experiencing homelessness from the general population. The odds of SARS-CoV-2 infection for PEH was 0.75 and 0.73 where the 'not-homeless' comparators were from the exact match and PSM samples, respectively.

Conclusion

Our analysis revealed that a year into the COVID-19 pandemic, the prevalence of SARS-CoV-2 amongst PEH in Wales was lower than the general population. A policy response to homelessness that moved away from communal accommodation may be partly responsible for the reduced SAR-CoV-2 infection amongst PEH.

Keywords

people experiencing homelessness; Wales; COVID-19; SARS-CoV-2

*Corresponding Author:
Email Address: ThomasIR2@cardiff.ac.uk (Ian Thomas)

Introduction

The onset of the Coronavirus (COVID-19) disease pandemic, caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) virus, prompted widespread concern about the potential impact of the virus on people experiencing homelessness (PEH) [1]. Concerns centred upon people who were literally roofless and those in communal forms of accommodation, such as shelters and hostels, where facilities and air space were shared. It was feared these environments could hamper a person's ability to adhere to public health instructions regarding hand hygiene, maintaining physical distancing, and isolation when symptomatic or following a positive SARS-CoV-2 test [1, 2]. The COVID-19 pandemic triggered unprecedented policy and practice responses, with homelessness being reframed as an urgent public health crisis for the very first time [3].

Several studies have begun to shed light on the different approaches taken in assisting PEH over the course of the pandemic [4, 5]. Three types of response have been identified internationally [5]:

(1) continued use and expansion of shelters, whilst introducing protective measures such as physical distancing and Personal Protective Equipment,

(2) continued use of shelters, supplemented by a broadened accommodation offer, including self-contained or single-room accommodation, and

(3) widespread commissioning of self-contained or single-room provision and the closure of communal forms of accommodation.

The focus of this article is Wales, a UK nation where the response to PEH during the pandemic followed the third of these response types. We aim to describe the prevalence of SARS-CoV-2 amongst PEH in Wales and consider the extent to which the response to homelessness may have led to the prevalence rates identified.

Background and study rationale

Homelessness can be conceptualised as a lack of housing which is habitable, secure, and private [6]. People experiencing homelessness (PEH) can include those living on the street ('roofless'), people living in temporary or crisis accommodation, such as hostels and shelters ('houseless'), and people living in insecure accommodation, for example 'sofa surfing', or severely inadequate accommodation ('housing excluded'). The extant literature on SARS-CoV-2 prevalence amongst PEH largely supports the initial concerns around high rates of infection amongst this group of people.

A systematic review of 37 mainly shelter-based studies of SARS-CoV-2 infection amongst PEH found that prevalence peaked in the context of shelter outbreaks; 31.5% compared to 2.3% without a local outbreak [7]. The same systematic review was also able to generate pooled estimates for SAR-CoV-2 prevalence amongst staff working in shelters, which were 14.8% and 1.5% in the context of an outbreak and no outbreak, respectively. Several studies, mainly from outside the US, have sought to compare SAR-CoV-2 prevalence amongst

PEH using shelters to other groups of PEH, and the general 'not-homeless' population.

In Atlanta, Georgia, testing for SARS-CoV-2 infection amongst people in shelters, those who were homeless and unsheltered (roofless), and staff members at shelters during April-May 2020, found prevalence rates of 2.1%, 0.5%, and 1.8%, respectively [8]. A study using random sample SARS-CoV-2 antibody testing in different forms of emergency/crisis service in Paris found prevalence rates of 27.8% at foodbanks, 50.5% of those in emergency shelters, and 88.7% in workers' residences—being temporary housing for migrants [9]. In a second European study, in Denmark, a cross-sectional nationwide examination of SARS-CoV-2 antibodies was conducted amongst people in shelters, shelter workers, and the general population in November 2020 [10]. The prevalence of antibodies was similar among homeless people in shelters (6.8%) and those working in shelters (6.3%), whilst prevalence rates of both groups were significantly higher than the general population (2.9%). Finally, in Ontario, Canada, population-based analysis using administrative data found that between January and July 2020, people with recent experiences of homelessness had higher rates of testing positive for SARS-CoV-2 than 'community-dwelling' people, 2.01 per 100 person-years compared to 0.39 per 100 person-years, respectively [11].

Current information on SARS-CoV-2 prevalence amongst PEH is based on a narrow definition of homelessness, largely drawing on houseless people in shelters and hostels. The high rates of infection found amongst PEH is therefore unsurprising given that the extant evidence base relates to PEH in high-risk communal accommodation, where there are large volumes of people in close contact, and a high degree of churn as clients enter and exit accommodation. However, as indicated previously, homelessness is more than people staying in shelters and hostels. Our study therefore sought to contribute to understanding SARS-CoV-2 prevalence amongst a wider population of PEH than those who are houseless staying in shelters and hostels.

Additionally, the literature on SARS-CoV-2 infection amongst PEH originates from countries and regions that continued to use communal forms of accommodation as a response to homelessness during the COVID-19 pandemic— response types 1 and 2—such as the United States. In the UK however, where self-contained and single-room provision constituted the primary response to homelessness during the pandemic, there is little published information on the prevalence of SARS-CoV-2 amongst PEH. One early UK study produced a prevalence estimate of 4.1% amongst PEH in England in May 2020 [12]. However, this estimate was based on modelling of possible infection rates amongst PEH in hostels, shelters, and living on the streets, rather than observational data. Therefore, the second contribution of this paper is to explore SARS-CoV-2 prevalence in a very different COVID-19 homelessness policy and practice setting to the extant literature, which we now discuss to situate our analysis.

Policy and practice context relevant to the study

The Housing (Wales) Act 2014 places a duty on local authorities to take reasonable steps to assist all households

at-risk of homelessness, or already homeless. Whilst this legislation results in positive outcomes for many people, there is ultimately no duty to provide temporary or settled accommodation unless a household is in 'priority need' for assistance—a legal category that primarily includes households with children [13, 14]. The COVID-19 homelessness response in Wales resulted in important changes to this legislative and policy context, consisting of two main phases.

Beginning in March 2020, the first phase of Wales' COVID-19 response included £10m of support and clear guidance from the Welsh Government, directing local authorities to ensure everyone was suitably accommodated – this included statutory guidance that extended priority need status to people sleeping rough. The local authority response primarily focused on sourcing additional temporary accommodation, including hotels, bed and breakfasts, holiday lets, university accommodation, and social housing. To a lesser extent, some existing temporary accommodation, such as night shelters, were decommissioned or adapted so that there were no shared rooms. A further £40m was made available by the Welsh Government in June 2020 to support the second phase recovery response. Crucially, guidance that accompanied this phase stated that short term accommodation must meet minimum expectations and it was no longer acceptable to offer emergency floor space, to provide tents, or to use sleeping 'pods'—portable self-contained cabins [15, pp. 9].

For people who were roofless, such as those living on the streets, changes to priority need meant that they now had access to emergency accommodation for which they may not previously have been eligible. Guidance provided to local authorities also indicated that those in insecure accommodation, such as sofa surfers or those leaving prisons, who may become roofless, were also considered vulnerable and eligible for temporary accommodation [16]. For people already in temporary accommodation, the closure of communal settings and the sourcing of self-contained accommodation meant that they now had access to improved housing that reduced household mixing and facilitated self-isolation. In summary, the changes instigated by Welsh Government during the pandemic were extensive and had far-reaching impacts on all segments of PEH in Wales. This paper provides the first empirical insights into the potential impacts of these changes on SARS-CoV-2 infection.

Methods

This study used administrative data sets made available via the Secure Anonymised Information Linkage (SAIL) databank. Data sets included: Emergency Department Data Set (EDDS); Patient Episode Database for Wales (PEDW); Primary Care (GP) data; Substance Misuse Data Set (SMDS); the Welsh Demographic Service (WDS); and SARS-CoV-2 pathology test results. GP data represent information from interactions with primary care providers in Wales, whilst the EDDS and PEDW reflect secondary healthcare interactions. The SMDS data relate to interactions with Welsh Government-funded substance use services in Wales. The WDS contains information on people registered with a General Practitioner in Wales.

All data sets were linked using a unique identifier for each person in Wales, the Anonymised Linkage Field (ALF), which is assigned to all data within SAIL [17]. ALFs are assigned either deterministically, based on NHS Wales numbers, or probabilistically based on personal identifiers, such as name, date of birth, postcode. We only retained records that had an ALF-match accuracy of greater than 90%, or where ALFs were matched deterministically.

The study period of interest was the 1st March 2020 up to and including 1st March 2021. Several processes were involved in preparing data prior to analysis, including: the creation of a population spine for Wales to provide a stable cohort of people for later analysis; linking to SARS-CoV-2 pathology testing data as the main outcome of interest; flagging people within the population spine who had experienced homelessness; and creating a series of 'not-homeless' comparison groups. Each of these processes is now discussed.

Creating a resident population spine

We used the WDS to create a population spine, within which PEH were flagged, and which was used to generate 'not-homeless' comparator groups for analysis. The WDS contains a unique residential identifier, known as the Residential Anonymised Linkage Field (RALF), alongside information about when a person registered and de-registered as living at that RALF, and a date of death. The cohort used in this analysis included all people who were alive and had an entry on the WDS which covered the 1st March 2020.

Main outcome measure

Having a positive test for SARS-CoV-2 was flagged in the resident population spine using both Polymerase Chain Reaction (PCR) and antibody serology test results. We included all positive test results where a sample was taken between 1st March 2020 up to and including 1st March 2021. As such, our analysis covers infection from the start of the SARS-CoV-2 outbreak—with the first recorded case in Wales being 28th February 2020 [18]—up to and including both the initial and second waves—which peaked in April 2020 and December 2020, respectively.

Though Lateral Flow Device (LFD) testing data are available within SAIL for linkage studies, we chose not to use LFD data. Meta-analysis of LFD testing has found that it has poor sensitivity in detecting SARS-CoV-2 infection [19], and potentially provides high false positive rates when community cases rates are low [20]. Furthermore, as SARS-CoV-2 testing policy in Wales is that a positive lateral flow test be followed by a confirmatory PCR test, then people with SARS-CoV-2 identified through positive LFD tests would, in most cases, be recorded under the more reliable PCR testing process [21].

From the original SARS-CoV-2 pathology testing data sets, 145,539 (5%) records were excluded from analysis due to either a lack of a linkage field (ALF), or because the linkage quality was low (fuzzy matching rate <90%). Most records (99.96%) were removed because there was no match. Where a person had multiple tests during the period of interest, if any of these tests were positive, then they were flagged as having had SARS-CoV-2.

Identifying people experiencing homelessness (PEH)

There is a lack of national individual-level homelessness data in Wales which was problematic given the study's aim of exploring the prevalence of SARS-CoV-2 amongst PEH. PEH in Wales were identified using non-housing related administrative data. We initially identified GP data and SMDS data as sources of information on PEH through exploration and analysis using data assets in the SAIL Databank [22]. We then cross-referenced and extended these initial data sources and codes, drawing on published peer reviewed studies from Canada, England, and Wales which use administrative data to identify PEH [23–25]. Based on this literature, we added further codes for extracting PEH from the GP data, along with additional sources (EDDS and PEDW). As our analysis represented a rapid response to COVID-19, we were unable to undertake further validation of data sources and codes beyond using those identified by studies that had undergone a peer review process.

Table 1 provides an overview of variables and codes within each data set that relate to experiences of homelessness ('homelessness event codes'), and the initial number of people identified as having experienced homelessness. Given the range of homelessness event codes, the population of PEH included in this analysis were people without housing living on the streets ('roofless'), people living in hostels, shelters, and other temporary accommodation such as bed and breakfasts ('houseless'), and people in insecure housing situations, such as 'sofa surfers' [6]. Though event codes related to inadequate housing were included, this accounted for a relatively small proportion of PEH. The population of PEH included in our analysis are therefore in line with international definitions of extreme homelessness [26].

Experiences of homelessness were identified in the health care data through its recording within diagnosis fields, with each health data set using a different coding system. PEDW uses the International Classification of Diseases 10th Revision which includes a series of codes related to factors influencing health status and contact with health services. General Practitioners use 'read codes' to record all aspects of interactions with patients, which can include social elements relevant to a diagnosis. Emergency departments use a specialised system of coding for diagnoses which allows up to six diagnosis codes per entry, with any homelessness related codes entered being eligible for inclusion. In the SMDS, people in receipt of support from a government-funded substance use service were explicitly asked their housing status on two occasions: upon initial assessment and during Treatment Outcomes Profiles conducted semi-regularly during receipt of treatment [27]. We drew on both these measures of homelessness from the SMDS data; this differs slightly to other uses of the SMDS to identify PEH in Wales [25] which drew on housing status from the initial assessment alone.

Several of the event codes, such as 'Z590 – Homelessness' in the EDDS, relate to a general state of homelessness without the ability to differentiate roofless, houseless, or housing excluded. However, most codes identify specific forms of homelessness, such as '13FL. – Living rough' in the GP data, or, as in the case of the SMDS, contained additional information for use by the clinician when coding the housing situation. In the SMDS, examples of situations that related to an 'urgent housing problem' were living on the streets, using hostel accommodation or sleeping in different accommodation each night; whilst for 'housing problems', examples provided to the clinician covered housing situations that were insecure, such as staying with others as a short-term guest, accessing temporary accommodation for short-term stays, or squatting [27].

For the purposes of this analysis, we retained all homelessness events that were recorded during the study period. As homelessness was recorded on a (health) event basis, rather than having a start and end date to the person's homelessness, we were not able to assess whether the person was experiencing homelessness at the time of infection with SARS-CoV-2. Our analysis therefore relates to people who had recently experienced homelessness, though we refer to them as people experiencing homelessness, or PEH, for ease. PEH were flagged in the resident population spine if they were recorded as having any homelessness event codes during the study period. There were an initial 7,006 event codes that related to homelessness, belonging to 4,049 unique individuals. Consistent with similar uses of these data sets to identify PEH in Wales [22, 25], the SMDS contributed the greatest number of potential homeless people to the analysis, with 73% of the initial number of unique individuals (n = 2,943) being present in the SMDS.

Creating the 'not-homeless' comparator groups

An initial comparison group included all people in the resident population spine who were not flagged as having experienced homelessness. However, consistent with what is known about homeless populations, there were significant differences between the homeless and not-homeless groups in terms of age, gender, and socio-economic status—discussed in detail in the findings section. We therefore adopted further methods of creating a comparison group that had similar characteristics as the homeless group. However, there were a limited number of characteristics that may have been important in positivity rates and were available in administrative data.

There is evidence to suggest that SARS-CoV-2 infection is higher amongst men, those from older age groups, and those living in more deprived areas [28–31]. Specific to Wales, there were also regional disparities in SARS-CoV-2 infection due to localised outbreaks which may have driven increases in positivity rates and increased testing through community testing drives. Though there is also consistent evidence for differences in SARS-CoV-2 infection and mortality by ethnicity [31], ethnicity data were not consistently available across all of the data sets provided and was therefore not included.

An initial matched comparator group retained people in the not-homeless population with the exact same characteristics as the homeless group, based on age, gender, local authority of residence and 2019 Welsh Index of Multiple Deprivation (WIMD) quartile. This group are referred to as the 'exact matched' comparator. We then adopted Propensity Score Matching (PSM) to create our final comparator group. This approach used a logistic regression predicting whether a person had experienced homelessness

Table 1: Variables and codes used to identify homelessness events, and initial number of people identified within each data set as experiencing homelessness (% of total number of unique people, N = 4,049)

Data set	Variables	Codes	Code description	n (%)
Emergency Department Data Set (EDDS)	Diagnosis Type code 1 to 6	31Z	Social Problems/Homelessness, other or unspecified	205 (5.1)
Patient Episode Database for Wales (PEDW)	Diagnosis code	Z590	Homelessness	366 (9.0)
		Z591	Inadequate housing	
		Z592	Discord with neighbours, lodgers, and landlord	
		Z593	Problems related to living in residential institution	
		Z598	Other problems related to housing and economic circumstances	
		Z599	Problem related to housing and economic circumstances, unspecified	
Primary Care GP data (GP)	Event code	13D..	Homeless/housing lack	706 (17.4)
		13D1.	Homeless family	
		13D2.	Homeless single person	
		13D3.	Tramp/vagabond	
		13D5.	Vagrant	
		13D7.	Sofa surfer - person of no fixed abode	
		13D8.	Length of time homeless	
		13DZ.	Housing lack NOS	
		13F40	Lives in sheltered housing	
		13F9.	Living in hostel	
		13FA.	Living in B&B accommodation	
		13FB.	Lives in lodgings	
		13FL.	Living rough	
		13FM.	Sleeping in night shelter	
		13FW.	Living in temporary housing	
		9k6..	Homeless - enhanced services administration	
		9k60.	Homeless - enhanced service completed	
		9Ngr.	Under care of homeless advocacy service	
		ZV600	Lack of housing/tramp	
Substance Misuse Data Set (SMDS)	Accommodation need (Main assessment)	01	No Fixed Abode - urgent housing problem	2,943 (72.7)
		02	Housing problem	
	Urgent housing problem status (TOP)	01	Client has urgent housing problem in the last 4 weeks	

TOP = Treatment Outcome Profile.

to generate propensity scores as an indicator of similarity between people. The pre-specified 'caliper' was set at 0.1, which is the maximum tolerated difference in propensity scores accepted as matching. PSM matching was based on the same variables used in the exact matching process. Furthermore, given there were many possible matches to choose from in the not-homeless population, we extracted five people from the general not-homeless population for every homeless person. The r package MatchIt [32] was used to undertake PSM.

To avoid matching to people who were likely to be of similar socio-economic characteristics to PEH by virtue of living at the same residence, but whose homelessness status could not be assumed, we excluded people living in the same RALF as a PEH on the 1st March 2020.

Analysis

We initially report the basic characteristics of the homeless and not-homeless comparison groups in terms of age (years), gender, local authority of residence on the 1st March 2020, 2019 WIMD quartile for the Lower Supper Output Area of their residence, and proportion of people who had been tested for SARS-CoV-2 at least once during the study period. Percentages and means, with standard deviations, are provided where appropriate. Standardised differences are reported to provide an indication of the difference in distribution of data between the homeless and each not-homeless comparator group separately [33]. Standardised differences with a value of 0.1 or more are taken to indicate imbalance between the homeless and not-homeless comparator [34].

Prevalence rates for testing positive for SARS-CoV-2 were calculated as the total number of people having a positive PCR or antibody test from 1st March 2020 up to and including 1st March 2021, divided by the total number of people within each group. Confidence intervals for point estimates of prevalence were calculated using the Agresti-Coull method, as appropriate for binomial data and larger sample sizes [35]. SARS-CoV-2 prevalence was calculated for the homeless, three not-homeless comparator groups, and the total resident population within the population spine.

Logistic regressions were run to explore the effects of having recently experienced homelessness on the probability of having a positive SARS-CoV-2 test. Three regressions were run including the different comparison groups—population, exact matched, and PSM. We adjusted for age, gender, local authority of residence and 2019 WIMD quartile to control for any remaining confounding from the matching variables. The odds ratios, 95% confidence intervals, and indicators of significance levels are provided for the characteristics included in the regressions—though the main interest is in the odds ratios for experiencing homelessness. As the regressions using the population comparator and the exact match comparator are 'population level', consideration of probability values is of lesser importance—the odds ratios reflect actual outcomes of the population rather than samples intended to represent a population.

Results

Of the initial 4,049 PEH during the study period, 3,153 had an entry in the population spine. Table 2 summarises the characteristics of the homeless and three not-homeless comparator groups. The homeless group were similar to PEH previously identified using administrative data sources [22–25], whilst conforming to known profiles of PEH from surveys conducted in the UK [36]. PEH were disproportionately male (67.9%), in their 30s to 40s (mean age of 38.5 years old ±14.0 years), and more concentrated in urban than rural local authorities—with Cardiff, Swansea and Newport representing the major urban authorities in Wales. There was a large imbalance (Std. diff >0.1) between the homeless and not-homeless population comparator in terms of area-level deprivation, as measured by the 2019 WIMD quartile. Homeless people were almost twice as likely to come from the most deprived areas than the not-homeless population (49.6% compared to 25.6%). Subsequent matching exercises reduced the imbalance between the homeless and not-homeless groups, as indicated by the reduction in the standardised differences for each of the main variables of interest in both the exact match and PSM cohorts.

For the total resident population in the population spine, the prevalence of SARS-CoV-2 was 5.6%, with 167,540 people testing positive by 1st March 2021 (Table 3). Published statistics, based on PCR testing data only, indicate that 193,415 people resident in a health board in Wales had tested positive for SARS-CoV-2 between 1st March 2020 and 1st March 2021 [37]. This difference in testing numbers between our cohort and published figures may largely be due to the removal of records within the SARS-CoV-2 testing data during

data processing—we discuss the implications of this in the strengths and limitations section.

The incidence of testing for SARS-CoV-2 was 43.6% for the homeless group and 30.0%, 33.3%, and 33.0% for the population, exact, and PSM match comparator groups, respectively. The prevalence of SARS-CoV-2 amongst the homeless group identified through non-housing administrative data was 5.0% (95% C.I.: 4.3-5.9). For the not-homeless population comparator group, the SARS-CoV-2 rate was 5.6% (95% C.I.:5.6-5.6). Prevalence of SARS-CoV-2 amongst the exact matched not-homeless comparator group was 6.9% (95% C.I: 6.9-7.0), and under PSM matching was 6.7% (95% C.I: 6.4-7.1).

Having recently experienced homelessness was found to be associated with a lower odds of testing positive for SARS-CoV-2 by 1st March 2021 in all three logistic regression models (Table 4). In the model based on population level data, having a homeless event during the study period was associated with a roughly 10% reduced odds of having a positive test result compared to not having a homeless event (O.R.: 0.90, 95% C.I.: 0.77–1.06, p = 0.209).

When using matched comparator groups, the effect of having recently experienced homelessness on testing positive for SARS-CoV-2 was lower than in the population level regression. There was a 25% reduction in the odds of testing positive if a person had experienced homelessness in the regression using an exact match not-homeless group (O.R.: 0.75, 95% C.I.: 0.64-0.88, p <0.001) and 27% for the PSM match group (O.R.: 0.73, 95% C.I.: 0.62-0.87, p <0.001f). This implies that people who were homeless showed an even lower odds of testing positive compared to people of a similar age, sex, area deprivation, and geographic breakdown as the homeless group—despite the higher incidence of testing amongst PEH.

Discussion

Our study found that the prevalence of SARS-CoV-2 amongst PEH was lower than that for the not-homeless population, 5.0% compared to 5.6%, along with a reduced risk of testing positive (O.R. 0.90). This finding runs counter to other studies that compared prevalence rates to the general population, which found that PEH were more likely to have a positive SARS-CoV-2 test result [10, 11]. Our contrasting findings may be attributable to the PEH in our analysis which potentially encompassed people who were roofless and in insecure housing situations. This definition is much broader than houseless PEH using shelters which dominates the literature. The inclusion of different groups of PEH, who may experience differing levels of risk of exposure to SARS-CoV-2 than those in shelters, may have diluted the prevalence rate, compared to if we had focused only on houseless PEH. The much lower use of communal forms of accommodation in Wales during the COVID-19 pandemic compared to other countries may also have contributed to a lower prevalence of SARS-CoV-2 amongst PEH in Wales.

There is a large body of evidence highlighting how communal forms of homelessness accommodation can lead to transmission of communicable diseases such as H1N1/swine flu [2], MRSA [38] and tuberculosis [39]. Furthermore, a recent

Table 2: Cohort characteristics

| | | Homeless | Not-homeless comparison group | | |
| | | | Population | Exact match[¥] | PSM[†] |
		n (% or ± S.D.)	n (% or ± S.D.)	n (% or ± S.D.)	n (% or ± S.D.)
Age, mean years		38.5 (±14.0)	42.4 (±23.8)	38.7 (±14.3)	38.5 (±14.0)
Standardised difference[§]		–	*0.198*	*0.017*	*−0.000*
Gender	Male	2,142 (67.9)	1,490,090 (49.9)	378,843 (63.4)	10,710 (67.9)
	Female	1,011 (32.1)	1,498,124 (50.1)	218,917 (36.6)	5,055 (32.1)
Standardised difference[§]		–	*0.374*	*0.096*	*0.000*
WIMD	Most deprived	1,565 (49.6)	764,617 (25.6)	275,423 (46.1)	7,825 (49.6)
	2nd	818 (25.9)	741,497 (24.8)	146,907 (24.6)	4,090 (25.9)
	3rd	476 (15.1)	741,564 (24.8)	94,702 (15.8)	2,380 (15.1)
	Least deprived	294 (9.3)	740,536 (24.8)	80,728 (13.5)	1,470 (9.3)
Standardised difference[§]		–	*0.611*	*0.138*	*0.000*
Local Authority	Blaenau Gwent	79 (2.5)	67,467 (2.3)	11,432 (1.9)	395 (2.5)
	Bridgend	104 (3.3)	138,646 (4.6)	18,543 (3.1)	520 (3.3)
	Caerphilly	244 (7.7)	174,908 (5.9)	46,543 (7.8)	1,220 (7.7)
	Cardiff	474 (15)	352,544 (11.8)	134,246 (22.5)	2,370 (15.0)
	Carmarthenshire	121 (3.8)	17,8014 (6.0)	27,004 (4.5)	605 (3.8)
	Ceredigion	46 (1.5)	65,576 (2.2)	5,070 (0.8)	230 (1.5)
	Conwy	90 (2.9)	109,973 (3.7)	12,170 (2.0)	450 (2.9)
	Denbighshire	89 (2.8)	92,979 (3.1)	98,71 (1.7)	445 (2.8)
	Flintshire	144 (4.6)	149,777 (5.0)	27,302 (4.6)	722 (4.6)
	Gwynedd	150 (4.8)	101,899 (3.4)	21,305 (3.6)	750 (4.8)
	Isle of Anglesey	72 (2.3)	62,975 (2.1)	6,383 (1.1)	360 (2.3)
	Merthyr Tydfil	73 (2.3)	60,232 (2.0)	8,038 (1.3)	365 (2.3)
	Monmouthshire	58 (1.8)	85,632 (2.9)	7,928 (1.3)	290 (1.8)
	Neath Port Talbot	117 (3.7)	139,090 (4.7)	23,151 (3.9)	585 (3.7)
	Newport	259 (8.2)	150,626 (5.0)	38,237 (6.4)	1,295 (8.2)
	Pembrokeshire	76 (2.4)	113,189 (3.8)	10,867 (1.8)	380 (2.4)
	Powys	57 (1.8)	118,433 (4.0)	7,613 (1.3)	285 (1.8)
	RCT	192 (6.1)	235,869 (7.9)	52,429 (8.8)	960 (6.1)
	Swansea	313 (9.9)	237,697 (8.0)	65,893 (11.0)	1,565 (9.9)
	Torfaen	107 (3.4)	91,994 (3.1)	14,301 (2.4)	535 (3.4)
	Vale of Glamorgan	133 (4.2)	126,996 (4.2)	24,793 (4.1)	663 (4.2)
	Wrexham	155 (4.9)	133,698 (4.5)	24,641 (4.1)	775 (4.9)
Standardised difference[§]		–	*0.306*	*0.300*	*0.001*
Tested for SARS-CoV-2 during study period		1,375 (43.6)	895,703 (33.0)	199,344 (33.3)	5,206 (33.0)
Base (N=)		3,153	2,988,214	597,760	15,765

WIMD = Welsh Index of Multiple Deprivation; RCT = Rhondda Cynon Taff; S.D. = Standard Deviation.
[¥]Exact match based on age (years), gender, local authority, and WIMD quartile.
[†]Using propensity scores. Ratio of intervention to control group 1:5, with a calliper for obtaining a match set at 0.1.
[§]Standardised difference compares distribution of data between homeless and not homeless group. Values >.1 indicate that there are differences between groups.

systematic review concluded that use of accommodation that satisfies basic needs such as a bed and food—as is the case with shelters—can do more harm than no intervention at all [40]. The Welsh policy response centring on the widespread commissioning of self-contained or single-room provision and the closure of communal forms of accommodation, may have had a protective effect in limiting the spread of SARS-CoV-2. As the majority of the PEH in our study were identified in the SMDS, and were therefore largely roofless, houseless, or in insecure accommodation, then we have a greater level of certainty that they would be beneficiaries of these policy changes. However, the role of communal accommodation alone is unlikely to explain lower prevalence rates, and does not explain the increase in this difference when comparing to not-homeless comparators with similar characteristics—the exact match and PSM groups. Here we offer two hypotheses that

Table 3: Summary of SARS-CoV-2 infection by 1st March 2021; counts and point estimate of prevalence (95% confidence interval)

| | | SARS-CoV-2 infection | | Total |
		No	Yes	
Not-homeless comparator	Population	2,820,833	167,381	2,988,214
		94.4% (94.4–94.4)	5.6% (5.6–5.6)	
	Exact match	556,388	41,372	597,760
		93.1% (93.0–93.1)	6.9% (6.9–7.0)	
	PSM	14,704	1,061	15,765
		93.3% (92.9–93.7)	6.7% (6.4–7.1)	
Homeless		2,994	159	3,153
		95.0% (94.1–95.7)	5.0% (4.3–5.9)	
Total resident population		2,823,827	167,540	2,991,367
		94.4% (94.4–94.4)	5.6% (5.6–5.6)	

Table 4: Effect of characteristics on the odds of SARS-CoV-2 infection, by not-homeless control group included in logistic regression model

		Population	Exact match	PSM
Homeless (Ref. cat.: Not homeless)		0.90 (0.77–1.06)	**0.75 (0.64–0.88)**	**0.73 (0.62–0.87)**
Age, years		**1.00 (1.00–1.00)**	**1.00 (0.99–1.00)**	1.00 (0.99–1.00)
Female (Ref. cat.: Male)		**1.28 (1.27–1.29)**	**1.42 (1.39–1.45)**	**1.41 (1.25–1.59)**
Local authority (Ref. cat.: Cardiff)	Blaenau Gwent	**1.39 (1.35–1.44)**	**1.41 (1.31–1.50)**	1.22 (0.84–1.77)
	Bridgend	**1.33 (1.30–1.37)**	**1.40 (1.33–1.48)**	*1.49 (1.09–2.04)*
	Caerphilly	**1.17 (1.15–1.20)**	**1.20 (1.16–1.25)**	1.17 (0.92–1.50)
	Carmarthenshire	**0.82 (0.80–0.84)**	**0.91 (0.86–0.96)**	0.96 (0.69–1.35)
	Ceredigion	**0.35 (0.33–0.37)**	**0.33 (0.27–0.40)**	0.60 (0.32–1.11)
	Conwy	**0.46 (0.44–0.47)**	**0.47 (0.42–0.52)**	**0.46 (0.28–0.77)**
	Denbighshire	**0.58 (0.56–0.61)**	**0.61 (0.55–0.67)**	*0.56 (0.35–0.90)*
	Flintshire	**0.89 (0.87–0.92)**	**0.84 (0.80–0.89)**	0.84 (0.60–1.17)
	Gwynedd	**0.34 (0.32–0.35)**	**0.33 (0.30–0.36)**	**0.30 (0.19–0.49)**
	Isle of Anglesey	**0.41 (0.39–0.43)**	**0.44 (0.39–0.51)**	*0.49 (0.29–0.85)*
	Merthyr Tydfil	**1.74 (1.69–1.79)**	**1.75 (1.63–1.88)**	1.29 (0.88–1.88)
	Monmouthshire	**0.69 (0.67–0.72)**	**0.69 (0.62–0.76)**	0.58 (0.33–1.02)
	Neath Port Talbot	**1.20 (1.17–1.23)**	**1.27 (1.21–1.34)**	*1.41 (1.05–1.90)*
	Newport	**1.07 (1.04–1.10)**	**1.09 (1.04–1.13)**	1.21 (0.94–1.54)
	Pembrokeshire	**0.39 (0.37–0.41)**	**0.42 (0.37–0.47)**	**0.39 (0.22–0.71)**
	Powys	**0.43 (0.41–0.45)**	**0.53 (0.47–0.60)**	*0.55 (0.31–0.98)*
	RCT	**1.42 (1.39–1.45)**	**1.45 (1.40–1.50)**	1.28 (0.99–1.66)
	Swansea	**1.07 (1.05–1.09)**	**1.08 (1.04–1.12)**	1.25 (1.00–1.56)
	Torfaen	**1.04 (1.01–1.07)**	*1.09 (1.02–1.16)*	1.05 (0.75–1.48)
	Vale of Glamorgan	**0.85 (0.83–0.88)**	**0.84 (0.79–0.89)**	0.89 (0.64–1.24)
	Wrexham	**1.17 (1.14–1.20)**	**1.21 (1.15–1.27)**	1.25 (0.94–1.66)
WIMD quartile (Ref. cat.: Least deprived)	Most deprived	**1.06 (1.05–1.08)**	1.02 (0.99–1.06)	1.00 (0.79–1.25)
	2nd	**1.08 (1.07–1.10)**	1.04 (1.00–1.08)	1.16 (0.92–1.48)
	3rd	*1.02 (1.00–1.03)*	1.03 (0.99–1.08)	1.11 (0.85–1.44)

RCT = Rhondda Cynon Taff; WIMD = Welsh Index of Multiple Deprivation.
Bold text = Significant at 1% level; *Italicised text* = Significant at 5%; Greyed text = Not significant.

relate to the isolation of homeless people and homelessness as a lack of privacy and control.

PEH can experience diminished social networks, largely the result of withdrawing or being pushed away from friends and family members [41–44]. This isolation may have acted to reduce the number of social contacts and interactions amongst PEH, thereby slowing the transmission of SARS-CoV-2 amongst this group to a greater extent than people who were not-homeless. In addition to reduced social connections, homelessness can be conceptualised in part as a lack of privacy and choice to engage in social relations and for people to conduct their lives as they see fit [6]. PEH are often subject to

a higher degree of surveillance and control over their private lives than not-homeless people, by virtue of having to interact with systems to meet their most basic needs such as shelter and food [45]. During the COVID-19 pandemic, this control manifested in a concerted effort by local authorities and third sector organisations to enforce and encourage PEH to conform to physical distancing requirements [5]. This may have led to a greater (forced) adherence to COVID-19 guidance amongst PEH, when compared to the not-homeless population, to continue to be permitted access to accommodation and basic necessities.

Study benefits and limitations

The use of non-housing data has meant that a national cohort of PEH could be identified, thereby compensating for the lack of a national individual level homelessness data collection in Wales. However, there are limitations to our analysis arising from the construction of the cohort of PEH. Homelessness was identified through data related to the recording of *events*, rather than periods of homelessness. Some of the people identified as PEH may therefore have been housed at the time they were infected. Use of data from health and substance use services to identify PEH may also mean that the homeless population described in this analysis have complex (health) needs and experiences. PEH identified through health data may have been accessing these services for SARS-CoV-2 related reasons. Furthermore, as the majority of PEH were identified in the SMDS, they presumably also had some level of substance use issue. Combined, these factors may have led to an inflation of prevalence rates amongst the PEH included in this study compared to PEH identified through housing services data, for example.

As the majority of PEH were present in the SMDS, PEH in our analysis were therefore likely to be predominantly people who were at least initially roofless, houseless, and 'sofa surfers' living in insecure accommodation—as these are the groups defined as homeless/experiencing housing issues in the SMDS [26]. Though not as broad a definition of homelessness as possible [6, 26], the PEH in our study combine several groups not included in predominantly shelter-based studies of SARS-CoV-2 prevalence amongst homeless populations. Due to the generalised nature of some of the homelessness event codes, such as 'Z590 – Homelessness', we were unable to drill down into different sub-groups of PEH, who may have had different prevalence rates.

The construction of the population spine using GP registration data meant that people not registered with a GP on 1st March 2020 were excluded from our analysis. This had the effect of reducing the number of PEH from 4,049, to 3,153 people. Missingness from the population spine may be because people have moved into Wales after the 1st March 2020, or they may have moved into Wales prior to 1st March 2020 and not de-registered from their GP elsewhere in the UK, or their GP may have cancelled their registration—meaning that the person has no entry for the snapshot date [46]. Alternatively, not being registered with a GP may be a sign of barriers to accessing public services. Given the range of reasons for missingness, we cannot adequately hypothesise whether presence/absence of people

in the population spine is associated with risk of SARS-CoV-2 infection, or the potential direction of any bias from this missingness.

Though we have been able to provide an estimate for the prevalence of SARS-CoV-2 amongst PEH, our outcome measure is based on administrative testing data and therefore largely reflects testing practices. In the initial peak in SARS-CoV-2 infection in the UK, testing focused on symptomatic people who were hospitalised, and key health and social care workers, with increasing community testing in response to local outbreaks in later phases of the pandemic. There will therefore be an under-estimation of the 'true' prevalence of infection from earlier in the pandemic where people were symptomatic but not hospitalised. However, the prevalence of SARS-CoV-2 infection will be lower than the actual rates as administrative data, by definition, will not contain asymptomatic-untested people.

Related to national testing practices, previous research has found that PEH can face barriers to accessing healthcare, for example due to the need for an address when accessing primary care [47]. These barriers may have led to reduced access to SARS-CoV-2 testing amongst PEH, and as a result potentially lower prevalence of SARS-CoV-2 than might have been seen without barriers to services—or at least equal barriers to those faced by the housed population. However, we observed that PEH in our study had higher rates of testing, and lower SARS-CoV-2 prevalence. The fact that our population of PEH were identified through healthcare and substance use services may have biased our estimate towards people able to access services, and therefore who were more able to access SARS-CoV-2 testing.

A final limitation comes from the quality of linkage for the SARS-CoV-2 pathology testing data, where 145,539 testing records were unavailable for linkage, of which 13,532 (9%) were positive for SARS-CoV-2 infection. The reduction in testing records potentially led to a reduced ability to detect SARS-CoV-2 infection. Unfortunately, removing these records was unavoidable where there was no ALF assigned—as linkage was not possible—and was necessary where the ALF match rate was low (<90%) as retaining these records may have led to increased false matches. As this missingness is likely to affect other studies using SARS-CoV-2 testing data with hard to reach (small) populations, then further exploration is required to understand the exact nature of the lack of ALF matches.

Conclusion

This study provides new insights and adds to a nascent international evidence base on the prevalence of SARS-CoV-2 amongst PEH in the UK. Its findings have important policy implications. Globally, the message is clear: the avoidance of communal living spaces will help reduce transmission of SARS-CoV-2 and its possible future variants, and help control other infectious disease outbreaks. Taken together with evidence on the wider harms of such accommodation, a rapid shift away from the use of hostels and shelters in our response to homelessness is called for. In Wales there is a commitment not to return to such provision and it will be important to monitor the impacts of such an important policy shift.

Acknowledgements

Conflict of interests

The authors declare that they do not have any conflict of interest.

Ethics statement

This study was reviewed by the independent Information Governance Review Panel (IGRP) of the SAIL Databank and approved under ID:0937. This study has Cardiff University School of Social Science Ethical Committee approval as part of a broader package of working on homelessness and health being conducted by Administrative Data Research Wales.

References

1. Tsai J, Wilson J. COVID-19: a potential public health problem for homeless populations. The Lancet Public Health. 2020;5(5):e186–e187. https://doi.org/10.1016/S2468-2667(20)30053-0

2. Baral S, Bond A, Boozary A, Bruketa E, Elmi N, Freiheit D et al. Seeking shelter: homelessness and COVID-19. FACETS. 2021;6:925–958. https://doi.org/10.1139/facets-2021-0004

3. Parsell C, Clarke A, Kuskoff E. Understanding responses to homelessness during COVID-19: an examination of Australia. Housing Studies. 2020; Online first:1–14. https://doi.org/10.1080/02673037.2020.1829564

4. Pleace N, Baptista I, Benjaminsen L, Bush-Geertsema V, O'Sullivan E, Teller N. European Homelessness and COVID-19. Accessed: 2021-08-03. https://www.feantsa.org/en/other-observatory-publications/2021/04/01/european-homelessness-and-covid-19

5. Fitzpatrick S, Mackie P, Pawson H, Watts B, Wood J. The COVID-19 Crisis Response to Homelessness in Great Britain. Accessed: 2021-08-03. https://housingevidence.ac.uk/publications/the-covid-19-crisis-response-to-homelessness-in-great-britain/

6. Amore K, Baker M, Howden-Chapmen P. The ETHOS Definition and Classification of Homelessness: An Analysis. European Journal of Homelessness. 2011;5(2):19-37. No doi available

7. Mohsenpour A, Bozorgmehr K, Rohleder S, Stratil J, Costa D. SARS-CoV-2 prevalence, transmission, health-related outcomes and control strategies in homeless shelters: Systematic review and meta-analysis. EClinicalMedicine. 2021;Online:1–20. https://doi.org/10.1016/j.eclinm.2021.101032

8. Yoon JC, Montgomery MP, Buff AM, Boyd AT, Jamison C, Hernandez A, et al. COVID-19 Prevalence among people experiencing homelessness and homelessness service staff during early community transmission in Atlanta, Georgia, April-May. Clinical Infectious Diseases. 2020;Online first:1–7. https://doi.org/10.1093/cid/ciaa1340

9. Roederer T, Mollo B, Vincent C, Nikolay B, Llosa AE, Nesbitt R, et al. Seroprevalence and risk factors of exposure to COVID-19 in homeless people in Paris, France: a cross-sectional study. The Lancet Public Health. 2021;6(4):e202–209. https://doi.org/10.1016/S2468-2667(21)00001-3

10. Röthlin Eriksen AR, Fogh K, Hasselbalch RB, Bundgaard H, Nielson SD, Jorgense CS, et al. SARS-CoV-2 antibody prevalence among homeless people, sex workers and shelter workers in Denmark: a nationwide cross-sectional study. 2021;Preprint:1–26 https://doi.org/10.1101/2021.05.07.21256388

11. Richard L, Booth R, Rayner J, Clemens KK, Forchuck C, Shariff SZ. Testing, infection and complication rates of COVID-19 among people with a recent history of homelessness in Ontario, Canada: a retrospective cohort study. CMAJ Open. 2021;9(1):E1–E9. https://doi.org/10.9778/cmajo.20200287

12. Lewer D, Braithwaite I, Bullock M, Eyre MT, White PJ, Aldridge RW, Story A, Hayward AC. COVID-19 among people experiencing homelessness in England: a modelling study. Lancet Respir Med. 2020;8(12):1181–91. https://doi.org/10.1016/S2213-2600(20)30396-9

13. Mackie PK. Homelessness prevention and the Welsh legal duty: lessons for international policies. Housing Studies. 2015;30(1):40–59. https://doi.org/10.1080/02673037.2014.927055

14. Mackie PK, Thomas I, Bibbings J. Homelessness prevention: reflecting on a year of pioneering Welsh legislation in practice. European Journal of Homelessness. 2017;11(1):81–107.

15. Welsh Government. Phase 2 – Planning Guidance for Homelessness and Housing Related Support Services. Accessed: 2021-08-05. https://gov.wales/sites/default/files/publications/2020-06/planning-guidance-for-homelessness-and-housing-support-services-coronavirus.pdf

16. Welsh Government. Guidance for local authorities in supporting people sleeping rough – Covid-19 outbreak. Accessed: 2021-10-19. https://gov.wales/sites/default/files/publications/2020-04/guidance-for-local-authorities-in-supporting-people-sleeping-rough-covid-19_0.pdf

17. Jones KH, Ford DV, Jones C, Dsilva R, Thompson S, Brooks CJ, et al. A case study of the Secure Anonymous Information Linkage (SAIL) Gateway: A privacy-protecting remote access system for health-related research and evaluation. Journal of Biomedical Informatics. 2014;50:196–204. https://doi.org/10.1016/j.jbi.2014.01.003

18. Welsh Parliament. Coronavirus timeline: Welsh and UK governments' response. Accessed: 2021-08-03. https://research.senedd.wales/research-articles/coronavirus-timeline-welsh-and-uk-governments-response/

19. Mistry DA, Wang JY, Moeser M, Starkey T, Lee LYW. A systematic review of the sensitivity and specificity of lateral flow devices in the detection of SARS-CoV-2. BMC Infectious Diseases. 2021;21:1–14. https://doi.org/10.1186/s12879-021-06528-3

20. Mytton OT, McCarthy N, Watson J, Whiting P. Interpreting a lateral flow SARS-CoV-2 antigen test. BMJ. 2021;373:1–5. https://doi.org/10.1136/bmj.n1411

21. Welsh Government. Lateral flow testing for people without symptoms Accessed: 2021-08-03. https://gov.wales/lateral-flow-testing-people-without-symptoms

22. Thomas I, Mackie P. Using data linkage to improve estimates of homelessness. Accessed: 2021-11-22. https://www.adruk.org/fileadmin/uploads/adruk/Documents/Data_Insights_Using_data_linkage_to_improve_estimates_of_homelessness_February_2021.pdf

23. Richard L, Hwang SW, Forchuck C, Nisenbaum R, Clemens KK, Wiens K, Booth R, Azimaee M, Shariff SZ. Validation study of health administrative data algorithms to identify individuals experiencing homelessness and estimate population prevalence of homelessness in Ontario, Canada. BMJ Open;9(10):1–11. https://doi.org/10.1136/bmjopen-2019-030221

24. Kaushal R, Jagpal P, Khanal S, Vohra N, Lowrie R, Johal J, et al. Representation of persons experiencing homelessness and coding of homelessness in general practices: descriptive evaluation using healthcare utilisation data. BJGP Open. 2021;No volume number:1-9. https://doi.org/10.3399/BJGPO.2021.0050

25. Song J, Grey CNB, Davies AR. Creating an e-cohort of individuals with lived experience of homelessness and subsequent mortality in Wales, UK. Journal of Public Health. 2021;Online:1–4. https://doi.org/10.1093/pubmed/fdab180

26. Busch-Geerstema V, Culhane D, Fitzpatrick S. Developing a Global Framework for Conceptualising and Measuring Homelessness. Habitat International. 2016;55:124–132. https://doi.org/10.1016/j.habitatint.2016.03.004

27. Digital Health and Care Wales. Substance Misuse Documentation. Accessed: 2021-08-03. https://dhcw. nhs.wales/information-services/information-and-statistics/substance-misuse/substance-misuse-documentation/

28. Abate BB. Kassie AM, Kassaw MW, Aragie TG, Masresha SA. Sex differences in coronavirus disease (COVID-19): a systematic review and meta-analysis. BMJ Open. 2020;10(e040129):1–10. http://doi.org/10.1136/bmjopen-2020-040129

29. Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JIR, Franssen GHL, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. BMJ Open. 2021;11(e044640):1–10. http://doi.org/10.1136/bmjopen-2020-044640

30. Public Health England. Disparities in the risk and outcomes of COVID-19. Accessed: 2021-08-05. https://www.gov.uk/government/publications/covid-19-review-of-disparities-in-risks-and-outcomes

31. Niedzwiedz CL, O'Donnell CA, Jani BD, Demou E, Ho FK, Celis-Morales C, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. BMC Medicine. 2020;18:1–14. https://doi.org/10.1186/s12916-020-01640-8

32. Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software. 2011;42(8):1–28. http://doi.org/10.18637/jss.v042.i08

33. Zhang Z, Kim HJ, Lonjon G, Zhu Y. Balance diagnostics after propensity score matching. Annals of Translational Medicine. 2019;7(1):1–11. https://doi.org/10.21037/atm.2018.12.10

34. Austin P. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. Communications in Statistics – Simulation and computation. 2009;38(6): 1228–1234. https://doi.org/10.1080/03610910902859574

35. Brown LD, Cai TT, DasGupta A. Interval Estimation for Binomial Proportion. Statistical Science. 2001;16(2):101–133. https://doi.org/10.1214/ss/1009213286

36. Fitzpatrick S, Bramley G, Johnsen S. Pathways into multiple exclusion homelessness in seven UK cities. Urban Studies. 2013;50(1):148–168. https://doi.org/10.1177%2F0042098012452329

37. Public Health Wales. Rapid COVID-19 virology Dashboard. Accessed: 2021-08-03 https://public.tableau.com/app/profile/public.health.wales.health.protection/viz/RapidCOVID-19virology-Public/Headlinesummary

38. Leibler JH, Liebschutz JM, Keosaian J, Stewart C, Monteiro J, Woodruff A, et al. Homelessness, Personal Hygiene, and MRSA Nasal Colonization among Persons Who Inject Drugs. Journal of Urban Health. 2019;96(5):734–740. https://doi.org/10.1007/s11524-019-00379-9

39. Beijer U, Wolf A, Fazel S. Prevelance of turberculosis, hepatitis C virus, and HIV in homeless people: a systematic review and meta-analysis. The Lancet Infectious Diseases. 2012;12(11):859–870. https://doi.org/10.1016/S1473-3099(12)70177-9

40. Keenan C, Miller S, Hanratty J, Pigott T, Hamilton J, Coughlan C, et al. Accommodation-based interventions for individuals experiencing, or at risk of experiencing, homelessness. Campbell Systematic Reviews. 2021;17(2):1–93. https://doi.org/10.1002/cl2.1165

41. Hawkins RL, Abrams C. Disappearing acts: the social networks of formerly homeless individuals with co-occurring disorder. Social Science & Medicine. 2007;65:2031–2042. https://doi.org/10.1016/j.socscimed.2007.06.019

42. Joly L, Cornes M, Manthorpe J. Supporting the social networks of homeless people. Housing, Care and Support. 2014;17(4):198-207. https://doi.org/10.1108/HCS-06-2014-0014

43. Sanders B, Brown B. 'I was all on my own': experiences of loneliness and isolation amongst homeless people. Accessed: 2021-08-05. https://www.crisis.org.uk/ending-homelessness/homelessness-knowledge-hub/health-and-wellbeing/i-was-all-on-my-own-2015/

44. Neale J, Stevenson C. Social and recovery capital amongst homeless hostel residents who use drugs and alcohol. International Journal of Drug Policy. 2014;26(5):475–483. https://doi.org/10.1016/j.drugpo.2014.09.012

45. Johnsen S, Fitzpatrick S, Watts B. Homelessness and social control: a typology. Housing Studies. 2018;33(7):1106–1126. https://doi.org/10.1080/02673037.2017.1421912

46. Smallwood, S., Lynch, K. An analysis of patient register data in the Longitudinal Study – what does it tell us about the quality of the data? Accessed: 2021-11-09. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/patientregisterqualityassuranceofadministrativedatausedinpopulationstatisticsdec2016

47. O'Carroll A, Wainwright D. Making sense of street chaos: ethnographic exploration of homeless people's health service utilisation. International Journal for Equity in Health. 2019;18(113):1–22. https://doi.org/10.1186/s12939-019-1002-6

Abbreviations

ALF:	Anonymised Linkage Field
EDDS:	Emergency Department Data Set
GP:	General Practitioner
PCR:	Polymerase Chain Reaction
PEDW:	Patient Episode Database for Wales
PEH:	People experiencing homelessness
PSM:	Propensity Score Matching
RALF:	Residential Anonymised Linkage Field
SAIL:	Secure Anonymised Information Linkage
SMDS:	Substance Misuse Data Set Wales
WDS:	Welsh Demographic Service
WIMD:	Welsh Index of Multiple Deprivation