International Journal of Population Data Science

Journal Website: www.ijpds.org

Exploring Loneliness and Living Alone Through Linked Administrative and Survey Data in Wales, UK

Rhodri D Johnson[1,*], Stuart Bedston[1], Rowena Bailey[1], Richard Fry[1], and Lucy J Griffiths[1]

Submission History

Submitted:	18/02/2026
Accepted:	22/08/2026
Published:	10/11/2026

[1] Population Data Science, Swansea University Medical School, Swansea, SA2 8PP

Abstract

Background

Loneliness is a public health priority, yet population-scale evidence that distinguishes types of loneliness and links with living alone circumstances is scarce. Population-wide linked administrative data within the Secure Anonymised Information Linkage (SAIL) Databank provide an opportunity to investigate these relationships across Wales.

Objectives

(1) Describe long-term living alone trends; (2) develop indicators for reason for living alone; and (3) examine associations between living alone (including derived indicators) and emotional, social and overall loneliness, alongside demographic, health and place-based factors.

Methods

We conducted a two-phase observational study in Wales. Phase 1 used annual, population level snapshots (2008–2024) to quantify living alone trends and correlates. Phase 2 linked individual level data in SAIL to National Survey for Wales loneliness data (De Jong Gierveld 6-item) to model adjusted associations. We inferred living alone reasons (partner loss, separation, unassigned) from life events data. Models adjusted for age and gender interaction, health, tenure, housing type, deprivation, and built environment indicators.

Results

Living alone households increased from 27.0% (314,806) in 2008 to 29.6% (363,133) in 2024. In adjusted models, living alone was strongly associated with loneliness risk, especially emotional loneliness. These associations were modified by reason for living alone, with the largest effects observed following recent partner loss. Younger adults reported highest overall loneliness; men showed higher social loneliness around midlife; women showed higher emotional loneliness. Built environment measures showed little evidence of independent association after adjustment.

Conclusions

Population-scale linkage can generate robust, policy-relevant indicators of living alone dynamics. Such indicators support monitoring and targeting within health and social care, complementing national statistics and enabling future linked data studies.

Keywords

Data linkage, loneliness, living alone, administrative data, survey data, built environment, population health

*Corresponding Author:
Email Address: r.d.johnson@swansea.ac.uk (Rhodri D Johnson)

https://doi.org/10.23889/ijpds.v11i1.3431
November 2026 © The Authors. Open Access under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/deed.en)

Introduction

Loneliness is a significant public health and wellbeing concern associated with adverse health outcomes including cardiovascular disease, cognitive decline, depression, and mortality [1–4], resulting in substantial societal and economic costs [5, 6]. Loneliness is a subjective state of perceived social deficit typified by feelings of isolation; social isolation is an objective lack of social contacts and relationships [2, 3, 7]. Due to the subjective nature of loneliness it is captured using various measurement tools and outcome classifications, such as emotional and social loneliness [8, 9]. Prevalence of loneliness can vary according to measurement tools, loneliness types, and across multiple factors including individual characteristics and life circumstances, such as gender, and age, which may influence social aspirations across the life course. Markers of social isolation include reduced social networks, fewer social contacts, and living alone. International research has associated living alone with both loneliness and adverse health outcomes including elevated mortality and suicide risk, particularly among middle-aged men [3, 10–13].

This dual aspect of living alone, as a correlate of loneliness and an independent health determinant, underscores potential value of including living alone measures in population research. Increased loneliness is also associated with living arrangements and partner bereavement [14, 15]. However, few studies have examined loneliness using population level administrative data including the inferred reason for why someone lives alone.

Additional loneliness risk factors include age, with mixed evidence; large coordinated analyses suggest a U-shaped curve (higher in younger and older adults) [16], whereas recent UK monitoring reports highest prevalence in younger adults [17], likely reflecting measure and context differences. Gender gaps are generally small across the life course [18]. Evidence from Wales and England reports increased loneliness for those from an ethnic minority background, those who live in material deprivation, and for those living in social or privately rented properties [14, 19]. Broader UK analyses report lower education and lower socioeconomic status being linked to higher loneliness and social isolation [20]. Neighbourhood greenness, walkability, and housing type have been linked with greater social connection and lower loneliness in certain contexts [21–23]. However, systematic review evidence emphasises variation across populations and setting [24–26].

Loneliness is likely to arise through the interaction of multiple individual, social and environmental factors. Living arrangements, health status and place-based characteristics may each influence opportunities for social connection and experiences of loneliness. Understanding these factors together may therefore provide a more comprehensive picture of loneliness risk within populations. Measurement of loneliness and living alone can be challenging - large surveys are costly, may include non-representative or modest samples, and repeated measures are needed to capture longitudinal trends, with measures often limited to binary classifications [27]. With advances in access to population level linked administrative data (generated through individuals' interactions with public services [28]) it is possible to improve our understanding of these circumstances by including place-based factors which may impact on, or shape experiences, of loneliness. With a statutory loneliness national indicator [14, 29], Wales provides a useful setting for this research using the mature Secure Anonymised Information Linkage (SAIL) Databank which supports privacy-protecting linkage of person-based health, social care and demographic data at population level [30]. Wales has an ageing population [31], has distinct small-area deprivation across diverse urban, rural, and post-industrial communities [32].

The SAIL Databank enables linkage of longitudinal administrative and survey data alongside demographic, health, and built environment measures at population-scale. This study used administrative data to develop richer, longitudinally consistent indicators of living alone as an exposure variable along with other risk factors and investigated associations with loneliness. By linking longitudinal administrative and survey data, we (1) explored the scale, temporal trends, and demographic characteristics of living alone in Wales, United Kingdom (UK), from 2008–2024 (Phase 1); (2) developed novel living alone reason and duration indicators (Phase 1); and (3) assessed associations between emotional, social, and overall loneliness and living alone as well as other potential risk factors including health, demographic, and built environment factors (Phase 2).

Methods

Research Environment and DLinkage

Analysis was conducted in the Secure Anonymised Information Linkage (SAIL) Databank which contains multi-sectoral anonymised data for residents of Wales, UK. The data linkage methods, security protocols, and governance criteria of SAIL are described in detail elsewhere[30, 33–35]. In brief, individuals are assigned a unique linkage key (Anonymised Linking Field – ALF [30]) that allows individual level linkage across a range of administrative and survey data. Additional household level linkage (via residential linkage RALF [36, 37]) and area-level linkage allow construction of flexible population-based cohorts and analytical datasets incorporating health, housing, and environmental metrics.

Data Sources

Welsh Demographic Service (WDS): individual level administrative register of all GP-registered residents in Wales, providing longitudinal demographic and address records. Extracted: gender, week of birth, week of death, and address registration dates.

National Survey for Wales (NSW) [38]: annual survey of adults, providing demographic and lifestyle measures. Extracted: age, gender, loneliness scores (De Jong Gierveld 6-item tool including emotional, social, and overall scales), education, general health, long-term illness, housing tenure. Education and long-term illness were not collected in wave 2020-21, we created a category assigned as 'not collected' as a structural placeholder (not for interpretation).

Census 2021 (CEN21)[39] used to assign ethnicity for those linked (previous work estimates linkage rate of WDS to CEN21 at 83%) [58].

Ordnance Survey, AddressBase Premium (ABP) [40]: housing dataset to classify housing type. Categories: Detached, Semi-Detached, Terraced, Flats (Self-contained flat, HMO), or Other.

Care Inspectorate Wales (CIW): [41]: register of care home properties in Wales. CIW list supplemented with additional properties from: WDS properties with >25 residents and mean age ≥ 65, or ABP properties classed as care homes, Sheltered accommodation, communal residence, residential institution, or hospice).

Welsh Index of Multiple Deprivation (WIMD 2014) [42]: area-level lower layer super output (LSOA) area income domain deprivation quintiles.

Office for National Statistics (ONS) Rural-Urban (2011) classifications [43]: area-level rural/urban classification. Categories: urban, town, village.

Survey mode indicator: binary variable accounting for NSW methodological changes during Covid-19; including move from face-to-face to combined telephone and online survey methods and targeting repeat respondents from April 2020 [44]. Additional sensitivity analyses restricted to pre-pandemic and during-pandemic survey waves for the emotional loneliness model demonstrated broadly consistent associations across principal modelled factors, with no material impact on overall interpretation.

Normalised Difference Vegetation Index (NDVI) [45]: linked at household level with a 50m buffer around households, 2017 version. Categorised into greenness quartiles: Q1 (NDVI <0.35, least green), Q2 (≤0.40), Q3 (≤0.50), Q4 (>0.50) most green).

Active Living Environments (ALE) - Transit scale (ALE transit) [46]: linked at LSOA level. Categorised in four main clusters: very low, low, moderate, high.

Outcome Measure: Self-Reported Loneliness

Loneliness was measured using the 6-item De Jong Gierveld short loneliness scale [9, 47, 48], comprising three emotional loneliness (EL) questions: 'I experience a general sense of emptiness', 'I miss having people around', and 'I often feel rejected'; and, three social loneliness questions: 'There are plenty of people I can rely on when I have problems', 'there are many people I can trust completely', 'there are enough people I feel close to'. Subscale scores were dichotomised (≥2 = lonely), and overall loneliness was dichotomised with a ≥4 cut-off [49, 50]. Previous research indicates minimal information loss when dichotomising these scales [51].

Primary Explanatory Variable: Living Alone

Living alone factors included a binary indicator (living alone vs. non-living alone), and a novel 'reason' variable. Reasons were inferred from household composition and transitions of the individual and previous cohabitants. For these analyses we used the simplest derived version with three main categories: partner loss (due to death or moving to a care home), separated (previously cohabiting with ≥1 individual), unassigned (no historical data available to assign reason). Duration was calculated in years for each continuous living alone episode then combined with the partner loss reason variable to examine loneliness according to time since the loss of a partner. Final categorisations were: partner loss ≤2 years, partner loss 3-4 years previous, partner loss >5 years, separation, unassigned. Partner loss categories included transitions following death of a cohabitant or movement of a cohabitant into a care home. These were combined due to small numbers and similar observable household-loss transitions within the available administrative data. The descriptive label 'separation' does not necessarily capture partner type relationship issues only and will include other situations for previously cohabiting individuals.

Additional Risk Factors: Health, Demographic, and Built Environment Factors

Demographic factors: Age, categorised into bands (16-30 (reference), 31-45, 46-60, 61-75, 76+); Gender, male (reference), and female; Ethnicity, white (reference), minority ethnic background, not linked to census; Education: degree/higher (reference), A-level/apprenticeships, GCSE, other qualifications, no qualifications.

Socioeconomic and housing characteristics: Housing tenure, mortgage or owned (reference), private rental, social housing; Housing type, detached (reference), semi-detached, terraced, flat, other; Deprivation quintiles, Q1 most deprived (references) to Q5 least deprived.

Health status: General health, good/very good (reference), fair, bad/very bad; Long-term illness, no (reference), yes.

Geographical and rurality measures: Region, South-East Wales (reference), Mid-Wales, North Wales, South-West Wales; Rurality (Phase-1 measure), urban city and town (reference), urban city and town (sparse setting), rural town and fringe, rural town and fringe (sparse setting), rural village and dispersed, rural village and dispersed (sparse setting); Rurality (Phase-2 coarse measure), urban city and town (reference), rural town, rural village.

Built environment exposures: NDVI greenness, Active Living Environments - Transit scale.

Phases of Analysis

Analyses were conducted in two phases. Phase 1 examined the scale, temporal trends, and demographic characteristics of living alone in Wales, including development of novel living alone reason and duration indicators. Phase 2 used linked administrative and National Survey for Wales data to assess associations between emotional, social, and overall loneliness and living alone alongside other potential risk factors.

Phase 1

Population-Wide Living Alone Trends

Phase 1 investigated population-wide living alone trends in Wales using longitudinal data. We produced visualisations and fitted logistic regression models to examine associations between living alone and demographic characteristics and age-gender interactions. Analyses were conducted at the household level. Households were classified as "living alone" if a single adult resident (aged ≥16 years) was registered at a residential property, and "not living alone" if two or more adults were registered. For regression analyses, data were aggregated by

household characteristics (e.g., age group, gender, deprivation, housing type), and binomial logistic regression models were fitted using the counts of living alone versus not living alone as the outcome. The oldest person in the household was designated as the household reference individual. This approach enabled efficient modelling of population-wide trends while maintaining consistency with census-based definitions.

Annual Population Snapshot (APS)

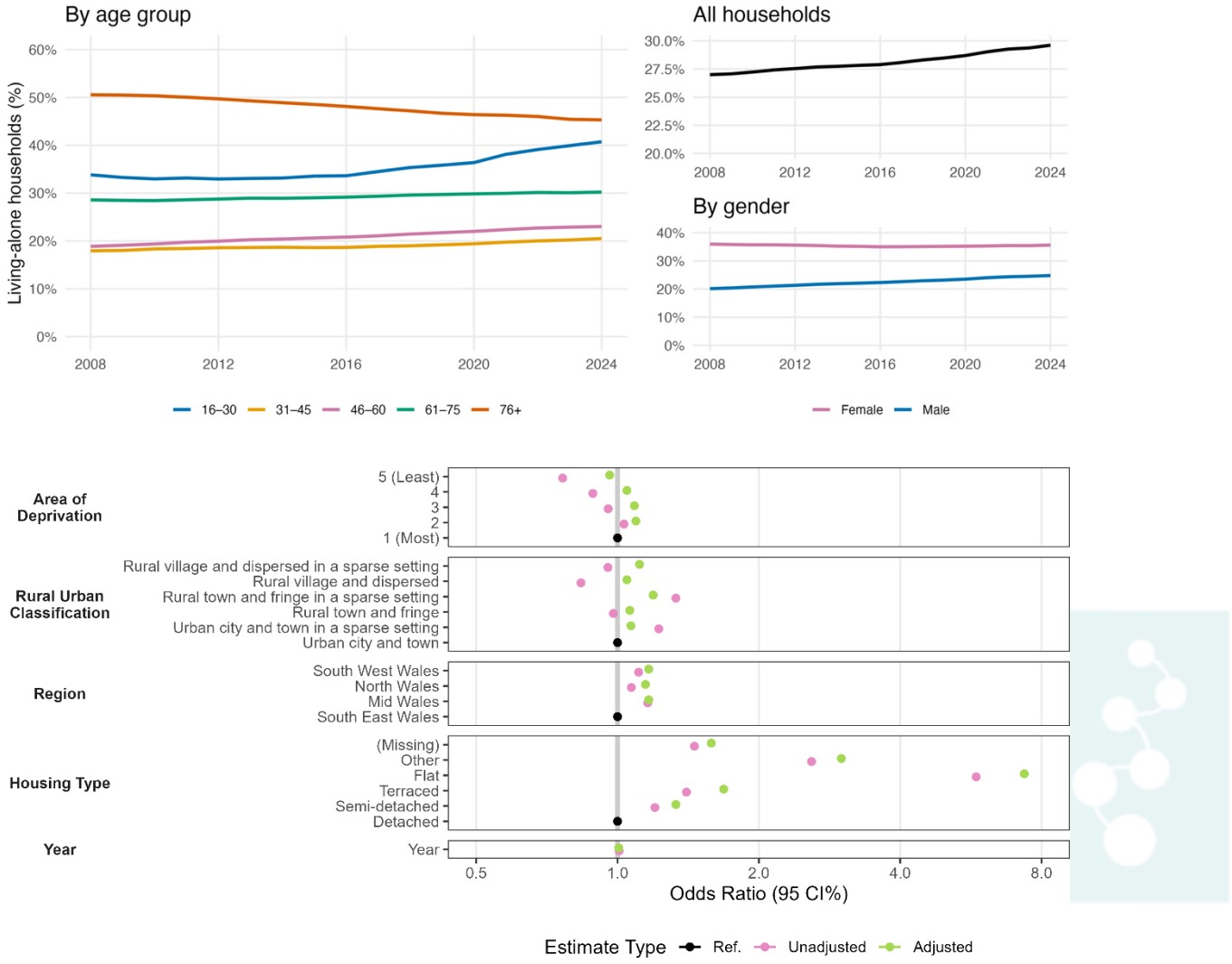
Annual population-wide, individual level snapshot (APS) data was created using the Welsh Demographic Service (WDS) for each year from 2008 to 2024, based on a mid-year census date of 30th June. We included records for individuals with valid age and gender data who were registered at a property with a residential anonymised linking field (RALF) located within Wales, where at least one adult resident (aged $\geq$16 years) was present, and where the total household size was nine residents or fewer. APS included age, gender, dates of death, ethnicity, care home residence, household composition (including living alone indicators), and area-level measures of deprivation and rurality. These data were used to examine the longitudinal scales and trends of living alone in Wales.

Phase 2

Final Loneliness Survey Study Cohort

Analyses Phase 2 used logistic regression modelling to examine associations between loneliness and factors including living alone, demographic characteristics, health, and aspects of the built environment. The main model was developed using emotional loneliness subscale and then applied to the social subscale and overall loneliness scale. Interactions between age and sex were examined and included. Models were repeated to obtain specific binary living alone results models, replacing the more granular reason-duration factor. Results are presented as odds ratios (OR) and adjusted odds ratios (aOR) with 95% confidence intervals (CI), and p-values. For descriptive statistics we used each participant's most recent record. For regression models, all available records, including repeat measurements, were included. Sensitivity analyses using cluster-robust standard errors clustered at the individual level were conducted to assess the impact of repeated observations on statistical inference. These analyses demonstrated no material impact on interpretation of the findings (Supplementary Table 5). Model adjusted R^2 and missing outcome data are reported.

Figure 1: Prevalence of Living Alone by Overall Household, Age Group and Gender, Wales 2008–2024



Results

Figure 1 shows trends in living alone households in Wales from 2008 to 2024. The proportion of households classified as living alone increased from 27.0% (314,806) in 2008 to 29.6% (363,133) in 2024, representing a 2.6 percentage-point increase (9.6% relative increase). Living alone was most common among those aged 76 years and over throughout the study period, although prevalence declined modestly over time. In contrast, the largest increase was observed among those aged 16–30 years, with smaller increases across the remaining age groups. Females consistently had a higher prevalence of living alone than males, although the proportion of males living alone increased slightly across the study period. Across the study period, the mean number of living alone households was 336,646 (SD=13,722), compared with 859,801 (SD=5,660) multiple-person households.

Phase 1: Population Level Living Alone in Wales

In unadjusted analyses, households in the least deprived areas were 24% less likely have an individual living alone than in the most deprived areas (OR 0.76, 95% CI 0.76-0.77). *Stronger associations were observed for housing type; flats had nearly six-fold higher odds of single occupancy compared with detached houses*. Following adjustment for all covariates, deprivation effects were small relative to the most deprived areas, those living in quintiles 2 to 4 being between 5 and 9% higher risk, and those living in the most deprived areas 4% less likely to live alone (OR 0.96, 95% CI 0.96-0.96). Living in regions outside of South-East Wales were associated with ∼16% increased odds (e.g. Mid-Wales OR 1.17, 95% CI 1.16-1.17). Living in a flat compared to a detached home was associated with seven-fold higher odds of living alone (OR 7.35, 95% CI 7.32-7.38). Full model results are presented in Supplementary Table 1.

Figure 1 Estimated unadjusted and adjusted odds ratios for risk of living alone by main effects (deprivation, housing type, region, and rurality), adjusted for age and sex interaction term (age-gender specific odds shown separately).

The age-gender interaction term is illustrated in Figure 2 (and Supplementary Table 2). For males, the risk of living alone was lower across most age groups, with only modest variation over the life course. For females, risk followed a general U-shaped pattern across the life course, declining through midlife before rising steeply in later life. By age >76, females had more than four times the odds of living alone compared with men (aOR 4.2).

Phase 2

Phase 2 analyses aimed to assess associations between living alone and emotional, social, and overall loneliness, as well as other potential risk factors including health, demographic, and built environment factors. We used survey data (NSW) linked to annual snapshot (APS) data. After applying linkage and exclusion criteria, the final analysis cohort comprised 49,167 records, for 43,196 individuals (Supplementary Table 3).

Cohort Characteristics

Table 1 summarises the number and percentage of unique individuals by binary living alone status. A quarter of the cohort were living alone and the assigned reasons for living alone were partner loss (22%), separation (42%), and unassigned (35%). Those living alone reported higher loneliness across all scales (emotional, social, and overall). In terms of demographics, individuals living alone were more often female, older, less qualified, had worse general health and more long-term illness. Individuals tended to reside in more deprived areas, more rural areas, in social housing, and in flats. There was minimal difference noted for built environment related factors (greenness and active living environments).

Loneliness Outcomes

Figure 3 shows each loneliness scale score distribution (excluding missing loneliness scores) by living alone status. For individuals living alone, emotional loneliness (32.0%) was higher than social loneliness (28.8%), however, for those not living alone emotional loneliness (19.4%) was lower than social loneliness (25.2%). Dichotomised and raw scores for each NSW survey wave are available in Supplementary Figure 1; notably, from April 2020 onwards missingness for loneliness scores was eradicated due to survey method changes.

Figure 4 shows model outputs for each loneliness score expressed as odds ratios adjusted for all main effects and an age-gender interaction.

Figure 2: Estimated age-gender specific adjusted odds ratios for risk of living alone by age and sex interaction term, adjusted for main effects (deprivation, housing type, region, and rurality)

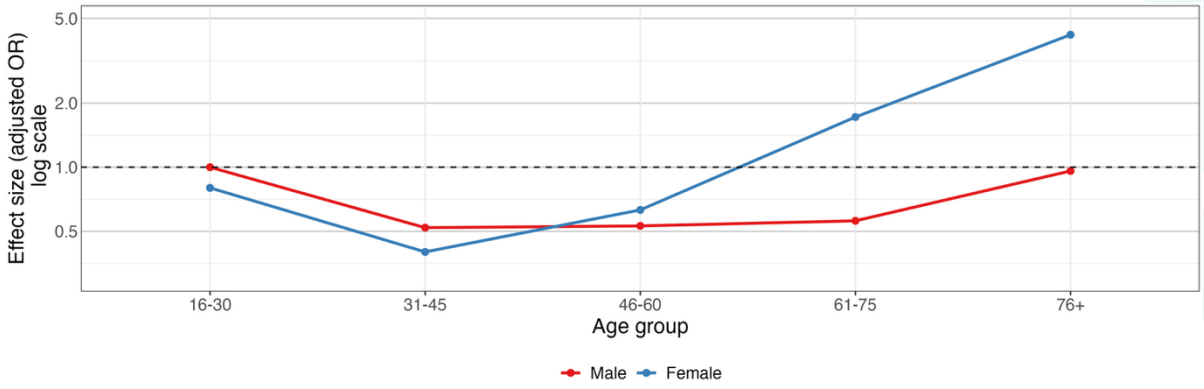


Table 1: Characteristics for full cohort and by living arrangements (living alone and not living alone), including outcome variables (emotional, social, and overall loneliness), and all other factors (total number of unique individuals and percentage)

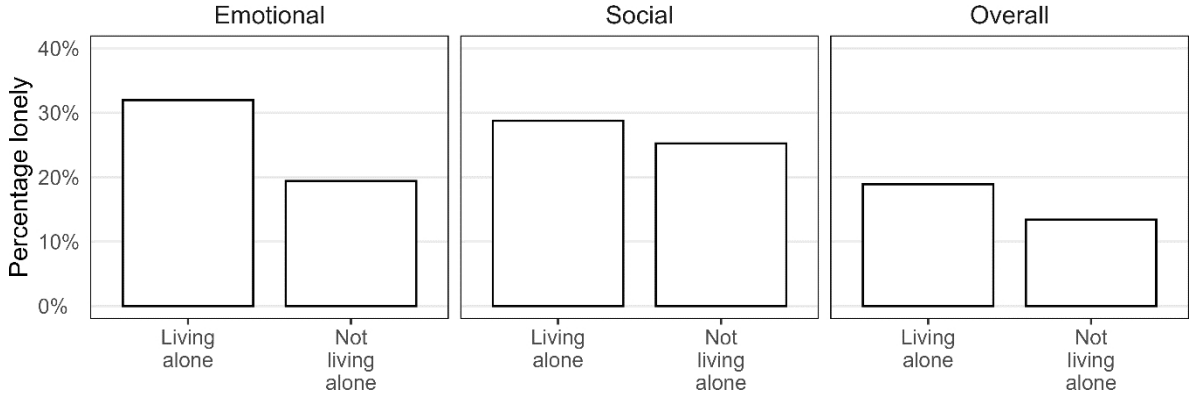
Characteristic	Category	Total	Living alone	Not living alone
Total	–	43,196 (100.0)	10,819 (100.0)	32,377 (100.0)
Emotional loneliness	Not Lonely	30,664 (71.0)	6,737 (62.3)	23,927 (73.9)
	Lonely	8,926 (20.7)	3,167 (29.3)	5,759 (17.8)
	Missing	3,606 (8.3)	915 (8.5)	2,691 (8.3)
Social loneliness	Not Lonely	30,294 (70.1)	7,243 (66.9)	23,051 (71.2)
	Lonely	10,705 (24.8)	2,923 (27.0)	7,782 (24.0)
	Missing	2,197 (5.1)	653 (6.0)	1,544 (4.8)
Overall loneliness	Not Lonely	33,405 (77.3)	7,942 (73.4)	25,463 (78.6)
	Lonely	5,794 (13.4)	1,853 (17.1)	3,941 (12.2)
	Missing	3,997 (9.3)	1,024 (9.5)	2,973 (9.2)
Living alone reason	Not Living Alone	32,377 (75.0)	-	32,377 (100.0)
	Partner Loss $\leq$2y	657 (1.5)	657 (6.1)	-
	Partner Loss $\leq$5y	830 (1.9)	830 (7.7)	-
	Partner Loss >5y	932 (2.2)	932 (8.6)	-
	Separation	4,585 (10.6)	4,585 (42.4)	-
	Unassigned	3,815 (8.8)	3,815 (35.3)	-
Age	16-30	4,524 (10.5)	402 (3.7)	4,122 (12.7)
	31-45	7,954 (18.4)	856 (7.9)	7,098 (21.9)
	46-60	10,761 (24.9)	2,130 (19.7)	8,631 (26.7)
	61-75	13,386 (31.0)	4,118 (38.1)	9,268 (28.6)
	76+	6,571 (15.2)	3,313 (30.6)	3,258 (10.1)
Gender	Female	24,504 (56.7)	6,302 (58.2)	18,202 (56.2)
	Male	18,692 (43.3)	4,517 (41.8)	14,175 (43.8)
Ethnicity	Minority Ethnic Background	897 (2.1)	113 (1.0)	784 (2.4)
	White	35,218 (81.5)	8,324 (76.9)	26,894 (83.1)
	Not Linked to Census	7,081 (16.4)	2,382 (22.0)	4,699 (14.5)
Education	Higher Degree	11,323 (26.2)	2,351 (21.7)	8,972 (27.7)
	A-Levels or Apprenticeships	3,998 (9.3)	806 (7.4)	3,192 (9.9)
	GCSE	5,883 (13.6)	1,248 (11.5)	4,635 (14.3)
	Other Qualifications	2,802 (6.5)	826 (7.6)	1,976 (6.1)
	No Qualifications	6,107 (14.1)	2,280 (21.1)	3,827 (11.8)
	Not Collected	13,083 (30.3)	3,308 (30.6)	9,775 (30.2)
Housing tenure	Owned/Mortgaged	31,643 (73.3)	7,324 (67.7)	24,319 (75.1)
	Private Rental	4,578 (10.6)	1,052 (9.7)	3,526 (10.9)
	Social Housing	6,462 (15.0)	2,281 (21.1)	4,181 (12.9)
	Other	513 (1.2)	162 (1.5)	351 (1.1)
Housing type	Detached	13,949 (32.3)	2,723 (25.2)	11,226 (34.7)
	Semi-Detached	12,359 (28.6)	2,876 (26.6)	9,483 (29.3)
	Terraced	13,170 (30.5)	3,411 (31.5)	9,759 (30.1)
	Flat	2,729 (6.3)	1,582 (14.6)	1,147 (3.5)
	Other or Missing	989 (2.3)	227 (2.1)	762 (2.4)
Deprivation	Most Deprived - Q1	7,051 (16.3)	1,948 (18.0)	5,103 (15.8)
	Q2	8,176 (18.9)	2,229 (20.6)	5,947 (18.4)
	Q3	9,298 (21.5)	2,434 (22.5)	6,864 (21.2)
	Q4	9,398 (21.8)	2,238 (20.7)	7,160 (22.1)
	Least Deprived - Q5	9,273 (21.5)	1,970 (18.2)	7,303 (22.6)
Rurality	Urban	26,950 (62.4)	6,742 (62.3)	20,208 (62.4)
	Town	8,130 (18.8)	2,220 (20.5)	5,910 (18.3)
	Village	8,116 (18.8)	1,857 (17.2)	6,259 (19.3)

Continued

Table 1: Continued

Characteristic	Category	Total	Living alone	Not living alone
General health	Good, Very Good	29,357 (68.0)	6,324 (58.5)	23,033 (71.1)
	Fair	9,535 (22.1)	2,955 (27.3)	6,580 (20.3)
	Bad, Very Bad	4,304 (10.0)	1,540 (14.2)	2,764 (8.5)
Long-term illness	No	16,528 (38.3)	3,063 (28.3)	13,465 (41.6)
	Yes	19,623 (45.4)	6,002 (55.5)	13,621 (42.1)
	Not Collected	7,045 (16.3)	1,754 (16.2)	5,291 (16.3)
Greenness category	Q1 - Least Dense	10,381 (24.0)	2,742 (25.3)	7,639 (23.6)
	Q2	9,551 (22.1)	2,455 (22.7)	7,096 (21.9)
	Q3	11,189 (25.9)	2,887 (26.7)	8,302 (25.6)
	Q4 - Most Dense	11,225 (26.0)	2,601 (24.0)	8,624 (26.6)
	Other or Missing	850 (2.0)	134 (1.2)	716 (2.2)
Active living environment	Very Low	14,715 (34.1)	3,723 (34.4)	10,992 (34.0)
	Low	10,599 (24.5)	2,738 (25.3)	7,861 (24.3)
	Moderate	14,194 (32.9)	3,395 (31.4)	10,799 (33.4)
	High	2,936 (6.8)	751 (6.9)	2,185 (6.7)
	Missing	752 (1.7)	212 (2.0)	540 (1.7)
Survey period	Pre-Covid	20,588 (47.7)	5,017 (46.4)	15,571 (48.1)
	During Covid	22,608 (52.3)	5,802 (53.6)	16,806 (51.9)

Figure 3: Crude prevalence of loneliness score by living arrangement (living alone vs not living alone) for each scale (emotional, social, and overall)



Bars show proportion classified as lonely.

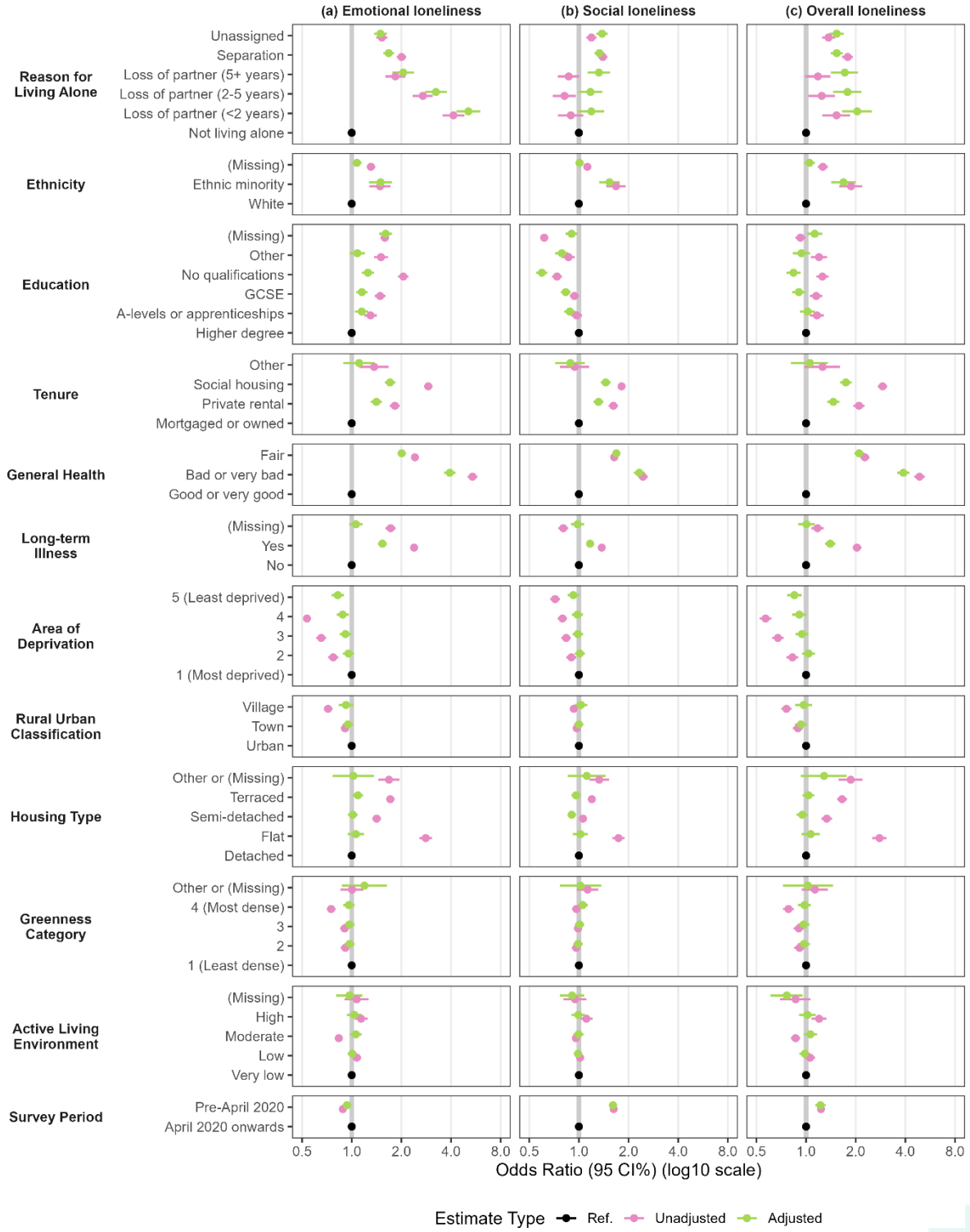
Living alone was associated with higher odds of loneliness across all scales, even after adjustment for demographic, health, and built environment factors, and the age-gender interaction terms. The effect size was largest for emotional loneliness.

In order to extract effect sizes for the binary living alone measure models were replicated with replacement of the reason-duration measure. For this set of otherwise identical models individuals living alone had 1.9 times higher odds of emotional loneliness (OR 1.9, 95% CI 1.8-2.0); 1.2 times higher odds of social loneliness (OR 1.2, 95% CI 1.1-1.3), and 1.5 times higher odds of overall loneliness (OR 1.5, 95% CI 1.4-1.6) compared to those not living alone. The main models included the reason-duration living alone measure and provided more detail. Those with a recent ($\leq$2 years) partner loss (bereaved or to a care home) had the highest odds of emotional loneliness (OR 5.1, 95% CI 4.3-6.0), with associations attenuating over time (2-4 years: OR 3.2, 95% CI 2.8-3.8); >5 years OR 2.1, 95% CI 1.8-2.4). Separation (OR 1.7, 95% CI 1.6-1.8) and unassigned reasons (OR 1.5, 95% CI 1.4-1.6) were also associated with emotional loneliness. Each living alone reason for social loneliness was associated with higher odds of loneliness compared to not living alone, but the same pattern was not observed with reduced effect sizes all below odds ratios of 1.5.

Minority ethnic background was associated with higher odds of both emotional and social loneliness (1.5, 95% CI 1.3-1.8). There was an indication that lower education levels were protective, specifically for those with no qualification, social loneliness (OR 0.6, 95% CI 0.6-0.6); but contrastingly a risk factor for emotional loneliness (OR 1.3, 95% CI 1.2-1.4).

Housing and deprivation were also significantly associated with loneliness. Social housing (1.7, 95% CI 1.6-1.8) and private rental (1.4, 95% CI 1.3-1.5) were associated with higher emotional loneliness compared with owner-occupied homes. Living in less deprived areas was associated with

Figure 4: Forest plot of unadjusted and adjusted odds ratios on a log-scale (with lower and upper 95% Confidence Intervals) for final model main effects for emotional, social, and overall loneliness scores



(Also adjusted for age-gender interaction terms).

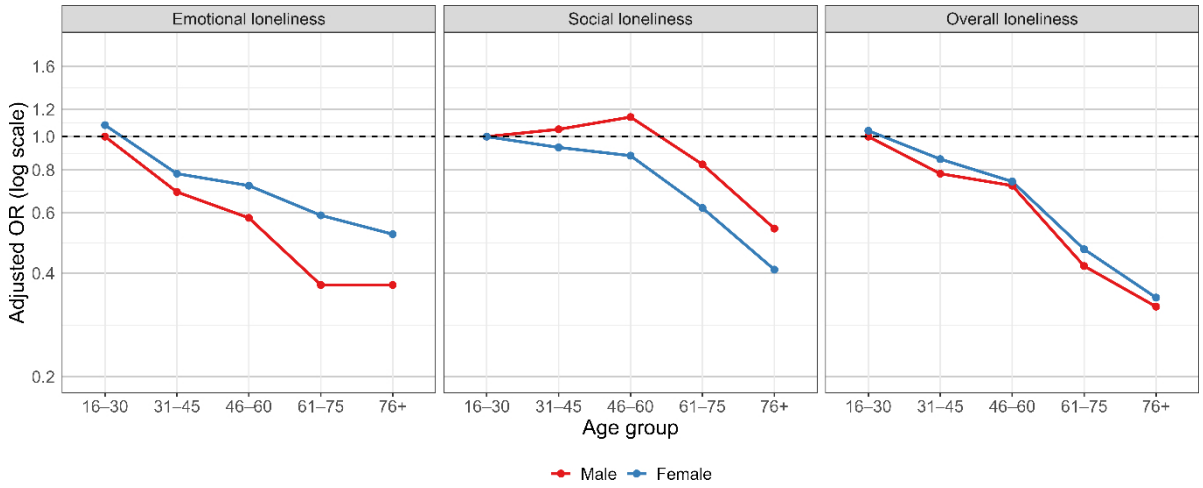
reduced odds of emotional loneliness (least deprived compared with most deprived quintile: OR = 0.8, 95% CI 0.8-0.9).

Health was strongly associated with loneliness. Compared with good/very good, fair health (2.0, 95% CI 1.9-2.1), and bad/very bad health (3.9, 95% CI 3.6-4.3) were strongly associated with emotional loneliness. Long-term illness (1.5, 95% CI 1.4-1.6) was also associated with emotional loneliness. Similar but smaller associations were observed for social loneliness.

Built environment factors showed no notable effects. Survey period was associated with social loneliness, higher before Covid-19 compared to during or after (1.6, 95% CI 1.5-1.7).

Adjusted age-gender interaction results are illustrated in Figure 5. Loneliness generally decreased with age with variation by age, gender, and across loneliness scales. Younger adults reported the highest emotional loneliness levels irrespective of gender; females had consistently higher odds

Figure 5: Adjusted odds ratios on a log-scale for age-gender interaction for emotional, social, and overall loneliness scores, adjusted for all other model covariates



than males across all ages, with the gap widening in later life. For social loneliness there was gender parity in the youngest age group and thereafter males reported higher loneliness; odds for males increased with age into midlife at age group 46-60 before showing a similar, parallel declining trend to females in older age groups. The overall loneliness scale showed a simpler linear decline with age and no clear gender interaction.

Full model outputs are shown in Supplementary Table 4, with clustered model sensitivity analyses shown in Supplementary Table 5. Model fit statistics (McFadden's adjusted R^2) and number of modelled records: emotional loneliness, $R^2 = 0.11$ ($n = 44{,}907$); social loneliness, $R^2 = 0.05$ ($n = 46{,}635$); overall loneliness, $R^2 = 0.09$ ($n = 44{,}405$).

Discussion

Within Phase 1 the study evidenced living alone patterns, with a slow continuously increasing trend in the scale by number and proportion of living alone households in Wales. Within modelled results, distinct living alone patterns by age and gender were observed across the life course; younger males showed the highest prevalence of living alone, while among females prevalence increased from midlife and remained elevated in older age groups. The largest adjusted association was observed for those living in flats compared with other housing types (detached, semi-detached and terraced). We also introduced a novel method for inferring and tracking longitudinal living alone measures using administrative data at population level.

Phase 2 of the study found that loneliness in Wales is strongly associated with factors related to living alone, poor health, and certain demographic factors, with the highest risk amongst those who have recently lost a cohabiting partner to death or entry to a care home. A linear negative association between age and overall loneliness was observed with younger adults reporting highest loneliness. Important age-gender differences were evident with contrasting patterns between emotional and social loneliness subscales. Males reported consistently higher social loneliness than females, especially around middle age (31-60), whereas females reported higher levels of emotional loneliness than males with a widening gap with age. Built environment factors showed little evidence of association.

The significantly increased odds of emotional loneliness for those recently bereaved aligns with previous findings indicating relationship loss is a key determinant of loneliness [15]. We quantified differences in measures of loneliness between living alone reasons: partner loss, separation, and unassigned cases, demonstrating that context matters beyond a simple binary indicator. This supports arguments that living alone is an important correlate of loneliness and health outcomes [3, 10, 11]. Our results confirm previous research linking living alone factors to loneliness and demonstrate the ability to apply such methods using administrative data [15, 49, 52].

We observed that modelled overall loneliness declined with age, with younger adults reporting the highest levels, and clear age-gender differences across the emotional and social subscales—a pattern consistent with large-scale survey evidence [53] and ONS monitoring [17]. This partly contrasts with large-scale studies reporting a U-shaped age pattern [16, 54]. Notably, NSW analysis found no independent associations of overall loneliness with either age or limiting long-term health conditions, alongside strong associations with general health and mental illness [14]. As noted previously [55], inconsistencies may reflect differences in measurement and modelling approaches; our examination of emotional and social subscales separately, and to include an age-gender interaction, may explain why we identified clearer declines with age alongside gender specific differences in line with previous research [17].

Age-gender differences between emotional and social subscales underline the value of analysing these dimensions separately, as determinants can differ by subscale [9, 51]. The UCLA-3 is widely used as an overall loneliness indicator and correlates with both De Jong Gierveld subscales [56].

Place-based effects were small in contrast to international studies reporting lower loneliness in more favourable neighbourhood and social environments [21–23, 57]. Our null findings are consistent with systematic reviews emphasising contextual variation and mixed results [24, 25]. This may reflect limitations of the available measures, or that, in Wales,

individual level factors such as health and living arrangements dominate over place effects.

Our results highlight several groups at elevated risk of loneliness. Young adults reported the highest levels of emotional loneliness, while recently bereaved adults showed particularly high risks of emotional loneliness persisting years after partner loss. Poor self-reported health and long-term illness were consistently associated with loneliness across subscales, suggesting potential implications for health and social care service use. We also observed higher social loneliness among midlife men, aligning with evidence of increased social isolation and suicide risk in this group [12, 13]. Taken together, these findings identify several groups that may benefit from further investigation and consideration within intervention and service-planning approaches, including bereavement support, community outreach and health-linked social prescribing.

Findings also suggest that loneliness may be particularly relevant for groups already in contact with health and social care services, such as those in poor health, living in social housing, and those recently bereaved; highlighting potential opportunities for identification and support.

This study has several strengths. In Phase 1, we created longitudinal population level snapshots for Wales, enabling analysis of living alone trends over time. We also developed novel living alone reason indicators which may prove useful for wider research using similar administrative data. In Phase 2, we drew on relatively large numbers by linking two survey sources to administrative data, applied validated multiscale loneliness scores, and demonstrated the practical use of novel living alone factors in modelling risk of loneliness. Inclusion of age-gender interactions across multiple loneliness scales provided more nuanced understanding than combining gender or limiting analyses to an overall loneliness score, avoiding the risk of masking potentially modifiable associations relevant for policy.

There are also several study limitations. As is standard with data linkage, some loss of cases occurred, for example when individuals were not assigned a linkage key (around 9%); or when NSW survey data were not successfully linked (around 6%) (Supplementary Table 3); this study therefore remains broadly representative of residents in Wales over 2016-2022. Use of annual snapshots may introduce temporal error due to factors such as discrepancies between survey dates, events such as death, and mid-year linkage points. The built environment ALE transit measure was available at LSOA level, which may be less sensitive when exploring associations. The NSW methodology also changed across the study period; not all variables were collected annually, and the COVID-19 pandemic may have influenced loneliness scores. Given the observational design and partial reliance on cross-sectional survey measures, associations should not be interpreted as causal. Relationships between loneliness, health status, housing circumstances and living arrangements are likely to be complex and potentially bidirectional.

Potential selection bias may have arisen through participant consent for linkage, successful creation and matching of anonymised linkage identifiers, and study-specific exclusions due to missing covariate or outcome information. However, comparisons between the initial eligible survey sample and final analytic cohort demonstrated broadly similar age distributions and only modest differences in sex distributions (Supplementary Table 6). Sensitivity analyses restricted to pre-pandemic and post-pandemic survey waves produced broadly consistent findings.

Finally, while our novel living alone reason measure provided additional insight, a simple binary indicator of living alone also showed clear associations with loneliness. For monitoring purposes, such binary indicators may therefore be sufficient, offering reduced complexity and easier implementation across large datasets. The more detailed measure highlights heterogeneity, especially for emotional loneliness, but comes at the cost of added derivation and interpretive complexity. Accepting this trade-off may be appropriate where increased simplicity and comparability are prioritised, though our findings suggest nuanced indicators add value in research contexts. Ability to derive and assign reason indicators varies over time as a natural data artefact related to address registration data; for individuals already living alone prior to the APS 2008 start date, reasons could not be inferred. These inferred categories represent administrative approximations of household transition circumstances rather than precise social or relationship histories and may therefore be subject to some misclassification. Population level living alone reason trends (Supplementary Figures 2,3) indicate more recent snapshots provide a clearer basis for understanding living alone dynamics.

Conclusion

We present longitudinal living alone trends and introduce a novel method to derive living alone reasons at the population level using large-scale, routinely collected data. Living alone, particularly when characterised by more granular inferred reasons, was strongly associated with emotional, social, and overall loneliness after adjustment for demographic and health factors, with the strongest associations for emotional loneliness. Including an age-gender interaction revealed nuanced differences across the life course. These findings extend the loneliness evidence base and offer scalable indicators for monitoring living alone dynamics. They also provide tools for future health and social care research using linked data and complement national statistics on loneliness and wellbeing.

Author Contribution

RDJ: Conceptualisation, methodology, formal analysis, data curation, visualisation, writing – original draft, approval of final submission.

RB, SB: Conceptualisation, supervision, methodology, writing – review & editing, approval of final submission.

RF: Conceptualisation, supervision, writing – review & editing, approval of final submission.

LG: Conceptualisation, supervision, methodology, writing – review & editing, approval of final submission.

Funding

This work is supported by Administrative Data Research (ADR) Wales (Grant ref: ES/W012227/1), part of the ADR UK investment, uniting research expertise from Swansea University Medical School and WISERD (Wales Institute of Social and Economic Research and Data) at Cardiff University with analysts from Welsh Government. ADR UK is funded by the Economic and Social Research Council (ESRC), part of UK Research and Innovation.

Acknowledgements

This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. We would like to acknowledge all the data providers and people who make anonymised data available for research. The authors would like to extend their gratitude and acknowledgement to the NHS, the SAIL Consumer panel as well as the IGRP who approved this project (SAIL project 1647).

Patient Consent for Publication

Not required.

Ethical Approval

The use of de-identified data in SAIL complies with National Research Ethics Service (NRES) guidance. Applications to use data held within the SAIL Databank, an ISO: 27001 and UK Statistics Authority (UKSA) Digital Economy Act (DEA) accredited Trusted Research Environment, must first be approved by the independent Information Governance Review Panel (IGRP). This panel contains individuals with expertise in data governance and protection, including the Chair of the Wales NRES Committee, Caldicott Guardians and members of the public. The IGRP approved SAIL project 1647.

Data Availability

This study makes use of anonymised, individual level data held in the SAIL Databank, a Trusted Research Environment, at Swansea University, Swansea, UK. Due to the nature and level of the data, data are not publicly available. All proposals to use SAIL data are subject to review by the independent IGRP. The IGRP gives careful consideration to each project proposal to ensure proper and appropriate use of SAIL data. If a project is approved, access to the requested data is gained through a privacy-protecting safe-haven and remote access system referred to as the SAIL Gateway. SAIL has established an application process to be followed by anyone who would like to access data via SAIL at: https://www.saildatabank.com/application-process/.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the author(s) used ChatGPT (OpenAI) in order to assist with language refinement and phrasing suggestions. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

1. Cacioppo JT, Cacioppo S. Older adults reporting social isolation or loneliness show poorer cognitive function 4 years later. Evidence-Based Nursing 2014 Apr.;17:59–60. https://doi.org/10.1136/EB-2013-101379

2. Hawkley LC, Cacioppo JT. Loneliness Matters: A Theoretical and Empirical Review of Consequences and Mechanisms. Annals of Behavioral Medicine 2010 Oct.;40:218–27. https://doi.org/10.1007/s12160-010-9210-8

3. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and Social Isolation as Risk Factors for Mortality. Perspectives on Psychological Science 2015 Mar.;10:227–37. https://doi.org/10.1177/1745691614568352

4. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul.;102:1009–16. https://doi.org/10.1136/heartjnl-2015-308790

5. Peytrignet S, Garforth-Bles S, Keohane K. Loneliness monetisation report. 2020;

6. Mcdaid D, Bauer A, Park A-L. Making the economic case for investing in actions to prevent and/or tackle loneliness: a systematic review A briefing paper. 2017;

7. National Academies of Sciences, Engineering and M. Summary. In: Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington, DC: The National Academies Press; 2020. p. 1–16. https://doi.org/10.17226/25663

8. Russell D, Peplau LA, Ferguson ML. Developing a measure of loneliness. Journal of Personality Assessment 1978;42:290–4. https://doi.org/10.1207/s15327752jpa4203_11

9. De Jong Gierveld J, Van Tilburg T. The De Jong Gierveld short scales for emotional and social loneliness: tested on data from 7 countries in the UN generations and gender surveys. European Journal of Ageing 2010 Jun.;7:121–30. https://doi.org/10.1007/s10433-010-0144-6

10. Stickley A, Koyanagi A. Loneliness, common mental disorders and suicidal behavior: Findings from a general population survey. Journal of Affective Disorders 2016 Jun.;197:81–7. https://doi.org/10.1016/j.jad.2016.02.054

11. Pompili M, et al. Loneliness and suicide risk in psychiatric disorders: A review. Clinical Neuropsychiatry 2010;7:227–39.

12. Shaw RJ, Cullen B, Graham N, Lyall DM, Mackay D, Okolie C, et al. Living alone, loneliness and lack of emotional support as predictors of suicide and self-harm: A nine-year follow up of the UK Biobank cohort. Journal of Affective Disorders 2021 Jan.;279:316–23. https://doi.org/10.1016/J.JAD.2020.10.026

13. Alothman D, Lewis S, Card T, Tyrrell E, Fogarty AW. The association between living alone, loneliness and suicide mortality and effect modification by age: A case:control study. Journal of Affective Disorders 2024 May;352:278–80. https://doi.org/10.1016/j.jad.2024.02.047

14. Welsh Government. Loneliness (National Survey for Wales): April 2022 to March 2023. 2023.

15. Utz RL, Swenson KL, Caserta M, Lund D, DeVries B. Feeling Lonely Versus Being Alone: Loneliness and Social Support Among Recently Bereaved Persons. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 2013;69B:85. https://doi.org/10.1093/GERONB/GBT075

16. Graham EK, Beck ED, Jackson K, Yoneda T, McGhee C, Pieramici L, et al. Do We Become More Lonely With Age? A Coordinated Data Analysis of Nine Longitudinal Studies. Psychological Science 2024 Jun.;35:579. https://doi.org/10.1177/09567976241242037

17. Office for National Statistics. Loneliness - What characteristics and circumstances are associated with feeling lonely? 2018.

18. Maes M, Qualter P, Vanhalst J, Van den Noortgate W, Goossens L. Gender Differences in Loneliness Across the Lifespan: A Meta-Analysis. European Journal of Personality 2019 Nov.;33:642–54. https://doi.org/10.1002/PER.2220

19. Department for Levelling Up Housing and Communities. Chapter 6: Well-being and loneliness [Internet]. 2023 [cited 2025 31]; Available from: https://www.gov.uk/government/statistics/chapters-for-english-housing-survey-2022-to-2023-headline-report/chapter-6-well-being-and-loneliness.

20. Kung CSJ, Pudney SE, Shields MA. Economic gradients in loneliness, social isolation and social support: Evidence from the UK Biobank. Social Science & Medicine 2022 Aug.;306:115122. https://doi.org/10.1016/J.SOCSCIMED.2022.115122

21. Maas J, van Dillen SME, Verheij RA, Groenewegen PP. Social contacts as a possible mechanism behind the relation between green space and health. Health and Place 2009;15:586–95. https://doi.org/10.1016/j.healthplace.2008.09.006

22. Astell-Burt T, Hartig T, Eckermann S, Nieuwenhuijsen MJ, McMunn A, Frumkin H, et al. More green, less lonely? A longitudinal cohort study. International Journal of Epidemiology 2022;51:99–110. https://doi.org/10.1093/ije/dyab089

23. Kearns A, Whitley E, Tannahill C, Ellaway A. Loneliness, social relations and health and well-being in deprived communities. Psychology, Health & Medicine 2015;20:332–44. https://doi.org/10.1080/13548506.2014.940354

24. Bower M, Kent J, Patulny R, Green O, McGrath L, Teesson L, et al. The impact of the built environment on loneliness: A systematic review and narrative synthesis. Health and Place 2023 Jan.;79. https://doi.org/10.1016/j.healthplace.2022.102962

25. Astell-Burt T, Hartig T, Putra IGNE, Walsan R, Dendup T, Feng X. Green space and loneliness: A systematic review with theoretical and methodological guidance for future research. Science of The Total Environment 2022 Nov.;847:157521. https://doi.org/10.1016/J.SCITOTENV.2022.157521

26. Klinenberg E. Social isolation, loneliness, and living alone: Identifying the risks for public health. American Journal of Public Health 2016;106:786–7. https://doi.org/10.2105/AJPH.2016.303166

27. Office for National Statistics. Measuring loneliness: guidance for use of the national indicators on surveys - Office for National Statistics [Internet]. 2018 [cited 2025 31]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/measuringlonelinessguidanceforuseofthenationalindicatorsonsurveys.

28. Administrative Data Research UK (ADR UK) – UKRI [Internet]. 2025 [cited 2025 31]; Available from: https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/administrative-data-research-uk-adr-uk/.

29. Welsh Government. Wellbeing of Wales: national indicators [Internet]. 2022 [cited 2025 31]; Available from: https://www.gov.wales/wellbeing-wales-national-indicators.

30. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke J-P, Ford D V, et al. The SAIL databank: linking multiple health and social care datasets. BMC Medical Informatics and Decision Making 2009 Dec.;9:3. https://doi.org/10.1186/1472-6947-9-3

31. Office for National Statistics. Population and household estimates, Wales - Office for National Statistics

[Internet]. 2022 [cited 2025 31]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimateswales/census2021.

32. Welsh Government. Welsh Index of Multiple Deprivation (WIMD): identifying groups of small areas based on deprivation indicators [Internet]. 2025 [cited 2025 31]; Available from: https://www.gov.wales/welsh-index-multiple-deprivation-wimd-identifying-groups-small-areas-based-deprivation-indicators-html.

33. Jones KH, Laurie G, Stevens L, Dobbs C, Ford D V, Lea N. The other side of the coin: Harm due to the non-use of health-related data. International journal of medical informatics 2017 Jan.;97:43–51. https://doi.org/10.1016/j.ijmedinf.2016.09.010.

34. Jones KH, Ford DV, Thompson S, Lyons R. A Profile of the SAIL Databank on the UK Secure Research Platform. International Journal of Population Data Science 2020 Oct.;4. https://doi.org/10.23889/ijpds.v4i2.1134

35. Ford D, Jones K, Verplancke J-P, Lyons R, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC Health Services Research 2009;9:157.

36. Johnson RD, Griffiths LJ, Hollinghurst JP, Akbari A, Lee A, Thompson DA, et al. Deriving household composition using population-scale electronic health record data - a reproducible methodology. PLOS ONE 2021 Mar.;16:e0248195. https://doi.org/10.1371/journal.pone.0248195

37. Rodgers SE, Lyons RA, Dsilva R, Jones KH, Brooks CJ, Ford D V, et al. Residential Anonymous Linking Fields (RALFs): a novel information infrastructure to study the interaction between the environment and individuals' health. Journal of Public Health 2009 Dec.;31:582–8. https://doi.org/10.1093/pubmed/fdp041

38. Welsh Government. National Survey for Wales [Internet]. 2025 [cited 2025 13]; Available from: https://www.gov.wales/national-survey-wales.

39. Office for National Statistics. Census [Internet]. Office for National Statistics 2021 [cited 2025 11]; Available from: https://www.ons.gov.uk/census.

40. Ordnance Survey. Classification scheme | OS Download Products' Documentation [Internet]. 2025 [cited 2025 13]; Available from: https://docs.os.uk/os-downloads/addressing-and-location/addressbase-fundamentals/classification-scheme.

41. Care Inspectorate Wales. Care Inspectorate Wales Home [Internet]. 2025 [cited 2025 13]; Available from: https://www.careinspectorate.wales/.

42. Welsh Government. Welsh Index of Multiple Deprivation (WIMD) 2019. 2019.

43. Office for National Statistics. Rural/Urban definition (England and Wales) - Office for National Statistics [Internet]. ONS National Archives 2016 [cited 2016 21]; Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/rural-urban-definition-and-la/rural-urban-definition–england-and-wales-/index.html.

44. Welsh Government. National Survey for Wales: quality report, January 2021 onwards [Internet]. 2021 [cited 2025 15]; Available from: https://www.gov.wales/national-survey-wales-quality-report-january-2021-onwards-html.

45. U.S. Geological Survey. NDVI, the Foundation for Remote Sensing Phenology [Internet]. 2020 [cited 2025 24]; Available from: https://www.usgs.gov/special-topics/remote-sensing-phenology/science/ndvi-foundation-remote-sensing-phenology.

46. Mah SM, Dasgupta K, Akbari A, Ross NA, Fry R. An international comparative study of active living environments and hospitalization for Wales and Canada. SSM - Population Health 2022 Jun.;18:101048. https://doi.org/10.1016/J.SSMPH.2022.101048

47. De Jong Gierveld J, Van Tilburg T. De Jong Gierveld 6-item Loneliness Scale [Internet]. 2006 [cited 2025 13]; Available from: https://www.nyc.gov/assets/dfta/downloads/pdf/about/dejong_gierveld_loneliness_scale.pdf.

48. de Jong-Gierveld J, Kamphuls F. The Development of a Rasch-Type Loneliness Scale. Applied Psychological Measurement 1985;9:289–99. https://doi.org/10.1177/014662168500900307

49. Franssen T, Stijnen M, Hamers F, Schneider F. Age differences in demographic, social and health-related factors associated with loneliness across the adult life span (19-65 years): a cross-sectional study in the Netherlands. BMC Public Health 2020 Aug.;20:1–12. https://doi.org/10.1186/S12889-020-09208-0/TABLES/2

50. Fierloos IN, Tan SS, Williams G, Alhambra-Borrás T, Koppelaar E, Bilajac L, et al. Socio-demographic characteristics associated with emotional and social loneliness among older adults. BMC geriatrics 2021 Feb.;21:114. https://doi.org/10.1186/s12877-021-02058-4

51. Von Soest T, Luhmann M, Hansen T, Gerstorf D. Development of Loneliness in Midlife and Old Age: Its Nature and Correlates. 2018; https://doi.org/10.1037/pspp0000219

52. Fierloos IN, van der Horst ML, Hoogendijk EO, van Groenou MIB. The association between loneliness and living arrangements in older adults: a comparison between emotional and social loneliness. BMC Geriatrics 2021;21:110. https://doi.org/10.1186/s12877-021-02058-4

53. Barreto M, Victor C, Hammond C, Eccles A, Richins MT, Qualter P. Loneliness around the world: Age, gender, and cultural differences in loneliness. Personality and Individual Differences 2021 Feb.;169:110066. https://doi.org/10.1016/j.paid.2020.110066

54. Luhmann M, Hawkley LC. Age differences in loneliness from adolescence to old age: A meta-analysis. European Journal of Ageing 2016;13:131–43. https://doi.org/10.1007/s10433-016-0361-4

55. Achterbergh L, Pitman A, Birken M, Pearce E, Sno H, Johnson S. The experience of loneliness among young people with depression: A qualitative meta-synthesis. BMC Psychiatry 2020;20:415. https://doi.org/10.1186/s12888-020-02818-3

56. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. Research on aging 2004 Nov.;26:655. https://doi.org/10.1177/0164027504268574

57. Lai KY, Kumari S, Webster C, Gallacher J, Sarkar C. Associations of neighbourhood housing density with loneliness and social isolation: a cross-sectional study using UK Biobank data. The Lancet 2021 Nov.;398:S60. https://doi.org/10.1016/S0140-6736(21)02603-9

58. Lyons J, Johnson RD, Edwards M, Turner S, Fry R, Griffiths LJ, Lyons RA. Administrative data linkage to Census 2021 in Wales, UK: A cross-sectional study examining completeness and representativeness for population analytics. Int J Popul Data Sci [Internet]. 2025 Nov 26;10(1). https://doi.org/10.23889/ijpds.v10i1.2994

Supplementary Table 1: Phase 1 Estimated unadjusted and adjusted odds ratios (OR), lower and upper 95% Confidence Intervals for risk of living alone by main effects (deprivation, housing type, region, and rurality), adjusted for age and sex interaction term (age-gender specific odds shown separately)

Term (main effects)	Unadjusted results				Adjusted results			
	OR	Lower CI	Upper CI	P-value	OR	Lower CI	Upper CI	P-value
Deprivation - 2	1.03	1.03	1.03	<0.000	1.09	1.09	1.10	<0.000
Deprivation - 3	0.96	0.95	0.96	<0.000	1.09	1.08	1.09	<0.000
Deprivation - 4	0.89	0.88	0.89	<0.000	1.05	1.04	1.05	<0.000
Deprivation - Least deprived	0.76	0.76	0.77	<0.000	0.96	0.96	0.96	<0.000
Rurality: Rural town and fringe	1.22	1.22	1.23	<0.000	1.06	1.06	1.07	<0.000
Rurality: Rural town and fringe in a sparse setting	0.98	0.98	0.98	<0.000	1.19	1.18	1.20	<0.000
Rurality: Rural village and dispersed	1.33	1.32	1.34	<0.000	1.05	1.04	1.05	<0.000
Rurality: Rural village and dispersed in a sparse setting	0.84	0.83	0.84	<0.000	1.11	1.11	1.12	<0.000
Rurality: Urban city and town in a sparse setting	0.95	0.95	0.96	<0.000	1.07	1.06	1.08	<0.000
Housing type: semi-detached	1.20	1.20	1.20	<0.000	1.33	1.33	1.33	<0.000
Housing type: terraced	1.40	1.40	1.41	<0.000	1.68	1.68	1.69	<0.000
Housing type: flat	5.80	5.78	5.83	<0.000	7.35	7.32	7.38	<0.000
Housing type: other	2.59	2.57	2.61	<0.000	3.00	2.98	3.02	<0.000
Housing type: missing	1.46	1.44	1.47	<0.000	1.58	1.57	1.60	<0.000
Region: Mid-Wales	1.16	1.16	1.17	<0.000	1.17	1.16	1.17	<0.000
Region: North Wales	1.07	1.07	1.07	<0.000	1.15	1.14	1.15	<0.000
Region: South-West Wales	1.11	1.11	1.11	<0.000	1.17	1.16	1.17	<0.000
Year	1.01	1.01	1.01	<0.000	1.01	1.00	1.01	<0.000

Supplementary Table 2: Phase 1 Estimated adjusted odds ratios for risk of living alone by age-gender interaction, adjusted for all main effects in Phase 1 model (deprivation, housing type, region, and rurality), reference group = 16-30 year old males

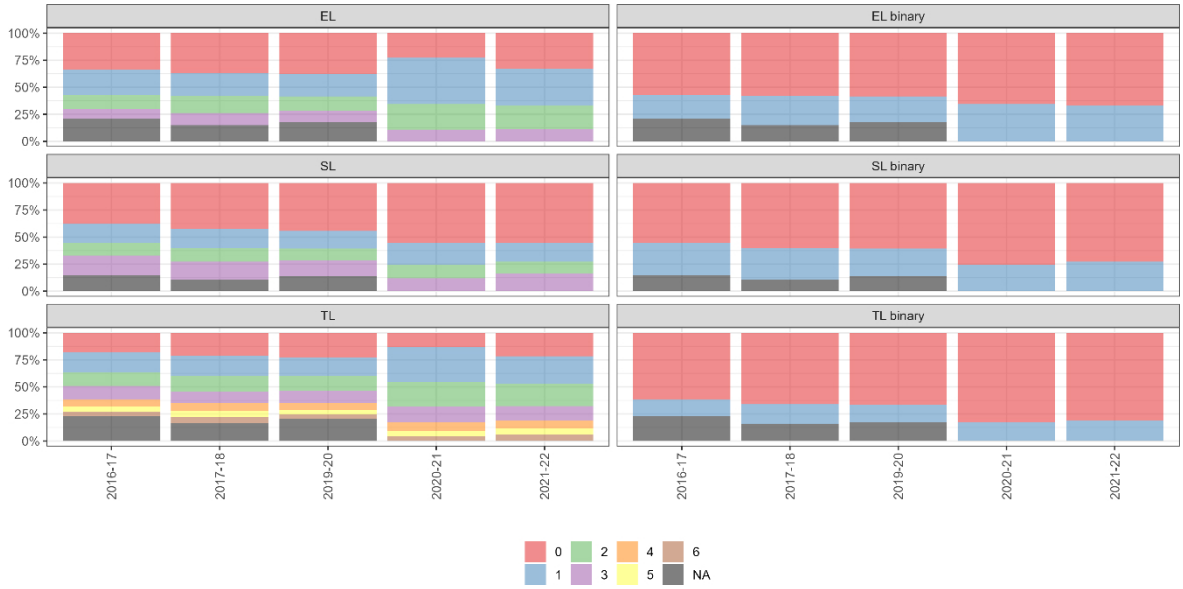
Gender	Reference group	16-30	31-45	46-60	61-75	76+
Male	Males aged	1.0	0.52	0.53	0.56	0.96
Female	16-30	0.8	0.40	0.63	1.72	4.19
Female to male odds ratio	Males within same age band	0.8	0.76	1.20	3.06	4.37

Within age female to male odds ratios included to show gender based risk for age band.

Supplementary Table 3: Phase 2 cohort construction

Stage	Description	Observations	Reduction
1	Initial data: National Survey for Wales	57,909	–
2	Excluded non-/low-confidence linkage matches (<90%)	52,836	5,073
3	Linked NSW to Annual Population Snapshot (APS, WDS living alone)	49,487	3,349
4	Excluded records with missing NSW covariates (education, health, long-term illness, housing tenure)	49,167	320
	Final analytical cohort (records (r) / individuals (n))	r=49,167 n=43,196	

Supplementary Figure 1: Dichotomised and full scores for each loneliness scale (EL – Emotional, SL – Social, and OL – Overall) for each wave of National Survey for Wales included in the study



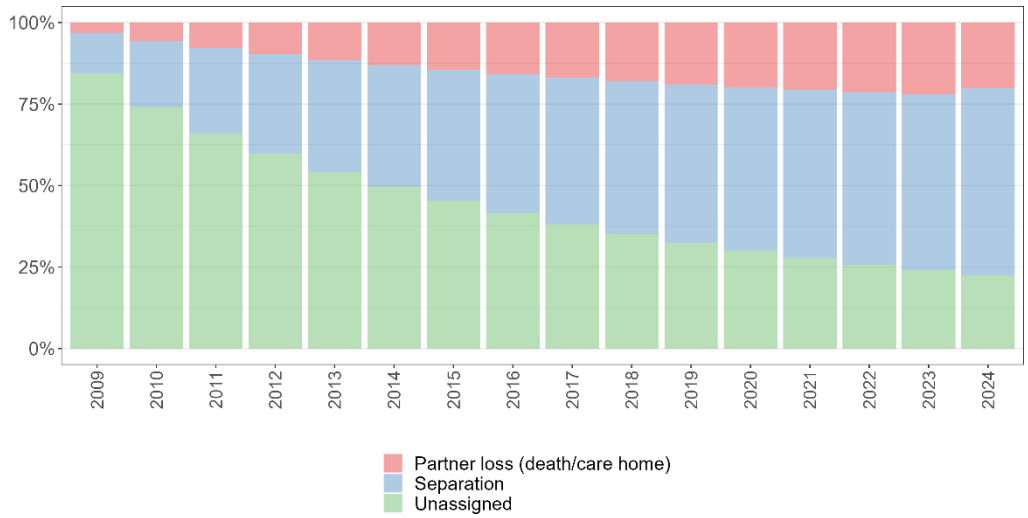
Notable change to missingness from 2020-21 wave onwards due to survey methodological changes.

Supplementary Table 4: Table of results of unadjusted and adjusted odds ratios (and lower and upper 95% Confidence Intervals) for final model main effects for emotional, social, and overall loneliness scores

Term	Emotional loneliness (Adjusted)	Social loneliness (Adjusted)	Overall loneliness (Adjusted)	Emotional loneliness (Unadjusted)	Social loneliness (Unadjusted)	Overall loneliness (Unadjusted)
Live alone reason: Partner loss ≤2y	5.1 (4.3-6.0)	1.2 (1.0-1.4)	2.0 (1.7-2.5)	4.1 (3.5-4.8)	0.9 (0.7-1.1)	1.5 (1.3-1.9)
Live alone reason: Partner loss ≤5y	3.2 (2.8-3.8)	1.2 (1.0-1.4)	1.8 (1.5-2.2)	2.7 (2.3-3.1)	0.8 (0.7-1.0)	1.2 (1.0-1.5)
Live alone reason: Partner loss >5y	2.1 (1.8-2.4)	1.3 (1.1-1.5)	1.7 (1.4-2.1)	1.8 (1.6-2.1)	0.9 (0.7-1.0)	1.2 (1.0-1.4)
Live alone reason: Separation	1.7 (1.6-1.8)	1.3 (1.2-1.4)	1.5 (1.4-1.7)	2.0 (1.9-2.1)	1.4 (1.3-1.5)	1.8 (1.7-1.9)
Live alone reason: Unassigned	1.5 (1.4-1.6)	1.4 (1.3-1.5)	1.5 (1.4-1.7)	1.5 (1.4-1.6)	1.2 (1.1-1.3)	1.4 (1.2-1.5)
Ethnicity: Minority ethnic background	1.5 (1.3-1.8)	1.5 (1.3-1.8)	1.7 (1.4-2.0)	1.5 (1.3-1.7)	1.7 (1.5-1.9)	1.9 (1.6-2.2)
Ethnicity: Not linked to census	1.1 (1.0-1.1)	1.0 (1.0-1.1)	1.0 (1.0-1.1)	1.3 (1.2-1.4)	1.1 (1.1-1.2)	1.3 (1.2-1.4)
Education: A-Levels or Apprenticeships	1.1 (1.0-1.3)	0.9 (0.8-1.0)	1.0 (0.9-1.1)	1.3 (1.2-1.4)	1.0 (0.9-1.0)	1.2 (1.1-1.3)
Education: GCSE	1.2 (1.1-1.2)	0.8 (0.8-0.9)	0.9 (0.8-1.0)	1.5 (1.4-1.6)	0.9 (0.9-1.0)	1.2 (1.1-1.3)
Education: Other qualifications	1.1 (1.0-1.2)	0.8 (0.7-0.9)	0.9 (0.8-1.1)	1.5 (1.4-1.7)	0.9 (0.8-0.9)	1.2 (1.1-1.3)
Education: No qualifications	1.3 (1.2-1.4)	0.6 (0.6-0.6)	0.8 (0.8-0.9)	2.1 (1.9-2.2)	0.7 (0.7-0.8)	1.3 (1.2-1.4)
Education: Not collected	1.6 (1.5-1.8)	0.9 (0.8-1.0)	1.1 (1.0-1.3)	1.6 (1.5-1.7)	0.6 (0.6-0.7)	0.9 (0.9-1.0)
Housing tenure: Private rental	1.4 (1.3-1.5)	1.3 (1.2-1.4)	1.5 (1.3-1.6)	1.8 (1.7-2.0)	1.6 (1.5-1.7)	2.1 (1.9-2.3)
Housing tenure: Social housing	1.7 (1.6-1.8)	1.5 (1.4-1.6)	1.7 (1.6-1.9)	2.9 (2.7-3.1)	1.8 (1.7-1.9)	2.9 (2.7-3.1)
Housing tenure: Other	1.1 (0.9-1.4)	0.9 (0.7-1.1)	1.1 (0.8-1.4)	1.4 (1.1-1.7)	0.9 (0.8-1.2)	1.3 (1.0-1.6)
Housing type: Semi-detached	1.0 (0.9-1.1)	0.9 (0.9-1.0)	0.9 (0.9-1.0)	1.4 (1.3-1.5)	1.1 (1.0-1.1)	1.3 (1.2-1.4)
Housing type: Terraced	1.1 (1.0-1.2)	1.0 (0.9-1.0)	1.0 (1.0-1.1)	1.7 (1.6-1.8)	1.2 (1.1-1.3)	1.7 (1.5-1.8)
Housing type: Flat	1.1 (0.9-1.2)	1.0 (0.9-1.1)	1.1 (0.9-1.2)	2.8 (2.6-3.1)	1.7 (1.6-1.9)	2.8 (2.5-3.1)
Housing type: Other/missing	1.0 (0.8-1.4)	1.1 (0.9-1.5)	1.3 (0.9-1.8)	1.7 (1.4-1.9)	1.3 (1.2-1.5)	1.9 (1.6-2.2)
Deprivation: Q2	1.0 (0.9-1.0)	1.0 (0.9-1.1)	1.0 (0.9-1.1)	0.8 (0.7-0.8)	0.9 (0.8-1.0)	0.8 (0.8-0.9)
Deprivation: Q3	0.9 (0.8-1.0)	1.0 (0.9-1.1)	0.9 (0.9-1.0)	0.7 (0.6-0.7)	0.8 (0.8-0.9)	0.7 (0.6-0.7)
Deprivation: Q4	0.9 (0.8-1.0)	1.0 (0.9-1.1)	0.9 (0.8-1.0)	0.5 (0.5-0.6)	0.8 (0.7-0.8)	0.6 (0.5-0.6)
Deprivation: Q5 - least deprived	0.8 (0.8-0.9)	0.9 (0.9-1.0)	0.8 (0.8-0.9)	0.4 (0.4-0.5)	0.7 (0.7-0.8)	0.5 (0.4-0.5)
Rurality: Town	0.9 (0.9-1.0)	1.0 (0.9-1.1)	0.9 (0.9-1.0)	0.9 (0.9-1.0)	1.0 (0.9-1.0)	0.9 (0.8-1.0)
Rurality: Village	0.9 (0.8-1.0)	1.0 (0.9-1.1)	1.0 (0.9-1.1)	0.7 (0.7-0.8)	0.9 (0.9-1.0)	0.8 (0.7-0.8)
General health: Fair	2.0 (1.9-2.1)	1.7 (1.6-1.8)	2.1 (2.0-2.2)	2.4 (2.3-2.5)	1.6 (1.6-1.7)	2.3 (2.1-2.4)
General health: Bad, very bad	3.9 (3.6-4.3)	2.3 (2.2-2.5)	3.9 (3.6-4.2)	5.4 (5.0-5.8)	2.4 (2.3-2.6)	4.9 (4.5-5.2)
Long-term illness: Yes	1.5 (1.4-1.6)	1.2 (1.1-1.2)	1.4 (1.3-1.5)	2.4 (2.3-2.5)	1.4 (1.3-1.4)	2.0 (1.9-2.2)
Long-term illness: Not collected	1.1 (1.0-1.2)	1.0 (0.9-1.1)	1.0 (0.9-1.1)	1.7 (1.6-1.8)	0.8 (0.8-0.9)	1.2 (1.1-1.3)
Greenness category: Q2	1.0 (0.9-1.0)	1.0 (0.9-1.1)	1.0 (0.9-1.1)	0.9 (0.9-1.0)	1.0 (0.9-1.0)	0.9 (0.8-1.0)
Greenness category: Q3	1.0 (0.9-1.0)	1.0 (0.9-1.1)	1.0 (0.9-1.0)	0.9 (0.8-1.0)	1.0 (0.9-1.0)	0.9 (0.8-1.0)
Greenness category: Q4 - most dense	1.0 (0.9-1.0)	1.1 (1.0-1.1)	1.0 (0.9-1.1)	0.8 (0.7-0.8)	1.0 (0.9-1.0)	0.8 (0.7-0.8)
Greenness category: Other/missing	1.2 (0.9-1.6)	1.0 (0.8-1.4)	1.0 (0.7-1.5)	1.0 (0.9-1.2)	1.1 (1.0-1.3)	1.1 (0.9-1.4)
Active living environment: Low	1.0 (0.9-1.1)	1.0 (0.9-1.0)	1.0 (0.9-1.1)	1.1 (1.0-1.1)	1.0 (1.0-1.1)	1.1 (1.0-1.1)
Active living environment: Moderate	1.1 (1.0-1.1)	1.0 (0.9-1.1)	1.1 (1.0-1.2)	0.8 (0.8-0.9)	1.0 (0.9-1.0)	0.9 (0.8-0.9)
Active living environment: High	1.0 (0.9-1.1)	1.0 (0.9-1.1)	1.0 (0.9-1.1)	1.1 (1.0-1.2)	1.1 (1.0-1.2)	1.2 (1.1-1.3)
Active living environment: Missing	1.0 (0.8-1.2)	0.9 (0.8-1.1)	0.8 (0.6-1.0)	1.1 (0.9-1.3)	0.9 (0.8-1.1)	0.9 (0.7-1.1)
Survey period: pre-Covid-19	0.9 (0.9-1.0)	1.6 (1.5-1.7)	1.2 (1.1-1.3)	0.9 (0.8-0.9)	1.6 (1.6-1.7)	1.2 (1.2-1.3)
Gender: Female	1.1 (0.9-1.2)	1.0 (0.9-1.1)	1.0 (0.9-1.2)	1.4 (1.3-1.4)	0.8 (0.8-0.9)	1.1 (1.1-1.2)
Gender: Female	1.1 (0.9-1.2)	1.0 (0.9-1.1)	1.0 (0.9-1.2)	1.2 (1.1-1.4)	1.1 (0.9-1.2)	1.2 (1.0-1.4)
Age group: 31-45	0.7 (0.6-0.8)	1.0 (0.9-1.2)	0.8 (0.7-0.9)	0.7 (0.7-0.8)	1.0 (0.9-1.0)	0.8 (0.7-0.9)
Age group: 31-45	0.7 (0.6-0.8)	1.0 (0.9-1.2)	0.8 (0.7-0.9)	0.8 (0.7-0.9)	1.1 (0.9-1.2)	0.8 (0.7-1.0)
Age group: 46-60	0.6 (0.5-0.7)	1.1 (1.0-1.3)	0.7 (0.6-0.8)	0.8 (0.7-0.9)	1.0 (0.9-1.1)	0.8 (0.8-0.9)
Age group: 46-60	0.6 (0.5-0.7)	1.1 (1.0-1.3)	0.7 (0.6-0.8)	0.8 (0.7-0.9)	1.2 (1.1-1.4)	0.9 (0.8-1.1)
Age group: 61-75	0.4 (0.3-0.4)	0.8 (0.7-0.9)	0.4 (0.4-0.5)	0.7 (0.6-0.7)	0.7 (0.7-0.8)	0.6 (0.5-0.6)
Age group: 61-75	0.4 (0.3-0.4)	0.8 (0.7-0.9)	0.4 (0.4-0.5)	0.6 (0.5-0.7)	0.9 (0.8-1.0)	0.6 (0.5-0.7)
Age group: 75+	0.4 (0.3-0.4)	0.5 (0.5-0.6)	0.3 (0.3-0.4)	0.8 (0.7-0.9)	0.5 (0.5-0.6)	0.5 (0.4-0.6)
Age group: 75+	0.4 (0.3-0.4)	0.5 (0.5-0.6)	0.3 (0.3-0.4)	0.7 (0.6-0.8)	0.6 (0.5-0.7)	0.5 (0.4-0.6)
Interaction: female, 31-45	1.0 (0.9-1.3)	0.9 (0.8-1.1)	1.1 (0.9-1.3)	0.9 (0.8-1.1)	0.8 (0.7-1.0)	1.0 (0.8-1.2)
Interaction: female, 46-60	1.2 (1.0-1.4)	0.8 (0.7-0.9)	1.0 (0.8-1.2)	1.0 (0.9-1.2)	0.7 (0.6-0.8)	0.9 (0.7-1.0)
Interaction: female, 61-75	1.5 (1.2-1.7)	0.8 (0.6-0.9)	1.1 (0.9-1.3)	1.3 (1.1-1.5)	0.7 (0.6-0.8)	1.0 (0.8-1.2)
Interaction: female, 75+	1.3 (1.1-1.5)	0.8 (0.6-0.9)	1.0 (0.8-1.3)	1.3 (1.1-1.5)	0.7 (0.6-0.9)	1.0 (0.8-1.2)

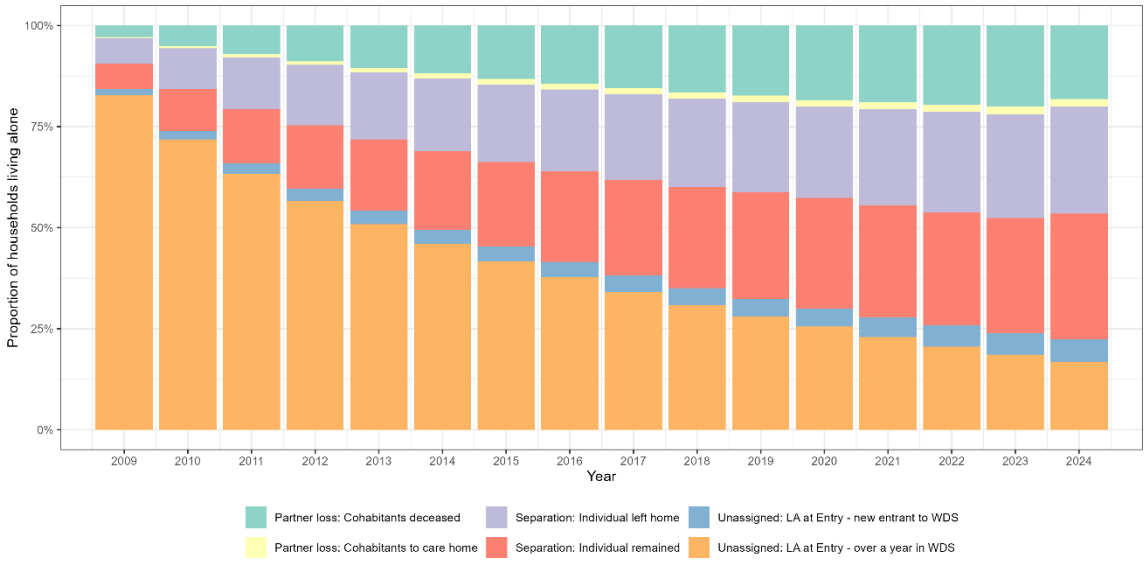
(Also adjusted for age-gender interaction terms).

Supplementary Figure 2: Phase 1 population level living alone reason for those living alone by year



Derived living alone reason annual trends; classified as partner loss (death or care home), separation, or unassigned. Distribution across the Phase 1 study period reflects increased ability to infer living alone reason in more recent years. In 2024, of 363,133 households classified as living alone, 20% (n=72,938) were due to partner loss, 58% (n=208,917) to separation, and 22% (n=81,278) unassigned.

Supplementary Figure 3: Phase 1 population level living alone reason for those living alone by year



Derived living alone reason annual trends with additional granularity; defined as: partner loss: care home entry and death; separation: leaving a previous residence, remaining.

Supplementary Table 5: Results of sensitivity analyses for emotional loneliness model using cluster-robust standard errors clustered at the individual level to assess the impact of repeated observations on statistical inference

Model term (Emotional loneliness model)	Original model estimate (adjusted OR)	Clustered model estimate (adjusted OR)	Original model standard error	Clustered model standard error	Ratio of standard errors	Original model p-value	Clustered model p-value
(Intercept)	0.17	0.17	0.08	0.09	0.97	0.00	0.00
Live alone reason: loss of partner (under 2 years)	5.09	5.09	0.08	0.09	0.95	0.00	0.00
Live alone reason: loss of partner (2-5 years)	3.24	3.24	0.08	0.08	0.94	0.00	0.00
Live alone reason: loss of partner (5+ years)	2.05	2.05	0.08	0.08	0.93	0.00	0.00
Live alone reason: separation	1.68	1.68	0.04	0.04	0.97	0.00	0.00
Live alone reason: Unknown	1.49	1.49	0.04	0.05	0.95	0.00	0.00
Gender: Female	1.08	1.08	0.07	0.08	0.98	0.30	0.31
Age group: 31-45	0.69	0.69	0.07	0.08	0.97	0.00	0.00
Age group: 46-60	0.58	0.58	0.07	0.07	0.97	0.00	0.00
Age group: 61-75	0.37	0.37	0.07	0.07	0.97	0.00	0.00
Age group: 75+	0.37	0.37	0.08	0.08	0.97	0.00	0.00
Ethnicity: Minority ethnic background	1.50	1.50	0.08	0.09	0.95	0.00	0.00
Ethnicity: Not linked to census	1.08	1.08	0.03	0.03	0.96	0.03	0.04
Education: A-Levels or Apprenticeships	1.15	1.15	0.05	0.05	1.01	0.00	0.00
Education: GCSE	1.15	1.15	0.04	0.04	1.02	0.00	0.00
Education: No qualifications	1.25	1.25	0.04	0.04	1.01	0.00	0.00
Education: Not collected in survey	1.61	1.61	0.05	0.05	1.01	0.00	0.00
Education: Other qualifications	1.08	1.08	0.05	0.06	0.99	0.15	0.15
Tenure: Other	1.11	1.11	0.11	0.11	0.98	0.35	0.36
Tenure: private rental	1.41	1.41	0.04	0.04	0.98	0.00	0.00
Tenure: social housing	1.71	1.71	0.04	0.04	0.97	0.00	0.00
General health: bad, very bad	3.93	3.93	0.04	0.04	0.98	0.00	0.00
General health: fair	2.01	2.01	0.03	0.03	0.98	0.00	0.00
Long-term illness: Yes	1.53	1.53	0.03	0.03	1.00	0.00	0.00
Long-term illness: Not collected in survey	1.06	1.06	0.05	0.05	0.99	0.22	0.22
Deprivation: 2	0.96	0.96	0.04	0.04	0.98	0.25	0.26
Deprivation: 3	0.92	0.92	0.04	0.04	0.96	0.03	0.04
Deprivation: 4	0.88	0.88	0.04	0.05	0.96	0.00	0.01
Deprivation: 5 - least deprived	0.82	0.82	0.05	0.05	0.98	0.00	0.00
Rurality: town	0.95	0.95	0.04	0.04	0.97	0.15	0.17
Rurality: village	0.92	0.92	0.05	0.05	0.97	0.12	0.13
House type: flat	1.06	1.06	0.06	0.06	0.95	0.31	0.33
House type: other or missing	1.02	1.02	0.15	0.15	0.97	0.88	0.88
House type: semi-detached	1.02	1.02	0.04	0.04	0.98	0.67	0.68
House type: terraced	1.09	1.09	0.04	0.04	0.96	0.02	0.02
Greenness category: 2	0.97	0.97	0.04	0.04	0.96	0.39	0.40
Greenness category: 3	0.97	0.97	0.04	0.04	0.97	0.37	0.38
Greenness category: most dense	0.96	0.96	0.04	0.04	0.96	0.31	0.33
Greenness category: other or missing	1.19	1.19	0.16	0.16	0.99	0.27	0.27
Active living environment: low	1.01	1.01	0.03	0.03	0.98	0.79	0.80
Active living environment: moderate	1.06	1.06	0.04	0.04	0.96	0.16	0.18
Active living environment: high	1.04	1.04	0.05	0.05	0.97	0.49	0.50
Active living environment: missing	0.97	0.97	0.09	0.10	0.97	0.73	0.74
Survey period: pre-April 2020	0.94	0.94	0.03	0.03	1.00	0.04	0.04
Interaction: female, 31-45	1.05	1.05	0.09	0.10	0.98	0.62	0.62
Interaction: female, 46-60	1.15	1.15	0.09	0.09	0.98	0.11	0.11
Interaction: female, 61-75	1.47	1.47	0.09	0.09	0.97	0.00	0.00
Interaction: female, 75+	1.28	1.28	0.10	0.10	0.97	0.01	0.01

Absolute differences are presented between models for key model metrics, i.e. adjusted odds ratios (estimates), standard errors (including ratios), and p-values.

Supplementary Table 6: Comparison of age and sex distributions between linked and non-linked participants at key stages of cohort construction

Stage comparison description	Group comparison rule	Percentage female	Mean age (standard deviation)
Stage 1: Initial data: National Survey for Walesversus Stage 2: Excluded non & low-confidence (<90%) linkage matches	Stage 1 and not Stage 2 Stage 2 only	48.0% 56.3%	54.2 (20.7) 56.0 (18.3)
Stage 1: Initial data: National Survey for Walesversus Stage 3: Linked NSW to Annual Population Snapshot (APS, WDS living alone)	Stage 1 and not Stage 3 Stage 3 only	49.8% 56.5%	53.5 (20.2) 56.2 (18.2)
Stage 1: Initial data: National Survey for Walesversus Stage 4: Excluded records with missing NSW covariates (education, health, long-term illness, housing tenure)	Stage 1 and not Stage 4 Stage 4 only	50.0% 56.6%	53.5 (20.3) 56.2 (18.2)