Methodology

1. Data and data linkage

Identifying unpaid carers

Unpaid carers were anonymously identified using local authority (LA) carer assessment data and general practice (GP) data sources.

Local authority data

Adult carers' assessment data was obtained from three Welsh LAs: Neath Port Talbot, Swansea, and Denbighshire. This LA data was uploaded into the Secure Anonymised Information Linkage (SAIL) Databank Trusted Research Environment (TRE) as a Core Restricted data source, using the split file process with SAIL's trusted third-party partner organisation Digital Health and Care Wales (DHCW) (1). The data items provided by each LA are listed in Appendix A.

GP data

The Welsh Longitudinal General Practice (WLGP) data source is held within the SAIL Databank. Read codes used to identify unpaid carers in this data source are listed in Appendix B (2).

Data linkage

Unpaid carer records were linked to anonymised individual-level, routinely-collected, population-scale electronic health record (EHR) and administrative data sources available within the SAIL Databank (3–5) (Appendix C) using an anonymised linkage field (ALF) to determine aspects of health and health-care services usage.

2. Definitions

Study period

The study period was defined per LA and reflected the date range of carers' assessment data available[1] (Table 1).

Table 1: Study period by LA.

Local Authority	Study period
Neath Port Talbot (NPT)	07/2017 - 05/2022
Swansea	04/2021 - 06/2022
Denbighshire	04/2020 – 03/2022

Inclusion criteria

1. Aged 18 and above.
2. Have a valid sex code.
3. Have a valid week of birth (WOB).
4. Residing in the LA of interest.

---

[1] For Neath Port Talbot, the data shared included all carers' assessments completed within the defined study period which were also linked to client assessments completed within this period. Internal sensitivity analysis conducted by NPT indicated that approximately 50 carers assessments per year were completed where no linked client assessment was available.

5. Resident at a Welsh address for a full year prior to identification date.
6. Registered with a Welsh GP providing data to SAIL for a full year prior to identification date.
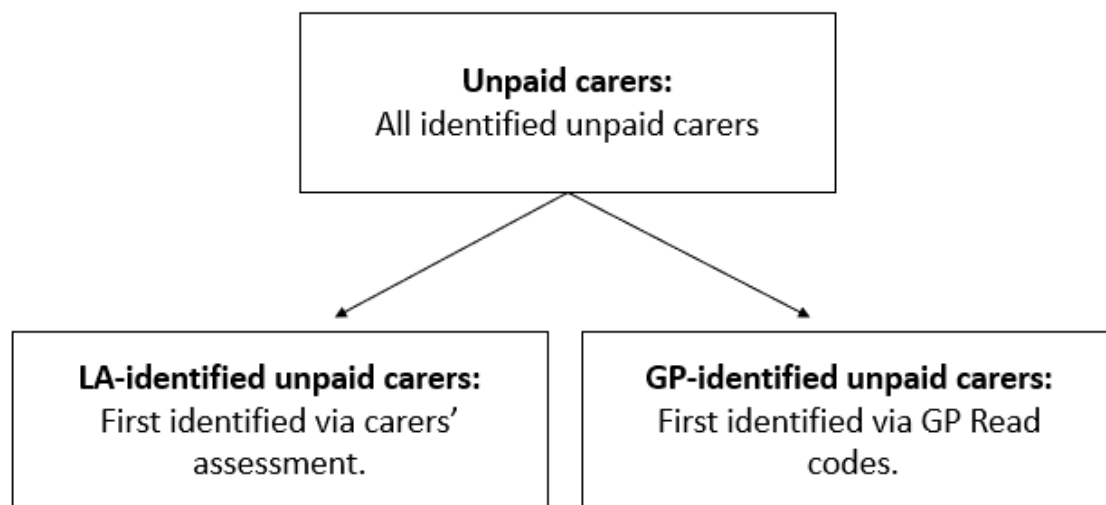7. Probabilistic linkage match of 90% or above.

Index date

The index date was defined as the earliest identification date as an unpaid carer within the study period where the inclusion criteria were met.

3. Cohorts

Unpaid carer cohorts

Three unpaid carer cohorts were created for each LA:

The sizes of the total cohorts for NPT, Swansea and Denbighshire were 2,024, 926 and 841, respectively. Table 2 shows the total number of unique LA records at each data processing step.

Table 2: Application of exclusion criteria for LA data. (Count for Swansea demographic criteria not met is included in identified earlier in GP)*

	NPT	Swansea	Denbighshire
Initial records	1,513	548	585
Duplicate records	576	0	135
Linkage issues	31	113	41
Client assessment outside study period	31	NA	NA
Carer assessment outside study period	340	NA	NA
Demographic criteria not met	22	*	33
Death before assessment	0	0	0
Not resident in Wales/not registered at a SAIL GP for 1 year prior to identification	73	115	147
Identified earlier in GP data	27	39*	16
LA-identified unpaid carers	413	281	213

Matched non-carers

A matching pool was created of individuals who were resident in the LAs of interest but were not identified in the unpaid carer cohort, nor able to be identified as an unpaid carer via Read codes during the two years prior to study start date. Using this pool and the R Library 'Matchit', a list of matches of non-carers for each unpaid carer on sex, age-group and local authority at study start date was generated (6). The index date of each unpaid carer was imputed to their matched non-carers as a pseudo-index date. The inclusion criteria were then re-applied to this list of matches based on the pseudo-index date, and a non-carer was selected at random from the remaining matches, resulting in a 1:1 exact match of carers and non-carers. We verified that the age and sex distribution of the unpaid carers and matched non-carers were similar using cross-tabulation functions in R (7).

4. Measures

Demographics

Demographic characteristics were defined as follows:

- Age group: Under 40 (i.e. 18-39), 40-49, 50-59, 60-69, 70-79, 80+.
- Sex: male, female.
- Welsh Index of Multiple Deprivation 2019 (WIMD) quintile (8): 1 (most deprived), 2, 3, 4, 5 (least deprived).
- Ethnic group (using methodology developed within SAIL Databank (9) and the 5 ONS Categories (10)): 1) Asian, Asian British or Asian Welsh, 2) Black, Black British, Black Welsh, Caribbean or African, 3) Mixed or Multiple ethnic groups, 4) White, 5) Other.

The age, sex, and deprivation quintiles were compared against the ONS 2020 mid-year estimates of adults aged 18 and above in the respective LAs to gain an understanding of the unpaid carers within the context of the local population (11). At the time of analysis, the full breakdown of 2021 mid-year estimates was not publicly available.

Long-term conditions (LTCs)

LTCs were identified using Read codes in the WLGP data source. Conditions of interest were defined by version 1.1 of the Cambridge Multimorbidity Score (12,13). Read code alternatives of conditions listed in the Cambridge Multimorbidity Score were used (14,15). The look-back period for each condition was as defined by Cassell (13). The full list of LTCs used in the analysis can be found in Appendix D. Chronic painful conditions should be interpreted with caution as they rely on prescription codes for identification, and pain medications can be used for other purposes (12). Multimorbidity rates (defined as two or more conditions) and most common conditions per cohort were evaluated.

Health Service Use

Health service use was determined over the year prior to index date. For ED attendances, inpatient admissions (both planned and emergency) and outpatient attendances, rates of the number of people with one or more attendances or admissions per 1,000 population were calculated. For GP interactions, the rate of interactions per person was calculated. Crude rates were compared between unpaid carers /non-carers and LA-identified/GP-identified using rate ratios and 95% confidence intervals. Additionally, age-sex standardised rates and rate ratios were calculated and compared for LA-identified and GP-identified using the ONS 2020 mid-year adult population estimates.

Carer-client relationship (Only included in feedback to Neath Port Talbot)

Data on carer-client relationships was only available in NPT carers' assessment data. Relationships were grouped into categories as described in Appendix E and evaluated as a proportion of all relationships. We also calculated the ratio of carers to clients and clients to carers.

5. Statistical analysis

This analysis strictly observed the statistical disclosure control (SDC) criteria established by the SAIL Databank TRE and data providers, including the Office for National Statistics (ONS) (16,17).

Differences in distribution of sex and deprivation between the LA and GP carers cohorts were tested using Chi-square and Mann Whitney U tests. T-tests were used to determine statistical significance of differences in age between the LA and GP carers cohorts. Rate ratios and confidence intervals were used to test differences in multimorbidity and health service use.

6. References

1. DIGITAL HEALTH AND CARE WALES PROVIDING THE NATIONAL TECHNOLOGY AND DATA SERVICES NEEDED BY HEALTH AND CARE PROFESSIONALS AND THE PEOPLE OF WALES HELPING TO ACHIEVE BETTER HEALTH AND WELL-BEING.

2. Huang F, Song J, Davies AR, Anderson C, Bentley L, Carter B, et al. Unpaid carers in Wales: The creation of an e-cohort to understand long-term health conditions amongst unpaid carers in Wales Suggested Citation Acknowledgements. 2021 [cited 2022 Sep 1]; Available from: www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

3. Home - SAIL Databank [Internet]. [cited 2022 Dec 16]. Available from: https://saildatabank.com/

4. Ford D v., Jones KH, Verplancke JP, Lyons RA, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC Health Serv Res [Internet]. 2009 [cited 2022 Dec 16];9:157. Available from: /pmc/articles/PMC2744675/

5. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke JP, Ford D V., et al. The SAIL databank: linking multiple health and social care datasets. BMC Med Inform Decis Mak [Internet]. 2009 [cited 2022 Dec 16];9(1). Available from: https://pubmed.ncbi.nlm.nih.gov/19149883/

6. matchit function - RDocumentation [Internet]. [cited 2023 Apr 6]. Available from: https://www.rdocumentation.org/packages/MatchIt/versions/3.0.2/topics/matchit

7. xtabs function - RDocumentation [Internet]. [cited 2023 Apr 6]. Available from: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/xtabs

8. Welsh Index of Multiple Deprivation (full Index update with ranks): 2019 | GOV.WALES [Internet]. [cited 2023 Jan 9]. Available from: https://www.gov.wales/welsh-index-multiple-deprivation-full-index-update-ranks-2019

9. Akbari A, Torabi F, Bedston S, Lowthian E, Abbasizanjani H, Fry R, et al. Developing a research ready population-scale linked data ethnicity-spine in Wales. medRxiv [Internet]. 2022 Nov 29 [cited 2023 Feb 27];2022.11.28.22282810. Available from: https://www.medrxiv.org/content/10.1101/2022.11.28.22282810v1

10. Ethnic group, England and Wales - Office for National Statistics [Internet]. [cited 2023 Jan 24]. Available from:

https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021

11. Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland - Office for National Statistics [Internet]. [cited 2023 Feb 22]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland

12. CPRD @ Cambridge - Codes Lists (GOLD) - Primary Care Unit [Internet]. [cited 2023 Jan 9]. Available from: https://www.phpc.cam.ac.uk/pcu/research/research-groups/crmh/cprd_cam/codelists/v11/

13. Cassell A, Edwards D, Harshfield A, Rhodes K, Brimicombe J, Payne R, et al. The epidemiology of multimorbidity in primary care: a retrospective cohort study. British Journal of General Practice [Internet]. 2018 Apr 1 [cited 2023 Jan 9];68(669):e245–51. Available from: https://bjgp.org/content/68/669/e245

14. Hanlon P, Jani BD, Nicholl B, Lewsey J, McAllister DA, Mair FS. Associations between multimorbidity and adverse health outcomes in UK Biobank and the SAIL Databank: A comparison of longitudinal cohort studies. PLoS Med [Internet]. 2022 Mar 1 [cited 2023 Jan 9];19(3). Available from: /pmc/articles/PMC8901063/

15. Payne RA, Mendonca SC, Elliott MN, Saunders CL, Edwards DA, Marshall M, et al. Development and validation of the Cambridge Multimorbidity Score. CMAJ [Internet]. 2020 Feb 3 [cited 2023 Jan 9];192(5):E107–14. Available from: /pmc/articles/PMC7004217/

16. Statistical disclosure control - Office for National Statistics [Internet]. [cited 2022 Dec 29]. Available from: https://www.ons.gov.uk/aboutus/transparencyandgovernance/datastrategy/datapolicies/statisticaldisclosurecontrol#policy-detail

17. Parker S, Scott J, Welpton R, Woods C. The Health Foundation) Arne Wolters (The Health Foundation). 2019;

7. Appendices

Appendix A: Unpaid carer data items by LA

LA_Table 1: Identifiable data Item	Swansea	Neath Port Talbot	Denbighshire
Carer ID	Yes	Yes	Yes
Carer first name	Yes	Yes	Yes
Carer surname	Yes	Yes	Yes
Carer address	Yes	Yes	Yes
Carer postcode	Yes	Yes	Yes
Carer date of birth	Yes	Yes	Yes
Carer NHS number	Yes	Yes	Yes
LA_Table 2: Identifiable data Item	**Swansea**	**Neath Port Talbot**	**Denbighshire**
Care recipient ID	No	Yes	No
Care recipient first name	No	Yes	No
Care recipient surname	No	Yes	No
Care recipient address	No	Yes	No
Care recipient postcode	No	Yes	No
Care recipient date of birth	No	Yes	No
Care recipient NHS number	No	Yes	No
LA_Table 3: Non-identifiable data Item	**Swansea**	**Neath Port Talbot**	**Denbighshire**
Carer ID*	Yes	Yes	Yes
Care recipient ID	No	Yes	No
Carer relationship to recipient	No	Yes	No
First carer assessment date (as available in data)	Yes	Yes	Yes
Last carer assessment date	Yes	Yes	No
Number of carer assessments completed	Yes	Yes	No

Appendix B: Unpaid carer Read codes

Row Number	Read Code	Description
1	918A.	Carer
2	918A0	Cares for a friend
3	918A1	Cares for a neighbour
4	918A2	Cares for a relative
5	918G.	Is a carer
6	918H.	Primary carer
7	918W.	Carer of a person with learning disability
8	918a.	Carer of a person with substance misuse
9	918d.	Carer of a person with mental health problem

10	918m.	Carer of a person with a terminal illness
11	918t.00	Carer from Black and minority ethnic group
12	918Y.	Carer of a person with sensory impairment
13	8IHE.	Carer health check declined
14	8O7..	Carer support
15	918b.	Carer of a person with alcohol misuse
16	8IEP.	Carer annual health check declined
17	8BAr.	Carer health check completed
18	13Wb.	Carer has sole parental responsibility
19	9NSS.	Carer health check offered
20	918J.	Carer - home telephone number
21	9Ngw.	Carer does not understand care plan
22	918W.	Carer of a person with learning disability
23	918y.	Carer of person with dementia
24	9Ngv.	Carer understands care plan
25	918X.	Carer of a person with physical disability
26	13VN.	Carer able to cope
27	69DC.	Carer annual health check
28	918M.	Carer - email address
29	9180.	Carer's details (also called 9180)
30	69DE.	Carer health check
31	918L.	Carer - mobile telephone number
32	918K.	Carer - work telephone number
33	388Q.	Carer strain index score
34	9d46.	Carer
35	918c.	Carer of a person with chronic disease
36	8HkA.	Ref for GP carer's assessment
37	8Hkb.	Referral for social services carer's assessment
38	13HH.00	Looks after chronically sick relative
39	13HH.	Looks after elderly dependent
40	13HH.11	Cares for mentally handicapped dependent
41	13HH.12	Looks after chronically sick father

42	13HH.13	Looks after chronically sick husband
43	13HH.14	Looks after chronically sick mother
44	13HH.15	Looks after chronically sick spouse
45	13HH.16	Looks after chronically sick wife
46	13HH.18	Looks after physically handicapped dependent
47	13VP0.	Receiving carer allowance
48	91800	Details of informal carer
49	918t.	Carer from Black and minority ethnic group
50	9NSW	Assessment of needs offered to carer
51	8IAk.	Referral to Princess Royal Trust carers centre declined

Appendix C: List of data sources used in SAIL

Dataset	Description	Dated provision version
Annual District Death Extract (ADDE)	From the Office for National Statistics (ONS) mortality register, containing data on all deaths in Wales within the study period	01/07/2022
Welsh Demographic Service Dataset (WDSD)	Contains linkable data on demographics of the Welsh population from the Census	04/07/2022
Wales Longitudinal General Practice (WLGP)	Patients' GP events in Wales	01/03/2022
Emergency Department Data Set (EDDS)	All emergency department attendances in Wales	01/06/2022
Patient Episode Database for Wales (PEDW)	All inpatient admissions in Wales	04/07/2022
Out-Patient Database for Wales (OPDW)	All outpatient appointments in Wales	01/05/2022

Appendix D: List of Long-term conditions

Obtained from source: Table A3.2 in https://www.cmaj.ca/content/cmaj/suppl/2020/01/28/192.5.E107.DC1/190757-res-3-at.pdf. Original study(15).

ID	Morbidity
Morbidities based on Read code ever recorded	
1	Alcohol Problems
2	Anorexia or bulimia
3	Atrial fibrillation
4	Blindness and low vision
5	Bronchiectasis
6	Chronic liver disease and viral hepatitis
7	Chronic sinusitis
8	COPD
9	Coronary heart disease
10	Dementia
11	Diabetes
12	Diverticular disease of intestine
13	Hearing loss
14	Heart failure
15	Hypertension
16	Inflammatory bowel disease
17	Learning disability
18	multiple sclerosis
19	Parkinson's disease
20	Peptic ulcer disease
21	Peripheral vascular disease
22	Prostate disorders
23	Psychoactive substance misuse (not alcohol)
24	Rheumatoid arthritis, other inflammatory polyarthropathies & systematic connective tissue disorders
25	Stroke & transient ischaemic attack

26	Thyroid disorders

Morbidities based on prescription in last 12 months

27	Constipation (treated)
28	Migraine

Morbidities based on combination of Read code ever recorded

and/or prescription in last 12 months

29	Epilepsy (currently treated)
30	Asthma (currently treated)
31	Irritable bowel syndrome
32	Psoriasis or eczema

Morbidities otherwise defined

33	Anxiety & other neurotic, stress related & somatoform disorders OR depression
34	Cancer - [New] Diagnosis in last 5 years (excluding non-melanoma skin cancer)
35	Chronic kidney disease
36	Painful condition
37	Schizophrenia (and other non-organic psychosis) or bipolar disorder

Appendix E: Carer-Client Relationship Categories

Relationship Category	Relationships Included
Partner	Married Partner Unmarried Partner
Child	Daughter Daughter-in-law Son Son-in-law
Parental	Father Foster Parent Mother Parent's Partner Step-Parent
Sibling	Brother Brother-in-law Sister Sister-in-law
Other	Extended Family (grandchild, cousin, niece or nephew) Friend Neighbour Non-specified family member

	Non-specific relationship