Version 1.0

The Networked Data Lab: Analysis plan for Topic 3: Unpaid Carers

Satellite analysis for NDL Leeds Lab

Ben Alcock, Alex Brownrigg, Helen Butters, Frank Wood

24/05/2022

v0.1

Contents

1. Research rationale and objectives

1.1 Rationale:

The Parliamentary Family Resources Survey (2021) estimated that in 2019/20 around 7% of the UK population were providing unpaid care, although their definition is 'informal care'. Similarly, the Carer's Week Research Report (2020) estimated that around 1 in 4 people in the UK could be providing some level of unpaid care, with around 74,000 estimated unpaid carers in the Leeds City area (2011 Census/Carers Leeds, 2020).

The needs of unpaid carers are known to be complex and varied, with surveys finding that around 83% of respondents reported a negative impact of caring on their physical health, and 87% responding negative impacts on their mental health (In Sickness and in Health, 2012). With so many people providing care across Leeds and such high levels of reported needs nationally, it is important to investigate the full health and social-care needs across Leeds, and identify potential gaps and inequalities in service use.

The Leeds NDL team has conducted and commissioned research to establish gaps in the knowledge of both the Leeds population and Leeds CCG and Leeds City Council teams regarding the needs of unpaid carers, and the utilisation of services by unpaid carers. Priorities for unpaid carers were determined initially via two insight reports, one conducted within team and one commissioned separately. A Task and Finish group comprised of professionals (council and third sector) who work with and on behalf of unpaid carers, and unpaid carers from different economic and ethnic backgrounds honed the key themes of the insight report, and pulled out questions on the most important topics both to carers themselves and to the council for service design.

1.2 Objectives:

Two key questions have been highlighted. These are:

Q1) What are the healthcare outcomes of unpaid carers, how does this compare with similar non-carers, and what is the impact of council services on health outcomes (such as respite services, financial aid)?

Q2) How equitably are support services being used? Are carers from all backgrounds aware of services, requesting access to services, and receiving services, and are outcomes proportional across all backgrounds or are there signs of different areas which present higher needs?

A number of additional supplementary questions have also been raised, with reduced feasibility and impact due to shortage of data in some areas. These are:

Q3) How well are health and social care services aware of carers, and what, if any, differences are there in the rate of carers recorded across different demographic backgrounds?

Q4) When do people stop providing care? Is it fully due to the circumstances surrounding the cared-for person, or are there any healthcare markers which are common indicators that show when someone is unable to provide care? Similarly, if such indicators exist, are there any health or social care pathways which feature these indicators but do not result in the person stopping care responsibilities?

Q5) What is the impact of providing care on a person's finances?

2. Data and data linkages

For this study, our analysis will rely on the following data sources:

- Primary care data (from EMIS & TPP), featuring patient demographics, appointments, events, and prescriptions. Key clinical data relating to patient conditions, as well as demographic information and patient carer status is recorded via both Read code (EMIS: v2, TPP: v3) and SNOMED codes, and will be used to both identify unpaid carers and to match patients with similar profiles. This dataset covers the entire Leeds registered patient population (both currently registered and previously registered, for the duration of each patient's registration to a Leeds GP Practice). After onboarding, this dataset will also contain a household pseudonym which will be used to identify people living with others who feature conditions likely requiring carer support. This data set is externally linkable via pseudonymised NHS number and internally linked through unique patient ID.
- Adult social care data extracts (from CIS), featuring referrals, reviews, assessments, and service provisions. After data onboarding, this will also contain flags for recorded carers, and will contain a link between a carer and cared-for person where it is known. This covers all people who have applied to Leeds City Council for access to adult social care services. This data set is externally linkable via pseudonymised NHS number, and internally linked through unique ASC ID.
- Secondary Uses Service (SUS), containing inpatient attendances, outpatient appointments, and A&E visits. This data set is linkable internally and externally via pseudonymised NHS number.
- Mental Health Services Data Set (MHSDS) for all non-IAPT mental health referrals and services accessed. This data set is externally linkable via pseudonymised NHS number and internally linked through unique patient ID.
- Improving Access to Psychological Therapies services (IAPT) for treatments relating to anxiety and depression disorders. This data set is externally linkable via pseudonymised NHS number and internally linked through unique patient ID.
- Yorkshire Ambulance Service (YAS) data containing 111 and 999 calls, primarily for investigation into use of services regarding to mental health needs. This data is linkable via pseudonymised NHS Number.
- External open data sources and/or APIs for the Index of Multiple Deprivation (IMD), clinical coding (Read/SNOMED lookups), population data (ONS census and mid-year estimates). This data is not linkable on a person-level, and so is usable at geographic level (LSOA/MSOA).

All internal data sets will be accessed from the Leeds Data Model, while external open data comes from a variety of sources listed above.

3. Statistical methods

3.1 Study design

This is an observational case-control study using routinely gathered data from health and social care systems.

3.2 Study period

Data for this study will cover 2016/04/01 – 2021/03/31. Primary care, secondary care, mental health (non-IAPT), and adult social care data covers this full range. IAPT referrals cover this full range but recorded appointments only start from 2016/06/01. Similarly, while 999 data covers the full range, consistent recording of NHS number does not start until 2018/04/01.

Council web analytics data is from October 2020, and DWP open data (carer's allowances) are from 2018 for 2011 LSOA boundaries.

3.3 Study population

The population will be split into different cohorts:

- One all-adult group, covering all 18+ people registered to GP practices in Leeds.
- One "GP Carers" group, consisting of all 18+ people registered to Leeds Practices who have been identified from primary care events as unpaid carers.
- One "Adult Social Care Carers" group, consisting of all people recorded in ASC dataset as being a carer/having a carer's assessment/using respite services.

After onboarding, supplementary questions will also look at:

- One "household care required" group, consisting of adults (18+) who live in a household with at least one other person requiring carer support.

Census data and ONS population estimators will also be used for comparisons on geographic levels.

3.4 Definitions of outcomes and exposures

The following outcomes and exposures will be used in this study.

Outcomes

- Q1a) Health conditions
    - This will look at the prevalence of conditions (including long-term conditions and shorter term conditions such as stress) of registered patients, using primary care records, comparing carers and non-carers.
- Q1b) Healthcare utilisation
    - Primarily looking into use of same-day/out-of-hours services, and non-elective and crisis services to determine unmet needs. This will be drawn from SUS (inpatient/A&E), MHSDS (crisis services), and primary care (same day appointments and out-of-hours services).
- Q2a) Knowledge of council services
    - Using the council analytics on page views around carers and ASC referrals.
- Q2b) Use of council services
    - Using ASC referrals, assessments, and service provisions.
- Q3) Probability of a person to be recorded as a carer a) in GP records, b) in ASC records
    - Using primary care and ASC records.
- Q4) Probability that someone will have a GP appointment containing a "no longer a carer"-type Read/SNOMED code, or will list an end date for care in ASC
    - Using primary care events and ASC records
- Q5) Proportion of people recording unemployment or featuring GP appointments relating to financial problems

Exposures

- Q1a) Segmenting by age, sex, ethnic background, and resident deprivation level (removing confounding health deprivation indices). Comparisons between carer and non-carer population will be made, and comparisons with ONS population estimates will also be made, to account for possible overreporting resident numbers from GP registrations.

- Q1b) Segmenting as in 1a), but also considering the prevalence of related conditions.
- Q2a) No segmentation will be possible due to lack of available data on information.
- Q2b) Segmented as in 1a)
- Q3a & b) Segmented as in 1a) – once household data onboarded also segmented using indicators of need for carer within household.
- Q4a) Segmented as in 1a) – after onboarding of household data and ASC carer/cared-for link this will be segmented by cared-for person's health need (grouping into high-level conditions such as neurological, physical, dependency-related etc).
- Q5) Standardised by sex, age, ethnic background, and deprivation.

3.5 Statistical approaches

Initially, summary statistics describing the cohorts (§3.3) will be produced, with rates of carers split by demographic factors listed. Similarly, for Question 1 prevalence of conditions (long-term and short-term) will be calculated and compared between carers, all non-carers, and demographic factors. Paired-patient cohorts will be also created to account for differences in the underlying carer population compared with the general population, and the above analysis will be re-calculated. The impact of council services on health outcomes will be analysed by pairing people (based upon their past healthcare utilisation and health needs) who use council services to carers who do not, and comparing rates of health events. Finally for question 1, logistic regression models will be run to estimate a person's probability of requiring healthcare assistance, split by demographic factors and a flag for council support.

Question 2 will be answered initially by comparing age-standardised rates of social-care service usage by demographic factors to the carer population profiles calculated before question 1. A language comparison of leeds.gov.uk web analytics on carers pages will be made to the language/ethnic make up of carers to determine the extent to which people are aware of services/accessing services, and the rate of service request rejection split by demographic factors will be calculated.

Question 3 will compare the "known" carer population demographics to those found in census data. Following this, a binomial generalised linear model (GLM) will be run using household information which predicts the probability that somebody is "known" to the health or social care services given their household make-up (flagging where they live with a person requiring a carer) and demographic factors. Odds ratios will be calculated for all covariates. If possible, positive-unlabelled learning methods will also be used on this dataset to estimate the rates of unreported care being provided.

Question 4 will use two flags for the end of care – firstly using the direct link in ASC between a carer and cared-for person, and secondly using an indirect link between a known carer and the person within their household who requires full-time care. In both cases, GLMs will be run predicting the likelihood of end of care given demographic factors, the healthcare needs of the cared-for person, and the access to services and support that the carer has experienced, and the importance of factors will be determined via odds-ratios.

Finally, for question 5 two things will be investigated. Firstly the rates of people reporting financial issues for the carer population will be compared to the non-carer population, and secondly the mobility of carers and non-carers will be investigated, looking at the change in area deprivation (both IMD and using housing/rental prices) for both populations.

3.6 Methods for addressing missing data

In some instances across most data sets no NHS number has been recorded, which will disallow for linkage to external data sets. In these cases, where combined data sets are required, these people will be removed from analysis.

In the case of missing deprivation data (where a person's LSOA is not recorded), the value will either be imputed from the average of all deprivation scores, or (if previous LSOAs are recorded for the person) the last known LSOA (or average of all last known IMD scores based upon LSOA history) will be used. For geographic analysis, these people will be discounted.

3.10 Known limitations

When using primary care flags for carers, it is unknown how long each person has been a carer for (i.e. whether they have been a carer for 20 years but only tell their GP upon change of registration, or due to another event such as COVID vaccine rollout), and it is unknown how long people remain carers (if someone is flagged as an unpaid carer in 2016, what is the likelihood that they are still providing unpaid care in 2020?). Similarly, there are uncertainties around the consistency of use of "unpaid carer" Read/SNOMED codes, especially around the end of caring responsibilities.

As stated in §3.2, there are different date ranges covered by different datasets. While most data (i.e. primary care, secondary care, adult social care, and non-IAPT mental health) is fully covered from the start of financial year 2016, other datasets are either not available until a later date, or are not consistently populated until a later date. This will not limit the scope of the study greatly, but it will make longitudinal comparisons more difficult when including supplementary data.

4. Governance

Availability of data and materials

The Leeds Data Model currently contains linked data from primary care, secondary care, mental health, community services, and adult social care services. Primary care data has been provided by all GP practices across Leeds, and permission has been given to link to other data sets and use for commissioning and service improvement purposes. Secondary care and mental health data sets are provided by NHS Digital, and are approved for secondary uses. Adult social care data is provided by Leeds City Council, and has been approved for linkage and secondary uses.

Additional data has been requested for this project. A data sharing agreement has been sent to all GP Practices across Leeds permitting use of pseudonymised household identifiers in analysis, and this will be pulled into the Leeds Data Model through primary care extracts. A data sharing agreement with Adult Social Care is being written which will permit the flow of data containing direct links between carers and cared-for people.

Ethics approval and consent to participate

Ethical approval has been given for all data contained within the Leeds Data Model to be linked and used for the purposes set out within this plan, both through data sharing agreements (for local data) and DARS agreements for national data sets.

5. Impact, dissemination and engagement

Answers to the questions listed above will be useful to local/national decision-makers in two ways: firstly, where gaps are found (or significant inequalities of service use or health

outcomes), these can be made aware of and accounted for in future service design. Secondly, while causal effects of ASC services it is not the key aim of this study and are unlikely to be definitively found, if any evidence of successful services are found then these can be investigated in future evaluations and can also feed into service design.

Several different groups will be engaged to disseminate results. A task and finish group comprised of unpaid carers, the lead on unpaid care at Leeds City Council, a representative from Carers Leeds (a charity which provides advice and assistance to unpaid carers in Leeds) is involved in discussing analysis, and will be used to pull key findings from results. This will mean that key-findings are known to the decision-makers within local government, as well as being informed by people with lived-experience. Further to this, the lead on unpaid carer for Yorkshire and the North East from NHSE/I is aware of this project and findings will be given to them, to inform decision-making at the wider level. Finally, commissioners across Leeds CCG will be made aware of this work where findings are relevant to their specific areas of work.

6. Appendix

Read code list for GP events relating to unpaid carers.

Read Code	Description
918A.	Carer
918A0	Cares for a friend
918A1	Cares for a neighbour
918A2	Cares for a relative
918G.	Is a carer
918H.	Primary carer
918W.	Carer of a person with learning disability
918a.	Carer of a person with substance misuse
918d.	Carer of a person with mental health problem
918m.	Carer of a person with a terminal illness
918t.00	Carer from Black and minority ethnic group
918Y.	Carer of a person with sensory impairment
8IHE.	Carer health check declined
8O7..	Carer support
918b.	Carer of a person with alcohol misuse
8IEP.	Carer annual health check declined
8Bar.	Carer health check completed
13Wb.	Carer has sole parental responsibility
9NSS.	Carer health check offered
918J.	Carer – home telephone number
9Ngw.	Carer does not understand care plan
918W.	Carer of a person with learning disability
918y.	Carer of person with dementia
9Ngv.	Carer understands care plan
918X.	Carer of a person with physical disability
13VN.	Carer able to cope
69DC.	Carer annual health check
918M.	Carer – email address
9180	Carer's details
69DE.	Carer health check
918L.	Carer – mobile telephone number

918K.	Carer – work telephone number
388Q.	Carer strain index score
9d46.	Carer
918c.	Carer of a person with chronic disease
8HkA	Ref for GP carer's assessment
918f.	No longer a carer
.918f	No longer a carer

Read code list for GP events relating to stress. From Masefield et al. (2021).

Read Code	Description
1B1L.	Stress-related problem
E28..	Acute reaction to stress (& [combat fatigue])
E280.	Acute panic state due to acute stress reaction
E281.	Acute fugue state due to acute stress reaction
E282.	Acute stupor state due to acute stress reaction
E283.	Other acute stress reactions
E2830	Acute situational disturbance
E2831	Acute post-trauma stress state
E283z	Other acute stress reaction NOS
E284.	Stress reaction causing mixed disturbance of emotion and conduct
E29..	Adjustment disorder
E290.	Brief depressive adjustment reaction
E290z	Brief depressive reaction NOS
E291.	Prolonged depressive adjustment reaction
E292.	Adjustment reaction with predominant disturbance of other emotions
E2921	Adolescent emancipation disorder
E2922	Early adult emancipation disorder
E2923	Specific academic &/or work inhibition
E2924	Adjustment reaction with anxious mood
E2925	Culture shock
E292y	Adjustment reaction with mixed disturbance of emotion
E292z	Adjustment reaction with disturbance of other emotion NOS
E293.	Adjustment reaction with predominant disturbance of conduct
E2930	Adjustment reaction with aggression
E2931	Adjustment reaction with antisocial behaviour
E2932	Adjustment reaction with destructiveness
E293z	Adjustment reaction with predominant disturbance of conduct NOS
E294.	Adjustment reaction with mixed disturbance of emotion and conduct
E29y.	Other adjustment reactions
E29y0	Concentration camp syndrome
E29y1	Other post-traumatic stress disorder
E29y2	Adjustment reaction with physical symptoms
E29y3	Elective mutism due to an adjustment reaction
E29y5	Other adjustment reaction with withdrawal
E29yz	Other adjustment reactions NOS
E29z.	Adjustment reaction NOS
Eu4..	[X]Neurotic
Eu40.	[X]Phobic anxiety disorders
Eu40y	[X]Other phobic anxiety disorders
Eu40z	[X]Phobic anxiety disorder Unspecified

Eu41.	[X]Other anxiety disorders
Eu413	[X]Other mixed anxiety disorders
Eu41z	[X]Anxiety disorder unspecified
Eu430	[X] Acute stress reaction (& [crisis reaction] or [crisis state] or [psychic shock] or [combat fatig
Eu432	[X] Adjustment disorders (& [culture shock] or [grief reaction] or [hospitalism in children])
Eu43y	[X]Other reactions to severe stress
Eu43z	[X]Reaction to severe stress unspecified
Eu46.	[X]Other neurotic disorders
Eu46z	[X]Neurotic disorder unspecified
K586.	Stress incontinence - female
Ry15.	[D]Undue concern and preoccupation with stressful events
Ryu58	[X]State of emotional shock and stress unspecified
Ua165	Feeling stressed
Ub014	Stress management
X00Sf	Post-traumatic stress disorder
X301U	Stress-related gastritis
X76AY	Work stress
Xa18j	Combat fatigue
Xa18v	Shell shock
Xa7mz	Carer stress syndrome
Xa8HQ	Tower block syndrome
Xa8HR	Wife unable to cope
XaA2F	Stress mapping
XaBUD	Stress monitoring call
XaEFB	Chronic post-traumatic stress disorder
XaI8j	Stress counselling
XaJgP	Stress-induced epilepsy
XaWye	Stress cardiomyopathy
XaX55	Acute post-traumatic stress disorder following military combat
XaX56	Chronic post-traumatic stress disorder following military combat
XaX58	Delayed post-traumatic stress disorder following military combat
XE0pM	Stress at home
XE0rR	Stress incontinence
XE1bo	Acute reaction to stress (& [post-traumatic] or [shell-shock])
XE1Ym	Acute stress reaction
XE1Yn	Acute stress reaction NOS
XE1Yp	Specific academic or work inhibition
XE1Zj	[X]Other specified anxiety disorders
XE2Nh	Stress monitoring admin.
XM0As	Stress and adjustment reaction
XM1aI	Stress at work
XM1Am	Undue concern and preoccupation with stressful events
XM1Q3	Specific work inhibition
ZV4B2	Stressful work schedule
ZVu4E	[X]Other stressful life events affecting family and household