# Data Explained

___

## SOURCE, Social care (care at home) recipients at the end of life

Author: Jan Savinc
Date: 21 August 2025 (updated 22 September 2025)

This Data Explained summarises experiences and learning from working with the SOURCE dataset containing social care data collected by Public Health Scotland, specifically data on care at home. This publication is intended to help guide future researchers using this data and to provide feedback into future dataset development and documentation.

# Introduction

The SOURCE dataset contains data on social care that is partly or fully funded by local authorities (LAs) in Scotland. Health & social care partnerships (HSCPs) submit data quarterly or annually to the Scottish Government, starting from 1$^{st}$ January 2018 onwards (i.e. Q4 of the 2017/18 financial year). Data are collected for all of Scotland (all 31 HSCPs), though there is some variation in how HSCPs report data and issues with completeness (see the Data limitations section further on). Most data are individual-level and can be linked to the Community Health Index (CHI) database via their demographic information, with a high degree of completeness (e.g. the 2020/21 data had a 99.5% CHI match rate (Scottish Government, 2022).

SOURCE data comprises several files, with examples of information contained below. Note that this list is not exhaustive; more detailed information is available from the references (Public Health Scotland; Scottish Government 2022):

- Demographics
  - Personal details, such as gender, ethnic group, postcode
- Client information
  - This contains yes/no flags for the client or service user groups the person belongs to: dementia, mental health problems, learning disability, physical/sensory disability, drugs, alcohol, palliative care, carer, elderly/frail, autism, other
  - Data on type of housing, whether person receives housing support, or has a social worker
  - Whether person lives alone and/or receives unpaid care
- Self-directed Support (SDS)
  - Category of SDS, start and end date, value and contribution of SDS
- Care at home (also known as home care)
  - Start and end dates
  - Service provider & type of service
  - Planned hours, actual hours, staff hours
  - Budget
  - Reablement
- Technology-enabled care (community alarms & telecare)
  - Service type, start and end date
- Care homes
  - Care home name, postcode, provider
  - Admission and discharge date
  - Type of admission
  - Nursing care provision
  - Type of funding, local authority (LA) contribution, service provision budget
- Indicator of Relative Need (IoRN)
  - Assessment date
  - Level of need
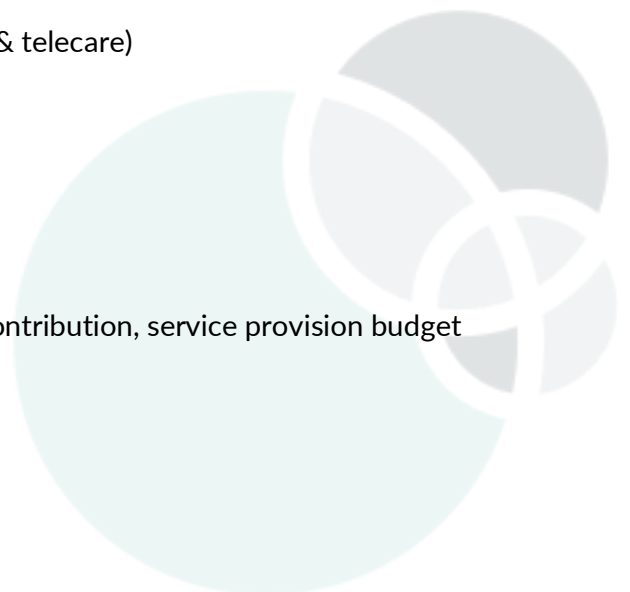
- Learning Disability Statistics Scotland (LDSS)
  - Learning disability or autism spectrum diagnosis
  - Out of area placement flag, area where person resides
  - Whether person receives employability support, whether they are in further education, and whether they take part in volunteering
  - Whether the person accessed advocacy support

# How is the data collected?

As of the latest Insights in Social Care Statistics for Scotland report (Public Health Scotland, 2023a), SOURCE data are available from Q4 of 2017/18 to Q4 of 2022/23.

SOURCE data are a mix of data collected annually and quarterly, and every record has a field for the financial year (and financial quarter) of the data collection period. Some data are mandatory to collect, while others are optional: the Indicator of Relative Need (IoRN), for example, was collected optionally, and is therefore recorded very sparsely; in fact, from Q1 2023 onwards, there is no requirement to collect it at all (Public Health Scotland, 2024b).

Public Health Scotland publishes an annual report titled *'Insights in Social Care', which is divided into separate reports for Care at Home and Care Home data, detailing* the degree of completeness. HSCPs sometimes cannot submit complete records in time for the quarterly/annual collection, and these data are usually updated retrospectively at the next collection period.

# SOURCE dataset in the 'Increased deaths at home in Scotland' project

The SOURCE dataset was linked to data in the 'Increased deaths at home in Scotland' project that had started in 2022. This project linked death registrations for all people who died in Scotland aged 18+ in the period from 23rd March 2015 to 22nd March 2021 to those people's use of health services in their final 12 months of life, including hospital admissions, community prescribing, cancer registrations, and unscheduled care (ambulance uses, A&E attendances, Out-of-hours GP attendances, and calls to NHS24, the Scottish 24-hour health helpline). In this project, we were interested whether people who received care at home were more likely to die at home, and whether this changed with the Covid-19 pandemic.

For this project, only the Client information, Demographics, Care at home, and IoRN files were extracted, for the period from Q4 of 2017/18 to Q4 of 2020/21, to match the study period used in the project (there is no SOURCE data before 1st January 2018).

For the project, we were interested in the following key variables from the SOURCE data:

| Dataset | Variables |
|---|---|
| Client information | Receiving palliative care<br>Living alone<br>Receiving unpaid care |
| Care at home | Start and end date of service |
| IoRN | Level of need |

In the project, we were interested in whether people who lived alone, received palliative care, or received unpaid care, were more likely to die at home, as well as whether people who received care at home were more likely to die at home.

# What can the data be used for?

SOURCE data are recorded on an individual level, and the demographic data for an individual can be used to link them to the CHI database, which can be used to link to other datasets such as the population spine, the census, or health data, for example.

The Client file contains data on what is known about social care clients, such as whether they live alone, whether they are receiving palliative care, and whether they are receiving any unpaid care – these aren't readily available in any of the Public Health Scotland datasets and are therefore an invaluable additional source of information, though they are only available for individuals receiving publicly funded social care.

The Indicator of Relative Need is similarly valuable because it contains data on needs assessments – in healthcare data we usually only have access to data on services used, and an assumption is usually made that these were the services needed. Unfortunately, the IoRN is scarcely recorded, and is no longer required in data collection as of Q1 2023.

Data on home care can be used to determine when people were in receipt of home care, how long for, how many hours in total they were receiving, and the number of staff hours that were required for home care. It is possible to compute the difference between the planned and actual hours of service, though both aren't always reported for all collection periods and all HSCPs, due to issues with data completeness.

Some of the data in SOURCE can be used to cross-validate or compare against similar data in other datasets – for example, the household composition variable in census data can be used to determine if someone lived alone, and this could be compared against the SOURCE flag for someone living alone.

Aside from the Health and Care Experience Survey (HACE) (Public Health Scotland, 2024a), SOURCE is the only dataset to record whether individuals received unpaid care, making it an invaluable resource, since it avoids costly and labour-intensive survey data collection.

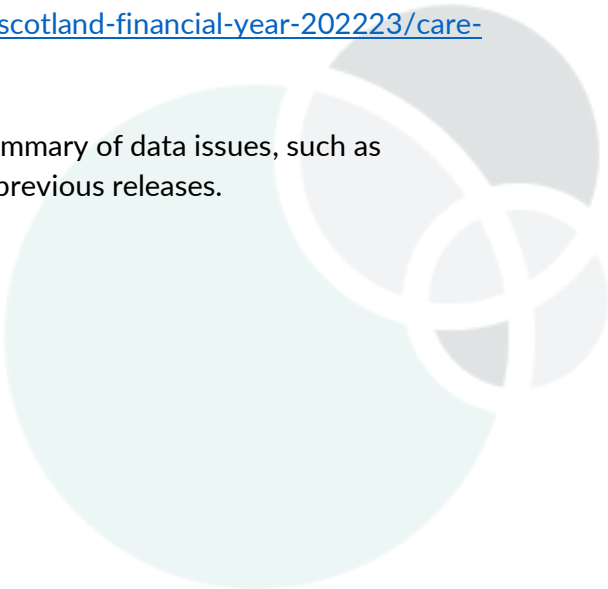# Existing research or examples of previous research

To my knowledge, the SOURCE dataset has not been used in previous research, though there have been studies looking at some of the datasets that are part of SOURCE or preceded it:

-   David Henderson and colleagues published a *Data Resource Profile* on the Scottish Social Care Survey and the Scottish Care Home Census, which are predecessors to the SOURCE dataset (Henderson et al., 2019)
-   Elizabeth Lemmon published a study looking at the influence of unpaid care on formal care use, also using the Scottish Social Care Survey (Lemmon, 2020)
-   David Henderson and colleagues also published a study of multimorbidity and social care use, also using the Scottish Social Care Survey (Henderson et al., 2021)
-   The *Mapping National Adult Social Care and Care Home Data Sources in Scotland* report outlines the sources of adult social care data in Scotland, including SOURCE and some of its predecessors (Scottish Government, 2022)

Aggregate-level reports of SOURCE data are published annually by Public Health Scotland, including dashboards showing top-level variable breakdowns; there are two annual reports, one focusing on Care at home data (Public Health Scotland, 2023a), and another on Care home data (Public Health Scotland, 2023b):

-   Care at home data, 2022/23: https://publichealthscotland.scot/publications/insights-in-social-care-statistics-for-scotland/insights-in-social-care-statistics-for-scotland-care-at-home-services-provided-andor-funded-by-health-and-social-care-partnerships-in-scotland-20222023/care-at-home-dashboard/
-   Home care data, 2022/23: https://publichealthscotland.scot/publications/insights-in-social-care-statistics-for-scotland/insights-in-social-care-statistics-for-scotland-care-home-services-funded-by-local-authorities-in-scotland-financial-year-202223/care-home-services-dashboard/

These also contain a detailed technical report and a summary of data issues, such as completeness, inconsistencies, and any changes from previous releases.

# Data limitations encountered

**Linkage to other datasets**

SOURCE data records contain a personal identifier; the *Social Care ID*. Care must be taken since this identifier is not required to uniquely identify a person across HSCPs. In other words, two different people can have the same identifier if they receive care funded by different HSCPs; conversely, the same person can have different identifiers if they receive care from different HSCPs, such as if they move areas, for example. Linkage to other datasets would generally happen via a person's demographic data rather than the Social Care ID, but it is important to bear the limitations of the identifier in mind when attempting to use it for linkage.

**Dates in Care at home data**

Social care data can be thought of as episodic by nature; that is, there is a start and end date recorded for each care at home arrangement. There are however, some complexities to how care at home is recorded. The service start date represents the date that home care was put in place for a new client, or the date of a significant change to the service. Similarly, the end date represented the date the service ened or was significantly changed in line with the care plan. In practice, a person with home care records spanning multiple reporting periods will often have the same start date recorded in every record. It is therefore possible to infer the start of home care receipt even if the data for that period are not available, and therefore to calculate the span of time during which a client was receiving a particular service. There were some records where the start date preceded the reporting period by several years or even decades – in these cases it is unrealistic that the service start date was recorded and instead more likely that there was a data entry error. For the purposes of the 'Increased deaths at home' study, we were only interested in the duration of home care during the final 12 months of life, so any start dates that were outside of the reporting period were simply changed to the first day of the reporting period.

Similarly, since most data in SOURCE are reported quarterly, care should be taken when aligning SOURCE records, especially when there isn't a start date included, with other data, to account for financial quarters. In the Increased deaths at home project, for example, the study period *almost* aligns with financial quarters (e.g. the pandemic cohort was defined as people who died between 23rd March 2020 to 22nd March 2021, which almost corresponds to Q1 to Q4 of 2020/21, which run from 1st April 2020 to 30th March 2021. This means that there were a small proportion of edge cases of missing social care data at the start of the study period, and another small proportion of edge cases with social care data after the end of the study period, which needed to be accounted for accordingly; such data were removed, or noted as missing in the limitations, respectively.

**Missing Care at home data**

Some Care at home data is missing as not all HSCPs are able to submit data on time, though these are usually updated retrospectively in future submissions. As of the latest *Insights in Social Care Statistics for Scotland* report (Public Health Scotland, 2023a), The following data were missing for the study period (March 2017- March 2021):

- Aberdeen City HSCP did not submit data for all of 2019/20, and for Q4 of 2020/21
- Comhairle nan Eilean Siar HSCP did not submit data for Q3 of 2020/21

**Inconsistencies in Care at home data variables**

There is also some variation between HSCPs and over time with how specific variables are reported, due to variation in care at home models. These are listed in detail in the *Insights in Social Care Statistics for Scotland* dashboard (see above). To summarise, some HSCPs include 24-hour care at home clients while others do not, with this also changing over time; some HSCPs provide weekly data on hours of care received rather than quarterly, to be consistent with previous reporting; start dates were missing for some clients, and there were various issues with reporting actual vs. planned hours, and personal care, across HSCPs.

# Suggested improvements & recommendations to data owners

The technical reports and dashboards alongside the annual Insights in social care statistics for Scotland are excellent and very detailed but contain a lot of information about missing data or special cases that may be difficult for a researcher to properly account for in analysis. One possible improvement would be to create a flag for the different considerations that apply to any data entry (e.g. 'A' meaning planned hours were submitted instead of actual hours of care at home service, 'B' meaning 24-hour care was excluded from the figure, etc.) as this would make it easier to interpret the data after processing.

Another issue to look into is the accuracy of the Care at home start dates, where they represent dates that are a very long time in the past and therefore unlikely to represent actual service start dates.

# Suggested future data linkages

Suggested future linkages include:

-   Health data such as hospital admissions, outpatient contacts, and unscheduled care contacts – this would enable the study of healthcare use by people who received social care, and a comparison to people who do not.
-   Census data – this would enable the study of demographic and household characteristics of people who received social care, and a chance to compare the living alone variable in SOURCE to the household composition variable in the census
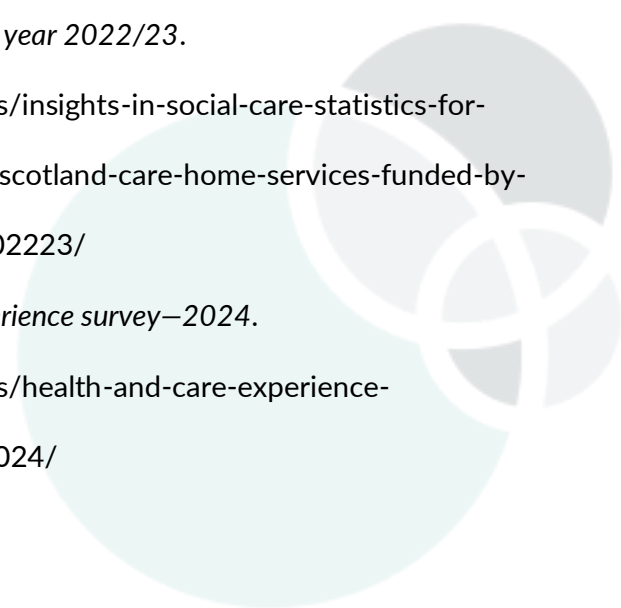
# Conclusion

In conclusion, the SOURCE dataset is a valuable resource for researchers interested in the population of adults receiving social care and allows for population-wide individual-level linkage to other datasets of interest.

# References

Henderson, D. A. G., Atherton, I., McCowan, C., Mercer, S. W., & Bailey, N. (2021). Linkage of national health and social care data: A cross-sectional study of multimorbidity and social care use in people aged over 65 years in Scotland. *Age and Ageing*, *50*(1), 176–182. https://doi.org/10.1093/ageing/afaa134

Henderson, D. A. G., Burton, J. K., Lynch, E., Clark, D., Rintoul, J., & Bailey, N. (2019). Data Resource Profile The Scottish Social Care Survey (SCS) and the Scottish Care Home Census (SCHC). *International Journal of Population Data Science*, *4*(1), Article 1. https://doi.org/10.23889/ijpds.v4i1.1108

Lemmon, E. (2020). Utilisation of Personal Care Services in Scotland: The Influence of Unpaid Carers. *Journal of Long-Term Care*, *2020*, Article 2020. https://doi.org/10.31389/jltc.23

Public Health Scotland. (2023a). *Insights in social care statistics for Scotland—Care at home services provided and/or funded by health and social care partnerships in Scotland 2022/2023— Insights in social care: Statistics for Scotland—Publications—Public Health Scotland*. https://publichealthscotland.scot/publications/insights-in-social-care-statistics-for-scotland/insights-in-social-care-statistics-for-scotland-care-at-home-services-provided-andor-funded-by-health-and-social-care-partnerships-in-scotland-20222023/

Public Health Scotland. (2023b). *Insights in social care statistics for Scotland—Care home services funded by local authorities in Scotland, financial year 2022/23*. https://publichealthscotland.scot/publications/insights-in-social-care-statistics-for-scotland/insights-in-social-care-statistics-for-scotland-care-home-services-funded-by-local-authorities-in-scotland-financial-year-202223/

Public Health Scotland. (2024a). *Health and Care Experience survey—2024*. https://publichealthscotland.scot/publications/health-and-care-experience-survey/health-and-care-experience-survey-2024/

Public Health Scotland. (2024b). *Social care dataset: Definitions and recording guidance*.

https://publichealthscotland.scot/publications/social-care-dataset-definitions-and-

recording-guidance/

Scottish Government. (2022). *Mapping National Adult Social Care and Care Home Data Sources in*

*Scotland: Baseline understanding & future potential: September 2022*. Scottish Government.

http://www.gov.scot/publications/mapping-national-adult-social-care-care-home-data-

sources-scotland-baseline-understanding-future-potential-september-2022/

## Glossary

- PHS: Public Health Scotland
- HSCP: Health & Social Care Partnership
- SDS: Self-directed support
- IoRN: Indicator of Relative Need
- LDSS: Learning Disability Statistics Scotland
- LA: Local Authority
- HACE: Health and Care Experience Survey

## Disclaimer

This work was produced using administrative data accessed through the Scottish National Safe Haven. The use of the data in this work does not imply the endorsement of the Trusted Research Environment or data owners in relation to the interpretation or analysis. This work uses research datasets which may not exactly reproduce National Statistics aggregates. National Statistics follow consistent statistical conventions over time and cannot be compared to linked datasets.

## Acknowledgements

## Contact

Name: Jan Savinc
Email: j.savinc@napier.ac.uk