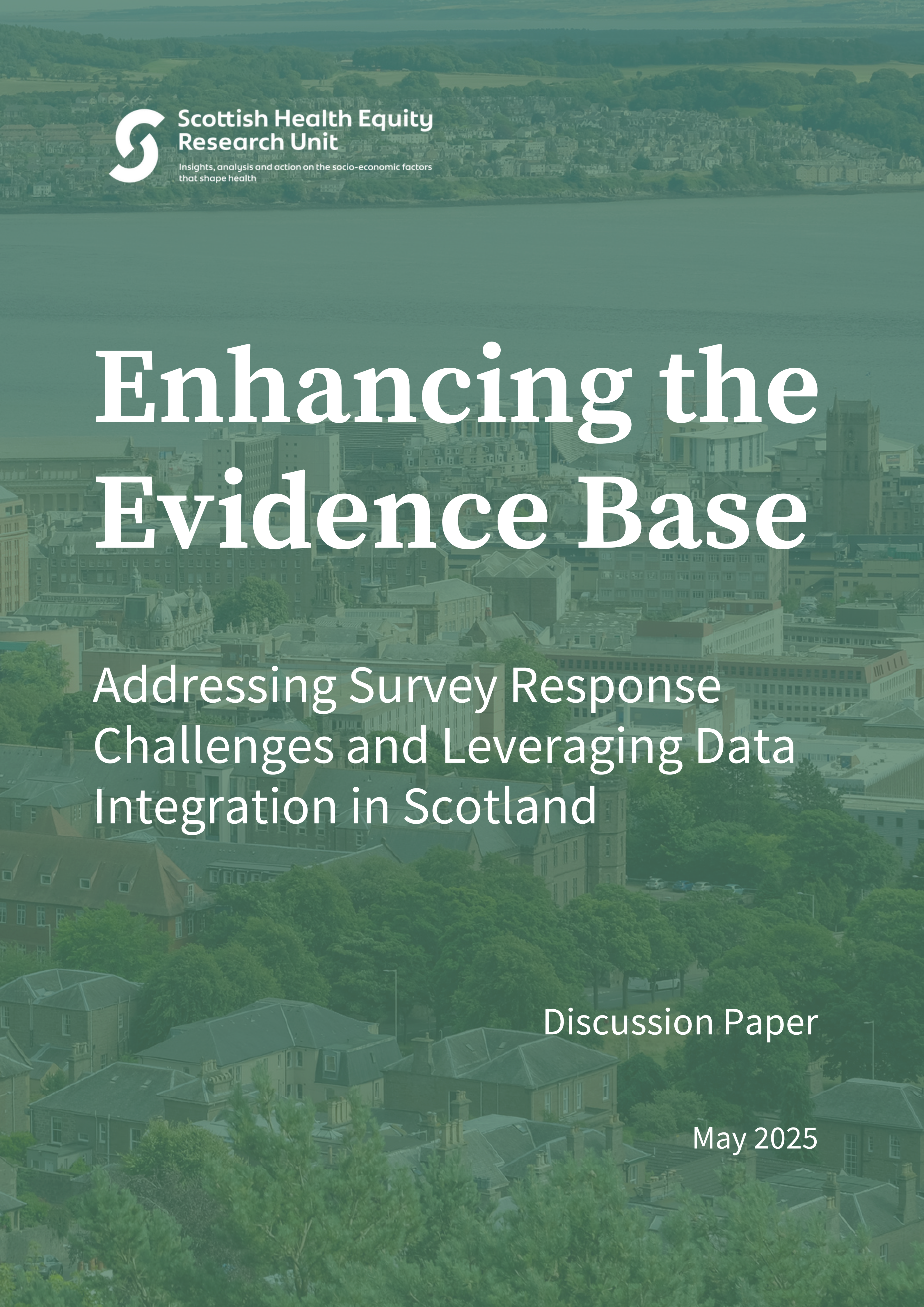


Scottish Health Equity
Research Unit
Insights, analysis and action on the socio-economic factors
that shape health

Enhancing the Evidence Base

Addressing Survey Response Challenges and Leveraging Data Integration in Scotland

Discussion Paper

May 2025

Executive Summary

As Scotland confronts rising health inequalities, longstanding data collection methods are under increasing strain. Falling survey response rates, digital exclusion, and public mistrust are undermining the reliability of core datasets, with serious implications for policy design and service delivery.

As the landscape evolves, there is an urgent need for smarter, more inclusive data strategies. This discussion paper outlines the case for modernising how we gather and integrate data to achieve a more comprehensive and precise understanding of the socioeconomic determinants of health inequalities in Scotland. Drawing on SHERU's work and wider developments across the UK, it identifies key shifts in the data terrain and highlights where action is most needed:

- **Safeguarding core social surveys**: With response rates declining in flagship surveys like the Labour Force Survey and Scotland's Census, there is an urgent need to strengthen participation, adapt methods, and supplement gaps with robust alternative sources.

- **Reducing exclusion in data collection**: Digital-only approaches risk leaving behind key populations—including older adults, low-income households, and rural communities—limiting the visibility of those most affected by inequality.

- **Maximising the value of administrative data**: Better use of existing records (e.g. benefits, health, housing) can reduce respondent burden and enhance coverage, while enabling more frequent, granular analysis.

- **Expanding data linkage and integration**: Linking datasets across sectors is essential for understanding the complex drivers of health inequality, especially where income, housing, and environmental factors intersect with health.

- **Innovating responsibly**: Emerging tools like Personal Data Stores and financial data platforms offer new insights, but require careful governance to ensure equity, quality, and public trust.

SHERU is contributing to the development of practical solutions in these areas—improving data accessibility, fostering collaboration, and exploring integrated approaches that connect data, particularly economic and health data, in more meaningful ways. This work supports a more resilient evidence base, equipping policymakers with the insights needed to drive meaningful action that addresses the root causes of health inequalities in Scotland.

Introduction

Reliable and timely data collection is fundamental to understanding and addressing socioeconomic determinants of health inequalities. However, recent declines in survey response rates and limitations in traditional data collection methods have created significant challenges for those aiming to design effective policy responses.

This discussion paper explores efforts to enhance survey research across the UK, with a particular focus on Scotland, ensuring social surveys continue to inform both understanding and actionable policies. Additionally, the paper examines how barriers can be mitigated while enhancing insights through alternative methods, data sources, and data linkage and integration—equipping researchers and policymakers with the evidence needed to drive effective interventions.

SHERU

The Scottish Health Equity Research Unit (SHERU) seeks to contribute to a future where health inequalities are tackled with greater precision and impact, guided by robust, evidence-driven insights. The unit is dedicated to improving the understanding of the socioeconomic factors that shape health outcomes in Scotland, with a particular focus on supporting the development of actionable solutions to address the complex drivers of inequality.

Funded by the Health Foundation, SHERU's efforts to strengthen the evidence base focus on three key areas. First, the unit works to make complex data more accessible and understandable to a wider audience. Second, SHERU explores innovative data approaches, such as Personal Data Stores (which allow individuals to securely manage and share their health, social care, and lifestyle data), while collaborating with data controllers to improve data quality through efforts like refining collection practices and standardising definitions. The third focus is maximising the value of existing data by championing the integration of both quantitative and qualitative insights, understanding that a fuller picture of health inequalities emerges through effective data linkage and the lived experiences and contextual nuances captured by qualitative research.

Throughout the discussion paper, we highlight the challenges and opportunities in strengthening data insights in Scotland—particularly around declining survey response rates and the integration of diverse data sources—and how these efforts can support more responsive and effective action on the socioeconomic determinants of health inequalities. We also reference SHERU's planned and ongoing work in this area, which focuses on deepening understanding of the drivers of health inequalities and informing more effective, evidence-based policy.

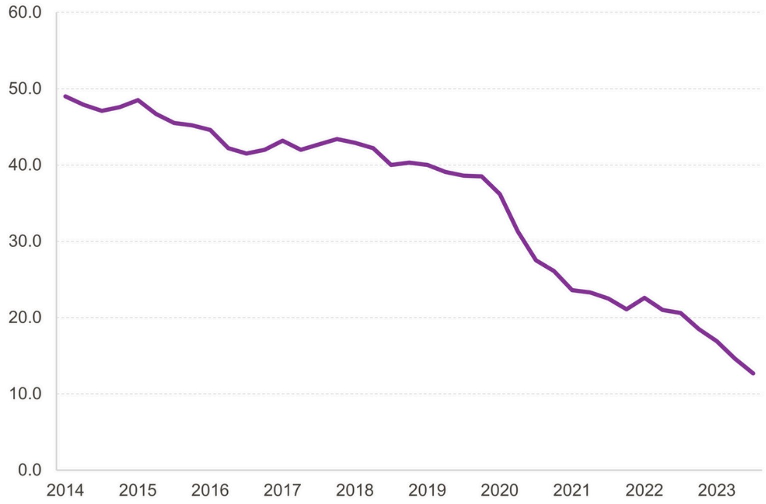
Challenges in Data Collection and Survey Response Rates

Declining response rates and shrinking sample sizes in key surveys have become a well-known issue. Various factors contribute to this trend, including survey fatigue, digital exclusion, the Covid-19 pandemic, public mistrust in data collection agencies, and difficulties in reaching marginalised populations. Additionally, rising survey costs, resource constraints, and challenges in recruiting and retaining fieldworkers further impact the ability to conduct high-quality, large-scale surveys.

Declining Response Rates in National Surveys – Labour Force Survey

The decline in survey response rates is a growing challenge for National Statistics, affecting the accuracy and reliability of key datasets. The UK's Labour Force Survey (LFS), a vital tool for tracking employment and economic trends, has experienced a drop in response rates, contributing to quality concerns. In November 2023, the Office for Statistics Regulation (OSR) removed the accreditation of ONS's LFS-based estimates and datasets, citing severe data quality issues [1]. This impacted multiple statistical outputs, including UK employment statistics, labour market reports for Wales and Scotland, and trade union membership statistics [2].

Figure 1. Total response rates for the Labour Force Survey (cumulative response rate for all waves, Great Britain, %)



Source: ONS: Labour Force Survey performance and quality monitoring reports [3]

In October 2024, OSR agreed to ONS's request to suspend the accreditation of outputs based on Annual Population Survey (APS) data, which utilises combined data from two waves of the main LFS [4]. The suspension affected APS-derived statistics, including those related to regional labour markets, population demographics, and personal well-being in the UK.

In response to these challenges, ONS is rolling out a long-term transformation plan. The Transformed Labour Force Survey (TLFS) is being introduced across England, Scotland, and Wales, with a focus on boosting participation through expanded use of online data collection [5]. While this move offers many benefits, one of the challenges is capturing complex variables that are typically easier to collect in an interviewer-led survey. The TLFS incorporates a ‘knock-to-nudge’ strategy, where field interviewers visit respondents' homes to encourage participation if they haven't already completed the survey. By increasing the sample size and improving representativeness, these efforts are aiming to ensure that survey participants more accurately reflect the broader population.

With issues compromising the reliability of the LFS data, which traditionally serves as the main source for understanding employment, unemployment, and economic inactivity in the UK - the Resolution Foundation used administrative tax data from HMRC’s Pay As You Earn (PAYE) system to suggest that the true employment rate in the UK is likely much higher than current LFS estimates [6]. The work also goes on to state that if the LFS is underestimating the employment rate, it must be “overestimating the unemployment rate and inactivity rate in some combination.”

SHERU applied the Resolution Foundation’s model specifically to Scottish data and found similar results [7]. The methodology indicates that since 2022, the LFS has likely been largely underestimating the volume of people in employment. It follows that the LFS must be overestimating the volume of people who are unemployed or economically inactive (neither in work nor looking for work). This clearly illustrates how alternative data sources, such as administrative data, can help address challenges with traditional surveys and provide additional insights— in this case, on the labour market.

Declining Response Rates in National Surveys – Scotland’s Census

Scotland’s Census 2022 had a final person response rate of 89.8% [8]. This was lower than the original target of 94% and lower than the 97% response rate for the 2021 Census in England and Wales. To address this, the National Records of Scotland (NRS) utilised oversampling in areas with anticipated lower responses, incorporated administrative data, and sought guidance from an International Steering Group. These efforts helped improve accuracy and maintain confidence in the census results.

The 2022 Scotland Census recognised that digital exclusion remains a consideration for certain groups, including those living in remote and rural areas. While digital connectivity and access have improved in recent years, rural communities were still identified as being less likely to complete the census online. As a result, additional support and alternative response options were put in place to ensure inclusivity. Alongside rural populations, other groups such as older adults, minority ethnic communities, people with physical, emotional or learning disabilities or difficulties, and those experiencing socioeconomic deprivation were also prioritised for targeted engagement, reflecting Scotland’s ongoing efforts to address potential gaps in digital participation and data collection.

National Records of Scotland (NRS) has an ongoing programme to enhance population statistics through greater use of administrative sources, aiming to better integrate these with census and survey data [9]. Their main goals are to inform future census recommendations, enhance the coherence of demographic statistics across the UK, and engage with data suppliers and stakeholders to refine these estimates and guide future developments. Although not an intended aim, this work supported NRS's response to the 2022 Census return rate, allowing them to incorporate administrative data quickly.

The NRS research aims to provide more frequent and accurate estimates by linking data from various administrative sources, including expanding to the Department for Work and Pensions (DWP) and HM Revenue and Customs (HMRC). Additionally, NRS is exploring the use of a Dynamic Population Model (DPM), which the Office for National Statistics (ONS) has implemented in England and Wales, to improve the accuracy of administrative estimates for Scotland. This approach highlights the potential of administrative data to complement traditional surveys and enhance the delivery of reliable, up-to-date population insights— an area of particular interest to SHERU as we continue analysing Scotland’s Census 2022 data and integrating its more detailed outputs into our work as they become available [10].

Challenges in Health-Related Surveys

Surveys critical for assessing health inequalities, such as those on economic inactivity and long-term health conditions, are amongst those facing declining response rates.

The shift towards online data collection has created a digital divide, where those without reliable internet access or digital literacy are less likely to participate in surveys. This exclusion particularly affects older adults, rural populations, and lower-income households, leading to gaps in data that are crucial for health inequality research.

Beyond digital access, trust in institutions also plays a role. Communities that have previously faced barriers to fair treatment within healthcare systems and government services may be less willing to engage with surveys. This reluctance further narrows the participant pool and risks excluding the very populations most affected by health inequalities.

The lack of participation from certain groups can result in data that fails to capture the full diversity of health experiences and outcomes across the population. This often leads to the underreporting of health issues among marginalised communities, who are already disproportionately affected by health inequalities. When policy decisions are based on incomplete data, there is a risk of reinforcing existing inequalities rather than addressing them. As a result, health policies, resource allocation, and service planning may overlook the needs of those most at risk, further widening the gap in health outcomes.

Alternative Approaches to Overcoming Survey Challenges

The survey community is struggling to implement traditional, pre-pandemic survey methods due to knowledge gaps in new techniques (like push-to-web, knock-to-nudge, and video interviewing) and a shortage of skilled professionals. There’s growing pressure to develop fit-for-purpose methods, manage rising fieldwork costs, and redefine the role of interviewers.

We have been following the work of Survey Futures which is a three-year initiative (funded by UKRI-ESRC from July 2023) designed to address these challenges [11]. It aims to transform how population surveys are conducted in the UK through rigorous research, practical guidance, and widespread dissemination of best practices. The project emphasises collaboration, training, and capacity building to support survey commissioners, designers, managers, and data users in adapting to new methods and maintaining high-quality, inclusive data collection.

Survey Futures now comprises nine research strands, each with its own focused programme of work. These strands span topics from enhanced sampling methods and the evolving post-pandemic role of interviewers to innovative areas such as video interviewing and the use of generative AI in questionnaire design. In addition to these research strands, Survey Futures also publish reports and working papers. One such paper discusses a framework for understanding and addressing the under-representation or exclusion of certain population subgroups in UK surveys, including mitigation measures for subgroups at risk of not receiving a survey invitation or not being able to participate.

While improving survey response rates remains a priority, it is equally important to explore alternative approaches that broaden participation and enhance data quality. Some of these approaches involve adapting survey methodologies, while others focus on complementary strategies beyond traditional data collection. The next few sections will delve into specific examples/options.

Mixed-Mode Surveys

To address digital exclusion, some researchers advocate for mixed-mode survey strategies—combining online, telephone, postal, and face-to-face methods. While more resource-intensive, these approaches can help reach underrepresented groups, improving inclusivity and the reliability of data collected for health inequality research. Ultimately, the most appropriate survey design will depend on balancing factors such as cost, stakeholder priorities, and data quality requirements.

The *Mixed Mode Research: Report to Inform the Scottish Government Long Term Survey Strategy* by Ipsos Scotland, the Ipsos Survey Methods Unit, and Peter Lynn at the University of Essex examined the potential benefits, risks, and trade-offs of adopting mixed-mode designs [12]. The study highlighted key considerations, including the impact of different modes on sample representation, data quality, and user needs. While mixed-mode surveys offer opportunities for cost savings and greater accessibility, they also present challenges such as understanding the likely nonresponse bias associated with different modes, reducing mode effects, and potential breaks in time series data.

The work emphasises the importance of transparent decision-making, stakeholder engagement, and ongoing investment in testing and questionnaire development to ensure a robust and defensible transition. Conducting parallel testing alongside questionnaire refinement is widely regarded as best practice, although this approach may increase survey delivery costs in the short term.

Deepening the Role of Administrative Data

As previously discussed, administrative data can assist in constructing more comprehensive sampling frames, thereby enhancing survey representativeness and addressing coverage issues. This approach can also reduce costs and respondent burden. Furthermore, administrative data sources such as healthcare usage records, benefits data, and educational qualifications offer valuable insights into health inequalities, enabling a clearer understanding without the need for additional survey data collection.

As noted earlier, NRS has been advancing the use of administrative data to support improved population statistics. Similarly, the *Mixed Mode Research: Report to Inform the Scottish Government Long Term Survey Strategy* highlights stakeholders' interest in leveraging administrative data to augment or potentially replace certain data collected by the Scottish Household Survey (SHS), Scottish Crime and Justice Survey (SCJS), and Scottish Health Survey (SHeS).

The Family Resources Survey is a vital tool for understanding income distribution, poverty, and economic well-being in the UK. It is now possible to link at least 95% of FRS respondents to their administrative records [13]. The Department for Work and Pensions (DWP) is undertaking the Family Resources Survey (FRS) Transformation project to integrate administrative data on benefits and earnings into the FRS, aiming to improve accuracy and reduce underreporting. These estimates have not yet been incorporated into the official Households Below Average Income (HBAI) statistics, which are derived from the FRS.

Concerns have been raised about the reliability of recent HBAI estimates, as survey-based statistics may not fully reflect reality due to lower response rates since the pandemic. The Institute for Fiscal Studies (IFS) have found that comparisons with HMRC tax data suggest that reported income declines, particularly in earnings, may be overstated, highlighting the need for more robust data integration [14].

DWP plans to make improvements to the 2024/25 FRS dataset by replacing survey responses for state benefits with information from administrative records. They believe this will help address the underreporting of benefits currently seen in FRS data. Additionally, DWP intends to use benefits administrative data dating back to 2018/19 to create a single, revised back series of estimates for FRS and HBAI grossing factors [15].

By exploring the feasibility of integrating Social Security Scotland and health data, we could gain a more nuanced and comprehensive understanding of the relationship between household income and health, as well as deeper insights into poverty dynamics. As work progresses to improve the Family Resources Survey (FRS) dataset, SHERU plans to develop proposals for linking to Scottish benefits data while collaborating with relevant stakeholders. Additionally, we aim to explore the potential for connecting this data with health records, further strengthening our understanding of how financial circumstances influence health outcomes across different population groups.

Looking ahead, administrative data is set to play an increasingly important role in survey research. As it becomes more integral, it offers the potential to enhance data quality and provide more comprehensive insights into complex social dynamics. We will revisit administrative data in a subsequent section, exploring how data linkage and integration initiatives are creating new opportunities to complement survey data and support more informed policy decisions.

Alternative Data Sources: Example 1 - Personal Data Stores

Data from Personal Data Stores (PDS) could, in the future, enhance the evidence base around health inequalities [16]. PDS allow individuals to securely control and share their own data—including health records, social care information, and lifestyle data—on a consent-driven basis. If widely adopted and ethically managed, they could offer researchers access to richer, more granular datasets directly from individuals, helping fill gaps left by traditional surveys and administrative sources.

Personal Data Stores also present opportunities for the collection and integration of qualitative insights—such as self-reported experiences, personal narratives, and feedback—directly from individuals. This could provide essential context to quantitative data, particularly in understanding how people experience health and care systems, and the lived experience of social determinants. For example, incorporating qualitative inputs from navigating multiple services could deepen the evidence base, illuminating gaps in service delivery that quantitative data alone may overlook.

By looking at research use cases, PDS could improve insights into the complex interplay of socioeconomic factors and health outcomes. This person-centred approach may offer a more nuanced understanding of how social determinants—such as housing, employment, and education—affect health inequalities. Moreover, data from people who engage with multiple services could support continuous improvement by providing a clearer picture of service interactions and outcomes across health and social care.

However, there are challenges. PDS data may be less consistently structured than traditional survey or administrative datasets, creating complexities for data harmonisation and analysis. There are also risks related to data quality, completeness, and potential biases, as early adopters of PDS technologies may not be representative of the wider population—particularly those most affected by health inequalities. Current pilot projects in Scotland are exploring the use of PDS with vulnerable groups, including people receiving cancer care, and those living with frailty or dementia.

Ensuring robust governance, data standards, and interoperability is essential, alongside maintaining trust through transparent, ethical data use. If these challenges are managed effectively, PDS could complement other data sources, supporting more inclusive, person-centred policy development and ultimately contributing to a better understanding—and reduction—of health inequalities in Scotland.

SHERU intends to revisit this after the completion of a pathfinder project, involving the Fraser of Allander Institute and partners, which is exploring the use of personal data stores to potentially support people with learning disabilities in engaging with health and social care services. By building on these findings, SHERU aims to assess how similar approaches could be applied more broadly to enhance data accessibility and enrich the evidence base, ultimately supporting better outcomes for vulnerable groups.

Alternative Data Sources: Example 2 – Smart Data Foundry

Smart Data Foundry (SDF) aims to make private sector financial data accessible and discoverable through a range of datasets and services designed to support research, innovation, and policymaking [17]. Its core strength lies in leveraging large-scale, anonymised financial data, primarily from banking transactions, to offer insights into individual and population-level economic behaviours. This data can help fill gaps left by traditional surveys, particularly in understanding income volatility, financial resilience, and broader socioeconomic trends.

A partnership between the Joseph Rowntree Foundation (JRF) and Smart Data Foundry has developed an Income Volatility Database that uses near real-time banking data to provide insights into fluctuations in income across the UK [18]. The data, drawn from NatWest current account holders is anonymised and made available through a public dashboard for general insights, as well as a Trusted Research Environment (TRE) for approved researchers to conduct more detailed analysis.

While there is a reliance on data from a single banking source—and it won’t capture individuals with multiple current accounts held outside of NatWest—this may raise questions about how representative the findings are for the broader population. Although those without bank accounts represent a small percentage of the population, they are often among the most marginalised. Research indicates that single adults and individuals from some minority ethnic groups are more likely to be in this category [19].

The dataset includes 1.2 million customer current accounts (which may include joint accounts) across Great Britain, with a slight overrepresentation in the Scottish Central Belt due to the inclusion of Royal Bank of Scotland data. The sample is randomly generated, and SDF states that it has sufficient coverage to provide representative insights, with data aggregated at the postcode area level. Even so, work is underway to develop weighting methodologies, which are important for further enhancing representativeness and strengthening the robustness of the insights.

Smart Data Foundry has also recently launched its Economic Wellbeing Explorer, a map-based tool that provides insights into financial wellbeing [20]. It is intended to support more informed approaches to resource planning, service design, and policy development, helping to guide evidence-led interventions. Available as a data-driven dashboard, the tool is underpinned by a dataset built on 5.3 million consumer bank accounts, with data dating back to November 2023. Users can explore financial behaviours—such as overdraft usage, living beyond one's means, and low emergency resilience—across different income and age bands. While these indicators provide valuable insights, it is important to recognise that they represent just a single aspect of a broader picture. To ensure a comprehensive understanding of financial wellbeing, they should be considered alongside other relevant evidence.

The resource holds strong potential for researchers and policymakers seeking to better understand financial vulnerability at both national and local levels, provided that the dataset's representativeness can be assured. Upcoming enhancements include additional indicators, such as high levels of benefits dependency, which will further expand its utility in exploring the links between income, resilience, and wellbeing.

As the number of research projects utilising these datasets continues to grow, it will be important to see how such resources complement existing data sources and enhance researchers' and policymakers' ability to understand income instability and its wider social impacts. A key focus for SHERU is the relationship between income, expenditure, and health outcomes. A notable project in this area is the *Homes, Heat and Healthy Kids Study*—in which Smart Data Foundry is one of the data partners. This work is developing a new anonymous dataset that links electronic health records with data on home energy efficiency and usage, smart meters, high street banking, air pollution, and climate [21]. SHERU anticipates that emerging insights from this work will help shape effective policy responses that improve outcomes for communities affected by socioeconomic inequalities.

While the datasets curated by Smart Data Foundry are primarily quantitative, they could be further enriched through the integration of qualitative data—such as community-level focus groups or lived experience interviews—especially when exploring the intersection of financial behaviour and wellbeing. For instance, pairing transactional insights with qualitative narratives could offer a fuller understanding of how individuals cope with financial instability, and how this affects their mental and physical health over time. SHERU is particularly interested in these mixed-method approaches as a way to build more holistic, actionable insights into the lived realities behind the data.

Data Linkage and Integration: Maximising the Utility of Existing Data

Beyond mitigating the challenges of survey-based research, there is a significant opportunity to enhance insights through more effective use of existing data. By linking administrative records, survey data, and other complementary sources, policymakers and researchers can develop a deeper understanding without the need to rely exclusively on new data collection.

The Digital Economy Act 2017 (DEA) provides a legal framework that enables public authorities to share de-identified data with accredited researchers for research that serves the public good [22]. This has paved the way for richer, cross-sector datasets that shed light on critical societal issues such as employment outcomes, social mobility, education, and poverty. One prominent example is the Longitudinal Education Outcomes (LEO) dataset, which links education and training records with HMRC tax and earnings data and DWP benefits data, enabling researchers to analyse how education impacts long-term employment and income outcomes [23].

While the potential of data linkage to generate valuable insights is becoming more widely recognised, realising this potential is far from straightforward. Despite dedicated funding, many projects have struggled to progress due to complex data access processes, governance barriers, and the time required to secure approvals. These challenges can delay or even prevent planned linkages from being carried out. However, important work is underway to address these issues. Research Data Scotland (RDS), for example, has launched the Researcher Access Service—a digital platform designed to streamline and simplify how approved researchers request and access public sector data for research [24]. Whilst the service currently provides access to health datasets, RDS is actively working to expand the service to include data from other areas such as education, social care, and justice.

An approved researcher is typically affiliated with a recognised organisation, such as a university, public body, or accredited research institution. They must have a legitimate, ethical research purpose that serves the public interest and are required to complete appropriate training, such as Safe Researcher Training, to ensure they handle data responsibly.

Data Linkage and Integration: Example 1 - Wage and Employment Dynamics (WED) project

The Wage and Employment Dynamics (WED) project—funded by ADR UK, brings together researchers from the University of the West of England (UWE), University College London (UCL), City, University of London, and the National Institute of Economic and Social Research (NIESR) [25] [26]. The project links key administrative datasets to examine labour market dynamics in Great Britain, focusing on the experiences of low-paid workers, gender pay gaps, and migrant labour. Its findings aim to inform policies that tackle in-work poverty and address structural inequalities.

A key advancement is the creation of newly linked datasets, notably combining the Annual Survey of Hours and Earnings (ASHE) with HMRC's Pay As You Earn (PAYE) and Self-Assessment (SA) data. This integration provides richer insights into earnings trajectories and employment patterns across Great Britain. Thus, enabling a deeper understanding of how individuals move between employee jobs and self-employment. The data also sheds light on wage progression across different groups over time, supporting the design of targeted interventions to improve job quality and reduce inequality.

The ASHE linked to Census 2011 (England & Wales) is now available in the Integrated Data Service (IDS), while the ASHE linked to PAYE and Self-Assessment (GB) remains accessible through the ONS Secure Research Service (SRS). The ONS is in the process of transitioning from the SRS to the IDS, with the migration currently underway.

The IDS, led by the Office for National Statistics (ONS), is a cross-government service which aims to enhance access to data on society and the economy for important research in the UK [27]. This new Trusted Research Environment (TRE) operates on a multi-cloud infrastructure, offering accredited researchers and analysts secure access to its growing catalogue of data.

Data Linkage and Integration: Example 2 – Healthy Households project

By bringing together datasets from multiple sectors, researchers are better equipped to explore the complex and interrelated factors that influence people's lives and wellbeing. Projects such as Healthy Households exemplify the value of linking health, housing, and socioeconomic data to better understand the complex factors influencing health outcomes [28]. Supported by ADR-UK, this initiative demonstrates how integrating administrative data can uncover actionable insights into the relationships between household environments and health [29]. For example, *Warm Homes, Brighter Futures* has explored the impact of housing quality on children's health and education, while the linked dataset has also supported research into how healthcare use and local environments influence social care needs.

Through this collaborative, cross-sector approach, Healthy Households demonstrates the power of integrated administrative data to shed light on the social determinants of health—insights that would be far more difficult to obtain without linked data.

Data Linkage and Integration: Looking Ahead

Expanding data linkages in Scotland to include household income and wealth measures, alongside health outcomes will be an important step forward. Much of our interest lies in connecting economic data with health data to build a more comprehensive understanding of the drivers behind health inequalities. SHERU's *Collaborative Research for Meaningful Impact* publication is inspired by the Areas of Research Interest (ARI) publications produced by various UK government departments [30]. This is designed to support SHERU's engagement with academic and policy experts, helping researchers identify key topics and questions that enhance our understanding of the socioeconomic determinants of health inequalities.

A current challenge is that, while Scotland's National Safe Haven holds valuable health (and education) datasets, economic data for research is primarily accessible through Trusted Research Environments (TREs) led by the ONS.

In this context, the work on a federated approach to data access and integration is key. The Connect 4 project—a one-year initiative exploring the federation of data services across the UK's four national Trusted Research Environments (TREs)—is a crucial part of this effort [31]. It focuses on linking services such as the Integrated Data Service (IDS), led by the ONS, and the Scottish National Safe Haven. To truly unlock the potential of administrative data, we need to enable secure, streamlined connections across TREs, ensuring that researchers can access and link data across regions to address complex policy questions and deliver meaningful insights.

With this in mind, SHERU also acknowledges the contributions of Health Data Research UK (HDRUK) and the Social and Environmental Determinants of Health Driver Programme. This aims to understand how social and physical environments contribute to health inequalities in the UK. By linking health data with social and environmental factors such as income, family structure, housing, and work patterns, the programme seeks to informs policies and interventions.

Its four main strategies include using pseudonymised national data to study population factors, creating time-varying datasets linked to Unique Property Reference Numbers (UPRN), generating new UPRN-linked data assets from administrative databases, and enhancing longitudinal cohort studies with UPRN-linked data [32]. SHERU recognises these efforts as important contributions to enhancing the evidence base and advancing our understanding of health inequalities.

In recent years, the growth of non-survey data sources—such as geospatial data, satellite imagery, and administrative records—has provided new opportunities to enhance survey programmes and address longstanding challenges. These data sources offer particular value in improving the evidence base on the socioeconomic determinants of health and health inequalities by providing richer contextual information about the environments in which people live and the services they access. However, effectively integrating these diverse data assets into surveys requires further development. Research Strand 7 (Data Integration) of Survey Futures aims to inform this by conducting a systematic literature review and developing a typology of non-survey data sources that have been, or could be, integrated with UK surveys to support research on complex social and health issues [33].

The project will assess the quality of these data sources, exploring their coverage, selection, and measurement properties, and will review quality indicators and correction methods. These efforts will help ensure that integrated datasets are robust enough to inform policies targeting health inequalities and the wider social determinants of health. The findings from this work will inform a report for survey practitioners, outlining data integration options, their quality implications, and potential strategies for improvement. Additionally, a series of case studies will demonstrate how non-survey data can be leveraged to evaluate and correct for nonresponse bias, and to support survey monitoring and intervention efforts.

One area where SHERU believes data integration offers significant potential is in exploring the relationship between household incomes and health outcomes in Scotland, complementing existing tools like the Scottish Index of Multiple Deprivation (SIMD). While SIMD remains a valuable and widely used measure for identifying geographic concentrations of socioeconomic disadvantage, its area-based focus limits its ability to capture deprivation as experienced by individuals and households—including poverty and its direct impact on health.

This limitation was recently highlighted by the Access Data Short-Life Working Group, established to explore additional data sources to complement SIMD in response to calls for more effective ways to measure and target widening access to higher education—a key priority for the Scottish Government and the sector [34]. While SIMD provides insight into area-based deprivation using postcodes and datazones, it cannot account for individual or household level, circumstances, such as income. As a result, individuals experiencing deprivation outside SIMD-defined areas may be overlooked, and vice versa. The working group recommended further consideration of supplementing SIMD with individual-level indicators—such as Free School Meals registration or receipt of the Scottish Child Payment—to more accurately identify and support those facing socioeconomic disadvantage. The work concluded that this approach would strengthen the evidence base available to universities when making contextualised admissions decisions.

While not a primary focus of the Working Group's report, this issue has wider relevance for understanding how inequality affects people differently depending on aspects of their identity, such as race, disability, age, or sex. Area-based measures like the SIMD may not fully capture the lived experience of discrimination or structural disadvantage, particularly when these do not align with geographic concentrations of deprivation. This can result in important inequalities being overlooked in both analysis and policy responses.

By integrating a broader range of datasets—including income, housing, health, employment, tax records, and social security data—at both the individual and household level, researchers can develop a more nuanced and comprehensive understanding of the complex factors driving health inequalities. This approach can help identify populations most at risk and has the potential to inform more targeted, effective interventions to address socioeconomic determinants of health.

As these examples demonstrate, data linkage and integration offer transformative opportunities to strengthen the evidence base for addressing complex societal challenges, including income-related health inequalities. By enabling cross-sector insights and interventions that tackle the underlying social determinants of health, integrated data approaches support the development of more effective strategies to reduce inequalities and improve population wellbeing. As Scotland advances its data integration capabilities, there is a significant opportunity to build a more inclusive, person-centred evidence base—one that drives smarter policy decisions, improves outcomes, and ensures interventions reach those most in need.

Conclusion

Scotland stands at a pivotal moment in its efforts to reduce health inequalities and, if policies are to be more effectively designed and targeted, we need more inclusive and effective data collection. Traditional methods—particularly large-scale surveys—are facing increasing challenges. Declining response rates and difficulties in reaching diverse populations are undermining the quality, reliability, and representativeness of essential social and health data.

Yet, there is significant potential for innovation. By embracing mixed-mode surveys, expanding the use of administrative data, and exploring new and alternative data sources, Scotland can begin to close critical evidence gaps—especially for populations historically underrepresented in conventional surveys. Increased data linkage and integration offer opportunities to build a richer, more comprehensive understanding of the factors driving health inequalities.

Realising this potential will require sustained investment in data infrastructure and safeguards that build and maintain public trust. Strengthening governance processes, ensuring the ethical use of data, and communicating transparently with the public are all essential to fostering confidence in these evolving approaches to data use.

The challenges are considerable, but the opportunities are greater. By addressing barriers to data collection and integrating diverse datasets, policymakers and researchers can gain deeper insights into the socioeconomic determinants of health. This evidence will be vital for designing more targeted and effective strategies and interventions to help reduce inequalities across Scotland. By embracing this approach now, we can lay the groundwork for more informed decision-making and improved health outcomes in the future.

What’s Next for SHERU

As outlined throughout this discussion paper, SHERU is focused on key areas that strengthen insight and drive action on health inequality in Scotland:

Innovative Data Sources: SHERU aims to consider Personal Data Stores (PDS) following the completion of a pathfinder project examining their potential in supporting individuals with learning disabilities. Additionally, SHERU will monitor the growing use of the Smart Data Foundry's data offering, exploring its potential to complement existing data sources and enhance understanding of income instability and economic wellbeing—particularly the relationship between income, expenditure, and health outcomes. SHERU is also interested in the possibility of combining near real-time quantitative data, such as those provided by Smart Data Foundry, with qualitative insights from direct engagement with individuals and communities. Emerging approaches like PDS could further support this direction, contributing to a more inclusive and contextually grounded evidence base for policy and practice.

Survey Futures: SHERU will continue to track the progress and outputs of *Survey Futures*, which aims to strengthen the relevance and resilience of social surveys. This programme will help ensure that evolving methodologies address the challenges of declining survey response rates and enhance data quality. It is particularly relevant to the work of ONS and their Transformed Labour Force Survey. SHERU is also closely monitoring the work of NRS on exploring the use of a Dynamic Population Model (DPM) while conducting our own analysis of the Scotland Census 2022 data.

Administrative Data: With the increasing use of administrative data, including research by NRS to provide more frequent and accurate population estimates, there is a growing focus on linking data from various administrative sources. This includes expanding partnerships with DWP and HMRC. SHERU will collaborate closely with ADR-Scotland and RDS to inform data acquisition and sourcing priorities, as well as the selection of datasets for linkage to enhance future research on socioeconomic determinants of health inequalities in Scotland. As work progresses to improve the FRS dataset, SHERU will also develop proposals for linking to Scottish benefits data and explore the potential for connecting to health data.

Data Linkage and Integration: SHERU is focussed on maximising the use of existing data and supports a federated approach to data access and integration, strengthening connections across TREs. By championing effective data linkages—particularly between household income and health outcomes— SHERU aims to facilitate deeper research on health inequalities in Scotland. A key focus is moving beyond SIMD by integrating diverse, evolving data sources, including those related to individual living standards. These are among the areas outlined in SHERU's *Collaborative Research for Meaningful Impact* output. This aims to strengthen the link between research and policy, ensuring that studies on socioeconomic determinants of health inequalities not only address current challenges but also proactively anticipate and help prevent future risks.

Building connections: SHERU aims to strengthen collaborations with organisations like HDRUK and its Social and Environmental Determinants of Health Research Driver Programme. This initiative is developing national infrastructure to integrate social and environmental data with health records—aligning closely with SHERU's goals. A key aspect of this work involves leveraging Unique Property Reference Numbers (UPRNs) to link social and environmental datasets with health records, enabling more precise analysis of how various factors contribute to health inequalities in the UK.

Sources

[1] Office for Statistics Regulation. (2023). Ed Humpherson to Alex Lambert and Darren Morgan: Labour market statistics derived from the ONS Labour Force Survey (LFS). Available at: https://osr.statisticsauthority.gov.uk/correspondence/ed-humpherson-to-alex-lambert-and-darren-morgan-labour-market-statistics-derived-from-the-ons-labour-force-survey-lfs/

[2] Scottish Government. (2025). The Scottish Government's Quality Assessment of the Office for National Statistics Labour Force Survey and Annual Population Survey data for Scotland. Available at: https://www.gov.scot/publications/quality-assessment-of-labour-market-data-sources-phase-1/

[3] Office for National Statistics. (2025). Labour Force Survey performance and quality monitoring reports. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/labourforcesurveyperformanceandqualitymonitoringreports

[4] Office for Statistics Regulation. (2024). Ed Humpherson to Michael Keoghan: Suspension of the accredited official statistics status for the estimates ONS produces from the Annual Population Survey. Available at: https://osr.statisticsauthority.gov.uk/correspondence/ed-humpherson-to-michael-keoghan-suspension-of-the-accredited-official-statistics-status-for-the-estimates-ons-produces-from-the-annual-population-survey/

[5] Office for National Statistics. (2024). Labour market transformation – update on progress and plans: December 2024. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/labourmarkettransformationupdateonprogressandplans/december2024

[6] Resolution Foundation. Corlett, A. (2024). Get Britain's Stats Working: Exploring alternatives to Labour Force Survey estimates. Available at: https://www.resolutionfoundation.org/publications/get-britains-stats-working/

[7] Scottish Health Equity Research Unit. (2025). A new model for economic inactivity in Scotland: Replicating alternative labour market projections. Available at: https://scothealthequity.org/a-new-model-for-economic-inactivity/

[8] National Records of Scotland. (2024). Scotland's Census 2022 – General Report. Available at: https://www.scotlandscensus.gov.uk/about/scotlands-census-2022-general-report/

[9] National Records of Scotland. (2025). Administrative data based population, household and ethnicity estimates (2016–2022): Statistical research. Available at: https://www.scotlandscensus.gov.uk/about/administrative-data-based-population-household-and-ethnicity-estimates-2016-2022-statistical-research/

[10] Scottish Health Equity Research Unit. Catalano, A., & Jack, D. (2024). Scotland's Census: Understanding changes in health and socioeconomic inequality since 2011. Available at: https://scothealthequity.org/scotlands-census-understanding-changes-in-health-and-socioeconomic-inequality-since-2011/

[11] Survey Futures. Available: https://surveyfutures.net/

[12] Scottish Government. (2024). Mixed mode research to inform the Scottish Government Long Term Survey Strategy: Summary report and framework to support decision-making. Available at: https://www.gov.scot/publications/mixed-mode-research-report-inform-scottish-government-long-term-survey-strategy/documents/

[13] Department for Work and Pensions. (2024). Family Resources Survey transformation: Integrating administrative data into the FRS. Available at: https://www.gov.uk/government/publications/family-resources-survey-transformation-integrating-administrative-data-into-the-frs

[14] Institute for Fiscal Studies. Ray-Chaudhuri, S., & Wernham, T. (2025). New data suggest dramatic decline in living standards, but it may just be the quality of the data that has declined. Available at: https://ifs.org.uk/articles/new-data-suggest-dramatic-decline-living-standards-it-may-just-be-quality-data-has

[15] Department for Work and Pensions. (Accessed April 2025). Statistical work programme. Available at: https://www.gov.uk/government/publications/dwp-statistical-work-programme/statistical-work-programme

[16] Medium. Hill, S. (2020, October 16). Introducing Macmillan My Data Store. Available at: https://medium.com/macmillan-my-data-store-pilot/introducing-macmillan-my-data-store-365359c692c8

[17] Smart Data Foundry. Available at: https://smartdatafoundry.com/

[18] Joseph Rowntree Foundation. Piazza, R. (2024). Income volatility (banking data) with Smart Data Foundry. Available at: https://www.jrf.org.uk/savings-debt-and-assets/income-volatility-banking-data-with-smart-data-foundry

[19] Adami, R. (2022). Financial exclusion in the UK: Evidence on ethnicity. Social Policy and Society, 23(3), 529–547. Available at: https://doi.org/10.1017/S1474746422000252

[20] Smart Data Foundry. Economic Wellbeing Explorer. Available at: https://smartdatafoundry.com/solutions/insights/economic-wellbeing-explorer

[21] Usher Institute, University of Edinburgh. Swann, O. (Principal Investigator). Homes, heat and healthy kids study. Available at: https://usher.ed.ac.uk/research/medical-informatics/homes-heat-and-healthy-kids-study

[22] Information Commissioner's Office. Data sharing across the public sector: The Digital Economy Act codes. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/data-sharing-a-code-of-practice/data-sharing-across-the-public-sector-the-digital-economy-act-codes/#dea

[23] Department for Education. (2024). Longitudinal Education Outcomes (LEO) data. Available at: https://www.gov.uk/government/publications/longitudinal-education-outcomes-leo-dataset/longitudinal-education-outcomes-leo-data

[24] Research Data Scotland. (2024). Researcher Access Service. Available at: https://www.researchdata.scot/work-and-impact/current-projects/researcher-access-service/

[25] Wage and Employment Dynamics. The WED Project. Available at: https://www.wagedynamics.com/

[26] ADR UK. (2023). Wage and employment dynamics in Britain. Administrative Data Research UK. Available at: https://www.adruk.org/our-work/browse-all-projects/wage-and-employment-dynamics-in-britain-143/

[27] Integrated Data Service. Available at: https://integrateddataservice.gov.uk/

[28] Healthy Households. How it works: Household analysis. Available at: https://healthyhouseholds.org.uk/

[29] ADR UK. (2024). Healthy Households. Administrative Data Research UK. Available at: https://www.adruk.org/our-work/browse-all-projects/healthy-households/

[30] Jack, D (2025) Enhancing the Evidence Base: Collaborative Research for Meaningful Impact. The Scottish Health Equity Research Unit (SHERU). Available at: https://doi.org/10.17868/strath.00092930

[31] Research Data Scotland. (2024). Connect 4. Available at: https://www.researchdata.scot/work-and-impact/current-projects/connect-4/

[32] Information Commissioner's Office. Pseudonymisation. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/pseudonymisation/

[33] Survey Futures. Research Strand 7: Data Integration. Available at: https://surveyfutures.net/research-programme/research-strand-7/

[34] Scottish Government. (2024). Final report of the Access Data Short-Life Working Group. https://www.gov.scot/publications/final-report-access-data-short-life-working-group/

Citation

Jack, D (2025) Enhancing the Evidence Base: Addressing Survey Response Challenges and Leveraging Data Integration in Scotland. The Scottish Health Equity Research Unit (SHERU) https://doi.org/10.17868/strath.00092931

Scottish Health Equity Research Unit

Insights, analysis and action on the socio-economic factors that shape health

---

The Scottish Health Equity Research Unit is supported by the Health Foundation, an independent charitable organisation working to build a healthier UK, as part of its Driving improving health and reducing health inequalities in Scotland programme.

Contact us at **sheru@strath.ac.uk**