

Homelessness Data England linked dataset

Key messages from stakeholders

April 2022



Abstract

This report records a roundtable discussion between third sector, policy and academic stakeholders and Homelessness Data England (HDE) project partners on the potential and use of the HDE dataset. It captures the main findings of the discussion with the intention to inform the future use of the dataset, including research priorities, scope for growth of the dataset, how to facilitate the voices of people with lived experience of homelessness in shaping the direction of research, and more.

Areas of research focus identified by stakeholders will inform expansion of the dataset and include the need for greater understanding of the incidences, drivers and length of homelessness. Stakeholders noted that those with lived experience of homelessness can provide insight on real, hidden and perceived needs and the impact of homelessness reduction policies; while those in the third sector can advise on dissemination of project findings. Ethical considerations include undertaking public engagement activities to mitigate project risks and possible perceived harms of this research.

This meeting was held on Tuesday 4 October 2021; a full list of attendees and areas for future research can be found in appendices 1 and 2, respectively.

1. Introduction

Collaboration across stakeholder groups is essential to ensure that research using public sector data is truly in the interests of those it hopes to benefit. Stakeholders include the intended users of research findings (for example, policymakers, service providers and public advocacy groups). All these groups have an interest in knowing whether the focus and methods of research using public sector data are ethical, robust and useful, and that any potential negative consequences have been considered and mitigated.

On 12 October 2021, a range of stakeholders came together to discuss the Homelessness Data England (HDE) dataset, a newly linked data resource (see Box 1) for research in the public interest. The purpose of the event was for stakeholders to find out more about the progress of the HDE project so far, the data that is being linked and the potential of this data resource. Discussion covered:

- priority research areas for the HDE project and scope for future linkages
- how people with lived experience of homelessness and the third sector could be engaged with this work
- what, if any, are the ethical considerations of this project and how to mitigate them where possible.

Stakeholders included policymakers working on housing policy and members of the third sector who work with, or on the behalf of, people who have experienced homelessness. A list of attendees, including stakeholder, the HDE research team and ADR UK can be found in Appendix 1. This report covers the key messages arising from discussions at the stakeholder event.

Box 1: Overview of the HDE linked dataset

The HDE project involves linking Homelessness Case Level Information Collections (H-CLIC) data from across local authorities in England, and to other administrative datasets. H-CLIC data records local authorities' actions under the 2017 Homelessness Reduction Act. The project has three elements, some of which will be phased:

1. Use linked H-CLIC data from across local authorities to better understand whether homelessness is resolved in the long term.
2. Match together H-CLIC data and the Rough Sleeping Evaluation Questionnaire to identify the effectiveness of homelessness interventions.
3. Match H-CLIC data to data gathered from other government departments or health agencies to determine the wider circumstances and outcomes of people who have experienced homelessness.

The de-identified HDE linked dataset will soon be made available to external researchers via the Office for National Statistics (ONS) [Secure Research Service \(SRS\)](#), to facilitate research on homelessness and the impact it has on people's lives. Researchers will need to be [accredited](#) and submit a successful application to access the data.

2. The research value of the HDE dataset as identified by the HDE team

The HDE project involves the creation of a de-identified, linked administrative dataset on homelessness in England, to enable researchers to develop a deeper understanding of its causes and impacts. This will allow decision makers to develop more effective policies to reduce homelessness and improve the lives of people across the country.

Data collection for the HDE project includes collecting personal identifiers from local authorities in England to create a de-identified look-up table. This facilitates the linking of Homelessness Case Level Information Collection (H-CLIC) data over time, across geographical boundaries, and to data from other government departments. The data includes all cases in which a homelessness application has been made and any legal duty of care accepted. This includes rough sleepers accommodated during the coronavirus pandemic. Some research questions the HDE ID look-up table aims to answer are described in Figure 1.

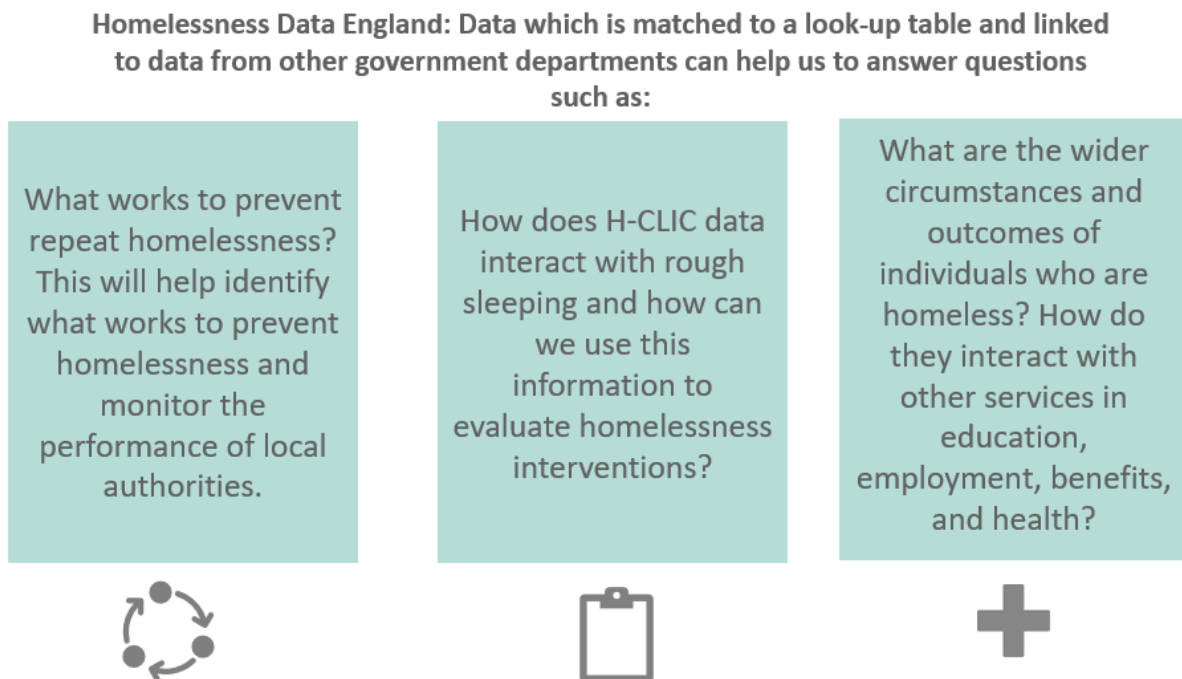


Figure 1: Research questions the HDE dataset aims to address

2.1 Project considerations

The key considerations for the HDE project, as identified by the HDE research team, include ethical issues, legal gateways and data security. The HDE project has three primary ethical considerations:

- households whose data are included in the ID spine need to be informed of this work
- whether privacy notices or informed consent are appropriate for the purposes of this project
- to ensure the data is de-identified and only used for research in the public benefit.

The HDE research team support people's right to be informed and have therefore developed privacy notices, with help from the Information Commissioner's Office (ICO), which have been shared with the local authorities who have contributed data to the project. Local authorities are expected to incorporate these notices into their existing privacy notices to ensure households are informed of this work. Data will also be checked before outputs are published to ensure it remains unidentifiable.

The HDE research team also identified that legal gateways and data security need to be considered. Legal gateways for the HDE project include compliance with data protection legislation, including UK General Data Protection Regulation, the Data Protection Act 1998, and existing legal gateways for data flow between government organisations.

Finally, to ensure the HDE project is in the public interest, the data must be protected. Data shared with researchers via ADR UK is de-identified, with rigorous safeguards in place to protect it from re-identification. When accessible to researchers, data does not include any personal identifiers, with any elements that could be traced directly back to individuals – such as names, contact details or any identifying numbers – having been removed first. What is left is a set of attributes of individuals along with recorded interactions with public services, allowing for relationships between these to be analysed. These safeguards are in line with the '[Five Safes](#)', as developed by the ONS.

2.2 Project benefits

This multidimensional dataset will help researchers and policymakers better understand experiences of homelessness. The addition of longitudinal data will assess the impact of public service use by people experiencing homelessness, in various areas of the country and across their lifetime. This will enable the analysis of interventions over time and across different local authorities during instances in which people travel across the country. Therefore, the extensive data linkage of the HDE dataset aims to close the gap between what people are seeing on the ground and existing evidence that currently informs policy decisions.

An aim of this project is to produce national-level and local-area reports to better understand the picture of homelessness across the country. To form this picture, participation from a large number of local authorities is required. Progress thus far, detailed in Figure 2, includes guidance written in partnership with ICO and data sharing agreements circulated to 37 local authorities in England. The next steps involve reviewing the engagement process with local authorities and determining which groups of local authorities should be prioritised. The fourth collection of local authority data began in January 2022.

Other government departments, such as the Ministry of Justice (MoJ) and Department for Work and Pensions (DWP), will be approached to contribute data to the project this year. Linking data from other government departments to the HDE database will provide opportunities to explore, for instance, the relationships between homelessness and the criminal justice system or the benefits system.

Groundwork, 2020

- Guidance document written with input from the Information Commissioner's Office
- Data Sharing Agreements and a Data Protection Impact Assessment written and circulated to local authorities

March 2021

- 10 local authorities provided Personal Identifier Data in the first data collection
- Office for National Statistics (ONS) downloaded and ingested data from Delta

April / May 2021

- ONS cleaned and created an ID spine to anonymise the Personal Identifiers for circa 3000+ records of the first data collection data

June 2021

- A second data collection opened and 17 local authorities submitted data, circa 2400 records submitted

September 2021

- The third data collection closed at the end of September, circa 7000 records submitted by a total of 20 local authorities

Figure 2: A timeline of progress from inception to September 2021

3. Key messages from stakeholders

Following a presentation by the HDE team on the research value of the dataset came discussion among stakeholders on priority areas of research; the value of engaging with those who have lived experience of homelessness and with the third sector; and remaining ethical considerations and ways to mitigate them.

3.1 Priority areas of research the HDE project should aim to address

Key to the success of the HDE project is filling in existing knowledge gaps, which can be done through the expansion of the dataset by linking in more data. Currently there remains a large number of local authorities yet to be engaged with this work, and it is anticipated the expansion of the project will incentivise more local authorities to get involved. Glen Bramley, Professor of Urban Studies, Heriot-Watt University, flagged the importance of involving more local authorities in this project. There is significant variation between local authorities in broadly the socio-economic and housing market context, and in the detailed policy implementation and practice around homelessness and related services. More data is therefore crucial for more detailed analysis of the effects of relevant policies and services. Engaging with more local authorities can also provide insight into why there are stark disparities between different ethnic groups and their risk of homelessness.

Once enough local authorities have submitted data to the project, government departments will be approached about sharing their data. Dr Peter Mackie, ADR Wales researcher from Cardiff University, highlighted the importance of linking in other government department data. For similar work occurring in Wales, researchers found that some people identify themselves as being homeless in health datasets but for various reasons don't appear in standard local authority datasets designed to measure homelessness. Similarly, Professor Bramley noted the [Health and Homelessness in Scotland data linkage project](#) demonstrates the value of linking multiple health datasets to better understand the effects of homelessness. Data linkage across other datasets, such as health, but also education, police data and others can help form a more comprehensive picture of homelessness.

In addition to counting more instances of homelessness, linking in more data can also provide insight into the causal links of homelessness. A representative from Crisis, a homelessness charity, flagged that information on homelessness isn't currently collected by DWP, for instance, which means data from other sources is needed to better understand housing stability and other areas of life, such as employment. People also tend to drop out of datasets aimed at capturing homelessness following successful interventions, so linking to housing history and location data can provide insights into longer-term outcomes of interventions.

A primary aim of HDE is better understanding drivers of homelessness to determine the effectiveness of interventions. Stakeholders suggested linkage to historical data from data sources such as the National Pupil Database (NPD), which records education data including school absences, and information from the Census, can produce a more holistic picture of the past circumstances of people who later experienced homelessness. Linking to education data has been considered as it can provide insights into the longer-term impact of homelessness on children, and Census data may be incorporated in the future. Linkage to longitudinal data can

also inform patterns of service use, which can help determine the length of time people experience homelessness.

3.2 How people with lived experience of homelessness and those working in the third sector should be engaged with this project

Stakeholders from third sector groups working with, or on the behalf of, people experiencing homelessness identified homelessness prevention as a knowledge gap, particularly concerning real, perceived and hidden needs. Understanding causal links of homelessness but incorporating data from other government departments, can provide a unique understanding of the effectiveness of homelessness reduction policies. While linking in more data can help fill in some knowledge gaps, lived experience of homelessness also offers valuable insights into understanding and addressing homelessness. Evidence on the ground from third sector groups offers an understanding of what happens during and after homelessness, but less about people's journeys into homelessness. It was identified that those who have a lived experience of homelessness are better able to flag the right questions the HDE dataset should focus on, thereby helping researchers better understand which datasets need to be included.

“The most helpful thing about understanding homelessness is knowing how [people] got there.”

Representative from Thames Reach

Government stakeholders expressed support for the publication of H-CLIC data into the public domain. H-CLIC is an information-rich data source, and it was felt findings of its analysis could support services, including third sector groups, to help inform their approach to homelessness.

3.3 Outstanding ethical considerations and how they might be mitigated

The discussion concluded with stakeholders and the HDE research team displaying agreement about the potential benefits of this project and need to mitigate ethical considerations.

“The real-world impact of this work is the fundamental case for doing this and the relevant mechanisms are in place not to infringe on people's personal data.”

Glen Bramley, Heriot-Watt University

Continued public engagement was noted in relation to mitigating risks. ADR UK [public panels](#) are available for all ADR UK-funded research, including the HDE project, to sense-check outputs, ensuring language is appropriate and socially sensitive before publication. Incorporating public engagement throughout the course of this work will be crucial to identifying potential adverse impacts of this work, including misrepresentation of this research. A suggested solution is working with the third sector to vet

research proposals seeking access to the HDE database to ensure research remains within the public interest.

4. Next steps

Signed data sharing agreements from all local authorities in England are needed to fulfill the aims of the HDE project but progress to date indicates that this will take some time. Demonstrating the value of the HDE project could be a key factor in securing local authority support and their agreement to share data. It is, therefore, proposed that the project is broken down into two stages to facilitate this.

4.1 Stage 1 (Jan-Mar 2022) Enhanced local authority engagement working with external partners

The first stage includes working with the anonymised H-CLIC data to understand more about repeat homelessness, identify the factors associated with repeated episodes and to identify potential interventions that can successfully reduce or mitigate homelessness. Once this work is complete, HDE colleagues within the Department for Levelling Up, Housing and Communities (DLUHC) will follow up on the themes discussed at this stakeholder event by working with the Centre for Homelessness Impact (CHI) to engage with targeted local authorities. This will initially be through the Local Government Association. This engagement work will aim to understand more about barriers to participation, enabling DLUHC to develop updated guidance materials incorporating learning over the duration of the project.

Other activities during this period include:

- Engagement with the Information Commissioner's Office (ICO). This includes assessing the feasibility of ICO webinars and a frequently asked questions document, co-written by HDE project partners.
- Scoping work on linking H-CLIC PI data with data from MoJ, Public Health England and DWP. This includes logistically determining how to extend existing agreements and curating new ones with other government departments to link their data to H-CLIC data.
- Further discussions with CHI on the development of the project.
- Making flows data in H-CLIC available in the ONS Secure Research Service. This includes assessing the feasibility and utility of the data, working with ONS colleagues, and developing documentation to support accredited researchers who wish to apply to use this data for research in the public interest.

4.2 Stage 2 (April-Oct 2022): Production of datasets and initiation of linkage process

Following enhanced engagement, the next step will be to link data collected from the DLUHC rough sleeping questionnaires and other smaller scale collections (including the Next Steps evaluation) to H-CLIC data.

The proposed linked datasets will be made available in the SRS and will follow the 'Five Safes' which are:

- *Safe Data:* Data held within ADR UK's trusted research environments is de-identified, meaning identifiable details such as names, addresses and identification numbers are removed before data is made available for any analysis.
- *Safe Person:* Any researcher accessing ADR UK data is assessed for their skills and suitability before being granted access to the data needed for their project.
- *Safe Place:* ADR UK data must be accessed in a safe and secure room within one of our trusted research environments, or otherwise via an assured connection at an accredited institution or SafePod.
- *Safe Project:* The research project itself is scrutinised and must be in the public interest.
- *Safe Output:* The researcher's actions whilst accessing the data are monitored using keystroke technology and all outputs are checked thoroughly, with any potentially re-identifiable outliers removed.

The internal linkage work under this stage is likely to highlight issues which will need to be resolved before embarking on the wider H-CLIC linked dataset. Significant work is required to identify the data to be linked, particularly in relation to specifying time periods for linking homeless data on access to wider public services. The ambition is to be able to use the H-CLIC look-up table to link to MoJ data on prison leavers and probation services, health data on hospital episodes and drug treatment data.

Ongoing work will also be required to refine use cases to test, working alongside local authorities, and mechanisms for sharing outputs. A more detailed project plan for this stage of the project would be produced in early Spring 2022.

4.3 Public engagement

Researcher and third sector engagement will be key to the development of the HDE project, especially in Stage 2, to make sure that the final dataset fits the needs of those experiencing homelessness and is informed by local authority input. The engagement work with CHI will help the HDE team identify research priorities and raise awareness of the project via presentations and discussions at CHI researcher networking events.

Acknowledgements

The Homelessness Data England project is a collaboration between the [Department for Leveling Up, Housing and Communities](#) and the [Office for National Statistics \(ONS\)](#). It is funded by [ADR UK \(Administrative Data Research UK\)](#). ADR UK is funded by the [Economic and Social Research Council](#) (part of [UK Research and Innovation](#)).

Author

Shayda Kashef, Public Engagement Manager, ADR UK Strategic Hub, shayda.kashef@ukri.org

Visit the [ADR UK website](#)



[@ADR UK](#)



Appendix 1: List of attendees, ordered alphabetically by organisation

Chair: [Ligia Teixeira](#), Chief Executive, Centre for Homelessness Impact

<i>HDE project partners</i>	Representatives
Administrative Data Research UK	Rosie French, Deputy Director
Administrative Data Research UK	Emily Oliver, Head of Training & Capacity Building
Administrative Data Research UK	Christine Boase, Strategic Lead for Campaigns & Communication
Administrative Data Research UK	Bogusia Wojciechowska, Research Manager
Administrative Data Research UK	Shayda Kashef, Public Engagement Manager
Administrative Data Research UK	Grainne Wrigley, Communications & Engagement Officer
Department for Education	Chris Douglass, Head of Data Sharing and Data Platform
Department for Levelling Up, Housing and Communities	Lan-Ho Man, Senior Research Officer
Department for Levelling Up, Housing and Communities	Catherine Barham, Head of Homelessness and Social Housing and HDE Project Lead
Department for Levelling Up, Housing and Communities	Jeremy Swain, Adviser on Homelessness and Rough Sleeping (Ex Thames Reach)
Department for Levelling Up, Housing and Communities	Valerie Vilorio, Statistician
Department for Levelling Up, Housing and Communities	Jo Beck, Homelessness Advice and Support Manager
Department for Levelling Up, Housing and Communities	Kirsty Hendry, Policy Adviser
Department for Work and Pensions	Maisie Payne, Senior Research Officer
Department for Work and Pensions	Graham Knox,
Office for National Statistics	Rhodri Williams, Data Acquisitions
Ministry of Justice	Kirby King, Head of Data Linking
Office for National Statistics	Nick O'Donnell, Head of Integrated Data Services
Office for National Statistics	Lorraine Ireland, Head of Strategic Data Curation
Office for National Statistics	Emma Atkinson, Senior Engagement Manager
Office for National Statistics	Alice Palmer-Ross, Homelessness Data Specialist Officer
Office for National Statistics	Hannah Garrington, Impact Officer
<i>External stakeholder</i>	Representatives
Cardiff University	Dr Peter Mackie, Reader (ADR Wales)
Centre for Homelessness Impact	Ligia Teixeira, Chief Executive
Centre for Homelessness Impact	Guillermo Rodriguez, Head of Evidence and Data
Crisis	Maeve McGoldrick, Head of Policy and Campaigns

Heriot-Watt University	Glen Bramley, Professor of Urban Studies (ADR Scotland)
Local Government Association	Charli Fudge, Policy Adviser (Homelessness)
St Basil's	Chris Harper, Director of Operational Support
Thames Reach	Bill Tidnam, Chief Executive

Appendix 2: Areas of interest for future research using the HDE linked dataset identified by stakeholders

Potential future linkages

- Can linking education, health, and/or employment data to the HDE database generate new insights into drivers of homelessness?
- Which historical datasets are needed to accurately capture instances of homelessness?
- What data is needed to best capture the length of time people experience homelessness?
- How can we work with local authorities to improve their data collection methods?

Engaging with people who have lived experience of homelessness and the third sector

- How can we better differentiate between real, perceived and hidden needs of people experiencing homelessness?
- How can lived experience of homelessness inform homelessness reduction policies?
- How can a better understanding of people's journeys into homelessness inform homelessness reduction policies?
- How can we best package findings from homelessness data sources, such as H-CLIC, to support work undertaken by the third sector?

Ethical considerations and suggested ways to mitigate them

- How can public engagement help to identify and mitigate risks related to this project?
- What are the possible harms of this project?
- Could the HDE dataset be used for predictive analytics? If so, how can this be mitigated? If not, how can we best communicate the project's safeguards to the wider public?