Document History

Version number	Revision date	Summary of changes
0.1	25-03-21	Creation of protocol document
0.2	01-04-21	Addition of references; submission as draft to NIHR
1.0	28-04-21	Finalisation following advisory group approval
2.0	17-05-21	Amended to reflect updated advice about ethical approval

Distribution This document has been distributed as follows

Name	Responsibility	Date of issue	Version
Co-I group	Opportunity for feedback	25-03-21	0.1
NIHR	Review by funder	01-04-21	0.2
Advisory group	Opportunity for feedback	15-04-21	0.2
Study sponsor	Signature	28-04-21	1.0
Study sponsor	Signature	17-05-21	2.0

NIHR | National Institute for Health Research

Project: **Unlocking data to inform public health policy and practice**

STUDY PROTOCOL

v2.0

Start date: 1st April 2021
End date: 31st January 2022

Purpose The purpose of the Protocol is to describe the study/project and provide information about the procedures for entering participants into the study/project. Every care has been taken in drafting this protocol; however, corrections or amendments may be necessary.

This protocol has been authorised by:

Name	**Role**	**Signature**	**Date**
Emily Tweed	Joint Principal Investigator		**17-05-21**

On behalf of project advisory group

Name	**Role**	**Signature**	**Date**
Debra Stuart	Study sponsor		**23/6/21**

MRC/CSO Social and Public Health Sciences Unit

MRC Medical Research Council · CHIEF SCIENTIST OFFICE · University of Glasgow

Table of Contents

Contact details – Chief Investigator & Co-Investigator

CHIEF INVESTIGATORS

Title and name	Emily Tweed	Peter Craig
Position	Clinical Lecturer	Prof Public Health Evaluation
Add1	MRC/CSO Social and Public Health Sciences Unit	
Add2	Berkeley Square, 99 Berkeley Street, Glasgow	
Postcode	G3 7HR	
Tel :	0141 353 7528	
E-mail :	emily.tweed@glasgow.ac.uk	peter.craig@glasgow.ac.uk

CO-INVESTIGATOR(S)

Title and name	Mirjam Allik	Title and name	Denise Brown
Position	Research Fellow	Position	Statistician/Epidemiologist
Add1	MRC/CSO Social and Public Health Sciences Unit	Add1	MRC/CSO Social and Public Health Sciences Unit
Add2	Berkeley Square, 99 Berkeley Street, Glasgow	Add2	Berkeley Square, 99 Berkeley Street, Glasgow
Postcode	G3 7HR	Postcode	G3 7HR
E-mail	mirjam.allik@glasgow.ac.uk	E-mail	denise.brown@glasgow.ac.uk
Title and name	David Henderson	Title and name	Charlie Mayor
Position	Research Fellow	Position	Safe Haven Manager
Add1	Usher Institute	Add1	
Add2	University of Edinburgh	Add2	
Postcode	EH8 9AG	Postcode	
E-mail	David.henderson@ed.ac.uk	E-mail	Charlie.mayor@ggc.scot.nhs.uk
Title and name	Petra Meier	Title and name	Jill Pell
Position	Prof Public Health	Position	Director of Institute
Add1	MRC/CSO Social and Public Health Sciences Unit	Add1	1 Lilybank Gardens
Add2	Berkeley Square, 99 Berkeley Street, Glasgow	Add2	Glasgow
Postcode	G3 7HR	Postcode	G12 8RZ
E-mail	Petra.meier@glasgow.ac.uk	E-mail	Jill.pell@glasgow.ac.uk
Title and name	Nick Watson		
Position	Chair of Disability Studies		
Add1	Adam Smith Building		
Add2	University of Glasgow		
Postcode	G12 8RT		
E-mail	Nicholas.watson@glasgow.ac.uk		

Sponsor: Debra Stuart, Research Governance Officer, MVLS Research Administration (debra.stuart@glasgow.ac.uk)

Funder: NIHR - 133585

1. Summary

Secondary data from non-health sources is vital to understanding how initiatives in policy areas such as education, housing, or social care affect population health and health inequalities.

At present, the potential of secondary data is not being fully realised due to challenges in governance, data access, and quality. Although these have been extensively documented in relation to healthcare data there has been less attention to the challenges of combining data from health and other sectors to inform healthy public policy. In particular, we know little about the perspectives of people who might use such cross-sectoral secondary data in their work as decision-makers in government, NHS public health teams, and third sector organisations.

Scotland's commitment to a 'whole-systems' approach to public health requires partnership working across sectors and new forms of data which are dynamic, can be extensively linked, and originate from real-world rather than experimental settings. Cross-sectoral administrative data is well placed to meet this need.

We propose to work collaboratively with decision-makers with diverse roles in healthy public policy to identify practical ways that secondary data can be more effectively used across sectors to support a whole-systems approach to public health, using case studies of existing linkage projects. We will ask:

RQ1. What is the current and potential future role of cross-sectoral secondary data in decision-making in local and devolved government, NHS, and third sector, in the context of a whole-systems approach to public health?

RQ2. What transferable lessons can be learnt from recent projects using cross-sectoral administrative data for public health research?

RQ3. What are the actions required to develop and maintain secondary data systems capable of supporting a whole-systems approach to healthy public policy at the local and national levels?

We will address these questions through workshops with decision-makers from local and national government, community planning partnerships, NHS public health teams, and the third sector, as well as data providers, analysts, governance panels, and public representatives. The workshops will draw on (a) a scoping review on models of evidence use for public health practice and policy and known barriers/facilitators to the use of administrative data in this context and (b) three case studies of existing cross-sectoral linkage projects, covering the policy areas of social care; care-experienced children and young people; and homelessness, criminal justice, and substance use.

The project will be guided by an advisory group including Public Health Scotland, the Convention of Scottish Local Authorities, the Improvement Service, the West of Scotland Safe Haven, NHS public health teams, and public representatives. The group will support recruitment of workshop participants to ensure representation across all relevant sectors; guide the development of workshop materials; and facilitate dissemination of results to practice and policy audiences.

2. Introduction

2.1 Background

The use of secondary data from non-health sources is vitally important in public health, given the centrality of social, political, environmental, and economic factors in influencing individual and population health (1-3). The development and evaluation of healthy public policies at the local and national level requires not only data from healthcare but also from sectors such as education, welfare, housing, and social care, in order to monitor and evaluate how initiatives in one policy area affect outcomes in others (2, 4, 5). In this proposal we use ‘secondary data’ as a broad term to refer collectively to the use, re-use, or novel combination of existing datasets, whether administrative data produced through routine service delivery; record linkage between datasets; and the re-use or linkage of primary research data.

Specific benefits of secondary data include large population sizes (especially for evaluating policies with subtle or heterogenous effects); multi-level linkages that integrate data from individuals, households, and wider environments; and the ability to include marginalised or disadvantaged populations often under-represented in primary research (3, 6-8). Secondary data provides rich opportunities for policy development, appraisal and evaluation, through natural experiments (4), embedded randomised controlled trials (9), and decision modelling (e.g. 10).

Yet progress in the use of secondary data has often been patchy and slow. A survey of health authorities across 29 European countries suggests relatively limited utilisation of cross-sectoral data for routine public health activities (11). Barriers to secondary data use can include: long, unpredictable timelines for approval; differing requirements across multiple data providers or jurisdictions; high governance burden for individual projects; the substantial investment required to ensure quality and usability; and uncertainty about the ethical or legal basis for data sharing. As a result, secondary data research is now often considered high-risk or poor value for money.

In preparing this proposal, we searched Web of Science, Medline/Embase, and ASSIA for studies examining barriers and facilitators of cross-sectoral use of secondary data in decision-making for healthy public policy since 2005, using a combination of terms relevant to secondary data; cross-sectoral or interdisciplinary approaches; and public health. We also searched The Knowledge Exchange grey literature database using similar terms and reviewed the websites of Administrative Data Research (ADR) UK, Health Data Research UK, National Centre for Research Methods, and Research Data Scotland. We reviewed titles and abstracts/executive summaries and followed up relevant items from reference lists.

We identified that most evidence to date relates to data from the healthcare system, rather than non-health sources, and to technical and ethical issues, such as public trust and data security (7, 12-14). We found little research on ways to maximise the value of secondary data from the perspective of evidence users. In particular, there is a lack of evidence examining institutional, cultural and political factors or the role of such data in the wider context of decision-making (13-15). Some research has suggested that challenges may be greater beyond the healthcare sector, due to lack of existing working relationships; lack of familiarity with the process of health research; and perceptions that benefits are less direct or immediate (1, 16). Various articles refer to a mutual lack of understanding between researchers and public authorities and ‘cultures of aversion’ within the latter that hinder research, but few have directly engaged with stakeholders outside academia or the health sector (5, 15-18). This gap is notable given the extensive literature on public attitudes to secondary data research (19), but is in keeping with a broader neglect of policymakers' views on the use of data and evidence in the policy process (20).

Where decision-makers use secondary data, they may have different priorities to researchers, data managers, or the public. A number of studies have reported a mismatch between evidence produced by academic researchers and the needs, preferences, and constraints of decision-makers (21-23). Public health policymakers across the UK consistently report that local data is the most highly valued and frequently used form of evidence and that there is an appetite to improve the quality and utilisation of routinely collected data (for instance, 22, 24-26).

Incorporating 'end-user' voices is especially important in the context of growing support for whole-systems approaches to public health (27-29). In Scotland, this is reflected in the national Public Health Priorities agreed by Scottish Government and the Convention of Scottish Local Authorities (COSLA), and in the mandate of Public Health Scotland (30, 31). Whole-systems approaches conceptualise public health challenges as complex adaptive systems emerging from dynamic interconnections between the heterogeneous components that together make up an intervention landscape (e.g. populations; social, economic and environmental conditions; the health and social care system; global and local political structures and institutions). The application of such approaches in public health has major implications for secondary data use. For instance, breadth, spatiotemporal disaggregation and near real-time release may be higher priorities than the depth and detail valued by traditional models of academic evidence (27, 29, 32).

2.2 Rationale

To realise the full potential of cross-sectoral secondary data use in this context, there is a need to better understand the specific requirements and challenges from the perspectives of decision-makers working in policy and practice areas relevant to public health. To address this gap, we will use participatory methods centred on three case studies of cross-sectoral secondary data projects, illustrating 'real-world' examples of benefits, barriers, and solutions, to co-produce a set of recommendations relevant to cross-sectoral secondary data research in the UK and beyond.

Scotland provides an ideal context in which to undertake this work, as a multi-level public health system encompassing multiple local authority and health board areas with diverse populations and geographies, which is also small enough to build on existing close relationships with key stakeholders. The establishment of Community Planning Partnerships and the development of Public Health Scotland as a single agency for public health has created new opportunities for joint working across NHS public health, local authorities, the devolved government, and other partners. Public Health Scotland's mandate commits to a 'whole-systems' approach to public health, which will require greater availability and use of data from beyond the health system (30). Scotland has a strong track record in research using linked healthcare data and is working to build capacity in cross-sectoral sharing of administrative data, via a network of five Safe Havens. There is therefore a window of opportunity to answer questions about access, use and value; to translate the results into impact on information systems, data infrastructure and decision-making; and to build networks and capacity with a view to optimal utilisation of initiatives such as Research Data Scotland.

Framing these questions through the lens of real-world case studies will help stakeholders engage with benefits, risks, obstacles, and trade-offs; articulate tacit knowledge; and propose concrete solutions. Previous reports have highlighted the importance of such case studies in documenting and disseminating experiences of cross-sectoral data analysis, to demonstrate the value of such projects and share lessons learned (1, 13, 15). Other case studies of healthcare and cross-sectoral linkage in the UK have been limited to examining the data systems themselves, and consist of researcher experiences or reviews of documentary evidence (16): a more granular approach which traces the process of individual studies, and which engages with evidence end-users, is likely to yield greater insight into specific challenges, barriers and facilitators.

2.3 Aims/Objectives/Research questions

Aim: to work with stakeholders to identify practical ways that secondary data can be more effectively utilised across sectors to support the implementation of a whole-systems approach to public health, using learning from real-world case studies.

Research questions:

RQ1. What is the current and potential future role of cross-sectoral secondary data in decision-making as part of a whole-systems approach to public health?

RQ2. What transferable lessons can be learnt from the benefits, challenges, and solutions encountered in recent projects using cross-sectoral secondary data for public health research?

RQ3. What are the actions required to develop and maintain secondary data systems capable of supporting a whole-systems approach to healthy public policy at the local and national levels?

3 Study Design/Methods

3.1 Study Design

Conceptual framework: This project draws on systems science approaches to public health (27, 28, 32) and transdisciplinary perspectives on the use of evidence in healthy public policy and practice (20, 33-35). We recognise that varied definitions of ‘evidence’ compete with multiple other factors in a complex system for decision-making across multiple institutions (e.g. government, NHS, third sector) and levels (e.g. local authorities, community planning, health boards, devolved and national government) (22, 33, 35). These conceptual foundations will underpin our work by informing development of the scoping review and systems map in WP1; conceptualisation of the benefits, risks, barriers, and facilitators elicited by the case studies in WP2; and the scope, targets, and content of the recommendations identified in WP3. For instance, WP1 will refine the broad focus of existing models of evidence use (identified in the scoping review) into a more granular understanding of the role of secondary data (as explored in the system mapping exercise).

Methods: The project will consist of three linked work packages (WP). Each will inform the overarching set of recommendations co-produced in WP3 but will also yield stand-alone outputs of wider interest and relevance, as illustrated in the flow diagram.

We will use participatory qualitative methods for stakeholder engagement via three sequential workshops linked to the research questions, with material for the workshops drawing on a scoping review and the three case studies described below. Our choice of participatory methods reflects the value placed on: collaboration throughout the research process; the experiential knowledge of practice and policy partners; and actionable outputs of real-world relevance (36). Our approach to planning has drawn on published design principles and recommendations for stakeholder engagement (37-39), involving partners from PHS, COSLA, the Improvement Service, and NHS public health teams, as well as representatives from public and patient panels.

To address uncertainties about the Covid-19 situation during 2021 and enhance accessibility of attendance to a broad range of stakeholders, we plan to undertake the engagement workshops virtually using collaborative tools such as Miro, Kumu, and Microsoft Teams. In doing so, we draw on extensive experience of undertaking stakeholder engagement using virtual methods from within the project team (for instance, PC's NIHR-funded work on evaluability assessments as part of the PHIRST project) and the host institution of SPHSU (which hosts a number of large stakeholder networks which have held exclusively virtual events for much of 2020, such as TRIUMPH, SHINE, and SIPHER) (40-42). Viability of a virtual-only approach is supported by reflections from other similar projects which suggest that online workshops can enhance and diversify engagement (42, 43), and will benefit from the familiarity with virtual meetings that participants are likely to have developed over the past year. Workshops will be up to 3 hours in length, divided into three sections with two 15 minute breaks; a mix of whole-group and small-group work; and opportunities for task- and group-switching. Our experience of conducting online evaluability assessment workshops during 2020 suggests this length and format is acceptable to a wide range of participants.

Workshops are intended to be sequential, involving the same group of participants throughout, to build on previous activities and relationships. The Covid-19 pandemic has resulted in unprecedented pressures on public and third sector organisations that create challenges for our ability to recruit, retain, and fully engage with key stakeholders. We have therefore worked closely with non-academic partners in developing this proposal and will continue to do so during delivery, to ensure the activities and outputs are designed in a way that addresses their needs, interests and priorities. Specifically, we will maximise engagement by: [1] making initial contact with potential workshop attendees via relevant advisory group partners for each sector or institution; [2] scheduling workshops to avoid school holidays when staffing pressures are greatest; [3] running virtual workshops which avoid the need to travel and enable attendance by a wider geographical diversity of participants; [4] while the expectation is that participants will attend all three workshops, we will offer those who can only make two an online meeting with the project team on an alternative date; [5] supplementing the workshops with other opportunities to contribute, including online engagement exercises and invited email contributions; [6] providing regular feedback to participants about how their input has been used, and seeking feedback from participants during each workshop about how the next can be improved.

We anticipate approximately 30 participants at each workshop. Participants will be provided with a concise briefing paper ahead of each workshop, explaining the purpose and format in the context of the project as whole. Workshops will begin with a brief presentation summarising the material to be discussed, before participants are divided into smaller breakout groups. Group discussions will use a variety of participation techniques covering ideas generation (such as Post-it note 'races'), mapping (using Miro and Kumu), and ranking or voting. Each workshop will conclude with dedicated time for attendees to reflect and feedback on their experience of participating, and to suggest ways in which the process could be improved: this feedback will be reviewed in detail at advisory group meetings and used to shape subsequent activities. We will maximise opportunities for stakeholder engagement by offering opportunities to contribute to polls, 'graffiti walls' and prioritisation exercises before and after each workshop, and through direct follow-up with key individuals to further explore or develop issues identified. In addition to data generated from participatory exercises such as mapping and voting, workshops will be recorded and transcribed, with key themes identified using framework analysis. Participants will be informed in advance of the recording and asked to consent to this prior to attending.

In the following sections, we describe in detail the activities planned for each WP based on initial engagement with key partners: however, we will also retain some degree of flexibility to accommodate feedback that arises through the engagement process (37).

Work package 1

Objective: to work with stakeholders from local and devolved government, the NHS, and the third sector to map the current and potential role of cross-sectoral secondary data (and research based on such data) in decision-making to support a whole-systems approach to public health.

The key deliverable from WP1 will be a systems map of how administrative data can interface with decision-making in practice and policy settings, encompassing both current and ideal practice as viewed by key stakeholders. The map will be used to frame stakeholder discussions around case studies and the process of developing recommendations, but will also be valuable as a stand-alone output which can be used to inform the development, operation, and evaluation of secondary data systems in other settings and contexts. To develop the map, we will first undertake a scoping review and targeted engagement with key informants via the project advisory group. We will use these outputs to inform a systems-mapping exercise with stakeholders at the first workshop.

The scoping review will bring together existing evidence from UK settings on:

1) Models for evidence use for healthy public policy and practice, such as (26): this will include the use of administrative data, where research exists

2) Known barriers and facilitators at the individual, organisational, and societal level to the use of administrative data in decision-making for public health and healthy public policy

Development and execution of the review protocol (which we shall publish on a suitable platform such as OSF) will be supported by an experienced information scientist at SPHSU with experience in rapid scoping reviews, and informed by the project advisory group. Searches will cover the main bibliographic databases (Medline, Web of Knowledge, ASSIA etc.,) plus manual searches of websites for key data providers and governance organisations. Data from studies and reports that meet the inclusion criteria will be extracted into a template (with a sample dual-extracted to check reliability) and synthesised narratively, drawing on SWiM guidance (44).

In advance of the workshop, we will undertake preparatory engagement with participants through asynchronous online activities, e.g. polling on key system drivers. We will then combine this preliminary data, scoping review outputs, and feedback from the project advisory group to develop a broad-brush map of the use of secondary data in decision-making for public health and healthy public policy, which will serve as a departure point for more detailed engagement and mapping at the first workshop. A detailed plan for the workshop will be developed in collaboration with the advisory group based on these emerging findings, but we envisage that we will use participatory exercises to explore questions such as: what attributes of data systems (e.g. timeliness, coverage, quality) are most useful in different decision-making contexts; how to identify and prioritise the most useful datasets; how institutions and actors at the local, regional, and devolved levels could work together most efficiently; and how administrative data can be augmented from other sources to address limitations. During the workshop, members of the research team will amend the systems map in real-time based on feedback and regular ‘sense-checks’. Following the workshop, we will produce a final version of the map which will be circulated to participants for agreement and reflection ahead of the second event. Whilst the primary focus of this exercise will be mapping current and ideal future practice, we anticipate that information will begin to emerge on facilitators and barriers, which we will take forward and build on during WP2 and the second workshop.

Work package 2
Objective: to use case studies of three existing projects and systems map produced during WP1 as the basis for more detailed engagement with stakeholders on practical aspects of using cross-sectoral data, to identify transferable lessons that can underpin our final recommendations.

In WP2, we will collate three case studies of using cross-sectoral data for public health research (Table 1) and use these as the basis for detailed and constructive engagement with stakeholders on the value, barriers, and enablers of such projects. Each case study focuses on different area of social policy relevant to health across the lifecourse, where policies are set by the devolved government and implemented by local authorities or health and social care partnerships. They have been chosen to reflect a diversity of policy topics and development journeys to date: each has encountered unique challenges in realising the cross-sectoral linkage, and in several cases have had to depart substantially from their original goals, all of which we will seek to learn from.

Study leads for each case study project are represented on the co-investigator team, so we can benefit from their first-hand insights and established working relationships over the course of the grant. We will draw on their experiences and documentary evidence from approvals processes, correspondence, and stakeholder meetings to produce for each case study: [1] a narrative report detailing the project process (including data flow diagrams, timelines, and governance milestones), estimates of time and resources used, and initial reflections on lessons learned; [2] summary engagement materials for the second stakeholder workshop.

In the second workshop, participants will choose one of three ‘streams’ based on the case studies, according to their area of policy interest. After introductions, each group will begin by discussing the systems map from the previous workshop in the context of the case study topic area; this will be followed by an interactive presentation by the case study lead using prepared engagement materials; followed by a facilitated discussion on the benefits, risks, barriers, and facilitators of these and similar projects. We will particularly focus on features of the data landscape necessary to enable systems-informed research, for instance in relation to timeliness, integration, and scope, in the context of ongoing and future initiatives such as ADR-UK, Research Data Scotland and equivalents in other settings. This discussion will be framed in the context of the systems map, using this to locate and inter-relate the multiple factors at work, and identify steps in the process with disproportionate impacts on outcomes which may be priority targets for recommendations.

From case study materials and workshop findings, the project team and advisory group will synthesise a common set of transferable lessons for the use of cross-sectoral secondary data in public health research. These will be informed by the map developed in WP1 and will go on to inform the development of recommendations in WP3.

Table 1. Case studies

	Case study 1	**Case study 2**	**Case study 3**
Focus of research	Health outcomes of care-experienced children (Children's Health in Care in Scotland project; CHiCS)	Health outcomes among people with overlapping experiences of homelessness, offending, and substance use	Multimorbidity and unscheduled care among people receiving community social care
Study design	Cohort study using linked administrative data from health and non-health sources		
Scope	Scotland-wide	Glasgow City Council area	Scotland-wide
Time period	2009/10 - 2016	2010/11-2018/19	2010/11 – 2015/16
Cohort size	~ 650,000 individuals (including ~13,000 with care experience)	~1 million individuals (including ~35,000 with ≥1 of the above experiences)	~1.1 million individuals
Non-health datasets involved (sources)	Pupil Census (ScotXed, Scottish Government) Children's Looked After Statistics (CLAS; ScotXed, Scottish Government)	HL1 applications for statutory homelessness support and Criminal Justice Social Work Reports (Glasgow City Health and Social Care Partnership) Prison records (Scottish Prison Service/ Scottish Government)	Recipients of social care (Scottish Social Care Survey)
Health datasets involved (sources)	Medication dispensing (Prescribing Information System) A&E attendances Hospitalisations (Scottish Morbidity Records) Birth and death registrations (National Records of Scotland)	Dispensing for opioid dependence (Prescribing Information System) A&E attendances (Trak A&E) Hospitalisations (Scottish Morbidity Records) Birth and death registrations (National Records of Scotland)	Medication dispensing (Prescribing Information System) Unscheduled care (Unscheduled Care Datamart)
Stage of development	Permissions obtained, datasets linked, analysis underway		
Lead	Mirjam Allik Denise Brown	Emily Tweed	David Henderson

Work package 3
Objective: to collaboratively develop a set of recommendations for actions that will support greater use of cross-sectoral data as part of a whole-systems approach to public health.

In WP3, we will bring stakeholders back together to identify a consensus set of recommendations for maximising the value of cross-sectoral administrative data in healthy public policy and practice. These will comprise actions at local and national level over the short, medium, and long term, each with nominated lead agencies. The scope will be defined by stakeholder priorities as identified during this and previous workshops but may include discoverability; governance processes; or prioritisation of target datasets based on decision-maker evidence needs. The first third of the workshop will briefly review outputs from previous events before focusing on eliciting potential recommendations, based on suggestions from participants using idea generation techniques such as brainstorming as well as any which have emerged from earlier events. The second third of the workshop will develop these recommendations into common themes, using the systems map as a guide; identify potential overlaps and contradictions; and refine their wording. The final third of the workshop will focus on allocating priority and resource implications to recommendations using techniques such as polls, ranking, and impact/resource matrices. Based on the outputs, the project advisory group will develop a set of recommendations, which will then be circulated to the wider participant group for a final round of feedback before finalisation.

3.2 Setting~~s~~

In this research, we will focus on the roles and relationships of local authority areas, territorial health boards, and devolved institutions in Scotland in undertaking and facilitating administrative data research for public health policy and practice. However, our unifying focus on cross-sectoral data utilisation means that many of the barriers and solutions identified by stakeholders in this study will be relevant elsewhere. We will maximise transferability through explicit attention to lessons from and for data systems in other parts of the UK and beyond: representatives from organisations such as ADR-UK, SAIL Databank, and the International Population Data Linkage Network will be invited to project workshops and offered dedicated time in the programme to reflect on commonalities and differences across systems and contexts.

3.3 Sampling

Our ‘community’ of interest in this project comprises decision-makers working in a variety of public health and social policy roles in local authorities; integrated health and social care partnerships; community planning; NHS public health teams; Scottish Government; and the third sector. We will seek to recruit individuals with broad roles relevant to healthy public policy as well as those with specialist remits relating to each of the case study topic areas. Examples might include directors of education, children’s services, or social work; elected members with roles in health and social care integration and community planning; civil servants in the devolved government; policy leads for third sector organisations; and consultants in public health. Whilst our primary focus is decision-makers in policy and practice roles, we will also invite representatives from across the data use process, including data providers, analysts, and governance leads. We will identify potential workshop participants through stakeholder mapping with the project advisory group and by drawing on existing professional networks accessible to the study team and advisory group, such as SOLACE (Society of Local Authority Chief Executives and Senior Managers), Scottish Health Promotion Managers network, Scottish Directors of Public Health, and Social Work Scotland, as well as topic-specific networks already established via each case study project, such as Homeless Network Scotland, Community Justice Scotland, NSPCC, the Care Inspectorate, and the Scottish Social Services Council. We will also undertake targeted recruitment where necessary, informed by published frameworks for identifying the optimal composition of participatory partnerships (36). Reflecting the importance of public trust and lay perspectives on this topic we will also invite a number of public representatives, comprising people with an interest in secondary data research (recruited via standing panels run by ADR Scotland and University of Glasgow) as well as people with lived experience of the case study issues (recruited via advocacy and support organisations with whom we have established relationships), as detailed in the sections on PPI.

3.4 Study Procedures

Described as above

3.5 Data Analysis

Analysis of workshop transcripts and other outputs such as maps, 'post-it' notes and polls will be undertaken using framework analysis in NVivo.

4. Research Governance and Regulatory Issues

4.1 Ethical issues

Research Ethics Committee: *University of Glasgow College of Social Sciences*

Research Ethics Committee Reference: *awaited*

We will seek approval from the University of Glasgow College of Social Sciences ethics committee. The NHS Health Research Authority decision tool indicates that NHS REC approval is not required (45). All data generated as part of workshops and associated engagement will be stored in accordance with SPHSU data management and archiving policy.

☐ If you **do not think that ethical approval is required for your study,** you should indicate by ticking this box that you have consulted the relevant ethics committees or spoken to unit staff with expertise in ethics, i.e. Marcela Gavigan, Mark McCann or Gillian Fergie to confirm this.

4.2 Data Monitoring/Quality Assurance

A detailed Data Management Plan is currently under development and should be referred to in parallel with this protocol. The Co-PIs are jointly responsible for data monitoring and quality assurance.

4.3 Data Management

A detailed Data Management Plan is currently under development and should be referred to in parallel with this protocol.

4.4 Data Storage and Retention

A detailed Data Management Plan is currently under development and should be referred to in parallel with this protocol.

5 Project Management

5.1 Project Manager

The Co-PIs will act as Project Managers with responsibility for the day to day management of the project.

5.2 Project Management Group

The Project Team consists of the following members:

Name	Division/Organisation
Peter Craig	SPHSU, University of Glasgow
Emily Tweed	SPHSU, University of Glasgow
TBC (research associate)	SPHSU, University of Glasgow

The Project Management Group will meet fortnightly, with monthly meetings of the project advisory group as detailed below.

Minutes of PMG meetings will be taken on the SPHSU template and a Decision Log will be created and maintained by the Project Manager.

5.3 Advisory Group / Steering Committee

The Project Advisory group will meet monthly throughout the project and will consist of the following members:

- Co-PIs, Co-Is, and project RA
- External stakeholders
    - Dermot O'Reilly, ADR-NI
    - Amy Tilbrook, DataLoch, University of Edinburgh
    - Bea von Wissmann, NHS Greater Glasgow and Clyde
    - Mark McAllister, COSLA
    - Iain McKay and Emily Lynch, Improvement Service
    - Others TBC

- Public representatives from ADR-Scotland and MVLS public panels – names TBC

5.4 Project Filing Structure

The electronic project files will be kept on: T:\projects\Unlocking data NIHR S00551

The paper project files will be kept: not applicable

6. Dissemination

6.1 Communication method

We will produce a report and slides describing the co-produced recommendations for optimising use of cross-sectoral secondary data research for public health research and practice, as per the project specification. Other outputs of stand-alone value will include the systems map (WP1) and the three detailed case study reports (WP2).

Outputs will be made available in a range of formats. As well as the final report and slides, we will produce for each RQ a stand-alone briefing accessible for a broad general audience. With the support of the SPHSU communications team, will work with project stakeholders to develop video and infographic summaries which can be shared widely via partner websites, blogs, social media, and meetings. We will draw on our existing partnerships to undertake broader dissemination activities specific to different audiences, such as presentations to Public Health Scotland Senior Leadership Team, the Scottish Directors of Public Health, SOLACE Scotland, the Community Planning Improvement board, Research Data Scotland, and via the Improvement Service's Knowledge Hub for local authorities and community planning. We will also run a webinar on the project and its recommendations, open to all with targeted invitations to key audiences, as well as featuring the project in the SPHSU's popular podcast, '15 minutes on Health Inequalities'.

Via devolved nation representatives on the advisory group, we will ensure that invitations to the dissemination webinar are targeted at key stakeholders from across the UK, and consider organising specific meetings with key stakeholders from outside Scotland such as other Administrative Data Research centres, Public Health England, and the Association of Directors of Public Health. We will also ensure that our blogs, videos, and infographic summaries are shared through UK-wide routes, such as the LSE Impact blog and The Conversation.

As well as the specific WP deliverables, we anticipate that the project will strengthen partnership working between academic, practice, and policy participants and identify opportunities for collaboration on future projects. Throughout, we will maximise engagement and share learning through blogs and short videos by members of the core research team, advisory group, and workshop participants. As illustrated in the flow diagram, these will describe and mark key milestones in the project such as project start-up, workshops, and progress on case study reports.

We will disseminate the work among academic audiences via methodological, topic-specific, and public health conferences (e.g., International Population Data Linkage Network and Society for Social Medicine and Population Health), as well as submission of papers to peer-reviewed journals (as described below in Section 6.2).

The key communications channels are:

- Advisory group and their networks
- NIHR website
- SPHSU website and social media
- Academic conferences and journals

6.2 Publication Policy

We will submit at least two papers to peer-reviewed journals with a focus on secondary data or healthy public policy, describing [1] the case studies and [2] the systems map and consensus recommendations. ICJME authorship criteria will be used. All publications and presentations relating to the project will be authorised by the Project Management Group.

6.3 Public Engagement and Knowledge Exchange

Our PPI activities reflect the focus of this proposal on the use of data in decision-making in local and devolved government, public health agencies, and the third sector. We will therefore ensure our involvement activities reflect both public perspectives, as the original source and owners of the data, and professional perspectives, as those whose work this project aims to support.

The key routes for involvement in this project are:

- Project advisory group - shaping the design, delivery, and dissemination of the study
- Workshop participation - contributing to data collection and development of outputs
- Targeted engagement with preparatory work (e.g. seeking feedback on workshop materials) and dissemination (e.g. invitation to contribute to reflective blogs over course of project)

Each of these will include both public and professional representatives. The advisory group – responsible for oversight and delivery of the project – will include two public representatives and up to six professional collaborators from partner organisations such as the Convention of Scottish Local Authorities, the Improvement Service, NHS Greater Glasgow and Clyde, Public Health Scotland, and the Scottish Safe Haven network. The participatory workshops – together comprising the central data collection activity for the project – will involve around 30 participants, with at least four public representatives alongside professional stakeholders from wider local authority, NHS, devolved government, and third sector roles.

Our public representatives comprise people with complementary perspectives, whether an interest in the use of data (recruited from existing public panels run by the University of Glasgow and ADR-Scotland) or experience of the issues featured in the case studies (recruited from existing lived experience networks supported by partner organisations with who we have established relationships, such as Homeless Network Scotland, Who Cares? Scotland, and Scottish Care). All public representatives will receive support to participate from both their ‘host’ organisation and the study PPI lead, and be fully remunerated for their time according to NIHR guidelines

7. Project Milestones / Timelines

See project Gantt chart in study master file for further details.

8. Project Risk Assessment

The risks relevant to the project are recorded in the risk assessment form and contained in the initial Project Risk/Issue log on: T:\projects\Unlocking data NIHR S00551

The Risk Log will be reviewed and updated at Project Management Group meetings.

9. References

1. Wellcome Trust. Enabling data linkage to maximise the value of public health research data: full report. London: Wellcome Trust, 2015.

2. Academy of Medical Sciences. Improving the health of the public by 2040: optimising the research environment for a fairer, healthier future. 2016.

3. McGrail K, Jones K, Akbari A, al. e. A Position Statement on Population Data Science: The science of data about people International Journal of Population Data Science. 2018;3(1).

4. Craig P, Katikireddi SV, Leyland A, Popham F. Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research. Annual Review of Public Health. 2017;38(1):39-56.

5. Ford E, Boyd A, Bowles JKF, Havard A, Aldridge RW, Curcin V, et al. Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. Learning Health Systems. 2019;3(3):e10191.

6. Connelly R, Playford CJ, Gayle V, Dibben C. The role of administrative data in the big data revolution in social science research. Social Science Research. 2016;59:1-12.

7. Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, et al. Challenges in administrative data linkage for research. Big Data & Society. 2017;4(2):2053951717745678.

8. Lyons RA, Ford DV, Moore L, Rodgers SE. Use of data linkage to measure the population health effect of non-health-care interventions. The Lancet. 2014;383(9927):1517-9.

9. Kim SYH, Flory J, Relton C. Ethics and practice of Trials within Cohorts: An emerging pragmatic trial design. Clinical Trials. 2017;15(1):9-16.

10. Meier PS, Holmes J, Angus C, Ally AK, Meng Y, Brennan A. Estimated Effects of Different Alcohol Taxation and Price Policies on Health Inequalities: A Mathematical Modelling Study. PLOS Medicine. 2016;13(2):e1001963.

11. Haneef R, Delnord M, Vernay M, Bauchet E, Gaidelyte R, Van Oyen H, et al. Innovative use of data sources: a cross-sectional study of data linkage and artificial intelligence practices across European countries. Archives of Public Health. 2020;78(1).

12. Sexton A, Shepherd E, Duke-Williams O, Eveleigh A. A balance of trust in the use of government administrative data. Archival Science. 2017;17(4):305-30.

13. Lugg-Widger F, Angel L, Cannings-John R, Hood K, Hughes K, Moody G, et al. Challenges in accessing routinely collected data from multiple providers in the UK for primary studies: Managing the morass. International Journal of Population Data Science. 2018;3.

14. Deeny SR, Steventon A. Making sense of the shadows: priorities for creating a learning healthcare system based on routinely collected data. BMJ Quality & Safety. 2015;24(8):505-15.

15. Jones K, Heys S, Tingay K, Jackson P, Dibben C. The Good, the Bad, the Clunky. International Journal of Population Data Science. 2019;4(1).

16. Mourby MJ, Doidge J, Jones KH, Aidinlis S, Smith H, Bell J, et al. Health Data Linkage for UK Public Interest Research: Key Obstacles and Solutions. International journal of population data science. 2019;4(1):1093-.

17. Ford E, Boyd A, Bowles JKF, Havard A, Aldridge RW, Curcin V, et al. Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. Learning Health Systems. 2019;3(3).

18. Jones K, Heys S, Daniels H, Ford D. Exploring barriers and solutions in advancing cross-centre population data science. International Journal of Population Data Science. 2019;4:22.

19. Aitken M, Jorre JdS, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Medical Ethics. 2016;17.

20. Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers. BMC Health Services Research. 2014;14(1):2.

21. Petticrew M, Whitehead M, Macintyre SJ, Graham H, Egan M. Evidence for public health policy on inequalities: 1: the reality according to policymakers. J Epidemiol Community Health. 2004;58(10):811-6.

22. Oliver KA, de Vocht F. Defining 'evidence' in public health: a survey of policymakers' uses and preferences. Eur J Public Health. 2017;27(suppl_2):112-7.

23. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. The Lancet. 2009;374(9683):86-9.

24. Evans BA, Snooks H, Howson H, Davies M. How hard can it be to include research evidence and evaluation in local health policy implementation? Results from a mixed methods study. Implementation science : IS. 2013;8:17.

25. McGill E, Egan M, Petticrew M, Mountford L, Milton S, Whitehead M, et al. Trading quality for relevance: non-health decision-makers’ use of evidence on the social determinants of health. BMJ Open. 2015;5(4):e007053.

26. Cheetham M, Redgate S, van der Graaf P, Hunter R, Ritson L. Local Authority Champions of Research Project: a report for the Health Foundation. 2019.

27. Rutter H, Savona N, Glonti K, Bibby J, Cummins S, Finegood DT, et al. The need for a complex systems model of evidence for public health. The Lancet. 2017;390(10112):2602-4.

28. Carey G, Malbon E, Carey N, Joyce A, Crammond B, Carey A. Systems science and systems thinking for public health: a systematic review of the field. BMJ Open. 2015;5(12):e009002.

29. Egan M., McGill E., Penney T., Anderson de Cuevas R., Er V., Orton L., et al. NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 1: Introducing systems thinking. London: National Institute for Health Research School for Public Health Research, 2019.

30. Scottish Government and COSLA. Public Health Priorities for Scotland. Edinburgh: Scottish Government, 2018.

31. Public Health Scotland. Supporting whole system approaches. URL: https://www.publichealthscotland.scot/our-organisation/about-public-health-scotland/supporting-whole-system-approaches/ [accessed 05/01/2021]

32. Meier P, Purshouse R, Bain M, Bambra C, Bentall R, Birkin M, et al. The SIPHER Consortium: Introducing the new UK hub for systems science in public health and health economic research [version 1; peer review: 2 approved]. Wellcome Open Research. 2019;4(174).

33. Oliver K, Lorenc T, Innvær S. New directions in evidence-based policy research: a critical analysis of the literature. Health Research Policy and Systems. 2014;12(1):34.

34. Haynes A, Rychetnik L, Finegood D, Irving M, Freebairn L, Hawe P. Applying systems thinking to knowledge mobilisation in public health. Health Res Policy Syst. 2020;18(1):134.

35. Cairney P. The Politics of Evidence-Based Policy-Making: Palgrave; 2016. 1-137 p.

36. Cargo M, Mercer SL. The Value and Challenges of Participatory Research: Strengthening Its Practice. Annual Review of Public Health. 2008;29(1):325-50.

37. Boaz A, Hanney S, Borst R, O'Shea A, Kok M. How to engage stakeholders in research: design principles to support improvement. Health Research Policy and Systems. 2018;16(1):60.

38. Laird Y, Manner J, Baldwin L, Hunter R, McAteer J, Rodgers S, et al. Stakeholders' experiences of the public health research process: time to change the system? Health Res Policy Syst. 2020;18(1):83.

39. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? Health Research Policy and Systems. 2019;17(1):33.

40. TRIUMPH: Transdisciplinary Research for the Improvement of Youth Mental Public Health. URL: http://triumph.sphsu.gla.ac.uk [accessed 05/01/2021]

41. SHINE: Scottish Schools Health and Wellbeing Improvement Research Network. 2020. URL: https://shine.sphsu.gla.ac.uk/ [accessed 05/01/2021]

42. Hassannezhad M, Gavens L. Participatory systems mapping in a socially distanced world. SIPHER blog: Systems science In Public Health and Health Economics Research. 2020. URL: https://sipher.ac.uk/participatory-systems-mapping-in-a-socially-distanced-world/ [accessed 05/01/2021]

43. Boaz A, Oliver K. Building new bridges between research and policy during a national lockdown. LSE Impact Blog. 2020. URL: https://blogs.lse.ac.uk/impactofsocialsciences/2020/11/27/building-new-bridges-between-research-and-policy-during-a-national-lockdown/ [accessed 05/01/2021]

44. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368:l6890.

45. UKRI Medical Research Council. Do I need NHS REC review? 2020. Available from: http://www.hra-decisiontools.org.uk.