

# Linking police data with health and education for public benefit

<b>Submission date</b> 27/04/2021	<b>Recruitment status</b> No longer recruiting	<input type="checkbox"/> Prospectively registered <input checked="" type="checkbox"/> Protocol
<b>Registration date</b> 05/07/2021	<b>Overall study status</b> Completed	<input type="checkbox"/> Statistical analysis plan <input type="checkbox"/> Results
<b>Last Edited</b> 08/07/2021	<b>Condition category</b> Other	<input type="checkbox"/> Individual participant data <input type="checkbox"/> Record updated in last year

## Plain English summary of protocol

### Background and study aims

Public health planning, crime prevention and health care are all greatly helped by data sharing between different organisations. The 1998 Crime and Disorder Act requires police, local government and NHS to collaborate on joint strategies. There are situations where data is shared on individuals, for example in high-risk domestic abuse cases, in the Multi-Agency Risk Assessment Conference (MARAC) which brings multi-agencies together at a local level. However, combining entire datasets of all data held by the police with health, education, social services data could greatly help give the whole picture and understanding of how early action could prevent serious illness/ serious crime. This study focuses on bringing together police data with NHS data to improve the ability to tackle domestic abuse in pregnancy.

### Who can participate?

Member of the police service living in Wales and dealing with public protection

### What does the study involve?

The researchers will examine in work package 1 if data from reports can be extracted by using automated text extraction methods. In work package 2 they will catalogue all the software systems used and how they might be combined together. In work package 3 they will undertake an example case study bringing together Public Protection Notification data (when a pregnant woman is present), with A&E, GP, hospital, maternal indicators dataset, national community child health data, substance abuse dataset. Finally, in work package 4, the researchers will examine how best to identify the same person in different datasets (e.g. in GP data and in police data). They will also hold two half-day multiagency workshops to discuss findings and how data sharing on a national level could be taken forward. They will work with members of the public in patient and public involvement teams. Finally, this work will complement work in England which brings together education and Ministry of Justice data and enable cross country comparisons. This work is the next step on the road to creating national multi-agency partnerships to improve the life chances for children growing up in disadvantage (such as situations of poor mental health /substance abuse/domestic violence/areas of high violence & crime).

### Where is the study is run from?

Swansea University (UK)

When is the study starting and how long is it expected to run?  
April 2021 to December 2021

Who is funding the study?  
National Institute for Health Research (UK)

Who is the main contact?  
Prof. Sinead Brophy  
s.brophy@swansea.ac.uk

## Contact information

**Type(s)**  
Scientific

**Contact name**  
Prof Sinead Brophy

**ORCID ID**  
<https://orcid.org/0000-0001-7417-2858>

**Contact details**  
Data Science Building  
Swansea University  
Swansea  
United Kingdom  
SA2 8PP  
+44 (0)7864836943  
s.brophy@swansea.ac.uk

## Additional identifiers

**Clinical Trials Information System (CTIS)**  
Nil known

**ClinicalTrials.gov (NCT)**  
Nil known

**Protocol serial number**  
Nil known

## Study information

**Scientific Title**  
Unlocking data to inform public health policy and practice: a scoping pilot study

**Study objectives**  
The problems of bringing together entire datasets at a national level (as opposed to sharing reports on the individual at a local level) include; the type of data held by the police is often

descriptive text and this is not easy to share or analyse, there are multiple different software systems used by different forces which are not compatible and so difficult to combine and share, there are few examples where entire databases have been shared and so it is difficult to demonstrate the advantages, the methods of identifying people are not the same in different organisations (e.g. NHS number is not recorded in the police dataset, pupil ID number is not recorded in police data). This project aims to overcome these barriers one by one and help unlock data to inform public health policy and practice.

### **Ethics approval required**

Old ethics approval format

### **Ethics approval(s)**

Approved 07/05/2021, Swansea University Medical School Research Ethics Committee (Swansea University Medical School, Institute of Life Science 2, Swansea University, Singleton Park, Swansea, SA2 8PP; +44(0)1792 60 2697; sumsresc@swansea.ac.uk), ref: 2021-0042

### **Study design**

Qualitative interviews, pilot study of data linkage and text mining

### **Primary study design**

Observational

### **Study type(s)**

Other

### **Health condition(s) or problem(s) studied**

Domestic violence/public protection

### **Interventions**

1. Qualitative interviews - interviewing members of the police service regarding their IT systems
2. Pilot study of data linkage - pilot study to link Public Protection Notification data with health (GP, A&E and hospital) data
3. Text mining - examining if relevant data can be extracted from text fields to be coded for information sharing and linkage

The researchers will examine in work package 1 if data from reports can be extracted by using automated text extraction methods. In work package 2 they will catalogue all the software systems used and how they might be combined together. In work package 3 they will undertake an example case study bringing together Public Protection Notification data (when a pregnant woman is present), with A&E, GP, hospital, maternal indicators dataset, national community child health data, substance abuse dataset. Finally, in work package 4, the researchers will examine how best to identify the same person in different datasets (e.g. in GP data and in police data). They will also hold two half-day multiagency workshops to discuss findings and how data sharing on a national level could be taken forward. They will work with members of the public in patient and public involvement teams. Finally, this work will complement work in England which brings together education and Ministry of Justice data and enable cross country comparisons. This work is the next step on the road to creating national multi-agency partnerships to improve the life chances for children growing up in disadvantage (such as situations of poor mental health /substance abuse/domestic violence/areas of high violence & crime).

### **Intervention Type**

Other

**Primary outcome(s)**

Evaluation of the added benefit of inclusion of Police data when profiling vulnerable families and evaluation of linkage quality when linking records from different organisations measured using qualitative interviews with the police service at 2 to 4 months

**Key secondary outcome(s)**

There are no secondary outcome measures

**Completion date**

31/12/2021

## **Eligibility**

**Key inclusion criteria**

Member of the police service living in Wales and dealing with public protection

**Participant type(s)**

Other

**Healthy volunteers allowed**

No

**Age group**

Adult

**Sex**

All

**Key exclusion criteria**

Not working in the Police service

**Date of first enrolment**

01/06/2021

**Date of final enrolment**

30/11/2021

## **Locations**

**Countries of recruitment**

United Kingdom

Wales

**Study participating centre**

**Swansea University**  
Data Science Building  
Swansea  
United Kingdom  
SA2 8PP

## Sponsor information

### Organisation

National Institute for Health Research

### ROR

<https://ror.org/0187kwz08>

## Funder(s)

### Funder type

Government

### Funder Name

NIHR Evaluation Trials and Research Coordinating Centre

## Results and Publications

### Individual participant data (IPD) sharing plan

The datasets generated during and/or analysed during the current study are not expected to be made available. The participant level data is interviews with the police force about data they collect and this is sensitive data and due to small numbers and the likely local context may enable individuals to be identified and so is not suitable for sharing in its raw form.

### IPD sharing plan summary

Not expected to be made available

### Study outputs

Output type	Details	Date created	Date added	Peer reviewed?	Patient-facing?
<a href="#">Participant information sheet</a>	Participant information sheet	11/11/2025	11/11/2025	No	Yes
<a href="#">Protocol file</a>	version v1.0	01/04/2021	08/07/2021	No	No