

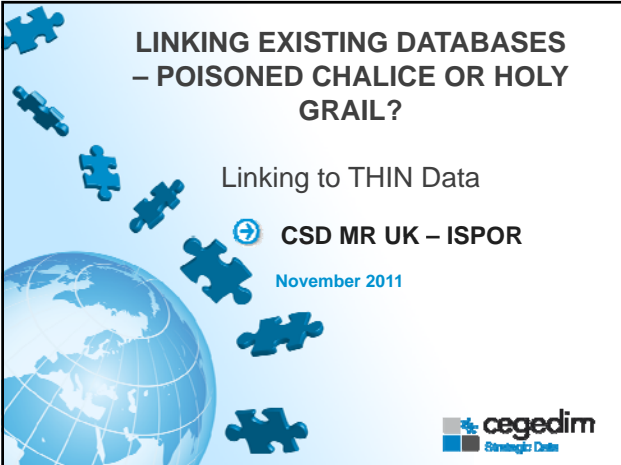


LINKING EXISTING DATABASES – POISONED CHALICE OR HOLY GRAIL?

Linking to THIN Data


 **CSD MR UK – ISPOR**

November 2011

Disclosure


I am paid as an employee of Cegedim, and act as a Board Director on THIN



This document should not be distributed without Cegedim authorization - Copyright 2011

Agenda

- What is THIN?
- What can we link?
- How do we do it?
- What are the issues?



This document should not be distributed without Cegedim authorization - Copyright 2011

Cegedim group



FOUNDED
1969

WORKFORCE
8,400

OPERATIONS
80 countries



This document should not be distributed without Cegedim authorization - Copyright 2011

THIN Data	
% of UK population	6%
Total Patients	10 million
Active patients	3.6 Million
Average length of follow up	>8 years (range 1-25 years)
Person Years of Computerisation	>67m

Patient Data Collected

- Demographics**
 - Year of birth, gender, registration dates
- Medical History**
 - Events, diagnosis, symptoms, risk factors, morbidities, referrals...
- Prescription**
 - Start dates, indications, therapeutic class, molecule, brand (dosage, form, duration)
- Clinical Data**
 - Height, weight, BP, laboratory results, immunizations, life habits...

*Pseudonymised
No name, address, postcode,
DOB, NHS/CHI number*

cegedim
Strategic Data

This document should not be distributed without Cegedim authorization - Copyright 2011 | 6

Agenda

- What is THIN?
- What can we link?
- How do we do it?
- What are the issues?

cegedim
Strategic Data

This document should not be distributed without Cegedim authorization - Copyright 2011 | 7

What can we link?

- **Data not already in THIN with a variable in common either within the THIN data or within the practice**
 - Based on patient location ie data related to postcode
 - New data ie additional data/sample collected from Practices or Patients
 - Existing data from other data sources relating to individual patients

cegedim
Strategic Data

This document should not be distributed without Cegedim authorization - Copyright 2011 | 8

Agenda

- What is THIN?
- What can we link?
- How do we do it?
- What are the issues?

Aim - To ensure that only the minimum personal data required to satisfy a purpose is collected and used and unnecessary information is stripped out.

This meets fundamental Data Protection Act 1998 principles as well as Caldicott principles and the Human Rights Act 1998.



How?

- **Depends on the data source**
 - Within the doctor's system/practice
 - Direct communication between GP and data custodian
 - Using Enhanced Trusted Third Party

How?

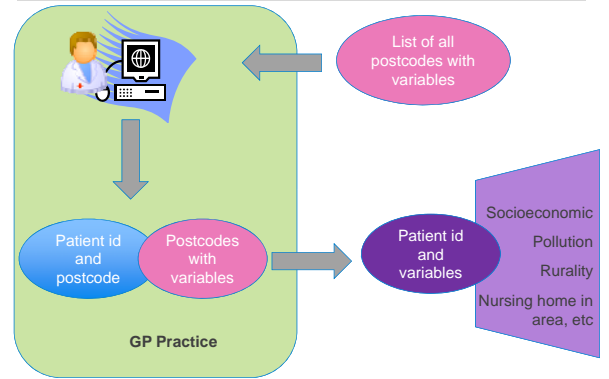
- **Depends on the data source**
 - Within the doctor's system/practice
 - Direct communication between GP and data custodian
 - Using Enhanced Trusted Third Party

How?

→ Within the doctor's system/practice

- Manual questionnaires to doctor/patient/carer linked by patient identifier (Additional Information Services – AIS)
- 'Pop-up' additional screens, triggered by criteria eg all patients consulting for depression
- Using handheld device compatible with Vision and synched to medical record
- Postcode data

Postcode Variable Indicator Data



How?

→ Depends on the data source

- Within the doctor's system/practice
- Direct communication between GP and data custodian
- Using Enhanced Trusted Third Party

Data Linked by Data Custodians

- GP runs utility to export linking variables to data custodian
- Linking carried out at data custodian site using linking variables available in both datasets – hard identifiers
- Linked data returned to THIN or directly to researcher without identifiers

How?

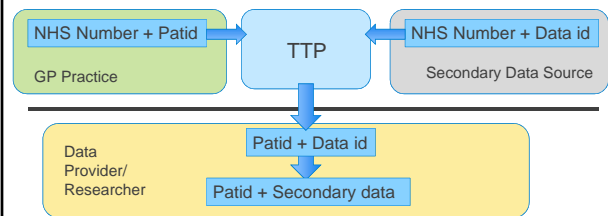
→ Depends on the data source

- Within the doctor's system/practice
- Direct communication between GP and data custodian
- Using Enhanced Trusted Third Party

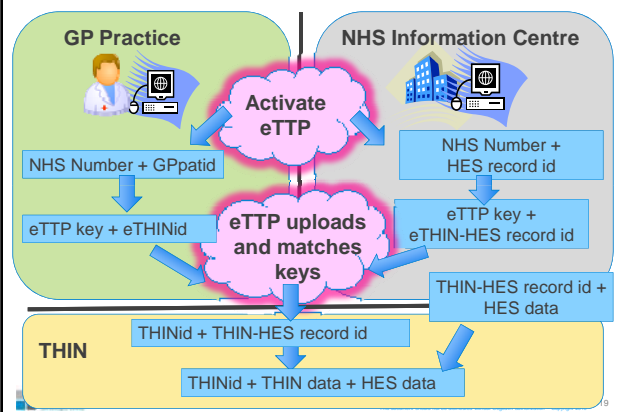
How?

Difference between Trusted Third Party model (TTP) and Enhanced Trusted Third Party (eTTP)

TTP



Link using Enhanced Trusted Third Party (eTTP) - Sapior



Agenda

- What is THIN?
- What can we link?
- How do we do it?
- What are the issues?

Progress?

→ Postcode Variable

- Several indicators already available ie socioeconomic, pollution, nursing homes etc
- Study specific indicators also collected
- **Issues – Small cost / Timing updates**

→ Additional Information Services (AIS) data

- Data are already available
- Need to code input files
- Need to write programs to locate and link
- **Issues – Resource**

When?

→ Pop-Up questionnaires for completion by GP during patients consultation

- System available now
- Need a pilot study so that formal tests with Vision can be completed
- **Timing and programming investment**

→ Questionnaires on mobile device for completion by Patients

- Available now
- One study on QoL completed
- **Investment and logistics**

Progress?

→ Direct communication between GP and Data Custodian

- THIN linked with Cancer registry data
- Pilot study completed
- Plans for full study
- **Inconsistencies between data sources**

→ eTTP linked

- National Information Governance Board for Health and Social Care (NIGB) review - began January 2011.....

New Ideas – eg Baby Duster



Progress?

→ eTTP linked

- NIGB review - began January.....
..... completed August 2011 – do not need NIGB approval as not identifiable data
- Loaded HES to eTTP
- Invite doctors
- Activate and collect linked data
- Samples to researchers in next 2 to 4 months
- **Quality of HES data**

Issues

- Priority
- Technical
 - Cost
 - Timing
- Governance
 - Uncertain authorities
 - Precedent

Linking existing databases – poisoned chalice or holy grail?



Like rare French wine – often expensive and hard to get but resulting data is rich and potent, often complex

Thank you

alison.bourke@cegedim.com