Establishing a routinely updated maternal-child linkage system in THIN database.

Linda Fiaschi
School of Community Health Sciences
Division of Epidemiology and Public Health
University of Nottingham, UK
Outline

- Motivations
- Objectives
- Data
- Method
- Results
- Conclusions
Motivations

- UK: 20,000 children born with malformation each year (3%)
- Leading cause of stillbirth and infant mortality (UK, USA)
- 1,500 children die before 1 month

- Are drugs responsible for congenital malformations?
- Diabetes, Epilepsy, Hypertensive disorders, Depression?
- Limitations on currently available data and studies.
Objectives (1)

Evidence for drugs impact in pregnancy using anonymised UK primary care data (THIN)

Validation of THIN:

- Recover pregnancies timing and outcomes (live births, stillbirths, terminations etc.) from THIN

- Validate these findings with national statistics
Objectives (2)

Create a subset of THIN with Pregnancy History

- Live events matched
- Link with all other tables and info (Therapy, townsend.. )

Pregnancy dataset for any type of study
Data
UK GP Data Strengths

- The General Practitioner (GP) represents mostly the first access point for health service requirement.

- Pregnant women in the UK are assigned with a qualified health professional who supports, cares and advises them during pregnancy.

- UK GP clinics records are internationally recognised as comprehensive and competitive data for epidemiological studies.
Data Limitations

1. **THIN data are anonymised**: mother–child link is lost

2. **Repetition and Historical Records**
The Health Improvement Network dataset: THIN

Demographic

Medical

Additional Health data

Therapy

Social and Environmental Factors
THIN0907 Numbers

- Updated every 4 months
- Primary care data from 446 UK GP practices
- Version 0907: 9,190,757 patients (5% population)

- ~400 mil. medical records: 99.7% acceptable
- ~600 mil. AHD records :99.7% acceptable
- ~ 650 mil. therapy records : 98.6% acceptable
### Patient file: patflags and registration status

<table>
<thead>
<tr>
<th>Patid</th>
<th>Patflag</th>
<th>YOB</th>
<th>Famnum</th>
<th>Gender</th>
<th>Regdate</th>
<th>Reg status</th>
<th>Xferdate</th>
<th>Regrea</th>
<th>Deathdate</th>
<th>Deathinfo</th>
<th>Accept</th>
<th>Institute</th>
<th>Marital</th>
<th>Dispensing</th>
<th>Prsc</th>
<th>exempt</th>
<th>System date</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001</td>
<td>A</td>
<td>19300000</td>
<td>009292</td>
<td>1</td>
<td>19990721</td>
<td>02</td>
<td>00000000</td>
<td>00</td>
<td>00000000</td>
<td>3</td>
<td>N</td>
<td>06</td>
<td>00</td>
<td>00</td>
<td>20100112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A</td>
<td>19220000</td>
<td>003618</td>
<td>2</td>
<td>19840430</td>
<td>99</td>
<td>20011101</td>
<td>01</td>
<td>20011013</td>
<td>0</td>
<td>N</td>
<td>02</td>
<td>03</td>
<td>20011113</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000a</td>
<td>S</td>
<td>19820000</td>
<td>003797</td>
<td>1</td>
<td>19790627</td>
<td>05</td>
<td>19860314</td>
<td>13</td>
<td>00000000</td>
<td>0</td>
<td>N</td>
<td>06</td>
<td>00</td>
<td>19990310</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000b</td>
<td>R</td>
<td>19470000</td>
<td>007658</td>
<td>2</td>
<td>19910611</td>
<td>05</td>
<td>19910611</td>
<td>03</td>
<td>00000000</td>
<td>0</td>
<td>N</td>
<td>06</td>
<td>00</td>
<td>19990310</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000d</td>
<td>D</td>
<td>19850000</td>
<td>004290</td>
<td>2</td>
<td>20060406</td>
<td>04</td>
<td>00000000</td>
<td>00</td>
<td>00000000</td>
<td>0</td>
<td>N</td>
<td>06</td>
<td>00</td>
<td>20060406</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000c</td>
<td>A</td>
<td>19780000</td>
<td>010360</td>
<td>2</td>
<td>19970210</td>
<td>05</td>
<td>19970911</td>
<td>13</td>
<td>00000000</td>
<td>0</td>
<td>Y</td>
<td>06</td>
<td>00</td>
<td>20050331</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000e</td>
<td>C</td>
<td>19220000</td>
<td>003618</td>
<td>1</td>
<td>19840430</td>
<td>99</td>
<td>20021106</td>
<td>01</td>
<td>00000000</td>
<td>0</td>
<td>N</td>
<td>02</td>
<td>04</td>
<td>20021119</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Acceptable**: Indicates the patient is still active in the system.
- **Permanent**: Marked when the patient status is permanent.
- **Died**: Marked when the patient has died.
- **Transferred out**: Marked when the patient has been transferred out of the system.
- **Death date**: Indicated by the date when the patient died.
### Medical records: diabetes entry

<table>
<thead>
<tr>
<th>patid</th>
<th>eventdate</th>
<th>enddate</th>
<th>datatype</th>
<th>medcode</th>
<th>medflag</th>
<th>staffid</th>
<th>source</th>
<th>episode</th>
<th>nhspec</th>
<th>locate</th>
<th>textid</th>
<th>category</th>
<th>priority</th>
<th>medinfo</th>
<th>inprac</th>
<th>private</th>
<th>medid</th>
<th>consult</th>
<th>sysdate</th>
<th>modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>04ne20050821</td>
<td>000000000</td>
<td>19</td>
<td>9N79.00</td>
<td>R</td>
<td>001J</td>
<td>0</td>
<td>0</td>
<td>000</td>
<td>A</td>
<td>0000001</td>
<td>5</td>
<td>Y</td>
<td>N</td>
<td>01uH</td>
<td>5Ana</td>
<td>20050821</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04ne20050821</td>
<td>000000000</td>
<td>02</td>
<td>8CAL.00</td>
<td>R</td>
<td>001J</td>
<td>0</td>
<td>4</td>
<td>000</td>
<td>A</td>
<td>2YNd</td>
<td>4</td>
<td>3</td>
<td>Y</td>
<td>N</td>
<td>2YNd</td>
<td>5Ana</td>
<td>20050821</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04ne20050821</td>
<td>000000000</td>
<td>12</td>
<td>C10F.00</td>
<td>R</td>
<td>001J</td>
<td>0</td>
<td>0</td>
<td>000</td>
<td>A</td>
<td>0000001</td>
<td>3</td>
<td>Y</td>
<td>N</td>
<td>007n</td>
<td>5Ana</td>
<td>20050902</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04ne20060621</td>
<td>000000000</td>
<td>44</td>
<td>1A59.00</td>
<td>S</td>
<td>R</td>
<td>0</td>
<td>ARB</td>
<td>I</td>
<td>09GR</td>
<td>4</td>
<td>0</td>
<td>Y</td>
<td>N</td>
<td>09GR</td>
<td>3wve</td>
<td>20060621</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04ne20070926</td>
<td>000000000</td>
<td>01</td>
<td>N224500</td>
<td>R</td>
<td>001I</td>
<td>6</td>
<td>4</td>
<td>SUR</td>
<td>I</td>
<td>0000007</td>
<td>3</td>
<td>3</td>
<td>Y</td>
<td>N</td>
<td>1jFq</td>
<td>3cEl</td>
<td>20070926</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04ne20091210</td>
<td>000000000</td>
<td>01</td>
<td>7H37000</td>
<td>R</td>
<td>000G</td>
<td>D</td>
<td>4</td>
<td>000</td>
<td>L</td>
<td>0000006</td>
<td>4</td>
<td>3</td>
<td>Y</td>
<td>N</td>
<td>2wwv</td>
<td>5LLL</td>
<td>20091210</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04ne20091130</td>
<td>000000000</td>
<td>01</td>
<td>9OX..11</td>
<td>R</td>
<td>000@</td>
<td>0</td>
<td>4</td>
<td>000</td>
<td>A</td>
<td>2W7q</td>
<td>6</td>
<td>4</td>
<td>Y</td>
<td>N</td>
<td>2W7q</td>
<td>5KJw</td>
<td>20091130</td>
<td>N</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Diabetic register**

**Type 2 Diabetes mellitus**

**Diagnosis**
## Maternity outcome

### Acceptable record

<table>
<thead>
<tr>
<th>PatID</th>
<th>EventDate</th>
<th>Ahdcode</th>
<th>Ahdflag</th>
<th>Data1</th>
<th>Data2</th>
<th>Data3</th>
<th>Data4</th>
<th>Data5</th>
<th>Data6</th>
<th>Medcode</th>
<th>Source</th>
<th>Nhsspec</th>
<th>Locate</th>
<th>Staffid</th>
<th>Textid</th>
<th>Category</th>
<th>Ahdinfo</th>
<th>Inprac</th>
<th>Private</th>
<th>AhdId</th>
<th>Consultid</th>
<th>Sysdate</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>06ll</td>
<td>20070917</td>
<td>1003040000</td>
<td>Y</td>
<td>Y</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>137P.00</td>
<td>0 0 A</td>
<td>000E</td>
<td>0000001</td>
<td>Y N</td>
<td>Y 02jr</td>
<td>4trs</td>
<td>20070917</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07lp</td>
<td>20080115</td>
<td>1066000000</td>
<td>Y</td>
<td>20100115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8B31400</td>
<td>0 0 I</td>
<td>001D</td>
<td>0000001</td>
<td>6</td>
<td>03I?</td>
<td>4x4A</td>
<td>20090115</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07lp</td>
<td>20080917</td>
<td>1090000000</td>
<td>Y</td>
<td>C10.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1252.00</td>
<td>0 0 I</td>
<td>001D</td>
<td>0005332</td>
<td>3</td>
<td>Y N</td>
<td>02X8</td>
<td>40NN</td>
<td>20080917</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07lp</td>
<td>20080925</td>
<td>1003040000</td>
<td>Y</td>
<td>D</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>137S.00</td>
<td>0 0 I</td>
<td>001D</td>
<td>0000001</td>
<td>2</td>
<td>03I?</td>
<td>4x4A</td>
<td>20080925</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01kq</td>
<td>20080915</td>
<td>1048000000</td>
<td>Y</td>
<td>LMP001</td>
<td>20090630</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1513.00</td>
<td>0 0 A</td>
<td>000x</td>
<td>0000001</td>
<td>3</td>
<td>Y N</td>
<td>00Go</td>
<td>4wwr</td>
<td>20080915</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01kq</td>
<td>20090716</td>
<td>1015000000</td>
<td>Y</td>
<td>20090716</td>
<td>LIV001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63...00</td>
<td>0 0 M</td>
<td>000n</td>
<td>0000001</td>
<td>6</td>
<td>Y N</td>
<td>008n</td>
<td>4rse</td>
<td>20090724</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Maternity outcome**: Live birth (outcome) and Secretary.
- **Acceptable record**: Discharge Date (outcome).
- **(Discharge Date)**: Letter from outpatients.
Method
Data Extraction

THIN
9,190,757

526,348
Potential Mothers:
• 15-45 Age
• Info on pregnancy or delivery

594,945
Children:
Under 20 registered within 1 year from birth
**1st THIN Limitation**

**THIN data are anonymised**: mother–child link is lost:

- Living in the same household
- Child birth date close to delivery date taken from pregnancy/delivery mother records
- Applied appropriate inter-pregnancy time (Matches minimal distance is 32 weeks)
Matches Steps

1. Children registered within 3 months
2. Children registered between 3 and 12 months
3. Children registered after 12 months

- Delivery match
- Pregnancy match
Match Example

Birth and THIN Practice Registration

Same Household

Caesarean Section

Birth

THIN Practice Registration

Exit THIN Practice

Mother Record

Child Record
2nd THIN Limitation

Repetition and History Records

Selecting only the actual events
Record Cleaning: Repetition

Live Delivery

Emergency Caesarean Section

2 days

3 days

Twins

Initial THIN Record

Final Mother Record

Time Line
Record Cleaning: History (1)

Spontaneous Abortion

Initial THIN Record

Live Birth Code

2 months

Final Mother Record

Time Line
Results
Results 1990 - 2009

Pregnancies = 590,532
women = 388,260

25%  75%

Non live outcomes 148,613

Miscarriages (69,364) 12%
Terminations (77,789) 28/1000 women
Stillbirths (1,460) 3.5/1000 births

13%¹

LIVE outcomes 441,919

Matched 405,102
Multiple (7,005) 12/1000 Preg.
Infant Deaths (1,300) ~3/1000 Live births

12%
16²
5.4³
92%
8%

ONS Abortion Statistics 1990-2008 England and Wales - aged 15-45
ONS Birth Statistics 1998-2008 England and Wales – all ages
ONS Infant and Perinatal Mortality 1998-2008 England & Wales

¹ Tommy’s Charity for Pregnancy Research Funding - within 15-45 age
² ONS Abortion Statistics 1990-2008 England and Wales - aged 15-45
³ ONS Birth Statistics 1998-2008 England and Wales – all ages
⁴ ONS Infant and Perinatal Mortality 1998-2008 England & Wales
Fertility Rate 1999 - 2009

Live Births per 1000 person-year

Fertility Rate

FR from THIN 0907
FR from ONS 2008

Maternal Age Groups

<20 20-24.9 25-29.9 30-34.9 35-39.9 >=40 tot

18 54 71.9 76 89.1 84 89 45 49.1 10 10.3 54 58.7

FR from THIN 0907
FR from ONS 2008
Delivery Type

Percentage of Live Deliveries

Assisted: 6.6 (THIN 1990-2009), 10.7 (HES 1990-2009)
Breech Del.: 0.5 (THIN 1990-2009), 0.1 (HES 1990-2009)
Pregnancy Length – Live Outcomes

Live Births by Gestation at birth

1 Only pregnancy with information about the length
23.4% of children have the info about the exact weight
0.1% children have info on High or Low weight

**Birth Weight**

<table>
<thead>
<tr>
<th>Weight in Kg</th>
<th>THIN 2006-2009</th>
<th>HES 2006-2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-1.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5-1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-2.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5-2.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-3.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5-3.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-4.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.5-4.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;=5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Percentage of Children**
Conclusions

- Majority of children are matched
- Identification of women’s pregnancies
- Identification of end of pregnancy events
- Nationally-representative DB routinely updated

- THIN is valuable resource for epidemiological analyses of pregnancy, delivery and childhood conditions
- Further studies can consider Townsend, smoking and other relevant variables.
Thank you!

Project Funded by: wellcome trust

THIN dataset by: CSD EPIC

Research Institution: The University of Nottingham