Uses of Patient Data

Paul Taylor
UCL
Farr Institute: Health Informatics Research Centres

- **Scotland**
  - Dundee, Glasgow, Edinburgh, St Andrews, Aberdeen, Strathclyde, MRC HGU, NHS NSS

- **HeRC N8**
  - Manchester, York, Lancaster, Liverpool, Sheffield, AHSNs

- **CIPHER**
  - Swansea, Bristol, Cardiff, Exeter, Leicester, Sussex, NWIS, Public Health Wales

- **London**
  - UCL, LSHTM, Queen Mary, Public Health England

Map Source: www.m62.net
Our Vision

“To harness health data for patient and public benefit by setting the international standard in trustworthy reuse of electronic patient records and related linkable data for large-scale research.”
Astonishing amounts of overlapping activity ...
Prime Minister joins Sir Ka-shing Li for launch of £90m initiative in big data and drug discovery at Oxford University

- 'Big data' to revolutionise healthcare
- Radical boost for drug discovery
- PM: 'What’s happening today really matters ... I think it’s going to lead to huge breakthroughs'

The Prime Minister, David Cameron, was at Oxford University today to mark the launch of a centre that will put the UK at the forefront of revolutionising healthcare through 'big data' and better drug discovery.
Overview

• Current controversy around the sharing of GP data for research purposes
  – care.data

• Long-running controversy around the use of hospital mortality data as a measure of quality
  – Standardised Hospital Mortality Ratios
Coded data
In last year’s HES summary statistics 785,263 in-patient episodes were coded under Obstetrics, – of which 16,992 were recorded for male patients

• Y34 Unspecified event, undetermined intent
Overview

• Current controversy around the sharing of GP data for research purposes
  – care.data

• Long-running controversy around the use of hospital mortality data as a measure of quality
  – Standardised Hospital Mortality Ratios
We used to think of patient data as either identifiable or not.

Patient identifiable data should be disclosed only with consent or under statute and accompanied by a contractual agreement.

De-identified patient data data from which patients could be identified only with unreasonable effort and which can be disclosed.
Now have a third category of data …

- Identifiable data should be disclosed only with consent or under statute and accompanied by a contractual agreement.
- De-identified patient data from which patients could be identified only with unreasonable effort and which can be publicly disclosed.
- Potentially re-identifiable data for limited disclosure.
Potentially re-identifiable data

- In 1997, a US agency sold data it believed to be anonymised describing the medical histories of 135,000 employees of the state of Massachusetts – data included zip codes, sex, and date of birth.

- Voter registration list – also available for purchase – contained those items associated with names.

- By linking the two datasets the medical histories of named individuals could be ascertained.
‘accredited safe havens’

• ‘unfinished business’ around the information governance of key elements of the new NHS

• Commissioning Groups are not providing care, and not allowed to access patient data

• Plan is to release pseudonymised data to ‘accredited save havens’ who will work with commissioners and researchers
care.data

- Detailed data on hospital episodes are returned to the DH and used in planning, research, audit

- NHS reforms obliges GPs to return ‘coded’ data to a single central database for use in commissioning but also in research including commercial research
  - linking GP records and HES data hugely exciting for researchers

- ‘GP Extraction Service Independent Advisory Group’ raised questions about this
Public Trust

• Patients can object to confidential data being submitted to the Health and Social Care Information Centre or released by it. Law requires any reasonable objection to be respected

• “Section 251” can be invoked but there must still be due process
  – patients’ concerns must be considered,
  – reasons for over-ruling objections made public
• Bungled attempt to establish consent is worse than nothing
  – Leaflets were misleading and poorly distributed and therefore no basis for consent

  – Fact of leafleting reinforces argument that consent required
Where are we now?

- Confusion in the public mind about what data is being shared and about how it is to be used
  - Research suggests that people generally support the sharing of data for research purposes, or to assist the NHS in commissioning.
  - People less supportive of data being used for commercial purposes
Attitudes have changed 1

• In 1997, Source Informatics bought anonymised data on GPs’ prescribing to sell to drug companies

• Government tried to use the law to prevent it
  – lawyers argued that to anonymise the data, the pharmacists first had to access it, which constituted using the not-yet-anonymised data for a purpose to which the patient had not consented
Attitudes have changed 2

• Judge ruled that “anonymisation of information obtained in confidence does not exonerate doctors from all further responsibility in relation to the use of such information”

• NHS, MRC, all major research charities appealed, Court of Appeal ruled that the individual has no stake in anonymised data
  – judges declined to rule on the question of whether trading in patient data in the public interest
Overview

• Current controversy around the sharing of GP data for research purposes
  – care.data

• Long-running controversy around the use of hospital mortality data as a measure of quality
  – Standardised Hospital Mortality Ratios
“I went home in tears I had seen and heard enough. The confused man in the next bay was once again being shouted at and told to stay in bed. I was exhausted, since my Mother’s fall she had not slept one night. Most nights she had needed a medic during the night as she had been gasping for breath. My daughter stayed with her that last night and I went home. In hindsight I should have demanded my Mother was seen by a doctor but I didn’t.”
... we then conducted a statistical analysis of the SMRs to examine to what extent they could have been due to random variation... We concluded that, for the three years we examined, there was a less than 5% probability that the high mortality rates at the trust for patients admitted as emergencies aged 18 or over were due to chance.
Hospital Standardised Mortality Ratio

- From NHS Hospital Episode Statistics calculate risk of dying based on:
  - Age group (<1, then 5-year bands to 90+)
  - Sex
  - Admission method/type (emergency, elective etc)
  - Admission source (home, transfers etc)
  - Deprivation quintile (based on postcode)
  - Diagnosis subgroup
  - Comorbidity
  - Emergency admissions in previous 12 months
  - Palliative care (any episode that has a treatment function code 315 or any Z515 code)
  - Month of admission
  - Year of discharge
- Can then calculate the ‘Expected Mortality’ for a hospital, using data about the patients it cares for
Hospital Standardised Mortality Ratio

HSMR = (Actual Mortality ÷ Expected Mortality) x 100

- If expected death rate the same as actual death rate, trust scores 100
- Actual scores tend to range between 75 and 120
- Between 1997 and 2008 the HSMR of Mid-Staffs peaked at 127 and never dropped below 108
- In all but two years the lower end of the 95 per cent confidence interval was above 100
Dr Foster Good Hospital Guide

Choose from the tabs above to view information on this procedure at your selected hospital. You can also view information on other procedures performed at this hospital.

HSMR

What is this trust’s overall level of mortality?
- Green: Less than expected
- Blue: As expected
- Red: More than expected

For those diagnoses that account for the majority (80%) of deaths in hospital, the Hospital Standardised Mortality Ratio (HSMR) measures the number of patients dying at a trust compared with a national benchmark.

Statistical models enable us to estimate the number of deaths at a trust. By dividing the number of actual deaths by the expected number, our modelled number, we create a ratio (shown on the graph with a coloured circle). Our model accounts for factors that can affect mortality but are beyond the control of the trust, for example, a patient’s age.

Control limits (the vertical blue box) measure a trust’s performance in relation to a national standard. If the ratio is within the vertical blue box it is consistent with national performance the trust is performing as expected. A ratio outside the blue box is inconsistent with national performance.

We used data from the last financial year for the one year HSMR; the three year HSMR is calculated with data from the last three financial years. HSMR trend shows the HSMR for each of the last seven financial years, benchmarked against the 2002/03 model. Compare your chosen hospital provider with the five nearest providers performing this procedure. Read more about this indicator.
Problems with HSMR 1: using simple outcomes to measure complex processes

• Expect ‘avoidable deaths’ to be a measure of the quality of care
  – but 98% of hospital visits don’t result in death
  – ~75 and ~90% of in-hospital deaths are unavoidable
  – hospitals are heterogeneous environments

• Expect to catch some poor hospitals, will miss some and wrongly stigmatise others
Figure 2: Distribution of patient numbers by total time in A&E, April-December 2007
Problems with HSMR 2: constant risk fallacy

- If hospital A has more emergency admissions than hospital B, it will probably have higher mortality rates.
  - can standardise for this if we know relative proportions of emergency vs non-emergency admissions.
  - but this assumes that emergencies at A and B are equally risky.
Problems with HSMR 3: gaming the metric

Coding of palliative care at Mid-Staffs
Where are we now?

Dr Foster still promotes measures
CQC uses a revised measure: SHMI
Hospitals assessed on a variety of metrics
Conclusions on Big Data

• Need to argue the case for the research that we want to do, and to win public trust

• All researchers using Big Data should constantly be asking where the data comes from and whether it means what it says