Big data and nursing research

Chris Hill, Royal Free London NHSFT chris.hill@nhs.net

What is big data?

'Big data' doesn't have an actual size, essentially it's data sources that are so large they're challenging to work with.

For example: The 'Mimic III' database contains records from 58,000 hospital admissions to critical care units of the Beth Israel Deaconess Medical Centre in Boston between 2001 and 2012 (Johnson et al, 2016).

Mimic III contains physiological measurements, lab results, prescriptions, freetext notes (nursing / medical / radiology and others), waveforms and more. All names have been removed, and all dates randomly changed to ensure anonymity.
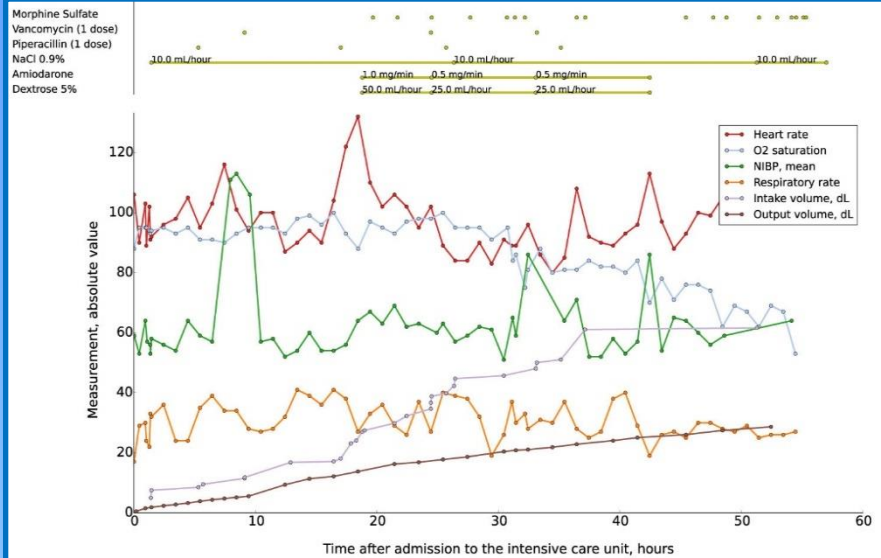
It consists of routinely collected clinical data, and therefore has data quality issues. Along with the fact that it is essentially observational data, this can present challenges.

The promise of big data is the ability to look at a larger picture of healthcare than is possible for humans, and recognise patterns, and correlations that will enable advances in healthcare.

Nurse should seize this opportunity to advance nursing care.

Some of the data that is available for each of the 58,000 patients in the Mimic III database (from Johnson et al, 2016).

Working with big data

Working with big data almost inevitably needs a team: to connect different data sources, and deal with missing or poor quality data, needs computing expertise, which is informed by data science / statistics input, and with clinical expertise to develop the research question and interpret and apply the results.

Research examples

Nurse continuity and hospital acquired pressure ulcers

Stift et al (2015) used data on 42,403 episodes of care to look at nurse continuity and HAPUs. Patient characteristics associated with HAPUs were nutrition, mobility, perfusion, hydration, patient age and skin problems on admission. But when these factors were controlled for, nurse continuity was not significantly associated with HAPU development.

Do 'Do not resuscitate' orders cause higher mortality?

Fuchs et al (2017) used the Mimic III database to look at the effect of a DNR order on 28 day mortality. They matched patients with a DNR in place on day one of ICU admission, and who survived at least 48 hours, with patients who did not have a DNR on day one.
Mortality at 28 days was significantly higher in the DNR group. This may reflect severity of illness not captured by other clinical factors, but the perceptions of the treating team related to do-not-resuscitate status could also be causally responsible for increased mortality in patients with do-not-resuscitate status.

References

Fuchs et al (2017) Quantifying the Mortality Impact of Do-Not-Resuscitate Orders in the ICU *Crit Care Med* 45:1019–1027

Johnson et al (2016) MIMIC-III, a freely accessible critical care database *Sci. Data* 3:160035

Stift et al (2015) Nurse Continuity and Hospital-Acquired Pressure Ulcers: A Comparative Analysis Using an Electronic Health Record "Big Data" Set *Nursing Research*, 64(5):361–371