



BCS Primary Health Care SG

CLICSIG 8/3/14 Analytics for Primary Health Care getting the best out of the data

Guyers House, Corsham

Report authors: Mary Hawking and Ian Herbert

CLICSIG Objectives

This CLICSIG was called to discuss Analytics in Primary Health Care. Information Governance, data quality and the Care.data program, although important, were ruled out of scope from the beginning of the meeting.

Questions:-

1. What data are we considering and where does it come from?
2. Analytics is not just about GP data: what other data sets and sources do we need to consider: what are the relationships between GP data and these and how do we get access to them?
3. How do we make the most of the data available and analytics to help manage the NHS, manage patients to improve outcomes, support clinicians and increase knowledge?
4. What are the constraints of analytics as applied to GP and other NHS and non-NHS data?

Summary & conclusions

There is general agreement that the rich data in general practice could and should be used to improve patient care¹ and also that the Big Data/single generic database approach has considerable problems. "Analytics" is a broad field including - but not confined to - "crunching" (there has to be a better term for this!) patient 'data' in aggregate, pseudonymised, and in some circumstances and/or on rare occasions identifiable form.

Primary care data is very rich but not comprehensive: the value might be increased by linking to other data, from both health and elsewhere (population, deprivation, ONS, economic among others) but consideration needs to be given to:-

1. Quality of information in the different data sets (GP data known/suspected to be variable- see discussion)
 - a. Reason for data being recorded and its effects on what is recorded and in what form
 - i. Direct patient care and life-long record(GP) vs post-discharge Coding for financial claims (Acute Trusts, PBR, ICD-10) vs minimum datasets (applies to MH and possibly SS)
 - ii. Coding - Read, CTV3, SNOMED-CT vs ICD-10 & OPCS
 - iii. Coding on data entry (during patient encounter for patient management) or from records (hospitals Coding in ICD-10 from notes for PBR and HES)
 - b. Comparability of data:
 - i. there are many more Read/CTV3/SNOMED-CT terms than ICD-10
 - ii. Read Code uses synonyms - which may not have same meaning as Preferred Term
 - iii. Mapping between different Coding sets may not be reliable e.g. between different versions of Read Code and Read/SNOMED-CT²
 - iv. Internal inconsistencies in SNOMED-CT and in other Coding sets.
2. Confidentiality, patient consent and risk of reidentification
3. Distribution of data collected/linked to be analysed: however strict the conditions, once data is outside the control of an organisation, possibility of abuse (cloud storage will be outside UK/EU law if controlled by non-EU company or held by an EU company in some countries, e.g. the USA, outside the EU)
4. Purpose of analytics e.g. service management (precise data quality less important: aggregate/securely de-identified data probably sufficient) whereas direct patient care needs identification, or in case of e.g. risk assessment requires the ability for responsible HCP to re-identify patient. Research - somewhere between. Decision support applications different again.
5. Risk of re-identification was in scope - information governance out of scope: we noted that new approaches e.g. Watson & Autonomy plus greater analytical capacity such as Google Analytics will change ability to "crunch" data, increase risks, and it is reasonable to expect even greater analytical capability in the future. There was general support for limiting distribution of data to that relevant for purpose of each analysis
6. Use of applied analytics in front-line situations e.g. QOF prompts and, in future, pathways, especially where care is delivered across a number of different organisations e.g. incorporating NICE guidance prompts into management of conditions such as DM and Dementia. Possibly Map of Medicine. We took examples of existing applications - Risk Stratification - and conditions with guidelines/extensive general practice data requirements - DM and Dementia - to tease out problems including those of incomplete or differently recorded/specified data and adherence to guidelines.

¹ <http://phcsg.org/publications/fair-shares-for-all-final/>

² Comment by Roger Weeks <http://www.bmj.com/content/335/7615/330?tab=responses>

http://www.ask.com/wiki/SNOMED_CT?o=2802&qsrc=999&ad=doubleDown&an=apn&ap=ask.com

Discussion

The discussions fell into different areas.

Purpose and scope of analytics in primary health care

Selection of data e.g. to enable maximum use of data from various sources e.g.:-

- only give me patients with diabetes & treated by Guideline X
- if using Read Code, what Codes should be extracted
- another example was grouping medication treatment together e.g. patients on a statin

Variations in data may give pointers to either data/recording quality/completeness or need to take local population/local circumstances into account e.g. :-

- Dementia recording in London showed much lower than National prevalence: on further investigation, in one PCT a lot of dementia is recorded in hospital records & by medication, not by diagnosis in GP records (also found some misdiagnoses): once realised this increased prevalence by 20-30%
- In Wandsworth, the population is mainly age 30-40 + children, so low prevalence as expected
- Bristol had a cigarette factory with free cigarettes for its employees - so high prevalence COPD explicable.

Financial and administrative incentives (e.g. QOF) affect both what - and how comprehensively - data is recorded and how it is entered. Practices create records for the purposes of the practice, not to create data for others, and there is wide variation and little guidance.³

Data requires context and interpretation.

Example given:-

- Cancer mortality high
- Is this due to late presentation?

Many cancers are first identified in A&E even though the patient had been seen in GP practice previously - often for a different problem, leading to an erroneous assertion that GPs were missing diagnoses or bad at identifying early cancers.

After considering the purpose and scope of analytics in primary health care we looked at the specific situations of:-

Risk Stratification - work already done, what data is needed and what it is/could be used for.

Disease domains - Diabetes Mellitus and Dementia - where there are a great many NICE and other guidelines and existing QOF and other data collection requirements, and where care crosses organisational boundaries

Implications of data collection methods on the type, completeness, comparability and quality of the data, and the implications for analysing such data.

Interpretation of available data

New storage and data manipulation techniques including text recognition/interpretation such as Watson and Autonomy: Big Data, pseudonymisation/de-identification/reidentification: distribution of data vs single database with queries run against it.

³ <http://www.ingentaconnect.com/content/bcs/ipc/2011/00000019/00000004/art00007>

Risk Stratification

Risk stratification was presented by a participant involved in active use of the process in a PCT/CCG.

- Originally the risk being examined was that of readmission to hospital (PARR+), and later readmission within 30 days (PARR30). These were based purely on in-patient data.
- Risk Stratification now includes general practice data .
- There are two main objectives in Risk Stratification: to improve service management - where there is no need to identify individuals - and for direct patient care - to allow concentration of support and resources on individuals at increased risk of admission/readmission.

The group noted that many patients at risk of readmission are extremely ill and recurrent hospital admissions may be necessary and not preventable: probably, as far as benefits and finances go, concentrating on patients before they reach this stage would produce better returns.

As yet data from social services and community is not included in the risk stratification calculations - although with the development of more care in the community, such data may/will be important. Savings and improved care have been reported, but these depend on following up on the information collected, co-operation between care providers and new ways of working.

Diabetes was taken as an example of a common condition with many NICE and other guidelines on management and Pathways, where the responsibility for care lies across many organisations..⁴

- There are implications for data management as well as organisation, finance and outcome assessment.
- Classification of DM into types 1 & 2 was required by QOF: did this lead to misclassification in some patients?
- Sharing medical records across different organisations e.g. GP: hospital services such as diabetic clinics, ophthalmology and vascular surgery: community (DSN - Diabetic Specialist Nurse - and specialised dietary and foot services etc): voluntary organisations supplying e.g. DESMOND services and Diabetes UK: screening and National Audits.
- There are also problems because medical record and coding (if any) in practices may be vastly different in the different organisations and are driven by different record requirements and practices.
- NB Sharing information for direct care is different from information shared for analysis and secondary purposes. For direct care of the individual patient, all the information has to be accurate and either complete or complete within reasonable and recognised parameters, whereas for secondary uses and analysis, some inaccuracy/incompleteness does not destroy the usefulness of the data and can be allowed for.
- Sharing data/patient records supports patient care: it does not replace it: clinical governance and responsibilities need to be agreed between all partners in integrated/fragmented care and mechanisms (which will include the record) put in place to ensure vital information is not overlooked or ignored.
- QOF incentivises general practice to record the data specified in QOF in a way recognised by QOF: it is not clear whether it has led to better record keeping, better adherence to accepted management guidelines, both or neither.
- QOF is only applicable to general practice, and it is not clear whether (or not) the reduced number of QOF Indicators will affect either the amount or quality of data entered. It seems clear that removal of QOF indicators will also remove the incentive for system suppliers to prompt for action(s) as well as the incentive at practice level to take such actions.

⁴ <http://pathways.nice.org.uk/pathways/diabetes>

Decision support

Some time was spent on considering the future possibilities for decision support being incorporated into clinical systems: DM was considered to be a suitable vehicle for looking at this, due to the large number of guidelines related to it and the need to provide integrated care across different care providers with the patient as a major player in self-care.

Computer assisted decision support depends on:-

- checking an individual patient's record to see whether the guidance applies to them (e.g. recognition that this patient has a diagnosis of DM)
- then checking to see whether the guidelines/pathways in the decision support software have been followed, values are within acceptable parameters (both time and value), and appropriate actions/medication have been taken/prescribed (unless there is a recorded contraindication)
- if any action needs to be taken, the system needs to alert the care provider/patient/person responsible for the action.
- Computerised decision support depends on the information in the patient record
- It follows that for integrated care across different providers there needs to be either a single record of prime entry or (at least some of) the data in every record needs to be recorded in all the others in a form which would trigger the decision support
- Decision support exists in general practice systems and is a requirement of GPSoC: prompts have been developed related to QOF: all GP systems incorporate prescribing warnings of interactions and recorded adverse reactions.
- It is not clear whether, if different providers in an integrated pathway are using different record-keeping systems (which may not be entirely electronic and/or coded), it would be feasible to depend on decision support
- Guidelines and Pathways change, so any decision support system needs to be designed to accommodate changes.
- The potential for improving care and avoiding errors is considerable, but it is not easy to apply except in limited areas such as prescribing

This is an area we would expect to develop in the future - but is very difficult, and, if the organisation both of the record and the integrated care scenario is not taken into consideration probably won't deliver full benefits.