Routine quality assurance of cancer registration data - a role for re-abstraction

NHS
Northern and Yorkshire
Cancer Registry and Information Service

Carol Lister, Caroline Brook,
Val Saunders, Christine Head

September 2005

Stage 1 Why we started the project

- Registration Standard 4.2 of the Cancer Registry Standards for England (2001) states that: *'The Cancer Registry should have in place the quality assurance programme as agreed by CRAG and recommended by the UKACR'*
- In practice, a national standard programme was not agreed, due to substantial differences in registration processes between cancer registries; local interpretation of the standard was therefore required
- The NYCRIS data acquisition processes are largely paper-based, relying on trained data abstractors for high quality, consistent data
- The re-abstraction process was designed to:
  - Ensure consistency of data abstraction/ interpretation
  - Identify training issues (for individuals or in general)
  - Highlight any issues requiring clarification or further investigation (e.g. new treatments/new definitions)

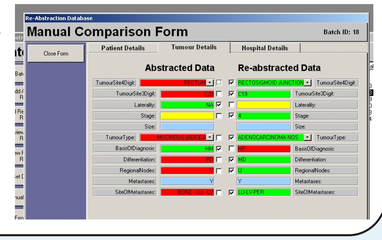
Stage 2 How did we design the project?

- A very small scale feasibility study was undertaken to test out the concept of the project
- Two methods were tested, re-abstraction on the same day and re-abstraction on a different day after initial abstraction. Each involved collection of a common dataset from approximately 50 sets of casenotes
- Data abstraction forms were cross-checked for differences
- This study showed that collecting data at a later date made it difficult to measure the differences as data can change in the casenotes over time
- As a result of the feasibility study it was decided to undertake a pilot exercise involving 6 abstractors and 3 re-abstractors

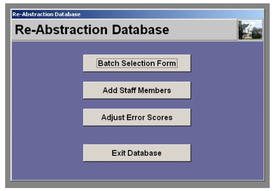
Stage 3 The Pilot study

- n = 207
- 6 abstractors/3 re-abstractors
- 6 different hospitals
- Abstractors and re-abstractors accessed hospital casenotes on the same day
- Abstractors and re-abstractors both abstracted the same dataset from the hospital casenotes using the NYCRIS CR2 form. Each record contained 51 data items
- Abstractors were using the partially completed CR2 from the initial notification, re-abstractors were starting with a blank CR2

- Data from both CR2s were manually input into the database by a temporary member of staff. This was independent of the routine registration process (which continued as normal)
- Once input, each pair of records was manually reviewed and scored by the Registration Quality and Development Manager. Errors could be made by either the abstractor or re-abstractor. It was not always possible to decide without the source material



- A stand-alone database was developed in Microsoft Access to hold and compare the datasets.



The database reports on "batches" per abstractor/re-abstractor combination and reports on the differences at individual data item level as well as for the batch as a whole. "Weightings" of 1, 2 or 3 can be assigned to the data items depending on importance

Stage 4 Results of the Pilot study

- The database produced individual reports for each abstractor/re-abstractor combination, i.e. per "batch"
- Table 1 shows the results for all 6 pilot batches combined, before and after the manual review process
- Categories PPD, PA and AP were identified for manual review. These equated to 12.3% of the data items in the pilot.
- Following review of the differences, PPD, PA and AP accounted for 8.3% of all data items abstracted. Reasons for changes during manual review are included in Stage 6
- This rate varied between batches from 5.3% to 11.0%
- In 3 out of the 6 batches the abstractors were assigned the highest number of errors and in the other 3 batches this was reversed with the re-abstractors scoring the highest number of errors. Some of the reasons for this are identified in Stage 6
- At an individual data item level, Table 2 shows the top/bottom 5 data items with the lowest/highest error rates, either by an abstractor or re-abstractor, and their assigned weightings. (NB. These exclude Address Line4 and details relating to hospital visit 4 or 5, as these were mostly blank)

Table 1

	% Before manual scoring	% After manual scoring
Present-Present Same (PPS)	54.5	57.7
Present-Present Different (PPD)	7.3	4.2
Present-Absent (PA)	3.1	2.4
Absent-Present (AP)	1.9	1.7
Absent-Absent (AA)	33.2	34.0

Table 2

Data items with lowest error rates			Data items with highest error rates		
	Errors	Weighting		Errors	Weighting
Surname	0	3	Ethnic Origin	64	1
Forename	0	3	Marital Status	42	1
Date of Birth	2	3	Hospital2 Date	35	2
Tumour Size	2	2	First Treatment Date	34	2
Other Surname	5	1	Regional Nodes	32	1

Stage 5 Feedback of results

Individually to abstractors

- Enter the data in the correct place on the form (e.g. staging in stage box)
- Where appropriate mark data items 'not known'; do not leave blank
- Review the data collection form on completion to ensure that all fields are completed and correct
- The correct patient pathway is very important - e.g. record all episodes and appropriate treatment
- Use only acceptable standard abbreviations
- Write legibly

Common issues for training

- GP referral date = date of letter, not date letter received
- If primary site unknown and multiple secondary sites are present then site must be carcinomatosis e.g. secondary peritoneum with liver mets. Site = Catosis not secondary peritoneum
- The difference between regional nodes and metastatic nodes
- Record precise 4 digit sub-sites e.g. lower lobe bronchus/ upper outer quadrant breast
- Diagnosis and treatment date = date of excision and not date specimen received

"It did show up some differences on some fields"

"It was interesting comparing someone else's abstraction"

"Interesting that two people could look at the same set of notes and see different things, or miss things"

"It was apparent that you know your own hospital better and are more aware of the pathways and which clinician does what"

Stage 6 Lessons learned

- Established abstractors collecting data routinely from the same hospital make fewer mistakes-although interpretation of definitions can be consistently wrong
- Experienced data inputter needed to eliminate input errors
- The re-abstraction process is very time-consuming

PPD Data items unable to be scored without further investigation

Postcode	e.g. BD23 4EE v BD23 3NT
Address	e.g. Ingber Bank v Inger Bank
Marital status	e.g. Married v Widowed
GP	e.g. S Whitehead v S A Hunter
Site	e.g. Inner canthus v nose
Clinical types	e.g. Clinical carcinoma M80103 / malignancy M80003 v non small cell carcinoma

PA/AP Data items consistently missed

Marital status
Ethnic origin
Laterality
Complete address

AA data items consistently absent

e.g. laterality for breast missing on both abstracted and re-abstracted data

Stage 7 Continuing the process

- Fewer data items to be collected, and other adjustments made (e.g. GP practice, not GP; omit consultant initials)
- Inconsistencies to be minimised through very small numbers of re-abstractors
- Re-abstractor to start with the same base information (i.e. the partially completed CR2 from the initial notification)
- Re-abstraction to be carried out the following day (less intimidating for abstractor)
- Experienced data inputter to be used

Design by Medical Illustration Services Cookridge Hospital