

Use of Natural Language Processing to Identify Significant Abnormalities for Follow-up in a Large Accumulation of Non-delivered Radiology Reports

Michael Hurrell^{1*}, Alan Stein² and Sharyn MacDonald¹

¹Department of Radiology, Canterbury District Health Board, University of Otago, Christchurch, New Zealand

²Department of Radiology, Hewlett Packard Enterprise, Palo Alto, USA

Abstract

Objective: A radiology information system failure affected too many radiology reports (13,601) for manual review and detection of findings requiring clinical action, and required a semi-automated screening system to find such patients in a timely manner.

Materials and methods: A novel SNOMED CT based healthcare platform was used to automatically find reports with actionable findings requiring clinical intervention. Record triage and abstraction was accomplished through a process which included data ingestion, user configuration, filter construction, and radiologist team review workflow. A lead radiologist optimised filters for American College of Radiology Category 3 actionable findings and against various exclusion criteria through a visual query construction interface and observed cohort results through a variety of graphical display renderings. A random sample of excluded reports was checked in order to confirm a statistically significant confidence level.

Results: The computer filtered subset of 2878 reports was then reviewed by a team of radiologists through a computer assisted chart abstraction process leading to 12 records for follow-up, and a single patient requiring semi-urgent imaging.

Discussion: This project used standard software that was interactively configured by the investigating radiologist to interrogate big data, rather than requiring specialised query design by nonclinical experts.

Conclusion: This project illustrates the practical application of a generic ontology based big-data healthcare analytics system to address a specific clinical challenge. Benefits included rapid processing, reduced human workload, and improved workflow.

Keywords: Radiology; Healthcare; Optimisation; Ontology; Cancer; Electronic medical record

Introduction

A radiology information system (RIS) printer setup error resulted in delivery failure of more than 13,000 inpatient and emergency department radiology reports between 2011 and 2015. Although softcopy reports were also available online, there was no tracking available to determine if the referrer had read and signed off on these reports. This had the potential to affect thousands of patients and involve months of time for radiologists and clinicians to review and check that patient care was not compromised.

However, it also created an opportunity to trial a natural language processing system to see if it could efficiently find reports that required further investigation, compared to the manual review method.

Materials and Methods

The text analytics system used in this study was HPE Healthcare Analytics Solution (HCAS, Hewlett Packard Enterprise, Palo Alto, CA). The overall process comprised 4 phases: data ingestion from the RIS into the HCAS indices; configuration of the radiologist reviewer interface; interactive filter configuration and optimisation; and radiologist review and report tagging including review of the electronic medical record and paper clinical notes.

Data ingestion

A static data extract of patient information for the 13601 non-distributed radiology reports was created. This extract included relevant patient information (name, National Health Index (NHI) number,

date of birth, date of death if stated, gender), as well as all subsequent radiology reports for each of these patients. The HCAS user interface was configured for 12 data elements (Table 1) within the extract schema and this data was ingested into HCAS. Unstructured (free-text) and semi-structured data was passed through the ontology tagging and context processing pipeline, which enabled searching for specific disease entities and their synonyms such as malignancy, carcinoma, cancer and so on, but ignoring negative statements such as “no evidence of malignancy”.

Technical description and query construction

The HPE HCAS platform accommodates structured and free-text healthcare data through underlying component big-data platform technologies, HPE IDOL and HPE Vertica. IDOL provides natural language capabilities to process unstructured text data and apply ontologies at scale. Vertica is a columnar database which stores structured data and provides sub-second query response times for

***Corresponding author:** Michael Hurrell, Department of Radiology, Christchurch Hospital, Canterbury District Health Board, University of Otago, Private Bag 4710, Christchurch 8011, New Zealand, Tel: (+643) 364-0770; Fax: (+643) 364-0907; E-mail: mikeh@cdhb.health.nz

Received December 13, 2017; **Accepted** December 28, 2017; **Published** December 30, 2017

Citation: Hurrell M, Stein A, MacDonald S (2017) Use of Natural Language Processing to Identify Significant Abnormalities for Follow-up in a Large Accumulation of Non-delivered Radiology Reports. J Health Med Informat 8: 297. doi: [10.4172/2157-7420.1000297](https://doi.org/10.4172/2157-7420.1000297)

Copyright: © 2017 Hurrell M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

many complex queries. HCAS uses a UIMA based text processing framework including an ontology tagger populated with Systematized Nomenclature of Medicine (SNOMED-CT) and International Classification of Diseases (ICD) mappings to annotate clinical concepts prior to negation classification. The user-interface enables visual query construction including operators, nesting, load/save, sub-query macros,

and query filters including sectionization to discriminate between different sections of the clinical reports. Figure 1 shows the initial screen before the setting of filters, and Figure 2 after filters are applied.

Aggregate query results are displayed in a variety of renderings including bar-charts, heat-maps, and tables. Full Electronic Medical Record (EMR) rendering of the radiology report, as well as configurable Comma Separated Values (CSV) extracts for any number of data fields, was available for every record within the result set. In this application, the lead-radiologist built a cohort of at-risk patients through an interactively developed query which combined parametric and conceptual terms. Criteria included specific SNOMED concepts e.g. “neoplastic disease”, text-based regular expressions e.g. “contains ‘cirrh’”, and structured parameter filters e.g. “exclude studies coded as accident-related”.

Radiologist reviewer interface configuration

After query construction, the resultant cohort was made available through the HCAS user interface so that a team of radiologists could determine follow-up status and required action, if any. HCAS included several mechanisms to facilitate this task. The system displayed all EMR radiology reports for a given patient chronologically in a single view, which meant that a minimum of user-interaction was required to access all available radiology studies for each patient. Computer-assisted chart abstraction was made possible through automatic highlighting of the

Data Element
Name
National Health Index number (NHI)
Date of birth
Deceased flag
Referrer
Visit Site e.g. Women’s Hospital
Visit type e.g. Outpatient
Examination description e.g. Ultrasound liver
Examination code e.g. D06 (for Chest radiograph single view)
Examination type e.g. CT
Body part e.g. Head
Report accession number
Report (unstructured text)
Date of radiologist report verification

Table 1: Data elements.

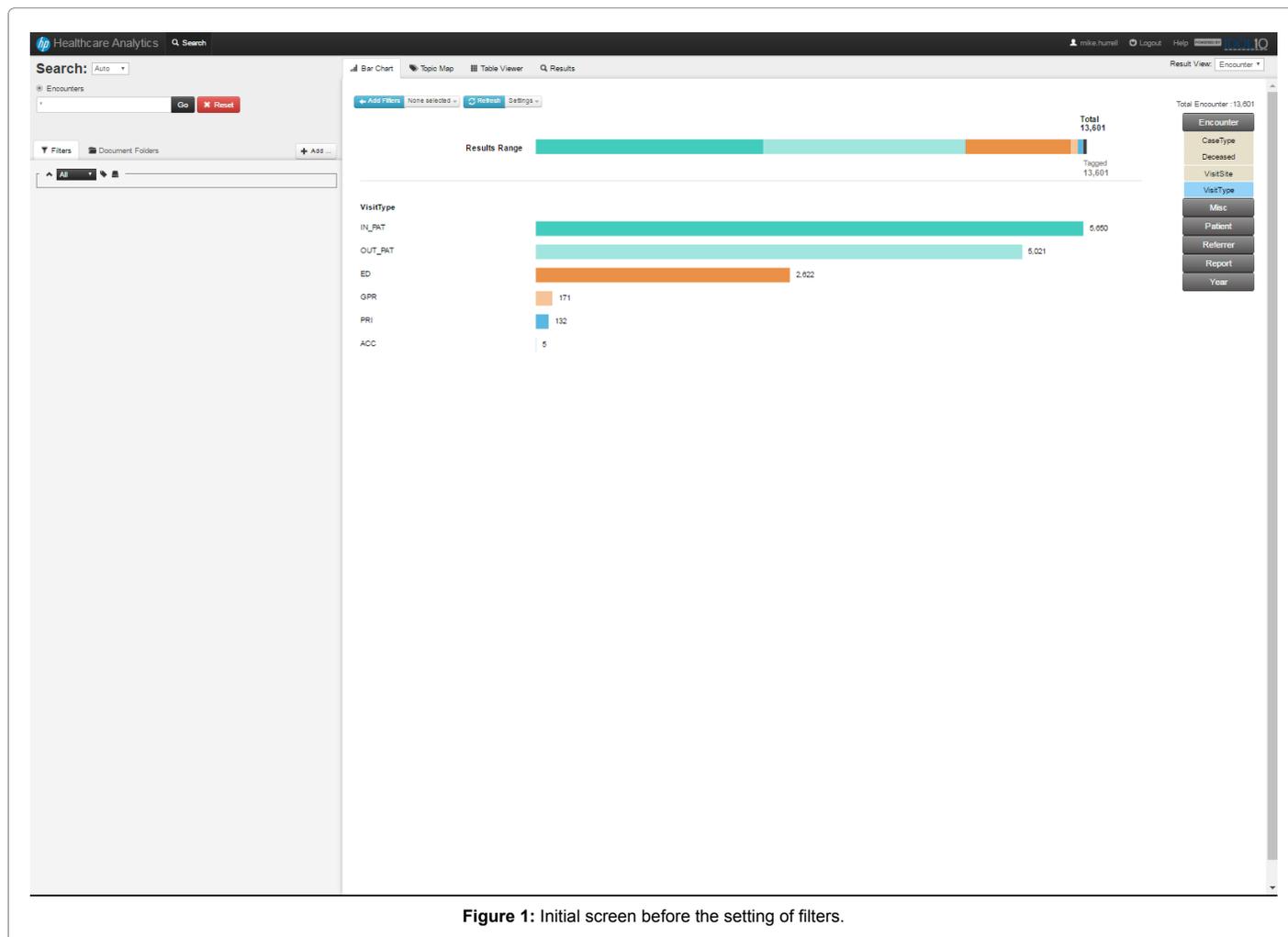


Figure 1: Initial screen before the setting of filters.

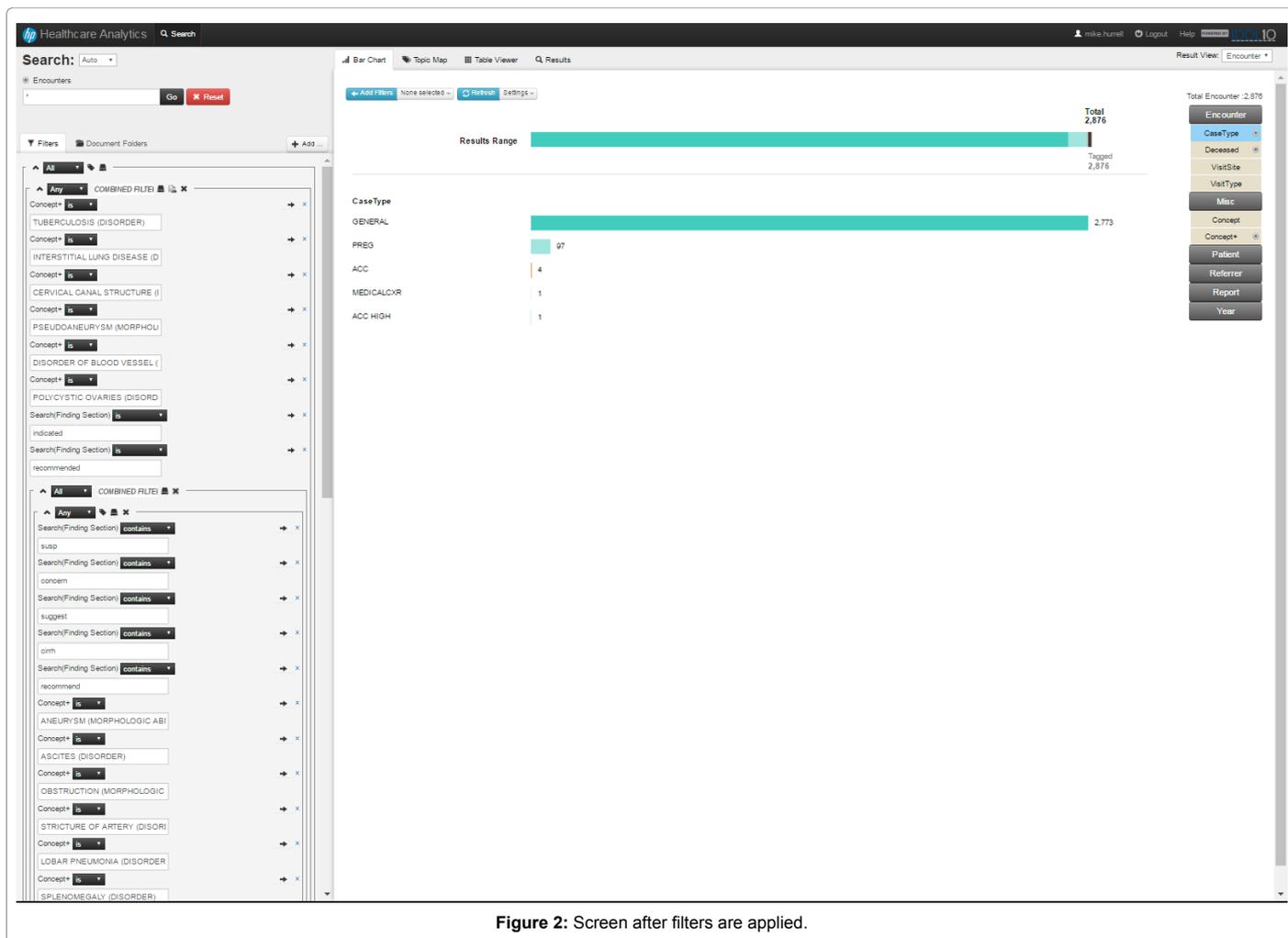


Figure 2: Screen after filters are applied.

query concepts, which allowed for rapid navigation of patient records. For example, any hits for malignancy terms could be highlighted. Figures 3 and 4 show how individual results were viewed.

Furthermore, a file-folder workflow allowed collaboration between radiology team members. The at-risk cohort to be manually vetted was evenly divided between the 4 available radiologists. Upon review, the clinician could categorize the patient record through the user interface. This enabled assessment of the abstraction task as the lead-radiologist could observe the flow of reports as they migrated from the initial queue into the various categories, such as “no further action” or “for review by clinician”.

Filter configuration and optimisation

The report database could be interactively filtered using user-defined filters, configured using standard Boolean logic and nested criteria, and searching variables including: the patient identification number, name, date of birth, date of death (if known), patient type (general, accident, pregnancy etc.), site, visit type (inpatient, outpatient, general practitioner etc.), referrer name, report verification date, examination details (body part, code, description, modality), and the report itself.

Each iteration of the filter took just a few seconds to show report totals in bar graph format. Bar graph displays themselves could be stratified according to the above variables.

However, much of the power of HCAS was its ability to set up searches using the SNOMED classification of clinical terms. The SNOMED nosology allows searching using terms in a branching tree hierarchy. All terms below that point are included in the search. For example malignancy would include carcinoma, lymphoma, metastasis and other hyponyms, and it also allows synonyms such as cancer and carcinoma (Table 2). Simple text searches were added where SNOMED terms did not provide coverage e.g. “recommend” to detect “recommended” or “recommendation” when the report recommended something to be done (Table 3).

Inclusion criteria corresponded with the American College of Radiology (ACR) Category 3 actionable imaging findings [1] i.e. findings that do not require immediate or acute treatment but may cause harm over time, such as possible malignancy, arterial stenosis or aneurysm, cirrhosis.

Acute abnormalities were not included because this task was performed 5 months after the end of the period the distribution problem occurred over. Acute abnormalities that should have healed or recovered were ignored unless raising the possibility of underlying malignancy e.g. lobar pneumonia. For similar reasons, pregnancy findings were not included unless ACR category 3 maternal findings were present i.e. findings still relevant post-delivery. Accident cases are all followed up independently by the New Zealand Accident Compensation Corporation, so these did not need inclusion.

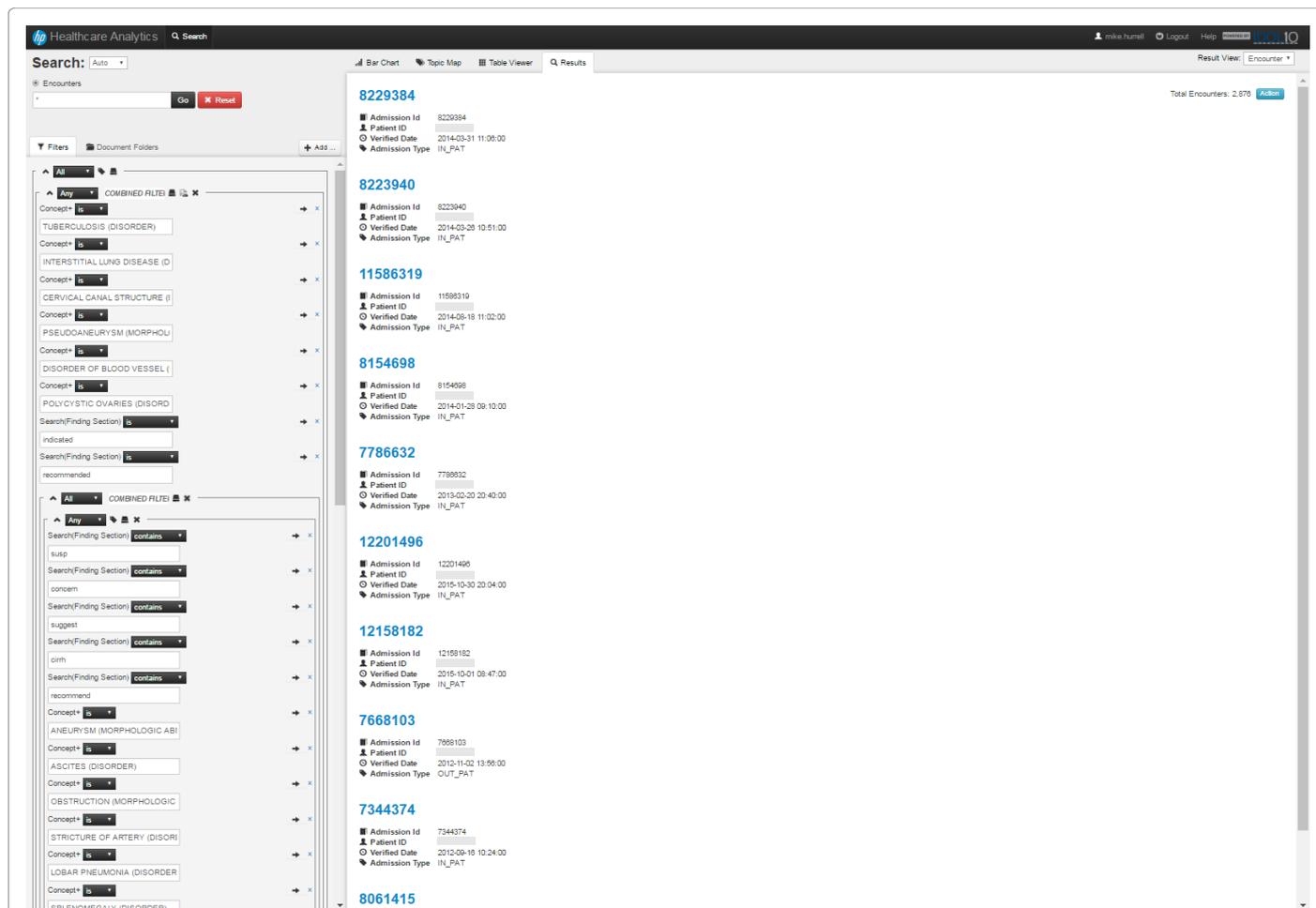


Figure 3: Individual results list.

Reports from deceased patients were dealt with in a separate but similar review process, to check that the report distribution error did not contribute to their death.

Validation of filter

There was the potential to exclude significant findings using this filter, so a random sample of excluded reports was checked. Consultation with a statistician indicated that a sample size of at least 608 records was required to have a 99% confidence level that these exclusions were adequate, assuming a confidence interval of 5%.

Radiology residents (SL, SM) then reviewed a sample and identified any cases that were incorrectly excluded. The lead reviewer (MH) reviewed these, modified the search criteria accordingly and increased the sample numbers again to compensate for those removed by the modified filter.

Radiologist review process

The filter matched 2878 undistributed reports. These were divided up for review by a team of 4 experienced radiologists (MH, AM, VM, RT) to assess whether findings were significant.

Each selected report and associated subsequent reports for that patient could be reviewed on a scrolling screen page. This allowed the reviewer to check if salient findings in the index report were seen in

the clinical information field of subsequent reports, indicating that the original report had been read by the clinical team. Because reports were also available electronically, most reports would have been read online, with the hardcopy only used for clinical backup.

Each report could be categorised with user-defined tags e.g. "Incidental abnormality", "Abnormal but communicated", "Abnormal but not communicated". These were not mutually exclusive, so a case could also be earmarked for discussion with a "For discussion" tag. The tag could be used like a folder so that a particular group of tagged records could be selected for review or further filtering.

Results

The reviewed reports, which had significant findings but no communication confirmed (848) in the RIS, were reviewed with online medical record data such as ward discharge letters or clinic letters, and further reduced to 71 patients requiring clinical review of the paper clinical notes. There were 12 reports that showed potentially significant abnormalities that had no record of being reviewed clinically. These were subsequently reviewed by a clinical team, revealing one patient who required semi-urgent imaging of a possible renal tumour, which was performed 6 months after the non-distributed report and showed slight interval growth; this patient had disseminated malignancy from another primary and in this instance the delay in reporting did not influence the patient outcome.

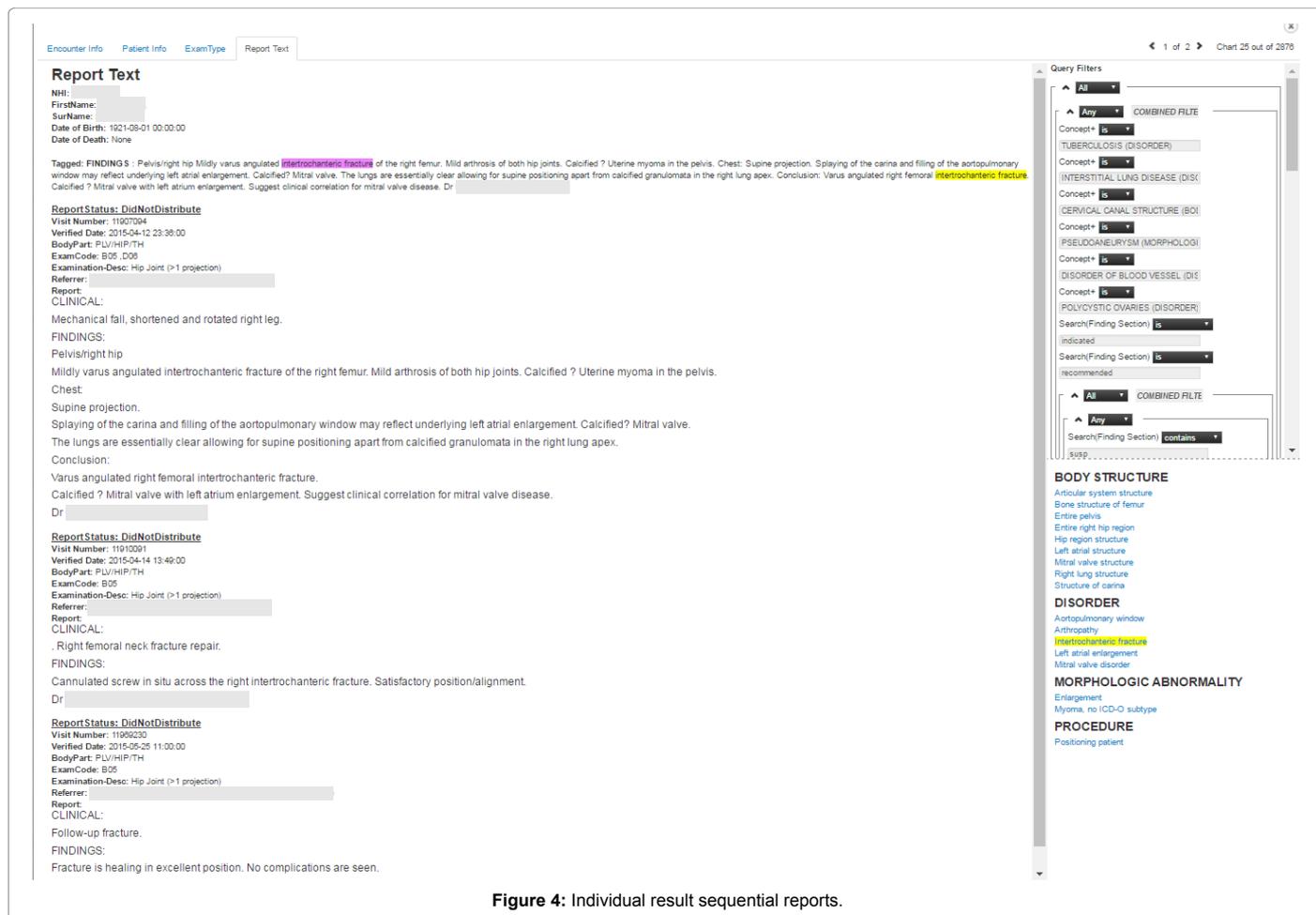


Figure 4: Individual result sequential reports.

Review of deceased patients was done to determine if non-printing of the report contributed to any patient death. The original 13,601 reports were filtered using the hospital information system database of deaths as a filter input to HCAS, yielding 547 deceased patients for review using follow-up radiology clinical information and clinical data.

These reports were assessed to see if there were any significant reported findings that could have contributed to the patient's death, and if so, had they been reviewed by the clinical team. This was confirmed either by clinical information in a follow-up imaging request or by information in the online patient record. Exclusions included no change or improving appearances for known abnormalities, and abnormalities that were not relevant to the patient's death e.g. aortic aneurysm in a patient dying of end-stage heart failure after treatment had been withdrawn. No death was attributable to report non-distribution. See Table 4 for a summary of numbers of patients left at each process stage.

Discussion

Cai et al. [2] describes a variety of approaches by which natural language processing tools can analyse unstructured clinical text. Promising results have been demonstrated with machine learning techniques, syntactic, and semantic approaches [3,4]. Typically, analytics tools are customized for specific applications such as follow-up recommendation detection [4,5] or identification of specific clinical conditions [6,7] such as venous thromboembolism [8]. The power of this approach increases when different EMR databases are linked,

such as when using both radiology and pathology reports for clinical decision support [9]. It is easy to see how analytics technology might become more widespread through more generalized applications such as reducing chart abstraction burden [10]. When this ubiquitous task is combined with a common use-case such as identifying actionable findings [11,12], broad applicability is apparent.

What sets this work apart is that in almost all research to date, the specific clinical application influenced development and configuration of the analytics tool. However, in this work, the analytics tool was developed in an environment that was completely abstracted from the clinical users. The entirety of SNOMED CT was used for annotation without regard to the intended clinical application. A flexible graphical interface permitted filter construction and displayed results in real-time, which enabled users to refine their filters interactively to achieve the accuracy desired for their specific project. And in-depth knowledge of the underlying ontology (e.g., SNOMED CT) was not required for effective system use. The authors believe that the combination of demonstrating value in specific clinical tasks while maintaining broad applicability throughout the healthcare system is important to accelerate adoption of healthcare analytics technologies.

This project illustrated the practical application to clinical practice of a big data filtering tool that used natural language processing. The total number of records scanned by radiologist eyes (including residents) was approximately 3500, amounting to about 25% of the total. Even with HCAS installation and setup time, this greatly

Body System	Term
General	Neoplastic disease (Disorder) (55342001)
	Proliferation (Morphologic abnormality) (30217000)
	Aneurysm (Morphologic abnormality) (85659009)
	Stricture of artery (Disorder) (68109007)
	Lymphadenopathy (Disorder) (30746006)
Chest	Tuberculosis (Disorder) (56717001)
	Cardiomegaly (Disorder) (8186001)
	Lobar pneumonia (Disorder) (278516003)
	Collapse (Morphologic abnormality) (16277007)
Abdomen	Interstitial Lung Disease (Disorder) (233703007)
	Ascites (Disorder) (389026000)
	Splenomegaly (Disorder) (16294009)
	Upper urinary tract dilatation and obstruction (Disorder) (236591007)
	Dilatation of ureter (Disorder) (95576001)
	Kidney stone (Disorder) (95570007)
	Calculus (Morphologic abnormality) (56381008)
	Malformation of urachus (Disorder) (205016000)
	Intestinal obstruction (Disorder) (81060008)
	Obstruction (Morphologic abnormality) (26036001)
Musculo-Skeletal	Polycystic Ovaries (Disorder) (26036001)
	Congenital skeletal dysplasia (Disorder) (105986008)
	Metabolic bone disease (Disorder) (50279003)
	Osteoporosis (Disorder) (64859006)
	Avascular necrosis (Morphologic abnormality) (86217007)
	Bone structure of acetabulum (Body structure) (37783008) AND Dysplasia (Morphologic abnormality) (25723000)

Table 2: SNOMED-CT search terms with associated hierarchy. Concept identifiers are shown for reference but not visible through the UI.

Text term	Target examples
susp	suspicious, suspected
concern	concern, concerning
suggest	suggest, suggested
cirrh	cirrhosis, cirrhotic
recommend	recommend, recommended, recommendation

Table 3: Free text search terms.

	Living	Deceased
Total reports	13601	-
Included in filter	2878	547
Abnormal but no communication noted in subsequent report	848	-
Abnormal but no communication found in any electronic medical records	71	12

Table 4: Numbers at each process stage.

shortened the total time required to review these records. In retrospect, additional linking of report patients with their clinical EMR data would have provided further efficiencies, avoiding an additional manual step. This project took 3 months to process over 13000 reports, including setting up HCAS from scratch. Although difficult to compare, a similar recent report review exercise conducted manually took 6 months to process over 5000 reports (private communication).

The final outcome was favourable only one patient with serious issues had been overlooked clinically and required further imaging, with no significant consequences, probably due to a combination of good luck and

some redundancy in the system. The online availability of reports allowed multiple members of the clinical team to access reports, making it more likely that significant issues were followed up, and many patients were followed up at planned clinic appointments that would prompt online report review.

Despite this, a number of patients had the potential to be harmed by this incident. A subsequent root cause analysis identified poor vendor and client supervision of error log files, which is why the non-printing of reports remained undetected for so long, especially as clinicians had the option of viewing the reports online. The subsequent root cause analysis report has recommended changes in how the RIS vendor handles new installations, and has recommended changing the report distribution system to a fully computerised system with closed loop report acknowledgement.

Acknowledgements

Sam Kalbag and the HPE implementation team for technical design and implementation.

CDHB Radiologists Vicky Mann, Alex Mchaffie and Rennae Thiessen for assisting MH with review of the processed group of radiology reports for significant findings that required further action. Radiology residents Steven Lim and Siska Muljadi for assisting MH with filter validation.

Alice Earnshaw-Morris (CDHB Project Coordinator) for coordinating the clinical aspects of the project.

Matthew Reid (CDHB statistician) for elucidating statistical sampling requirements for filter validation.

The University of Otago for publishing support.

References

- Larson PA, Berland LL, Griffith B, Kahn C, Liebscher LA (2014) Actionable findings and the role of IT support: report of the ACR Actionable Reporting Work Group. *J Am Coll Radiol* 11: 552-558.
- Cai T, Giannopoulos AA, Yu S, Kelil T, Ripley B, et al. (2016) Natural Language Processing Technologies in Radiology Research and Clinical Applications. *Radiographics* 36: 176-191.
- Kimia AA, Savova G, Landschaft A, Harper MB (2015) An Introduction to Natural Language Processing: How You Can Get More From Those Electronic Notes You Are Generating. *Pediatr Emerg Care* 31: 536-541.
- Lacson R, Sugarbaker N, Prevedello LM, Ivan I, Mar W, et al. (2012) Retrieval of radiology reports citing critical findings with disease-specific customization. *Open Med Inform J* 6: 28-35.
- Oliveira L, Tellis R, Qian Y, Trovato K, Mankovich G (2015) Follow-up Recommendation Detection on Radiology Reports with Incidental Pulmonary Nodules. *Stud Health Technol Inform* 216: 10-28.
- Lakhani P, Kim W, Langlotz CP (2012) Automated detection of critical results in radiology reports. *J Digit Imaging* 25: 30-36.
- Warden GI, Lacson R, Khorasani R (2011) Leveraging terminologies for retrieval of radiology reports with critical imaging findings. *AMIA Annu Symp Proc* 2011: 1481-1488.
- Tian Z, Sun S, Eguale T, Rochefort CM (2017) Automated Extraction of VTE Events From Narrative Radiology Reports in Electronic Health Records: A Validation Study. *Med Care* 55: 73-80.
- Patel T, Puppala M, Ogunti RO, Ensor JE, He T, et al. (2017) Correlating Mammographic and Pathologic Findings in Clinical Decision Support Using Natural Language Processing and Data Mining. *Cancer* 123: 114-121.
- Carrell DS, Halgrim S, Tran DT, Buist DS, Chubak J, et al. (2014) Using natural language processing to improve efficiency of manual chart abstraction in research: the case of breast cancer recurrence. *Am J Epidemiol* 179: 749-758.
- Rink B, Roberts K, Harabagiu S, Scheuermann RH, Toomay S, et al. (2013) Extracting actionable findings of appendicitis from radiology reports using natural language processing. *AMIA Jt Summits Transl Sci Proc* 2013: 221.
- Kadom N, Doherty G, Solomon A, Close M, Friedman S, et al. (2017) Safety-Net Academic Hospital Experience in Following Up Noncritical Yet Potentially Significant Radiologist Recommendations. *Am J Roentgenology* 209: 1-5.