Visual Summarisation of Text for Surveillance and Situational Awareness in Hospitals

Hanna Suominen
NICTA, The Australian National University, and University of Canberra
Locked Bag 8001
Canberra, ACT 2601, Australia
firstname.lastname@nicta.com.au

Leif Hanlen
NICTA, The Australian National University, and University of Canberra
Locked Bag 8001
Canberra, ACT 2601, Australia
firstname.lastname@nicta.com.au

ABSTRACT

Nosocomial infections (NIs, any infection that a patient contracts in a healthcare institution) cost 100,000 lives and five billion dollars per year for 300 million Americans alone. Surveillance in hospitals holds the potential of reducing NI rates by more than thirty per cent but performing this task by hand is impossible at scale of every appointment, examination, intervention, and other event in healthcare. Narratives in patient records can indicate NIs and their automated processing could scale out surveillance. This paper describes a text summarisation system for NI surveillance and situational awareness in hospitals. The system is a cascaded sentence, report, and patient classifier. It generates three types of visual summaries for an input of patient narratives and ward maps: cross-sectional statuses at the same point of time, longitudinal trends in time, and highlighted text to see the textual evidence leading to a given status or trend. This gives evidence for and against a given NI in the levels of hospitals, wards, patients, reports, and sentences. The system has excellent recall and precision (e.g., 0.95 and 0.71 for reports) in summarisation for the subset of NIs from fungal species on 1,880 authentic records of 527 patients from 3 hospitals. To demonstrate the system design, we have developed a mobile iPad compatible web-application and a simulation with eighteen patients on three medical wards in one hospital during one month with 61 records in total. The design is extendable to other summarisation tasks.

Categories and Subject Descriptors

H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*current awareness systems*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*natural language, screen design*; H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*organizational design*; J.3 [**Computer Applications**]: Life and Medical Sciences—*medical information systems*

ADCS '13, December 05 - 06 2013, Brisbane, QLD, Australia

http://dx.doi.org/10.1145/2537734.2537739.

General Terms

Design

Keywords

Information Storage and Retrieval; Medical Informatics; Medical Records; Nosocomial Infections

1. INTRODUCTION

Health surveillance can reduce *nosocomial* (a.k.a hospital-acquired) *infection* (NI) rates significantly. NI refers to any infection that a patient contracts in a healthcare institution [20]. It is a serious burden on healthcare (Table 1). Surveillance means ongoing observation and response to the health status of the population and the factors that may affect it through the systematic *monitoring*, that is, collection, analysis, and interpretation of health data [30]. Simply monitoring incidence rates and outcomes regularly can reduce the NI rates by more than thirty per cent [9, 14, 34]. As such, surveillance could release 280 million AUD in Australia every year and 1.5 billion USD in the USA.

In addition to reducing NI rates through better policies and education [30], surveillance can also contribute directly to a given patient's health through healthcare workers' improved *situational awareness*. This is defined as the workers' perception of environmental elements in time and place, the comprehension of their meaning, and the projection of their status in the near future [8]. With NI, it could mean

1. healthcare managers' awareness of hospitals, wards, and beds that have higher or lower NI rates;
2. clinicians' awareness of their patients' current NI-status and risk level; and

Table 1: NI in Australia and the USA

	Australia	USA
Population [million (M)]	20	300
Annual NI rate [M]	0.18	1.7
Annual NI cost:		
- Lives	–	100,000
- Dollars [M]	950 AUD	4850 AUD
	≈ 979 USD	≈ 5,000 USD
- Extra hospital days [M]	0.85	1.7–51
References	[6, 10]	[33, 4, 17, 23]

Table 2: Realistic simulation from 1 to 31 Jan 2013

No. of hospitals (wards) [patients]	1 (3) [18]
No. of patients with 2, 3, 4, 5, 6, and 7 reports, respectively	6, 4, 5, 2, 0, 1
No. of IFI positive (neutral) [negative] reports in the total of 61 reports	21 (16) [24]
Ratio of a patient's IFI positive (neutral) [negative] reports to all reports	
- average — standard deviation	0.29 (0.25) [0.47] — 0.28 (0.22) [0.33]
- minimum — maximum	0.00 (0.00) [0.00] — 0.67 (0.67) [1.00]
No. IFI positive (neutral) [negative] sentences in the total of 485 sentences	38 (48) [62]
No. words (unique words)	8,023 (1,249)


```
- <PatientReports>
 - <Report ReportID="NICTAhospital_1_1_r" ReportType="r" ReportingTime="[2013, 01, 01, 14, 23, 15]" ReportGrade="medium" PatientGrade="medium"
   IFIDiagTime="none">
    <field name="Hospital name">NICTA hospital</field>
    <field name="Location">[Medical, 1, 1]</field>
    <field name="Patient ID">1</field>
    <field name="Result type">CT Chest Hi Resolution</field>
    <field name="Result date">[2013, 01, 01, 14, 23, 15]</field>
    <field name="Result status">Modified</field>
    <field name="Result title">CTCHEHI</field>
    <field name="Performed by">System 1, on 01 January 2013 12:30</field>
    <field name="Verified by">System 1, on 01 January 2013 12:30</field>
    <field name="Encounter info">NICTAhospital, Inpatient, 2012/12/23 -> </field>
    <field name="Final report">Yes</field>
    <field name="Reporting date">[2013,01,01]</field>
    <field name="Reporting time">[14,23,15]</field>
    <field name="Reported by">Dr Marcus Scott</field>
  - <field name="Ctchehi_ct_chest_hi_resolution">
   - <sentences>
       <sentence sentenceno="1">CT chest and upper abdomen. </sentence>
       <sentence sentenceno="2">HI RESOLUTION. </sentence>
     </sentences>
    </field>
  - <field name="Exam">
   - <sentences>
       <sentence sentenceno="1" sentencecolor="medium">? fungal infection. </sentence>
     </sentences>
    </field>
  - <field name="Findings">
   - <sentences>
     - <sentence sentenceno="1">
         Non-contrast axial scans demonstrate no abnormality.
       </sentence>
     - <sentence sentenceno="2">
         Marked volume loss is seen on the right particularly in the upper thorax with distortion of the chest wall and marked crowding of the ribs.
       </sentence>
       <sentence sentenceno="3" sentencecolor="medium">? fungal infection. </sentence>
     - <sentence sentenceno="4">
         A spiculated density at the left lung apex and peripherally in the left lung in the region of the upper margin of the oblique fissure are not significantly altered in appearance from
         the previous CT of 07/11/2002.
       </sentence>
     - <sentence sentenceno="5">
         Well positioned Hickman line. Technique. No other focal lesion is seen.
       </sentence>
     </sentences>
    </field>
  - <field name="Conclusion">
   - <sentences>
       <sentence sentenceno="1">Baseline CT prior to chemotherapy. </sentence>
       <sentence sentenceno="2" sentencecolor="low">No features of fungal infection. </sentence>
     </sentences>
    </field>
  </Report>
 - <Report ReportID="NICTAhospital_1_2_r" ReportType="r" ReportingTime="[2013, 01, 02, 13, 59, 02]" ReportGrade="green" PatientGrade="green" IFIDiagTime="none">
    <field name="Hospital name">NICTA hospital</field>
    <field name="Location">[Medical, 1, 1]</field>
    <field name="Patient ID">1</field>
    <field name="Result type">CT Chest Hi Resolution</field>
    <field name="Result date">[2013, 01, 02, 13, 59, 02]</field>
    <field name="Result status">Modified</field>
    <field name="Result title">CTCHEHI</field>
    <field name="Performed by">System 1, on 02 January 2013 12:20</field>
    <field name="Verified by">System 1, on 02 January 2013 12:20</field>
    <field name="Encounter info">NICTAhospital, Inpatient, 2012/12/23 -> </field>
    <field name="Final report">Yes</field>
```


Figure 1: XML document

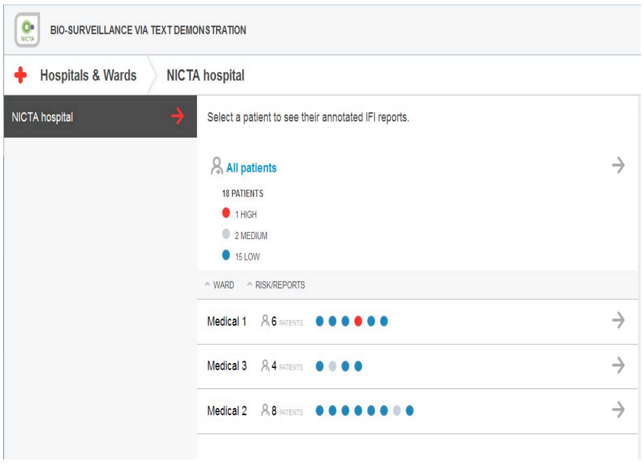

Figure 2: Cross-sectional status for a given hospital

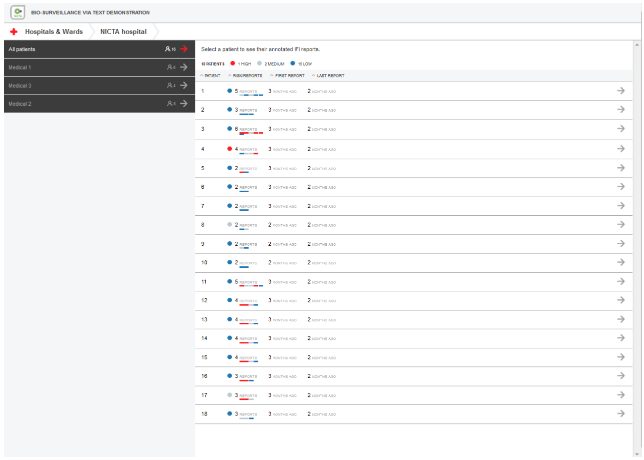

Figure 3: Longitudinal trend of all patients

3. infectious disease specialists' awareness of NI-patients' pathways in hospitals.

This potential of surveillance for reducing NI rates has been officially recognised and consequently, healthcare service providers are increasingly being required to perform NI surveillance. For example, the US Congress addressed this in the Deficit Reduction Act of 2005 by removing, in 2009, the right to charge NI cases not present on admission from hospitals [32]. Also, a surveillance program for NIs was launched in the USA [7] and similar programs exist both in other countries [35] and internationally [27]. However, meeting this surveillance requirement by hand is difficult, if not impossible, at scale of every appointment, examination, intervention, and other event in healthcare. An automated system that analyses NI information in *patient records* enables systematising and scaling out surveillance. The records refer to birth certificates, consent forms, death certificates, dental records, hospital records, medical records, nursing records, patient-maintained health records, and other documentation of patients' clinical care [20]. They serve the purpose of health communication, typically from a professional to another and often in a narrative, free-text form. For example, in a radiology report, a paediatrician might write *Clinical history: a 7-month old boy with wheezing* before sending the patient to an X-ray in order to inform the radiologist performing the examination. The radiologist's response to the paediatrician, after interpreting the X-ray, could be *Impression: Borderline hyperinflation with left lower lobe atelectasis vs. pneumonia. Clinical correlation would be helpful. Unless there is clinical information supporting pneumonia such as fever and cough, I favor atelectasis.* The methodology of reviewing records by hand to investigate NI cases has been used [24, 2, 21]. However, this does not scale to every record in healthcare nor support healthcare workers' situational awareness in real time; though difficult, early detection of NI is key for the rate reduction [36, 28, 12, 25].

The goal of this paper is to design a text summarisation system for supporting NI surveillance and healthcare workers' situational awareness automatically in real time. The system is a cascaded sentence, report, and patient classifier. It generates three types of visual summaries for an input of patient narratives and ward maps:

1. cross-sectional statuses at the same point of time on levels of hospitals, wards, patients, and reports;

2. longitudinal trends in time on levels of hospitals, wards, patients, reports, and sentences; and

3. highlighted text to see the textual evidence leading to a given status or trend.

This gives evidence for and against a given NI on the levels of hospitals, wards, patients, reports, and sentences.

The contributions of this paper include

1. a system design that is novel in terms of its input and output, when compared to the text summarisation and visualisation systems that are used to support NI surveillance and situational awareness in hospitals;

2. a demonstration of this design as a mobile iPad compatible web-application; and

3. its use as a simulation with eighteen patients on three medical wards in one hospital during one month with 61 records in total.

2. MATERIALS AND METHODS

2.1 Evaluation Materials

Statistical performance evaluation of processing methods was conducted on expert annotated authentic reports from three different hospitals in Melbourne, VIC, Australia. This considered the subset of NIs caused by molds and yeasts, that is, invasive fungal infections (IFI). These molds and yeasts are common everywhere in the world (incl. homes and hospital buildings). Normally, when they enter the body through, for example, inhaling the spores or via contaminated surgical instruments, our immune system destroys them. However, this is not the case with chemotherapy, bone marrow transplant, and other immunocompromised

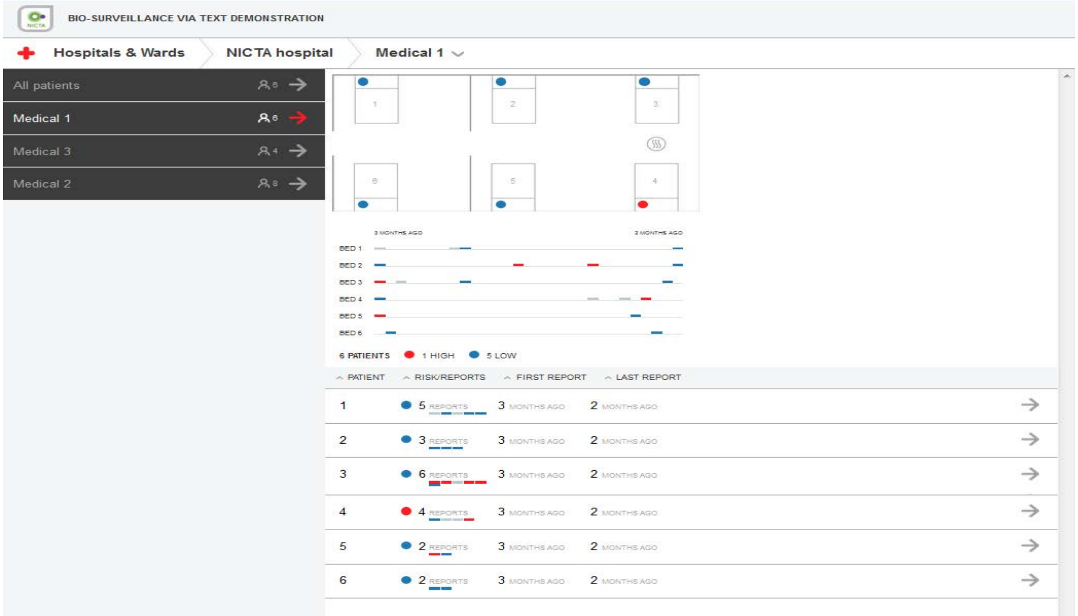

Figure 4: Cross-sectional status of a ward and longitudinal trends in terms of beds and patients

patients, whose compromised immune system cannot effectively fight off the infection. IFIs can occur anywhere in the body and cause serious and often life-threatening conditions. Their early and accurate diagnosis is shown to increase the treatment effectiveness while decreasing its intensity and costs. In Australia, IFI costs over a thousand lives and a hundred million dollars a year [29].

In our previous study [18], we have shown that words in patient reports can give early indications of IFI, and be reliably used to predict the patient's current IFI status (recall of 0.95 and precision of 0.71). The evaluation used $1,880$ reports from 527 patients, of whom 270 had IFI. The reports represented predominantly the genre of radiology. 25 per cent (20 %) of positive (negative) patients' sentences were chosen for expert annotation. The inter-annotator agreement on these $7,083$ sentences was substantial for sentences and almost perfect for reports (Cohen's κ of 0.64 and 0.83, respectively).

2.2 Processing Methods

The processing methods identified the evidence for and against a given NI in patients' free-text reports. This evidence was then used to solve the text summarisation and visualisation problem in the levels of hospitals, wards, patients, reports, and sentences.

The methods classified each patient, report, and sentence as positive, neutral, or negative as a cascade of multi-class classification problems. First, a sentence classifier assigned each sentence into one of the three categories or into the category of other. Second, based on the sentence assignments for all sentences of the report together with other report-level features, a report classifier assigned each report into one of the three categories. Third, based on the sentence and report assignments for all reports and sentences available for a given patient at a given time, a patient classifier assigned each patient into one of the three categories.

The methods were experimented with different compilations of features, statistical classifiers, and decision rules [18]. The feature experiments included bag-of-words, -phrases, and -concepts; their optimised subsets; and negation analyses generated by using the MetaMap program developed at the US National Library of Medicine to map biomedical text to the meta-thesaurus of Unified Medical Language System (UMLS). The classifier experiments included support vector machines (SVMs), Naïve Bayes, Random Forests, and Bayesian Nets. The decision rules, experimented both as features of the statistical classifiers and as classifiers themselves, for the report classifier [patient classifier] were

1. Conservative: If any sentence in the report [any of the patient's available reports] is classified as positive, then assign positive. Else if any sentence in the report [any of the patient's available reports] is classified as neutral, then assign neutral. Else assign negative.

2. Proportional: compare the amount of positive, neutral, and negative sentences in the report [reports available for the patient] and determine the assignment as the majority vote.

The best compilation of processing methods had the sensitivity of 90 per cent (95 % confidence interval (CI) from 86 % to 93 %) and the specificity of 77 per cent (95 % CI from 70 % to 82 %). Of the 25 (5.6 %) missed positive reports, only 4 (0.9 %) were clinically significant. To simplify the experiments, two categories were considered — the neutral category was pooled into the positive category. Ten-fold cross validation on the training set of 366 annotated reports was used to optimise the classifiers, which were then tested on the evaluation set of 83 annotated reports. The sets were mutually exclusive. As expected, recognising negation was beneficial and the bag-of-(words+phrases+concepts) resulted in a better performance than the bag-of one or two approaches. SVM was the best statistical classifier and with it, using all features performed better than their optimised

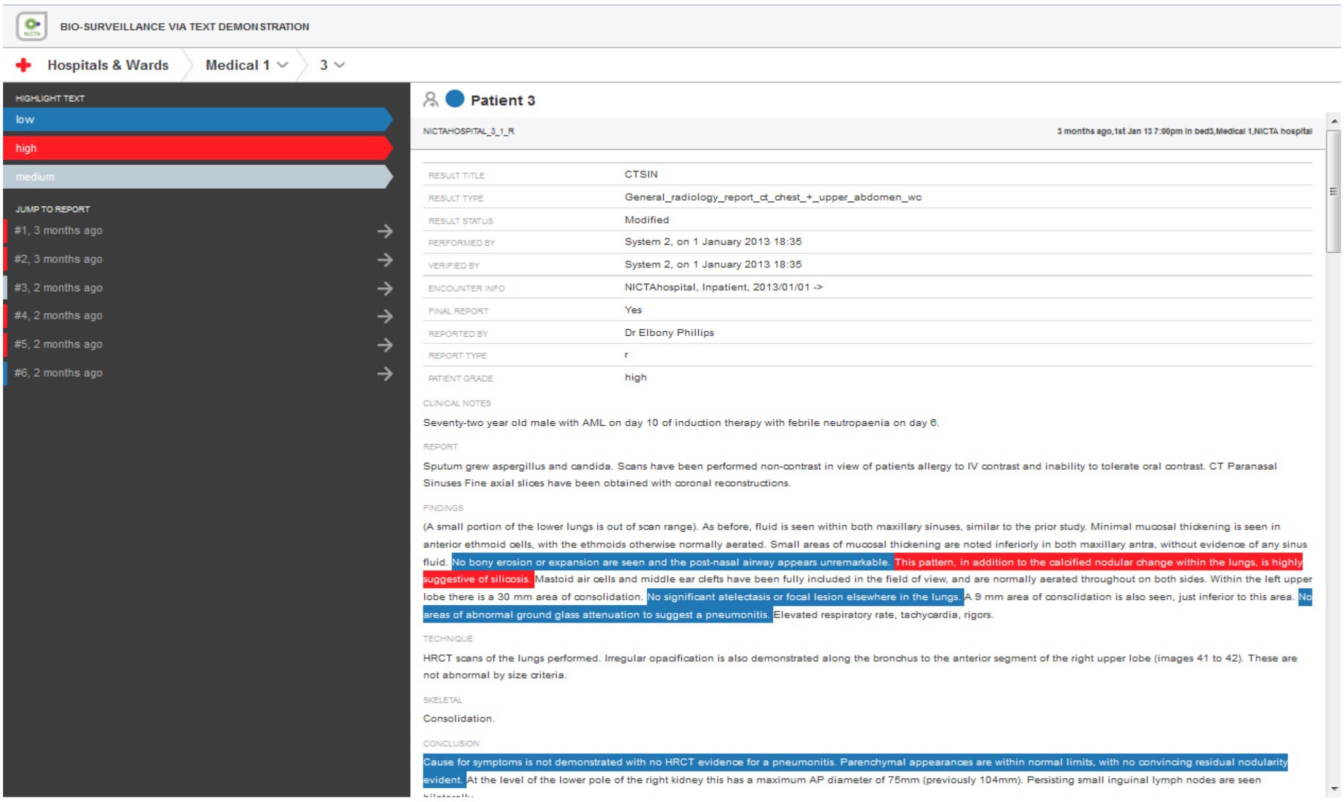

Figure 5: Evidence in the records of a given patient

subset. The best compilation was a hybrid of statistical classifiers and decision rules.

2.3 Demonstration Materials

To demonstrate the system design, we constructed a simulation with eighteen patients on three medical wards in one hospital during one month (Table 2). We acknowledge that this duration may not be realistic, but this allowed demonstrating a ward, which is undergoing an IFI outbreak (i.e., Medical 1), a ward, which has experienced but recovered from an outbreak (i.e., Medical 2), and a ward with strong evidence against an outbreak (i.e., Medical 3).

3. RESULTS

We have designed a text summarisation system for supporting NI surveillance and healthcare workers' situational awareness automatically in real time. The visualisation addresses evidence for and against a given NI in the levels of hospitals, wards, patients, reports, and sentences and supports mobile clinical use. In our demonstration, this NI is IFI but the design is extendable to other tasks.

To demonstrate the design, we implemented a web-application that is iPad compatible.[1] The application is written in HTML5 to allow any web-browser to use it. It uses a consistent colour-coding across all visualisation levels: red refers to positive, grey to neutral and blue to negative. A circle refers to a patient-level prediction, a rectangle to a report-level prediction, and a text highlight to a sentence-level prediction. As an input, the system receives ward maps (i.e., images) and an XML document, which includes all information (i.e., text and classifications in the levels of patients, reports, and sentences) to be visualised in all levels for all hospitals, wards, patients, and reports (Figure 1). The output consists of three types of visual summaries:

1. cross-sectional statuses at the same point of time in levels of hospitals, wards, patients, and reports,

2. longitudinal trends in time in levels of hospitals, wards, patients, reports, and sentences, and

3. highlighted text to see the textual evidence leading to a given status or trend.

The user experience of the system begins with a menu for selecting a hospital and the system visualising its current cross-sectional status (Figure 2). The demonstration includes only one hospital. The visual summary includes

1. an overview of all patients with a capability to select all patients for seeing the longitudinal trend of the hospital and

2. overviews of all wards with a capability to select a given ward for its cross-sectional status and longitudinal trends in terms of beds and patients.

In this level, the demonstration hospital has eighteen patients in total at the wards Medical 1, 2, and 3. The system predicts one out of the eighteen patients as IFI-positive (i.e., high risk of IFI), two as IFI-neutral (i.e., medium risk of IFI), and fifteen as IFI-negative (i.e., low risk of IFI). We

[1] available `http://goo.gl/QwZBln`

also learn that all wards are full. The capacity of the ward is provided after the ward name. Each circle on the right-hand-side of the capacity corresponds to one patient. The colour of the circle corresponds to the patient's current IFI status, predicted by the system. The ward listing can be sorted in descending or ascending order with respect to the ward name and report-level IFI-risk.

A clinician or administrator may select the longitudinal trend of the hospital (Figure 3). This opens a visual summary with

1. an overview of all patients in terms of their current IFI status and

2. a longitudinal trend across all patients in time with the up-to-date patient-level IFI prediction and retrospective report-level IFI predictions.

Each rectangle corresponds to a report. The colour of the rectange corresponds to the report-level IFI-status, predicted by the system. The listing can be sorted in descending or ascending order with respect to the patient ID, report-level IFI-risk, entering time of the first report, and entering time of the last report. Each patient, or a given ward, can be selected for further analysis. In this level, we learn that the demonstration hospital has had many IFI alarms in the report level, but the situation has improved during the past month. Also the red alarms tend to calm down to blue ones via grey in the report level and similarly, before a red report, we see early indicators as grey reports.

We continue the analysis of the high-risk case by selecting the ward of Medical 1 to be visualised. This opens a ward-level visualisation (Figure 4) with

1. the ward map of all patients at this moment with the up-to-date patient-level IFI predictions,

2. longitudinal visualisation across all beds in time with the retrospective report-level IFI predictions,

3. cross-sectional summary of the ward map, that is, the patient-level IFI predictions for all patients at the ward at this moment, and

4. longitudinal visualisation across all patients in time with the up-to-date patient-level IFI prediction and retrospective report-level IFI predictions.

The ward map and cross-sectional summary are intended to support healthcare workers' situational awareness and serve diagnostic purposes. The longitudinal visualisation across all beds allows taking patient flow into account. This is to serve epidemiological purposes. The longitudinal visualisation across patients is there to support research and the list can be sorted in descending or ascending order with respect to a patient number, report-level IFI-risk, date of the first report, and date of the last report. From the ward map we learn that the high-risk case is located on the bed four, next to the air ventilation duct of the ward. From the longitudinal visualisation across all beds we can conclude that the ward has had very many IFI alarms during the past month, but the situation seems to have improved. The alarms seem to be clustered around the duct on beds 2–5 but the beds 1 and 6 behind the space divider have had no alarms. The longitudinal visualisation across all patients allows monitoring report-level trends in time and also shows that our high-risk patient has had early indicators of IFI on previous two reports already.

If we wish to study the evidence causing the alarm for the high-risk case, we can select the records of this patient to be visualised (Figure 5). The left-hand-side column shows a longitudinal summary across all reports of the patient with report-level IFI predictions. This menu allows selecting a particular report to be visualised. The right-hand-side column shows longitudinal details across all reports in time. The sentence-level IFI predictions are shown as text highlights. Text not seen as relevant to the IFI analysis by the system has no highlighting.

4. DISCUSSION

We have designed and implemented a system to provide text summarisation and visualisation in hospitals to support NI surveillance and healthcare workers' situational awareness. The visualisation provides supporting- and contra-evidence for a given NI in the levels of hospitals, wards, patients, reports, and sentences. The inputs of the system are free-form clinical text (patient records) and images (ward maps). The outputs are three types of visual summaries:

1. cross-sectional statuses at the same point of time in levels of hospitals, wards, patients, and reports;

2. longitudinal trends in time in levels of hospitals, wards, patients, reports, and sentences; and

3. highlighted text to see the textual evidence leading to a given status or trend.

The design can be extended from the case of IFI to other tasks on text summarisation and surveillance in hospitals — one system for every task.

Manual IFI surveillance on patient records has been performed on the Prospective Antifungal Therapy Alliance's registry (subset of 7,526 IFI cases in 6,845 patients from all inpatients at 25 medical centers in North America between 2004 and 2008) and Hema e-Chart registry (subset of 147 patients with mycoses from all inpatients at 26 haematology units in Italy between 2007 and 2009) [2, 21]. However, early diagnosis of IFI is crucial to facilitate treatment success, rigorous enforcement of surveillance (by means of blood cultures, antigens and diagnostic imaging) of all patients at risk for IFIs, and diagnostic delays explaining IFI outbreaks [28, 25]. Surveillance by hand and in real time is difficult, if not impossible, at scale of every appointment, examination, intervention, and other event in healthcare.

Natural language processing (NLP) and *text mining* have been recognised as ways to automate the analysis of patient narratives; this includes the pioneering studies [3, 1, 13, 5] from 1970's and over 12,000 more recent references in PubMed. Successful examples systems that have progressed from research to use in hospitals include MedLEE, Medical Language Extraction and Encoding System and Autocoder: MedLEE is routinely used in the New York Presbyterian Hospital in New York, NY, USA to parse English patient records and map them to UMLS [19]. Autocoder assigns diagnosis codes to patient records at the Mayo Clinic in Rochester, MN, USA which has changed the coding personnel's duties from code assignment to verification and reduced workload by eighty per cent [22].

We are not aware of text summarisation or visualisation systems that are used to support surveillance and situational awareness related to IFI in hospitals. Based on our scientific literature search, the most similar language processing system to ours that has been used for surveillance in hospitals is TheraDoc [11], a rule-based system that was used three times a day during the 2002 winter Olympics in Salt Lake City, in UT, USA to monitor influenza, pneumonia, and twenty-two other concerns for public health on radiology reports and tests. NLP is used to convert free text to incident rates and then the system visualises the rates as longitudinal graphs of rates in time. However, our geographic modality of maps is not addressed and textual evidence is not given, visualised, nor summarised. Based on our scientific literature search, the most similar language processing system to ours that has been used for text summarisation and visualisation in hospitals is MAGIC [15], a rule-based system used in one cardiothoracic intensive care unit of the Columbia University Medical Center in New York, NY, USA. MAGIC is designed to support healthcare workers' workflow, information needs, and decision during the peri-operative period of cardiac surgical patients. It summarises individual patient's raw medical data into procedural milestones, illness severity, and care therapies as well as generates two displays: a briefing about a patient's operative course and a depiction of a patient's clinical status at a given time point with respect to illness severity. We believe that extending this system to surveillance would be straightforward by using our design if the depictions, briefings, and input raw data were combined with ward maps.

Based on our search of patents, the most similar system to ours is for analysing information related to public health from structured and unstructured sources and visualising these results on a temporal and geo-spatial display [16]. These patent applications explicitly mention NLP of medical records as well as temporal and geo-spatial visualisations to support surveillance and situational awareness. However, the visualisations are generated in a global level as opposed to our hospital-specific perspective. Based on our search of patents, two other related systems are for graphically displaying a health status of geographically dispersed hospitalised patients with no connection to language processing [26] and for assigning diagnosis codes to patient records based on language processing but with no connections to surveillance, visualisation, geographic modality of maps, textual evidence, or text summarisation [31].

5. CONCLUSIONS

NLP could enable systematic NI (incl. IFI) surveillance at scale and in real time in hospitals. This automation would scale the surveillance up from manual and often delayed reviews of most suspicious cases to systematic analyses of every report of any patient for any evidence for or against a given NI immediately after the report has become available. This holds the potential for improving healthcare management and clinicians' situational awareness in real time and is applicable to both epidemiological or diagnostic purposes.

Acknowledgments

NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program. NICTA is also funded and supported by the Australian Capital Territory, the New South Wales, Queensland and Victorian Governments, the Australian National University, the University of New South Wales, the University of Melbourne, the University of Queensland, the University of Sydney, Griffith University, Queensland University of Technology, Monash University and other university partners. We thank Andrea Lau and Jack Zhao from Small Multiples, David Martinez, Lawrence Cavedon, and Mike Wu from NICTA and Sumithra Velupillai from DSV-Stockholm University for their contribution to the system experimentation and implementation. We express our gratitude to our hospital partners Karin Thursky, Michelle Ananda-Rajah, and Monica Slavin from Alfred Health, Peter Mac, The Royal Melbourne Hospital.

6. REFERENCES

[1] B. Anderson, I. D. J. Bross, and N. Sager. Grammatical compression in notes and records: Analysis and computation. *American Journal of Computational Linguistics*, 2(4):68–82, 1975.

[2] N. Azie, D. Neofytos, M. Pfaller, H. Meier-Kriesche, S. Quan, and D. Horn. The PATH (Prospective Antifungal Therapy) Alliance registry and invasive fungal infections: update 2012. *Diagnostic Microbiology and Infectious Disease*, 73(4):293–300, 2012.

[3] H. Becker. Computerization of patho-histological ndings in natural language. *Pathologia Europaea*, 7(2):193–200, 1972.

[4] J. Burke. Infection control — a problem for patient safety. *The New England Journal of Medicine*, 348(7):651–6, 2003.

[5] M. F. Collen. Patient data acquisition. *Medical Instrumentation*, 12(4):222–225, 1978.

[6] DoHA. *Government working on $158 million transition beds — getting elderly into appropriate aged care*. Australian Government, Department of Health and Ageing, 2007.

[7] DoHHS. *Healthcare-Associated Infection (HAI)*. U.S. Department of Health & Human Services, 2010.

[8] A. Fore and G. Sculli. A concept analysis of situational awareness in nursing. *Journal of Advanced Nursing*, 25 Mar., 2013.

[9] P. Gastmeier, C. Geffers, C. Brandt, I. Zuschneid, D. Sohr, F. Schwab, M. Behnke, F. Daschner, and H. Ruden. Effectiveness of a nationwide nosocomial infection surveillance system for reducing nosocomial infections. *Journal of Hospital Infection*, 64(1):16–22, 2006.

[10] N. Graves, K. Halton, D. Paterson, and M. Whitby. Economic rationale for infection control in Australian hospitals. *Healthcare Infection*, 14(3):81–8, 2009.

[11] A. Gundlapalli, J. Olson, S. Smith, M. Baza, R. Hausam, L. Eutropius, S. Pestotnik, K. Duncan, N. Staggers, P. Pincet, and M. Samore. Hospital electronic medical record-based public health surveillance system deployed during the 2002 Winter Olympic Games. *American Journal of Infection Control*, 35(3):163–71, 2007.

[12] Y. Hayashi, K. Morisawa, M. Klompas, M. Jones, H. Bandeshe, R. Boots, J. Lipman, and D. Paterson. Towards improved surveillance: the impact of ventilator-associated complications (VAC) on length of

stay and antibiotic use in patients in intensive care units. *Clinical Infectious Diseases Advance Access*, 2012. (accepted).

[13] L. Hirschman, R. Grishman, and N. Sager. From text to structured information: automatic processing of medical reports. In *American Federation of Information Processing Societies: 1976 National Computer Conference*, volume 45 of *AFIPS Conference Proceedings*, pages 267–75, 1976.

[14] S. Huang, D. Yokoe, V. Hinrichsen, L. Spurchise, R. Datta, I. Miroshnik, and R. Platt. Impact of routine intensive care unit surveillance cultures and resultant barrier precautions on hospital-wide methicillin-resistant Staphylococcus aureus bacteremia. *Clinical Infectious Diseasess*, 43(8):971–8, 2006.

[15] D. Jordan and S. Rose. Multimedia abstract generation of intensive care data: the automation of clinical processes through AI methodologies. *World Journal of Surgery*, 34(4):637–45, 2010.

[16] A. Kass-Hout and M. Mirabito. *Global Disease Surveillance Platform, and Corresponding System and Method. WO 2008/013553 A2, EP 2052318 A2, US 2009/0319295 A1, and US 2012/0284039 A1. Patent applications*, 2008/2009/2012.

[17] R. Klevens, J. Edwards, C. Richards, T. Horan, R. Gaynes, D. Pollock, and D. Cardo. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Reports*, 122(2):160–6, 2007.

[18] D. Martinez, H. Suominen, M. Ananda-Rajah, and L. Cavedon. Biosurveillance for invasive fungal infections via text mining. In *Conference and Labs of the Evaluation Forum 2012, CLEF eHealth*, Rome, Italy, 2012.

[19] E. Mendonca, J. Haas, L. Shagina, E. Larson, and C. Friedman. Extracting information on pneumonia in infants using natural language processing of radiology reports. *Journal of Biomedical Informatics*, 28(4):314–21, 2005.

[20] MeSH. *Medical Subject Headings (MeSH) Browser*. U.S. National Library of Medicine, 2012.

[21] A. Nosari, M. Caira, M. Pioltelli, R. Fanci, A. Bonini, C. Cattaneo, C. Castagnola, S. Capalbo, P. D. Fabritiis, V. Mettivier, M. Morselli, D. Pastore, F. Aversa, G. Rossi, and L. Pagano. Hema e-Chart Registry of invasive fungal infections in haematological patients: improved outcome in recent years in mould infections. *Clinical Microbiology and Infection : the Official Publication of the European Society of Clinical Microbiology and Infectious Diseasesy*, 15, 2012.

[22] S. Pakhomov, J. Buntrock, and C. Chute. Automating the assignment of diagnosis codes to patient encounters using example based and machine learning techniques. *Journal of the American Medical Informatics Association: JAMIA*, 13(5):516–25, 2006.

[23] E. Perencevich, P. Stone, S. Wright, Y. Carmeli, D. Fisman, and S. Cosgrove. Raising standards while watching the bottom line: making a business case for infection control. perencevich en, stone pw, wright sb, carmeli y, fisman dn, cosgrove se; society for healthcare epidemiology of america. source raising standards while watching the bottom line: making a business case for infection control. *Infection control and hospital epidemiology : the official journal of the Society of Hospital Epidemiologists of America*, 28(10):1121–33, 2007.

[24] M. Petersen, M. Holm, S. Pedersen, A. Lassen, and C. Pedersen. Incidence and prevalence of hospital-acquired infections in a cohort of patients admitted to medical departments. *Danish Medical Journal*, 57(11):A4210, 2010.

[25] E. Repetto, C. Giacomazzi, and F. Castelli. Hospital-related outbreaks due to rare fungal pathogens: a review of the literature from 1990 to june 2011. *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology*, 31(11):2897–904, 2012.

[26] B. Rosenfeld and M. Breslow. *System and Method for Displaying a Health Status of Hospitalized Patients. US 7454359 B2 and US 7433827 B2. Granted patents*, 2008.

[27] V. Rosenthal, D. Maki, and N. Graves. The International Nosocomial Infection Control Consortium (INICC): goals and objectives, description of surveillance methods, and operational activities. *AJIC:American Journal of Infection Control*, 36(9):e1–12, 2008.

[28] M. Rüping, J. Vehreschild, and O. Cornely. Patients at high risk of invasive fungal infections: when and how to treat. *Drugs*, 68(14):1941–62, 2008.

[29] M. Slavin, J. Fastenau, I. Sukarom, P. Mavros, S. Crowley, and W. Gerth. Burden of hospital-ization of patients with Candida and Aspergillus infections in australia. *International Journal of Infectious Diseases*, 8(2):111–20, 2004.

[30] SpringerReference. *Encyclopedia of Public Health*. Springer, New York, NY, USA, 2013.

[31] N. Stefan, R. Bharat, L. Vlad, and Y. Shipeng. *System and Method for Large Scale Code Classification for Medical Patient Records. WO 2008/143865 A1 and US 2008/0288292 A1. Patent applications*, 2008.

[32] P. Stone. Changes in Medicare reimbursement for hospital-acquired conditions including infections. *AJIC:American Journal of Infection Control*, 37(9):A17–8, 2009.

[33] P. Stone, E. Larson, and L. Najib Kawar. A systematic audit of economic evidence linking nosocomial infections and infection control interventions: 1990–2000. *AJIC: American Journal of Infection Control*, 30(3):145–52, 2002.

[34] E. Sydnor and T. Perl. Hospital epidemiology and infection control in acute-care settings. *Clinical Microbiology Reviews*, 24(1):141–73, 2011.

[35] VICNISS. *VICNISS Healthcare Associated Infection Surveillance System*. VICNISS, Australia, 2002.

[36] T. Walsh, J. Hathorn, J. Sobel, W. Merz, V. Sanchez, S. Maret, H. Buckley, M. Pfaller, R. Schaufele, C. Sliva, E. Navarro, J. Lecciones, P. Chandrasekar, J. Lee, and P. Pizzo. Detection of circulating candida enolase by immunoassay in patients with cancer and invasive candidiasis. *The New England Journal of Medicine*, 324(15):1026–31, 1991.