# Pattern-based information extraction from pathology reports for cancer registration

Napolitano, G., Fox, C., Middleton, R., & Connolly, D. (2010). Pattern-based information extraction from pathology reports for cancer registration. Cancer Causes & Control: an international journal of studies of cancer in human populations, 21(11), 1887-1894. DOI: 10.1007/s10552-010-9616-4

**Published in:**
Cancer Causes & Control: an international journal of studies of cancer in human populations

Cancer Causes Control (2010) 21:1887–1894
DOI 10.1007/s10552-010-9616-4

ORIGINAL PAPER

# Pattern-based information extraction from pathology reports for cancer registration

**Giulio Napolitano · Colin Fox · Richard Middleton · David Connolly**

Received: 19 January 2010 / Accepted: 6 July 2010 / Published online: 23 July 2010


**Abstract**
*Objective* To evaluate precision and recall rates for the automatic extraction of information from free-text pathology reports. To assess the impact that implementation of pattern-based methods would have on cancer registration completeness.
*Method* Over 300,000 electronic pathology reports were scanned for the extraction of Gleason score, Clark level and Breslow depth, by a number of *Perl* routines progressively enhanced by a trial-and-error method. An additional test set of 915 reports potentially containing Gleason score was used for evaluation.
*Results* Values for recall and precision of over 98 and 99%, respectively, were easily reached. Potential increase in cancer staging completeness of up to 32% was proved.
*Conclusions* In cancer registration, simple pattern matching applied to free-text documents can be effectively used to improve completeness and accuracy of pathology information.

**Keywords** Surgical pathology · Automatic data processing · Cancer registries · Pattern matching · Information extraction · Text mining · Unstructured data management · Pathology report · Cancer registration · Regular expression

G. Napolitano (✉) · C. Fox · R. Middleton
Northern Ireland Cancer Registry, Centre for Public Health, Queen's University of Belfast, Mulhouse Building, Grosvenor Road, Belfast BT12 6BJ, Northern Ireland (UK)
e-mail: g.napolitano@qub.ac.uk

G. Napolitano
Centre for Statistical Science and Operational Research, Queen's University Belfast, David Bates Building, University Road, Belfast BT7 1NN, Northern Ireland (UK)

D. Connolly
Department of Urology, Belfast City Hospital, Lisburn Road, Belfast BT9 7AB, Northern Ireland (UK)

## Introduction

Knowledge representation systems, dedicated to the storage of knowledge items in a way which is suitable for subsequent retrieval, face a number of challenges: the collection of raw information; the manipulation of this information and its projection into a knowledge representation schema; the storage of such projections; the retrieval of knowledge items for any purpose.

These challenges may be seen both as successive steps in building working systems for use in research departments, although the order of the steps may vary, and as stages of information flow. Much of the current literature and research is focused on the optimal, rigorous way to represent biomedical information and standardise the communication. Biomedical ontologies [1] and HL7 Messaging Standard [2], for example, are instances of such efforts. When dealing with existing corpora of documents, once a suitable system of representation has been chosen, the burden to convert the existing information into the new format is not just a matter of translation. While conversion algorithms from one coding system into a new one can prove difficult to be decided and developed, e.g., between SNOMED [3] and ICD-10 [4], different problems arise when incorporating 'raw' data into a representation structure. While most of medical informatics and bioinformatics literature and mining tools are dedicated to genetic- or gene-related information retrieval from journal papers, there is a need for extracting usable information from clinical or pathology documents [5]. The main problem

researchers have to face is the different kinds of 'rawness' of the information they have to manipulate: clinical pathology documents have a layer of added ambiguity originating by the use of quasi-ungrammatical natural language, which may differ between different documents and even over a single document [6]. Several methods have been investigated to collect usable information from unstructured clinical documents, mainly by using Natural Language Processing techniques [7]. This approach, however, is not always justified when high precision is sought for very specific and urgent information extraction tasks [8]. This is what happens, for example, when term-value data, buried in free-text documents collected by cancer registries, are needed for epidemiological studies.

The present study investigates whether simple pattern matching can be effectively used to improve the completeness of pathology information collected for individual patients, reducing the number of records with partial or no pathology information, or to supply information not available through routine extracts from the main feeder database systems, in the domain of cancer registration. In particular, we investigate a method for extracting important prognostic features from pathology reports associated with two types of cancer, the Gleason score for prostate cancer [9] and the Clark level [10] and Breslow depth [11] for malignant melanomas of skin.

## Materials and methods

The study was set in the Northern Ireland Cancer Registry (NICR), a unit operating within the Centre for Public Health in the Queen's University of Belfast. Although the cancer registration dataset [12] in the UK is involved in the recent phase of projects for the enhancement and, hopefully, higher level of integration of IT systems within the National Health Service [13], at the moment the values of many clinical test results, staging information and other pathology-related data items are not captured by the main database system. As such information is not recorded at the source (laboratory computer systems), it is not received from the data providers as a specific item of the dataset. As a result, all such data are either not available or have to be obtained by human inspection of the free-text pathology reports or by the application of ad-hoc techniques.

### Data sources

Free-text pathology reports are received electronically by the NICR with a certain degree of regularity (monthly) and completeness (around 85% of all cancer registrations), and kept in a separate repository. Specifically, they are held in MSWord files in a designated folder on a server—each file representing a calendar year worth of biopsy reports. Over 30,000 reports are received annually, of which around 900 are reports of prostate cancer and 350 of melanoma biopsies.

### Gleason scores from prostate cancer pathology reports

The first objective was to develop a means of extracting the Gleason score for prostate cancers from the main body of the text reports, using software techniques for keyword search and pattern matching. There is a degree of variation in how the Gleason score is recorded within the pathology report (see typical examples below), and thus, the technique employed for extraction would require some flexibility in order to yield the correct values. A manual survey demonstrated that tertiary scores were rarely reported in the histopathological documents of the chosen time period. As a result, tertiary Gleason scores were not routinely extracted.

Report 1

CLINICAL HISTORY:-
PSA 5.8 … Is this pure prostatic carcinoma or is it metastatic.
PATHOLOGIST'S REPORT:-
… with the prostatic markers, it is most likely that this is indeed a prostatic primary adenocarcinoma rather than metastatic carcinoma to the prostate. The tumour demonstrates **Gleason grades 4 + 4, Gleason score 8**.

Report 2

CLINICAL DETAILS
Carcinoma prostate. Re-do TURP.
PATHOLOGIST'S REPORT
… Histology again shows high grade adenocarcinoma of **Gleason score 8–10**. Tumour is present in virtually every tissue fragment. …

Report 3

CLINICAL DETAILS
Trucut biopsy of prostate. ? Ca prostate.
PATHOLOGIST'S REPORT
… Some of these fragments, including the fragments of skeletal muscle, are infiltrated by prostatic adenocarcinoma. … The **Gleason score is 7 (3 + 4)**.

Report 4

CLINICAL HISTORY:
TURP for chronic retention
Previous TURP
PATHOLOGIST'S REPORT:
… Histological examination shows in four of the chippings features of an invasive prostatic adenocarcinoma **Gleeson grade 3. Total score 6.** The remaining

prostatic tissue shows features of benign nodular hyperplasia …

## Technique for Gleason score extraction

A Microsoft Visual Basic routine was developed to scan the whole content of the training set of files containing 323,905 reports for the years 1993 to 2004. For each occurrence of the words "Gleason" or "Gleeson" (the latter is the common misspelling of "Gleason"), the routine created a new line in an Excel file, storing the file path, the Pathology Number of the patient and the found word plus the following 30 characters from the body of the report. Firstly, the records were quickly checked for problems (for example in some cases, the sought word was a surname) or normalisation (for example, it was found a need for the conversion of roman numerals or plain English numbers into arabic numerals). Secondly, the most common textual patterns, used by the clinician to record the Gleason grades (*G1* and *G2*) and total score (*S*), were identified with some possible variants. For example,

''Gleason score $G1 + G2 = S$''or

''Gleason score is $G1 + G2 = S$''or

''Gleason [...] (S) $G1 + G2$''etc.

The patterns identified were coded into *Perl* regular expressions [14], used in a routine which, for each pattern, extracts the relevant parts of the Gleason staging (grades and score) where occurring. For example, the *Perl* regular expression for the first two patterns previously mentioned is as follows

$$(\backslash d+)\backslash s*\backslash +\backslash s*(\backslash d+)\backslash s*=\backslash s*(\backslash d+)$$

For each record, all the extracted values were appended to it into a new file. The new file generated was manually inspected. Records for which null or wrong values had been extracted were analysed, in order to identify new patterns to be coded in *Perl* or patterns whose regular expressions had been miscoded.

The improved *Perl* routine was run again against the file containing all the original records and the process was repeated, until an acceptable level of retrieval was achieved without overcomplicating the *Perl* routine (16 different patterns were eventually needed).

The final routine was then run against a training subset of 2,263 pathology reports of prostate cancer patients, which were manually inspected and independently given a Gleason score by a urologist (David Connolly). The final output from this subset was analysed to calculate preliminary values for recall and precision:

$$\text{Precision} = \#\text{ correctly and completely extracted scores}/\#\text{ total extracted scores}$$

$$\text{Recall} = \#\text{ correctly and completely extracted scores}/\#\text{ total scores}$$

To further investigate the behaviour of the *Perl* routine, their recall and precision rates were calculated by performing the pattern matching over a variable number of characters extracted after the sought keyword.

Finally, the *Perl* routine was run against a 'gold standard' test set of 915 reports, (year 2005, disjoint from training set) manually scored in order to obtain unbiased values for precision and recall.

As an additional exercise, the Gleason scores extracted from the year 2001 reports were also compared with the values stored, for the same patients, in a subsidiary database of the Registry. Gleason scoring is not routinely collected in the NICR, and an audit project performed in 2004 provided information for prostate cancer patients diagnosed in 2001, in the form of an Access database manually populated by trained personnel looking at hospital charts. These patients, however, may have had more than one histological sample (e.g., a prostate biopsy followed by radical prostatectomy), and those reports were assessed separately.

## Technique for Clark level and Breslow depth

The same technique was used for the extraction of grading values (Breslow depth and Clark level) from skin biopsy and excision biopsy reports. The keywords were identified and fed, together with the following 50 characters, to the *Perl* routine. Below are a few examples of how the values are expressed in pathology reports (note that 'Clark' is always spelt 'Clarke'):

*Breslow depth*

…Breslow thickness is 1.2 mm. The melanoma should b…
…Breslow thickness is 4.0 millimetres. …
…Breslow thickness is 8.0 millimetres. …
…Breslow depth 0.34 mm. …
…Breslow's depth: 0.98 mm Mitoses/HPF: …

*Clark level*

…Clarke's level of 3 and a Breslow thickness of 1.56 mm. …
…Clarke's leve 1, Breslow thickness 0. The picture howeve…
…Clarke level 4. As such it comes into the high risk categ…

…Clarke's level II. This falls within the good prognostic c…
…CLARKE LEVELIV. INVASION: …

In this case, our aim was to estimate how valuable the technique would be to increase the proportion of skin cancers with a recorded grade in the Registry database. Thus, the values extracted from the free-text reports were compared with the values already stored in the Registry central databases and all differences analysed.

Given the uniformity of melanoma staging representation in the reports, recall and precision were estimated by manual inspection of the *Perl* output files, and no further analysis was needed. Specifically, the portions of text mentioning the grading value were inserted in an electronic sheet alongside the value extracted by the *Perl* script. These were manually assessed and the accuracy verified.

## Results

### Gleason score

When compared to manual examination, up to 93% recall was initially achieved during the training sessions, depending on settings. 90% of the missed automatic extractions were due to the presence of 'confusing' numbers following the actual score. To be safe, the *Perl* patterns used would not extract any Gleason score if other numbers follow, unless they can be easily disambiguated, e.g., tumour-nodes-metastasis (TNM) classification values [15]. In some isolated cases (10% of missed scores), the summation performed by the pathologist was wrong and, again, the routine would not extract those values.

The results of running the extraction routine over strings of increasing length, following the keyword, on the test set are presented in Fig. 1. It is shown that, in the case of the extraction of Gleason score, recall and precision follow a quasi-step function of the number of characters parsed. There are thresholds above which precision and recall can be considered quite stable, although slightly increasing with the number of characters.

The drop in precision when considering around 11 characters after the keyword is explained as follows. "Gleason pattern $x$" alone is used by pathologists as the short form of "Gleason score $x + x = 2x$", and the *Perl* routine performs the appropriate extraction and calculation. Sometimes, however, the pathologist writes "Gleason pattern[s] $a + b = c$", which would give misleading information to the routine when truncated at around 11 characters after 'Gleason'.

For the test session, a test set of 915 reports potentially containing Gleason score for prostate cancer cases was identified for the year 2005. The final recall and precision achieved on this set were 98.4 and 99.8%, respectively, which is very encouraging (see Table 1).

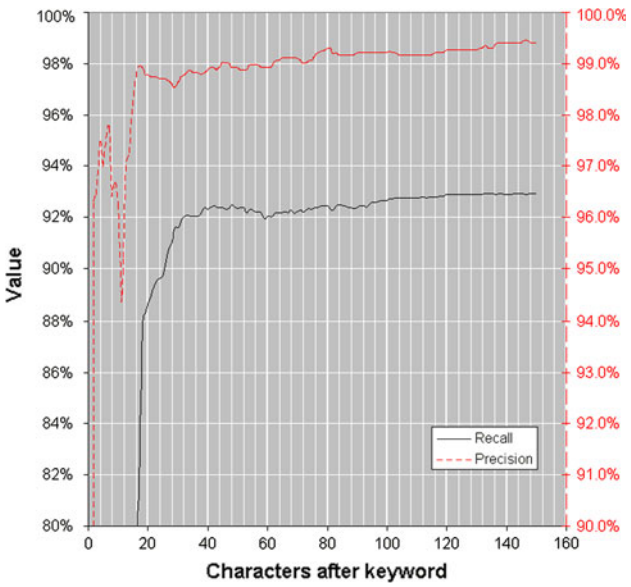


**Fig. 1** Recall and precision for pattern matching over string of increasing length after the keyword (training set)

**Table 1** Performance on manually graded test set (year 2005 reports)

| | Manually | *PERL* |
|---|---|---|
| Total extracted | 896 | 884 |
| Correctly extracted | | 882 |
| Recall | | 98.4% |
| Precision | | 99.8% |

**Table 2** Comparison of extracted Gleason scores with values manually collected during an audit (year 2001)

| | Manually | *PERL* |
|---|---|---|
| Total extracted | 282 | 276 |
| Total agreement | | 255 |
| Proportion | | 90.4% |

The comparison with the 2001 audit database is shown in Table 2. The disagreement of about 10% is due to a deficiency of the main registration system (mirrored in the audit database), which is designed to record only one grading score per patient, while grading values may evolve with the disease. The lower precision may be explained by patients who have more than one pathology report.

Finally, although the NICR does not routinely record the procedure employed to produce the histopathological tissue discussed in the report, manual checking indicated that lack of accuracy was not confined to reports associated to a

particular procedure, namely prostate needle biopsy versus transurethral resection of prostate.

### Clark and Breslow

The results for the mining of reports for malignant melanoma cancer cases are shown in Tables 3 and 4. In this case, a quick manual check for recall and precision was possible, because of the uniform way in which these values are recorded by the pathologists, and showed this uniformity allowed for final values of 100% recall and precision.

Melanoma staging values are routinely manually collected in the NICR, and a comparison with the values already stored in the Registry was possible. The results from the processing of 992 reports mentioning 'Clark' and 2,128 mentioning 'Breslow' were compared with the NICR database and, an estimate of the increase in completeness of melanoma staging data was also calculated, defined as the ratio between the number of melanoma records in the NICR database containing the information and the total number of melanoma records.

For Breslow and Clark scores, the majority (50–60%) of the non-matching values between the Registry and the reports are due to missing values in the NICR database. The set of results which are in real disagreement, around 10% of the whole set of reports, is again due to the presence of more than one report for the same patient. If the Breslow values are grouped into categories comparable to Clark levels [16], then this disagreement drops to 6.3%.

From the number of staging values extracted by our procedure, which were not already present in the NICR database, it turned out that our procedure increased the completeness of melanoma staging by 32 and 18% for Breslow depth and Clark level, respectively.

## Discussion

It is only in recent years that research of information extraction from clinical documents has reached acceptable levels, however, the application of this research into existing live systems is still very rare [17]. The aim of the current study was to assess the potential benefits of pattern-based information extraction to cancer registration and the challenges that such approach would pose. We found that over 90% recall and precision are achievable in the extraction of staging term-value(s) data, even when pathologists use several different ways of representing them. This compares very well with even much more complex techniques [17, 18]. In the case studied, the completeness of cancer staging was increased by up to a third. The approach illustrated is very easy to implement and adapt by individual registries or similar agencies, involves only a few trial-and-error cycles for the fine

**Table 3** Comparison of extracted Clark levels with values stored in the main Registry (all years)

| Clark level | | | | |
|---|---|---|---|---|
| Total no. of existing pathology reports | | | | 1,168 |
| Reports not linked to patients on NICR main system | | | | 176 |
| Values compared with NICR main system | | | | 992 |
| Comparison between values extracted from reports and values stored in the NICR | Real values | Values blank in both systems | Total agreement | Comments |
| Same values | 729 | 12 | 741 | (74.7% of total number of compared values) |
| Differing values | 251 | | | |
| Nothing extracted from reports | 19 | | | |
| Nothing stored on NICR main system | 153 | | | (61.0% of differing values) |
| Real different value | 79 | | | (8.0% of total number of compared values) |
| Number of Clark groups difference between reports and values stored in the NICR | | | | Number of reports |
| 1 | | | | 56 |
| 2 | | | | 17 |
| 3 | | | | 6 |

**Table 4** Comparison of extracted Breslow depths with values stored in the main Registry (all years)

| Breslow depth | | | | |
|---|---|---|---|---|
| Total no. of existing pathology reports | | | | 2,312 |
| Reports not linked to patients on NICR main system | | | | 184 |
| Values compared with NICR main system | | | | 2,128 |
| Comparison between values extracted from reports and values stored in the NICR | Real value | Values blank in both systems | Total agreement | Comments |
| Same value | 1,024 | 108 | 1,132 | (53.2% of total number of compared values) |
| Different value | 996 | | | |
| Nothing extracted from reports | 195 | | | |
| Nothing stored on NICR main system | 513 | | | (51.5% of differing values) |
| Real different value | 288 | | | (13.5% of total number of compared values) |
| Number of Clark-comparable groups difference between reports and values stored in the NICR | | | | Number of reports |
| 0 | | | | 154 |
| 1 | | | | 92 |
| 2 | | | | 31 |
| 3 | | | | 11 |

tuning of the regular expressions used and is now part of the data acquisition routine in the NICR.

We have used the same method in several other projects, for instance to differentiate the operative procedures performed from histopathology report data and to distinguish between metastatic and non-metastatic prostate cancer in isotope bone scan reports [19]. Even though the standardisation of pathology reporting is being encouraged, pathologists may be slow at fully embracing the practice and, in addition, there is generally an opportunity to add free-text on their reports to allow for narrative explanations and comments. The techniques illustrated here may be used to extract information from legacy documents, to quickly process the structured sections of new standardised reports—where these report have not been converted into electronic records at the source in the laboratory—and to extract potentially interesting information specified in the surviving free-text sections of the standardised reports.

The methods assessed in the current study concerned only three prognostic factors. Once the routines have been set up for the search of the keywords and the pattern-based extraction, it is very quick to extend the extraction to new items. However, at the present stage, our routines still need IT-aware personnel for most of the extraction cycle, the most specialised stage being the regular expression representations in *Perl*. Some familiarity is also needed with the relevant medical terminology, to be able to identify the linguistic constructs conveying the information being sought and most commonly used by the pathologists. This knowledge is then further refined and extended after some analysis of the corpus of reports, showing the 'live', actual usage of medical language by the pathologists. Another limitation of this method is represented by its reliance on reasonably well-defined search terms or expressions. For instance, an attempt to use a similar technique to extract the exact anatomical site from pathology reports of colorectal polyp tissue did not produce good results, while other pieces of information, such as grade, size and procedure, were successfully extracted. Additionally, some investigation is usually needed to determine the optimal number of characters in the proximity of the keyword to be submitted to the pattern-matching routine. Our tests have shown that this would be particularly important where the performance of the routine is non-monotonic. For instance, it could happen that when using complex patterns, which do not retrieve the sought value in the presence of 'noise'—such as other numbers—within the extracted string, the recall would decrease if too long strings are parsed and no disambiguation mechanisms are in place. To estimate the optimal number of characters

to be included in the pattern matching, it was necessary to run the extraction several times on a set of reports independently annotated by a human expert, varying the string length around the keyword between runs. Each run produces a set of extracted values, and the comparison of all sets with the human annotations is the source for calculating precision and recall as functions of the number of characters. This process was automated by a separate routine, but did require a small number of man hours to set up, test and run.

An additional benefit of this approach is that it enables one to rapidly structure a collection of unstructured pathology reports, which could be used for clustering and categorisation purposes. The building of indexes of the documents based on different prognostic factors, for example, would allow researchers to identify very quickly all pathology reports showing the value of a given factor within a given range. It would also allow identification of patterns of co-occurrence of abnormal values for sets of prognostic factors.

We also performed an informal estimate of the economic benefits of automatic extraction of prognostic factors from pathology reports. In the case of the test set of 915 reports potentially containing Gleason score, manual extraction was performed in about 30 person-hours, as opposed to around 4 person-hours for the modification and fine tuning of already existing extraction routines. The actual computer processing time required to perform the task was negligible in this context.

## Conclusions and future work

This study shows how simple pattern matching can be effectively used to improve completeness and accuracy of pathology information, at least in the domain of cancer registration, although our results can be easily generalised to other domains. In our specific case, it also pointed out that about 10% of patients are potentially associated with ambiguous grade for their tumour, because only one staging value is stored in the main DB while at least one pathology report shows a different value. The method used here, although very simple, shows that useful data can be extracted from free-text pathology reports with minimal effort on the part of staff. The next stage of our work will be the design of a generic tool for the extraction of term-value data which would be user friendly enough to require minimal or no intervention by IT personnel. Eventually, by expanding our search dictionary and enhancing the techniques, we hope to be able to reconstruct more complex information, such as metastatic/primary status or full staging of some tumours, and index the reports on this basis. Further experiments have also been performed on reports written in other languages (Italian and Spanish), producing comparable preliminary results.

## Availability

The *Perl* routines are freely downloadable from ftp://ftp.qub.ac.uk/pub/users/gnapolit/perl/

**Acknowledgments** The Northern Ireland Cancer Registry was funded by the Department of Health, Social Services and Public Safety Northern Ireland (DHSSPSNI), at the time this study was completed. It is now funded by the Public Health Agency. We also wish to thank Alejandra González Beltrán for her stimulating comments on this paper.

**Financial support** The Northern Ireland Cancer Registry was funded by the Department of Health, Social Services and Public Safety Northern Ireland (DHSSPSNI), at the time this study was completed. It is now funded by the Public Health Agency.

## References

1. Stevens R, Wroe C, Lord P, Goble C (2004) Ontologies in bioinformatics. In: Staab S, Studer R (eds) Handbook on ontologies. Springer, Berlin, pp 635–657
2. Health level 7. http://www.hl7.org/. Accessed Jan 2010
3. Systematized nomenclature of medicine. http://www.snomed.org/. Accessed Jan 2010
4. International classification of disease. ver. 10. http://www.who.int/classifications/icd/en/. Accessed Jan 2010
5. Collier N, Nazarenko A, Baud R, Ruch P (2006) Recent advances in natural language processing for biomedical applications. Int J Med Inform 75:413–417
6. Taira RK, Soderland SG, Jakobovits RM (2001) Automatic structuring of radiology free-text reports. Radiographics 21:237–245
7. Hotho A, Nürnberger A, Paaß G (2005) A brief survey of text mining. LDV Forum 20:19–62
8. Turchin A, Kolatkar NS, Grant RW, Makhni EC, Pendergrass ML, Einbinder JS (2006) Using regular expressions to abstract blood pressure and treatment intensification information from the text of physician notes. J Am Med Inform Assoc 13:691–695
9. Gleason DF (1977) The veteran's administration cooperative urologic research group: histologic grading and clinical staging of prostatic carcinoma. In: Tannenbaum M (ed) Urologic pathology: the prostate. Lea and Febiger, Philadelphia, pp 171–198
10. Clark WHJ, From L, Bernardino EA, Mihm MC (1969) The histogenesis and biological behavior of primary human malignant melanoma of the skin. Cancer Res 14:705–726
11. Breslow A (1970) Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. Ann Surg 172:902–908
12. NHS Information standards board, data standards: cancer registration data set, data set change notice (2005). http://www.connectingforhealth.nhs.uk/ dscn/dscn2005/092005.pdf
13. NHS connecting for health. http://www.connectingforhealth.nhs.uk/. Accessed Jan 2010
14. Friedl JEF (1997) Mastering regular expressions. O'Reilly & Associates, Cambridge (MA)
15. Sobin LH, Wittekind C (2002) UICC TNM classification of malignant tumours. Wiley-Liss, New York

16. SEER training modules, skin cancer: melanoma. U. S. National Institutes of Health, National Cancer Institute. http://training.seer.cancer.gov/melanoma/abstract-code-stage/staging.html. Accessed 19 July 2010
17. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF (2008) Extracting information from textual documents in the electronic health record: a review of recent research. Yearb Med Inform 2008:128–144
18. Coden A, Savova G, Sominsky I, Tanenblatt M, Masanz J, Schuler K, Cooper J, Guan W, de Groen PC (2009) Automatically extracting cancer disease characteristics from pathology reports into a disease knowledge representation model. J Biomed Inform 42:937–949
19. van Leeuwen PJ, Connolly D, Napolitano G, Gavin A, Schröder FH, Roobol MJ (2009) Metastasis-free survival in screen and clinical detected prostate cancer: a comparison between the European randomized study of screening for prostate cancer and Northern Ireland. J Urol 181(4)Suppl 1: 798