Final Report
Promoting and Expanding Structured Pathology Reporting of Cancer
June 2012
Introduction

This document is the final report as specified in clause 11.4 of the Standard Funding Agreement between the Department of Health and Ageing and The Royal College of Pathologists of Australasia dated Feb 2010 and addresses the elements specified in the above mentioned funding agreement for the Structured Pathology Reporting Standards for Cancer project and the Deed of Variation No. 3 dated January 2012.

FINAL report

11.4 (a) the Guidelines and Standards and the Aim of the Project and whether the Aim of the Project was achieved and, if not, why not.

The aims of the project, according to the Deed of Variation # 3 are to promote and expand the use of structured reporting of cancer by:

- undertaking a national program of education on the developed cancer protocols (breast, melanoma, lung, lymphoma, colorectal and prostate);
- developing further protocols in conjunction with international bodies;
- sharing the findings of the literature review, providing advice and feedback and participating as a member of the Collaborative Management Group and as a partner with the University of Melbourne and Dianella Community Health, on the project titled Effective Communication of Pathology Results to Requesting Practitioners and Consumers; and
- working collaboratively with Professor Jon Patrick at the University of Sydney to develop a model to automatically review protocol content and develop a baseline for the structured reporting of colorectal cancer.

The project has been successful in achieving these aims and continues to make excellent progress on all aspects of the project. A detailed review of project achievements and progress is described below under section 11.4(e).

11.4 (b) an audited detailed statement of receipts and expenditure in respect of the Funds prepared by an Approved Auditor in compliance with the Australian
Auditing Standards which must include a definitive statement as to whether the financial accounts are complete and accurate and a statement of the balance of the Funds in the bank account referred to in clause 10.1.

Please refer to Appendix A which is a statement of accounts for the project as at June 2012. A full audit is expected to be completed within 1 month and a full audited statement of accounts will be sent to DoHA at that time.

11.4 (c) a statement of how much (if any) the Participant needs from the final payment to meet current liabilities under legal commitments entered into by the Participant for the performance of this Agreement

The projected budget as included in the Funding Agreement is sufficient to meet current liabilities.

11.4 (d) a certificate provided by the Chief Executive Officer or Chief Financial Officer of the Participant, or a person authorised by the Participant to execute documents and legally bind it by their execution, confirming that:

(i) the Funds and Other Contributions received were spent for the purpose of the Project and in accordance with this Agreement and that the Participant has complied with this Agreement.

(ii) salaries and allowances paid to persons involved in the Project are in accordance with any applicable award or agreement in force under any relevant Law on industrial or workplace relations; and

(iii) at the time of the Final Report is provided to the Commonwealth, the Participant is able to pay all its debts as and when they fall due.

Please refer to Appendix B.

11.4 (e) any other requirements specified in Item D, which are as follows:

Item D

i. Report on all activities and outcomes that have been achieved throughout all stages of the project
**Education and Promotion**

The original six protocols published in February 2010 - Lung, Melanoma, Breast, Colorectal, Lymphoma and Prostate are posted to the RCPA website:


In addition, as of April 2012, another nine protocols have been published to this site after approval by the RCPA council:

- Renal Parenchymal Malignancy (Kidney)
- Endometrium Cancer
- Gastric Cancer
- Thyroid Cancer
- Soft Tissue Tumours
- CNS Tumours
- Testicular Tumours
- Oral Cancer
- Bladder Cancer

Each protocol is developed via a multi-disciplinary expert team and is then sent out for public consultation and approved via the RCPA Council before publication. A total of 22 authoring groups, residing under the 12 cancer committees who report to the Cancer Services Advisory Committee (CanSAC), have been convened totalling 164 experts.

After an initial development of the guide and form for the Melanoma protocol by a contracted Forms Designer, forms and guides were developed by the SPR project manager for each of the other published protocols and are now developed in conjunction with any new protocol. Hyperlinks have been added to the guides to provide quick, easy access to the information in the protocol. Forms and guides are published to the website with the protocols. Request information sheets based on the clinical information included in each protocol were developed more recently and are also published to the website.

Protocols are now in a continuous cycle of development with new protocols being kicked off at regular intervals. A procedure manual has been developed to document the publication process (attached as Appendix C).
Thirteen additional protocols are currently planned or in development:

- Gynaecology – vulva, cervix
- Gastrointestinal – oesophagus/gastro-oesophageal junction oesophageal ER; colorectal local resections
- Genitourinary – prostate – core biopsy
- Head, Neck and Endocrine – adrenal gland
- Bone and soft tissue – bone, GIST, salivary gland
- Haematopoietic – Bone Marrow biopsies
- Ophthalmic – SCC conjunctiva
- Paediatrics – Neuroblastoma

2nd editions of the first six protocols have been drafted. Where applicable eg Melanoma, Lung and Prostate the 2nd edition will include the internationally agreed elements from the International Collaboration on Cancer Reporting (ICCR) process (Refer to section on International Collaboration below).

A national education program on the initial six published protocols was conducted throughout 2010 and was based on a "train the trainer" approach. Champions were identified in each state and in New Zealand. These champions were flown to Sydney for an initial education event on Friday 16th April 2010. At this meeting A/Prof David Ellis provided a general background and overview on structured pathology reporting and then presentations were given by each of the chairs of the six protocol expert committees. Each presentation was videoed and made available on the structured pathology reporting website as well as via DVD.

Following this initial meeting, a series of meetings were conducted in each state at which presentations by A/Prof Ellis and the SPR project manager were given, and fliers, DVD’s, example protocols, guides and forms were distributed. Attendance at various conferences such as the IAP, NZ ASM, Trends, COSA and Pathology Update was also included in the education and promotion plan. Promotional material was included as a satchel inserts at IAP and fliers, forms, guides and DVD’s have been made available at the RCPA booths at the other venues.

Publication of new protocols is well advertised. Notification emails are sent to all stakeholders as well as fellows of the college and other interested parties. Additional notifications on the publications appear in Pathology Today and the Structured Pathology Reporting of Cancer newsletter.

Other promotional activities included:
• Publication of an article in the ANZMTG Australia and New Zealand Melanoma Trials Group newsletter in April 2010
• Publication of an article in Wongi Yabber the newsletter of the Australian Cancer Network in August 2010

The RCPA website for structured reporting protocols includes in addition to the published protocols, information in regard to the project, framework documents and release strategy, governance of the project; Frequently Asked Questions, compliance model, functional requirements, educational videos as well as the developed implementation aids (forms and guides).

A Structured Pathology Reporting of Cancer newsletter is sent out to all stakeholders and fellows of the college each quarter. The newsletter has proven to be a useful tool in communicating progress on the project and has been well received.

Authors of the protocols are encouraged to submit articles or other publications eg posters, on the evidence base for the protocols. Several articles have been published in Pathology, the official journal of the RCPA:

1. Prognostic factors in prostate cancer. Key elements in structured histopathology reporting of radical prostatectomy specimens.
2. Surgical pathology reporting at the crossroads: beyond synoptic reporting.
4. Optimising the management of soft tissue tumours

A poster for oral cancer was accepted for display at the Australian and New Zealand Head and Neck Cancer Society conference last September. Several additional articles are planned for the coming months.

A poster on Structured Pathology Reporting has been designed. The poster is aimed at pathologists and has been printed in 2 formats – a freestanding pull up banner and a 400x600mm version – both will be on display at IAP in June 2012.

**International collaboration**

Throughout 2010, the SPR project manager, A/Prof David Ellis and A/Prof Paul McKenzie engaged in discussion with Royal College of
Pathologists UK (RCPath), College of American Pathologists (CAP) and the Canadian Partnership Against Cancer (CPAC) about closer collaboration. This resulted in the signing of a Memorandum of Understanding (MOU) with CAP mid 2010 (see Appendix D). Further discussion on the benefits of collaboration lead to a teleconference between all four parties in October 2010. This was quickly followed by a face to face meeting on Nov 14th 2010 in Chicago. On February 27th 2011 a meeting of all four international parties was held in conjunction with United States & Canadian Academy of Pathology (USCAP) meeting, in San Antonio, Texas. A formal “Agreement to Collaborate” document was put forward by the RCPA SPR representatives to the international team which in simple terms outlined the goals and steps required to deliver an outcome from the collaboration. The result was a very significant step forward with all four parties signing the agreement and agreeing to work towards the standardization of cancer core data, metadata, naming conventions and value lists for all cancers, beginning with prostate, endometrium, melanoma and lung. (Refer to Appendix E)

The ICCR team comprises:

1. The Royal College of Pathologists UK (RCPath) - represented by Dr Lynn Hirschowitz, Pathologist and Chair of the Working Group for Cancer Services and Prof Michael Wells, Pathologist and Chair of the Histopathology Specialty Advisory Committee.

2. The College of American Pathologists (CAP) - represented by Dr Mary Kay Washington, Chair, Cancer Committee.

3. The Canadian Association of Pathologists (CAP-ACP) in association with the Canadian Partnership Against Cancer (CPAC) - represented by Prof John Srigley, Chair, National Pathology Standards Committee (CPAC).

4. The Royal College of Pathologists of Australasia (RCPA) - represented by A/Prof David Ellis; Pathologist Clinical lead, and Meagan Judge, Project Manager, Structured Pathology Reporting of Cancer project.

The ultimate goal of the collaboration is that cancer reports for all patients across the world will be of the same high quality content and format, underpinning cancer management at all levels, from individual patient care to international benchmarking and national cancer strategy.

Work on all four cancer datasets commenced in June 2011 with each country taking on the leadership of an international team working on the development of one cancer. The priority for each cancer specific
team was to agree the elements which are CORE. CORE elements are those which are essential for the clinical management, staging or prognosis of the cancer. Core was later updated to REQUIRED elements. Secondary to this aim was for each team to review and propose the value list or responses to the Required elements eg “present”, “not indicated” and then to establish a list of recommended elements. Assessment of evidence against the extended NH&MRC guidelines for evidence assessment (Merlin, T. et al) was part of the development process. Progress was assessed at a meeting of the ICCR at the August 2011 European Society of Pathologists (ESP) meeting in Helsinki.

The meeting of the ICCR in Helsinki was considered a ‘huge success’. Not only had all teams made significant progress but a number of important decisions were made as to how to proceed and a clear set of action items was developed including:

1. Preparation of a brief communiqué with key points to be used in communications, newsletters etc.
2. Preparation of a detailed report of the project (following normalisation of each approach)
3. Establishing a preferred and simplified process for further protocol development
4. Abstract submissions to USCAP, and to UICC 2012.
5. Submission of 4 cancer specific papers and one generic paper to lead clinical journals,
6. Further investigating and progressing our engagement with key International Cancer Organisations

Since the August meeting significant progress has been made:

- A format for the publication of the datasets was agreed and four draft international datasets have been developed – 2 are published to the following website:
  
  www.rcpa.edu.au/Publications/StructuredReporting/ICCR.htm

- The communiqué noted in point 1 above it also posted to the website, a second communiqué is in development

- The process for continued development has been agreed and is in being ‘tested’ through the work on the Kidney cancer dataset – see below for additional information.

- A publication on Endometrium Cancer has been completed and accepted for publication to the IJGP. Work on the Melanoma, lung and prostate articles has commenced.
- Harmonisation of terminology across the datasets has been an ongoing piece of work with significant progress being made.
- Seeking a long term governance model for the collaboration has also made significant progress and is described below:

The ICCR committee considered that the International Agency for Research on Cancer (IARC) was the best fit for work on international cancer datasets with their focus on cancer data. Looking at the organisational structure (http://www.iarc.fr/en/research-groups/org_chart.pdf) there are 2 groups of note: the Cancer Information (CIN) section which directs the collection of cancer data and hosts the IACR – International Association of Cancer Registries as well the section which manages the production of the WHO monographs (IARC Monographs (IMO)).

Calls with Helen Farrugia (Cancer Council Victoria) and David Roder (Cancer Australia, Cancer Council SA) were held and both agreed to recommend the project to David Forman head of the Cancer Information (CIN) section of IARC.

Ian Frazer, current Chairperson of the Scientific Advisory Council of IARC, was contacted at the recommendation of Ian Olver CEO, Cancer Australia. Ian was very supportive of the whole international cancer dataset idea and raised the matter with Chris Wild, Director of IARC.

A meeting was then held with Chris Wild, in Perth in November and he expressed great interest in the project and could see real opportunities in its implementation. The ICCR committee agreed to put forward a proposal to IARC that would include the completion (publication, open consult process) of the initial four cancer datasets under the IARC imprimatur as well as undertaking development of a new cancer dataset with an expanded ICCR to pilot the entire process. This proposal was developed and submitted to IARC for consideration. IARC have responded positively and a meeting in July with A/Prof David Ellis and Dr Chris Wild and David Forman of IARC is planned.

In order to expand the collaboration beyond the initial quadripartite group an approach to the European Society of Pathology (ESP) was made and has resulted in an invitation to present to the chairs of the working groups and advisory committee in Prague in September 2012.

As noted above, the ICCR agreed that the next cancer to be tackled internationally is kidney. This was considered an ideal choice in that it was a major point of discussion at the ISUP meeting in March in Vancouver and involvement of ISUP would make it a truly international effort. This process has commenced with the identification of an
expert group; research on kidney datasets undertaken and meetings planned.

Regular calls with the ICCR are held to agree next steps, discuss issues and ascertain progress.

**Implementation**

Laboratory Information Systems (LIS) in use in Australia are generally not currently capable of fully supporting structured pathology. Several areas of improving this situation have been explored:

1) To direct the type of development needed in our LIS to support structured reporting two documents have been produced:

   a) a SPR LIS Functional Requirements document which provides a background and discussion of general principles in relation to structured reporting eg standards and guidelines; and also describes specific functionality requirements eg “The pathologist must be notified and asked to complete those standards (mandatory fields) that have not been completed when attempting to verify the report (signout)”. The functional requirements are divided into mandatory and recommended items.

   b) a Universal Design Requirements document which contains a set of design principles which should govern all LIS development eg “Nothing should ever require human entry, which the computer should be capable of synthesizing from known information.”

   These documents were circulated and reviewed at a series of meetings with LIS vendors and other interested parties in 2011. Feedback from the review was incorporated into the documents. The documents will provide assistance to those acquiring new Laboratory Information Systems and also inform the LIS vendors of what functionality is required to fully support structured Pathology reporting.

   The documents are posted to:


2) The project developed an electronic form (e-form) which worked very well and was very useful for educational purposes however it is difficult to maintain for widespread/ every day use by the pathology community.
3) In conjunction with the development of the first six structured reporting protocols for cancer, NeHTA commissioned Ocean Informatics to develop archetypes for each of the protocols for the purpose of enabling HL7 messaging between applications. Ocean Informatics worked in parallel to the protocol development and in October 2009 the archetypes were at the stage to have key clinicians review them to ensure that the intent and logic of the protocols had been preserved. However, the review was very limited as we were unable to engage the protocol authors in reviewing the archetypes in CKM satisfactorily. CKM proved quite challenging and time consuming for the pathologist authors of the protocols, who were already time poor.

4) NEHTA recently adopted Ocean’s CKM for the development of its DCMs and templates. Following this, discussions with NEHTA were held regarding the potential local development of archetypes under the oversight of NEHTA. Local development of the archetypes was hoped to eliminate the complexity of review and issues introduced by third party development. The DCM tools were also hoped to provide a means to develop XMLs for implementation and facilitate HL7 messaging.

XML’s were initially raised by LIS vendors during the functional requirements reviews – see point 1 above. Vendors asked if it was our intention to provide XML’s to facilitate implementation. Most laboratory systems being implemented in Australia are American and the vendors are therefore familiar with the process of the CAP XMLs. Further investigation provided the information that US and Canadian laboratories implement their cancer checklist via XML’s distributed from the College of American Pathologists (CAP).

After some education in the tool it was determined that the level of expertise required to develop an archetype locally is quite prohibitive; resources able to do this in Australia extremely limited and review of the resulted archetype very challenging.

5) Discussions with the College of American Pathologists – SNOMED Technology Solutions (CAP STS) were undertaken to investigate the process that CAP use. CAP STS have a very simple tool to edit from the paper checklist to an electronic checklist (eCC) which includes the construct of a simple information model. From this, XMLs are created and when ready, these are distributed to vendors for installation – this ensures a level of conformity with the checklist that individual implementations at a local level from a paper model
will not achieve. The XML also includes terminology mapping such as SNOMED CT, ICDO etc.

As a result of these discussions a proposal was developed in which CAP STS have proposed to undertake development of XML’s for the RCPA Protocols. (Refer to Appendix F)

6) Cancer Control New Zealand (CCNZ) are revamping their cancer registry system and including a wider range of information such as the TNM staging. This additional range of information will be possible only if the RCPA protocols are adopted and therefore CCNZ is mandating their use. To facilitate implementation of the protocols they are progressing development of an e-form (similar to that noted in point 2 above). The beta test copy of the RCPA e-form was distributed in late 2010 and this is the version that CCNZ have had recommended by their pathologists. The SPR project has proposed a collaboration agreement to work on the development of the e-form together with CCNZ.

7) The project has also investigated the potential to pilot the implementation of structured reporting and messaging of cancer data from a laboratory to a registry. Cancer Council Victoria are extremely keen to progress this, however laboratory implementation has been slow to progress.

8) Another challenge to successful implementation is getting the macroscopic recorded as structured information as most macroscopics are dictated and will continue to be dictated. Therefore in order to encourage people to still record the data as structured, a web page has been designed which is designed to use a minimum of key touches either from an iPad or PC/laptop. The person accesses the macroscopic webpage and then clicks on the icon representing the cancer they are interested in. The checklist for the macroscopic from the specific structured pathology reporting protocol will open and provides all the information the pathologist needs to record (via dictation) during cutup. There is a word template which matches each of these checklists for the typist posted to the cancer protocol website. Help is available from the checklist via a question mark icon (?) including specimen handling information. The draft website can be accessed via the following link.

Compliance

To progress adoption of the protocols across Australia, a request to NPAAC was submitted requesting consideration of the development of a standard on the reporting of cancer against which laboratories would be accredited. A compliance matrix (available from the link below) has been developed to describe the significant stages of development that laboratories must transit to achieve a fully structured model.


Compliance at level 3 was requested as this will start laboratories on the process and to achieve the benefits of improved completeness of cancer reports. This request was discussed at the October NPAAC meeting. Following on from this in December there was a recommendation from the college to include “Structured pathology reporting, especially where consensus report templates are available, should be used wherever possible” after CC6.1 on page 16 of the draft Requirements for Medical Pathology Services Tier 2 document. A tier 3 document specific to Anatomical Pathology is planned and will describe recommendations on structured pathology reporting. Ratification of these documents will result in the mandating of structured pathology reporting of cancer.

Terminology

The SPR project is also working with the Pathology and Units Terminology Standardisation Project (PUTS) to assess what terminologies are being used elsewhere in the world for structured pathology reporting protocols, as well as to undertake an initial exercise to assign LOINC and SNOMED CT codes to an example protocol. LOINC codes have been assigned to the 2nd edition of Melanoma and this will be circulated for review.

Report Format Project/Literature Review

The MOU between University of Melbourne, Dianella and RCPA has progressed with the agreement of a Memorandum of Understanding (MOU) which was signed February 2012 (See Appendix G).

The first meeting of the steering committee was held on 19th October 2011.
A literature review was undertaken by the SPR Project Manager to investigate published evidence in relation to communicating report information with the aim of enhancing the understanding and readability of cancer reports. The literature review was completed in June 2011 and is attached as Appendix H.

Following on from the initial steering committee meeting in October several report prototypes were sent for review at the end of November 2011. These report prototypes were intended to be used with focus groups to seek feedback as to which style of report they prefer and how understandable they are to them. The college circulated the prototypes to a small group of pathologists (given the short timeframe) and comment was provided back to the University of Melbourne.

At this stage, the results of Dianella’s patient surveys are being compiled so that a final version of the proformas can be developed.

The next step is to review the proformas with relevant pathologists, the RCPA will assist in this process.

**NLP Colorectal Cancer reports**

A contract with Professor Jon Patrick of Health Language Laboratories (HLL) was signed in January 2012 (Refer to Appendix I). The objectives of the project were to:

1) Create an automatic processor that populates structured reports for colorectal cancer from sample anatomical pathology reports.

2) Assess the extent to which current colorectal pathology reports are already using the elements of structured reports. The rate of use should be reported as the percentage of the report(s) presented in a structured format ie question and answer format.

3) Assess the level of completeness of structured reporting for colorectal cancer. Completeness should be reported as a quantitative measure of adherence to the (1) standards only and (2) guidelines in the Colorectal Cancer Structured Pathology Reporting protocols 1st edition.

The stages in this project are

1. Acquire a corpus of 400 colorectal cancer reports.
2. Design a Tagset for annotating the reports.
3. Build a mapping between the Tagset and the Structured report fields.
4. Train staff in the use of the tagset.
5. Annotate the corpus using the tagset.
6. Build a machine learning model for computing the annotation tags;
7. Run a series of evaluations on different models;
8. Develop the programme code to extract the descriptive statistics;
9. Project report writing

38 reports (7 word documents and 31 pdf documents) were submitted from pathologists as a result of a request by QAP and through Pathology Today and SPR newsletters. These were supplied to Prof Patrick in Jan 2012 to initiate the project.

Reports were requested from Cancer Institute NSW (CI NSW) and Cancer Council Victoria for large numbers of colorectal cancer reports. Unfortunately the CI NSW had a problem supplying de-identified reports so the SPR Project Manager volunteered to undertake the de-identification process. CI NSW required ethics approval before this could be undertaken, therefore the ethics applications and supporting documentation were written and submitted in December 2011 in anticipation of signing of the contract with Prof Patrick.

An issue was reported by Prof Patrick in early February in that he was not able to use any scanned reports but required the reports to be in a txt or excel format. A specific format was not included in the contract nor in any pre-contract conversations. Fortunately Cancer Council Victoria was able to deliver 270 reports in an excel format so that the project could commence. Another 74 reports from Cancer Council Victoria sent to Prof Patrick were pdf or jpeg type documents. It was agreed that Prof Patrick would contract for these and the initial 31 reports to be OCR’d so that they could be used.

Cancer Institute NSW was contacted and clarification was requested on whether they were able to send any reports in an electronic format. They replied to say that they only store the extracted details they need in an electronic format and the whole of the report is only available in a scanned version. Therefore no reports were able to be obtained without considerable additional cost from Cancer Institute NSW.

Cancer Council WA was also contacted in hopes that as they had a more advanced cancer registry system they may be able to supply additional reports in the required format however they were unable to
do this in the required format as well. Cancer Control NZ were also contacted and were unable to supply reports in the required format. Cancer Council Victoria was asked to supply additional reports to make up the shortfall and a further 362 additional reports have been supplied to Prof Patrick.

Prof Patrick’s first progress report is attached as Appendix J. Stages 1-6 are complete.

ii. **Provide a report on the project with the Health Language Laboratories of the University of Sydney for Prof Jon Patrick to build a clinical language processor engine for colorectal cancer and the audit of approximately 400 cases.**

See above section 11.4(b) NLP Colorectal Cancer reports

iii. **Provide a summary of feedback on the uptake and use of the protocols where available**

There are 9 stages to the contracted audit process with Prof Patrick. Stages 1-6 are complete. The report on uptake and use of structured reporting is not available at this time – this will be completed as stage 9.

Prof Patrick’s first progress report is attached as Appendix X.

iv. **Provide a review of all tasks and their current stage of development in this project.**

See section 11.4(e) (i)
Structured Pathology Reporting of Cancer Protocols
Procedure Manual
Update to existing protocol

1. Compile feedback on the individual protocol into the feedback excel. This may be via:
   a. Form from web (see process below)
   b. Emails
   c. Notes from calls etc
2. Edit and include track changes in the protocol. Note any changes required for updates to copyright. Review from informatics approach eg changes to promote use of atomic responses over narrative etc.
3. Review by DE re any outstanding questions/concerns
4. Send out to lead author and chair for preliminary review and address outstanding questions. (See appendix A)
5. Update draft protocol and example report
6. Send out to lead author/chair for review and to address outstanding questions
7. Update draft protocol
8. Send out to protocol expert group for review
9. Incorporate any track changes/comments
10. Review and finalise with lead author /chair
11. Then follow publication process

New protocol

1. Format initial draft and resolve issues with lead author. Note any requirements for diagrams/copyright
2. Organise diagrams and copyrights – refer to SPR COPYRIGHT PERMISSIONS excel.
3. Develop request information sheet in Indesign – include in Appendix 1 of the protocol
4. Update draft protocol and example report – send to expert committee for review
5. Incorporate any track changes/comments
6. Send out to lead author/chair for review and to address outstanding questions
7. Update draft protocol
8. Send out to protocol expert group for final review
9. Review and finalise with lead author /chair
10. Then follow open consultation process

Open consultation process

1. Check open consultation review form on web has protocol listed in drop down. See feedback form notes Appendix C
2. Include draft in watermark on protocol
3. PDF protocol
4. Post to website under ‘initialreview’
5. Email APAC and CanSAC and independent review group requesting review with specific deadline. See Appendix D for example email.
6. Include feedback from Initial review into feedback xls.
7. Discuss with lead author/chair and resolve any changes – include responses in xls
8. Post finalised protocol and PDF
9. Post to website under ‘publicconsultation’ with 28 days expiry listed on the site
10. Send out specific notification emails (see example in Appendix D)
   o All general stakeholders (stakeholder xls) use previous notifications as a guide. Note on stakeholder xls who has been sent a notification.
   o Specific cancer related stakeholders
   o All AP (and Haem if required) fellows (from IMIS, IQA, ALL only fellows – sort by discipline.)
   o All expert committees
   o CAP representative
11. Post notification to Path Today and SPR newsletter
12. After open consultation, compile responses from submitted forms (See Appendix C on how to process forms) and send to lead author/chair.
13. Agree changes and update feedback xls and protocol
14. Send out to expert committee for final review.
15. Then follow publication process.

Publication process
1. Develop/update guide in Indesign – (see Appendix E for hyperlinking guides)
2. Develop/update form in Indesign
3. Update request information sheet in Indesign
4. Develop a macroscopic icon for webpage and both the word template and PDF.
5. Review and acquire any remaining copyrights - make sure SPR Copyright xls is updated with TRIM nbrs.
6. Check AJCC errata
7. Update feedback xls and include responses
8. Review and update protocol with any updates from ‘Record of updates” file (which contain framework updates)
9. Acquire publication numbers – Cancer Institute organise this for us – Request via Narelle Grayson.
10. Final audit – dates/page numbers, update Table of contents, edition number /publication date etc page formatting, remove DRAFT from watermark. Gridlines in Ch 6 (Grey ¼ pt),
11. Check example report, Appendix 1.
12. Unlink citations, save and PDF
13. Check the feedback form on web has protocol listed in drop down. See feedback form update notes Appendix C
14. Post protocol, guide, request information, macro pdf and macro word template, form and feedback Excel to web.
15. Notify ISBN people of web address for protocol:
   Corinne Stolzenhein, Corporate Librarian
   NSW HEALTH,
   73 Miller Street, North Sydney NSW 2060
   LMB 961 North Sydney NSW 2059
   Tel: 02 9391 9078
   Email: cstol@doh.health.nsw.gov.au
16. Post notification to Path Today and SPR newsletter
17. Send out email notification to all stakeholders – see example Appendix D
   a. All general stakeholders (stakeholder xls)
   b. Specific cancer related stakeholders (stakeholder xls)
   c. All AP (and Haem if required) fellows and trainees (from IMIS, IQA, All trainees and ALL only fellows – sort by discipline.)
   d. All expert committees
# Appendix A – Chairs/Lead authors

## Chairs

<table>
<thead>
<tr>
<th>Specialty</th>
<th>Chair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gynaecological</td>
<td>Dr Nicholas Mulvany (VIC)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>Prof Bob Eckstein (SYD)</td>
</tr>
<tr>
<td>Neurological</td>
<td>Dr Michael Rodriguez (SYD)</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>Prof James Kench (SYD)</td>
</tr>
<tr>
<td>Head/Neck &amp; Endocrine</td>
<td>Prof Jane Dahlstrom (ACT)</td>
</tr>
<tr>
<td>Bone and soft tissue</td>
<td>Dr Chris Hemmings (ACT)</td>
</tr>
<tr>
<td>Pulmonary and Mediastinum</td>
<td>Dr Jenny Ma Wyatt (WA)</td>
</tr>
<tr>
<td>Skin and adnexal</td>
<td>Prof Richard Scolyer (SYD)</td>
</tr>
<tr>
<td>Haematolymphoid</td>
<td>Dr Debbie Norris (QLD)</td>
</tr>
<tr>
<td>Paediatric</td>
<td>Dr Susan Arbuckle (NSW)</td>
</tr>
<tr>
<td>Ophthalmic</td>
<td>Dr Sonja Klebe (SA)</td>
</tr>
<tr>
<td>Breast</td>
<td>Dr Gelareh Farshid (SA)</td>
</tr>
</tbody>
</table>

## Lead authors

### Published

<table>
<thead>
<tr>
<th>Specialty</th>
<th>Leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometrium</td>
<td>Dr Nicholas Mulvany</td>
</tr>
<tr>
<td>Vulva</td>
<td>A/Prof Jim Scurry</td>
</tr>
<tr>
<td>Colorectal</td>
<td>A/Prof Bob Eckstein</td>
</tr>
<tr>
<td>Gastric</td>
<td>A/Prof Priyanthi Kumarasinghe</td>
</tr>
<tr>
<td>CNS tumours</td>
<td>Dr Michael Rodriguez</td>
</tr>
<tr>
<td>Prostate (Radical Prostatectomy)</td>
<td>Prof James Kench</td>
</tr>
<tr>
<td>Renal Parenchymal Malignancy</td>
<td>Prof Brett Delahunt</td>
</tr>
<tr>
<td>Thyroid</td>
<td>Prof Alfred Lam</td>
</tr>
<tr>
<td>Soft Tissue</td>
<td>Dr Chris Hemmings</td>
</tr>
<tr>
<td>Lung</td>
<td>Dr Jenny Ma Wyatt</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Prof Richard Scolyer</td>
</tr>
<tr>
<td>Breast</td>
<td>Dr Gelareh Farshid</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Dr Debbie Norris</td>
</tr>
</tbody>
</table>

### Not published

<table>
<thead>
<tr>
<th>Specialty</th>
<th>Leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Cancer</td>
<td>A/Prof Hedley Coleman</td>
</tr>
<tr>
<td>Prostate (Core Bx)</td>
<td>Prof James Kench</td>
</tr>
<tr>
<td>Testis</td>
<td>Dr David Clouston</td>
</tr>
<tr>
<td>Cervix</td>
<td>Dr Marsali Newman, Dr Kerryn Ireland-Jenkin,</td>
</tr>
<tr>
<td>SCC Conjunctiva</td>
<td>Dr Sonja Klebe</td>
</tr>
<tr>
<td>Bone</td>
<td>Dr Chris Hemmings</td>
</tr>
<tr>
<td>GIST</td>
<td>Dr Chris Hemmings, A/Prof Bob Eckstein</td>
</tr>
<tr>
<td>Bladder</td>
<td>A/Prof Hemamali Samaratunga</td>
</tr>
<tr>
<td>Colorectal polyps</td>
<td>A/Prof Bob Eckstein</td>
</tr>
<tr>
<td>Adrenal</td>
<td>Prof Alfred Lam</td>
</tr>
<tr>
<td>Oesophagus</td>
<td>A/Prof Priyanthi Kumarasinghe</td>
</tr>
<tr>
<td>Endoscopic resection (Gastric)</td>
<td>A/Prof Priyanthi Kumarasinghe</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>Dr Amanda Charlton</td>
</tr>
</tbody>
</table>
Appendix B  - Independent review group

Dr David Moffat (SA) david.moffat@imvs.sa.gov.au

Dr Angela Chong (Singapore) pychong1@gmail.com

Dr Cecily Metcalfe (WA) Cecily.Metcalf@health.wa.gov.au

Dr John Roberts (NSW) jcr@vicdoctor.com.au; tishomingo2@bigpond.com

Dr Sharon Wallace (VIC) Sharon.wallace@sjog.org.au
Appendix C – Feedback form

To edit the open consultation review form or the feedback form

1. Open the open consultation review form or the feedback form.
2. Go to FORM menu and edit forms in designer.

3. Highlight the protocol title field and on the right hand side edit the names of the protocols.
4. Save and exit – distribute the file but save a copy to the folder. This comes in as “file_distributed” and another called “file_responses”.
5. Open the original file again and go to advanced folder and click on extend features in adobe reader. This is to re-enable usage right after editing. Save to the same name.
6. Post this to web.

Compiling responses

1. As forms are completed save in protocol/web folder under initial review and OC or other applicable folder.
2. Open the file_responses file.
3. Click get started
4. Click add (on left navigator and then browse to add in completed files previously saved.
5. Then highlighted all the lines added and export to csv file.
7. Add in responses column and reformat to include all comments down the page as per feedback excel.
8. Add in another tab for any new protocol and copy in the feedback
9. Once updated and finalised the updated feedback excel can be posted to the web site.
Appendix D – example emails

**To APAC and CanSAC for initial review:**

Dear APAC and CanSAC members,

The structured reporting protocol for Gastric Cancer is now ready for your review prior to being posted for a period of public consultation.

Could you please undertake a review of the protocol in regard to both...

a) its fitness for purpose, general structure and approach, and
b) the number of standards included to ensure that only the critical elements pertaining to prognosis or treatment are included as standards.

Could I have your replies by **19th November 2010**.

The protocol has been published to a webpage – please click on the link below to access the protocol:


Thankyou,

Meagan

---

**Public Consultation notifications**

Dear Andrew,

This email is to notify you that the Soft Tissue tumours, CNS tumours and Thyroid Cancer protocols, developed as part of the structured pathology reporting project to establish protocols for different cancers, are available for review and feedback.

The protocols are available via the RCPA website as follows:


As a representative of COSA could you please notify other interested members of your organisation that these protocols are available for review?

There will be a 28 days review period and then the protocols will be removed from the website. The end date is posted on the site. Feedback on the protocols will be via an electronic form the website.

If you have any questions or have any issues accessing the protocols or completing the online form, please contact me on the details below.

**General notification re publication**
NEW STRUCTURED PATHOLOGY REPORTING PROTOCOLS AVAILABLE

On 3rd November 2011, the RCPA Council endorsed another two cancer specific structured pathology reporting protocols:

  - Testicular tumours
  - Oral cancer

This means we now have 14 cancer specific protocols available. Visit the link below to download a protocol or for easy access to the information in the protocol, download a hyperlinked guide instead.


Remember your feedback is very important to improving the protocols so if you have any thoughts or comments on these new protocols, or any of our other protocols, please visit the following link below and send in your feedback.

www.rcpa.edu.au/Publications/StructuredReporting/feedback.htm

For more information on the Structured Pathology Reporting of Cancer project please phone or email me,

Regards

Meagan
Appendix E – Guides

1. Standards – verdana 8pt bold and guidelines = verdana 8pt regular
2. PDF all standards and guidelines in the guide. File names must not include spaces. Include any reference info eg diagrams or WHO in Appendix. Name each pdf as follows: PX_S1_03 or PX_G1_11 etc where PX are as indicated below:
   - P01 Melanoma
   - P02 Colorectal
   - P03 Lung
   - P04 Prostate Rad Pros.
   - P05 Lymphoma
   - P06 CNS
   - P07 Endometrium
   - P08 Thyroid
   - P09 Gastric
   - P10 Soft Tissue
   - P11 Kidney
   - P12 Oral
   - P13 Testis
   - P14 Breast
   - P15 Bladder

   PDF specimen handling stds and guidelines or other commentary applicable at the section (ie macro, micro etc) using PX_Micro_note or PX_specimen_handling

3. Crop to appropriate size.
4. Upload the pdf files to Juice CMS (web editor) in structured reporting folder and publish.
5. In InDesign in the guide - highlight standard or guideline.
6. Go to Type menu and then hyperlinks & cross-references
7. Click on new hyperlink:
8. Unclick the shared hyperlink destination
9. Include url eg for 1st edition:
   (make sure there is no space after .pdf)
   or 2nd edition (add in edition number)
   http://www.rcpa.edu.au//static/File/Asset%20library/public%20documents/Publications/StructuredReporting/P11_e2_S3_01.pdf

10. Change the style to hyperlink or hyperlink bold or hyperlink white (depending on whether it is standard or guideline or other notes/commentary) (characteristics: underline offset 2pt, underline weight is 0.5pt, colour hyperlink blue or white)
11. To edit the url you need to delete the Standard or guideline number and start again.
12. When exporting to pdf – click “include hyperlinks”.

Macroscopics

1. Find or create an appropriate icon and include on the website – set the image border to 0.
2. PDF – include help icons – see help icon document for ‘blank’ versions. (note you cannot redo the hyperlink on the image once a hyperlink has been assigned you need to use a blank)
3. Set PDF as follows: File, properties, initial view 100%, full screen mode.
4. Crop pdf from bottom 0.625in
ICCR Guides.

1. Required are in verdana 8pt bold and recommended = 50% black bold verdana 8pt
2. Icon for hyperlinks - use the one in hyperlink icon blank form (otherwise will not be able to add in a hyperlink)
3. Highlight at smallest level and then add in url as above.
4. Notes – on required use “guideline notes” style; notes on recommended use “recommended unbold note”. (Colour hyperlink blue – no underline 8pt).
5. On notes version – pdf then add in links (right click and link) then add link to Back as well.
Literature Review

“Enhancing the understanding and readability of cancer reports”
## Document History

<table>
<thead>
<tr>
<th>Version</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 0.1</td>
<td>Initial draft</td>
<td>March 2011</td>
</tr>
<tr>
<td>Version 0.2</td>
<td>Review by SPR Project Group</td>
<td>May 2011</td>
</tr>
<tr>
<td>Version 1</td>
<td>Submitted to DoHA</td>
<td>June 2011</td>
</tr>
</tbody>
</table>
# Table of Contents

Introduction ................................................................................................................................ 4

Search criteria/methodology ........................................................................................................ 5

Results ......................................................................................................................................... 6

  Formatting Pathology Reports ..................................................................................................... 6
  Review of Existing Standards ....................................................................................................... 7
  Readability ..................................................................................................................................... 7
  Standardisation ............................................................................................................................ 8
  Topography/Legibility .................................................................................................................. 10
  Eye Tracking ................................................................................................................................ 13

Summary ................................................................................................................................... 16

References ................................................................................................................................... 17
Introduction

Cancer is a growth industry. The incidence rate for all cancers combined has increased by 27% since 1982\(^1\) with the number of people being diagnosed with cancer projected to increase by over 3,000 extra cases per year as a result of the ageing population.

The pathology report is the foundation of a patient’s cancer journey. Most significant treatment decisions are based on the pathology report which:

- provides the definitive diagnosis
- includes information necessary for staging,
- evaluates the adequacy of the surgical excision,
- gives a determination of appropriate chemotherapy and/or radiation therapy
- and identifies important prognostic indicators.

The magnitude of the surgical intervention and the extent and type of treatment is all based on the pathology report details. It is the pathology report that dictates management and prognosis by detailing the disease process. In addition, the pathology report is the most significant source of information for cancer registries providing information for cancer care planning and resource distribution at a national level.

A pathologist’s role is to provide information. The pathology report is the one forum in which they are able to document and communicate that information. The challenge faced by pathologists, is to create a document that is simple to read, streamlined, comprehensive, useful, and that contains sufficient documentation to meet all requirements. However, the style in which this data is presented eg layout, font, tables, is often constructed solely for aesthetic and marketing (branding) purposes with little regard to its impact on readability and comprehension.

The aim of this literature review is to identify the evidence relating to the style (text layout) of pathology reports and the impact that this has on enhancing communication, understanding and readability. While the focus is cancer reports, by necessity this evaluates the form/layout of pathology reports in general and in a wider context, other healthcare documentation.
## Search criteria/methodology

The criteria for the literature review were determined as follows using the setting population intervention comparison and evaluation (SPICE) approach:

<table>
<thead>
<tr>
<th>Category</th>
<th>Response</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>setting</td>
<td>pathology; healthcare</td>
<td>pathology, healthcare, medical</td>
</tr>
<tr>
<td>population</td>
<td>medical/scientifically capable readers</td>
<td>pathologists, oncologists, clinicians</td>
</tr>
<tr>
<td>intervention</td>
<td>format/layout</td>
<td>format, layout, design, configuration, user views, aesthetic, typography, legibility</td>
</tr>
<tr>
<td>comparison</td>
<td>compared with existing report formats</td>
<td>patient reports; pathology reports; cancer reports; medical reports</td>
</tr>
<tr>
<td>evaluation</td>
<td>speed to acquire information &amp; accuracy/comprehension</td>
<td>readability, accuracy, speed, comprehension, communication, eye tracking, eye movement, cognitive psychology</td>
</tr>
</tbody>
</table>

A wide variety of PUBMed, SCIRUS, and GOOGLE search strategies were undertaken using the combinations of keywords above. Further review of likely sources of information was pursued from bibliography in the above searches.

In addition, specific review of bibliography in the following two articles was undertaken:


Note: Non-English publications were excluded from this literature review.
Results

No comprehensive reviews into style (text layout) on readability and comprehension of pathology reports in general, and anatomical pathology reports specifically, were identified from the literature.

The results of this literature review have been divided into topics suggested by the literature itself and the areas they have explored.

Formatting Pathology Reports

There are two publications which discuss the impact of formatting of anatomical pathology reports. The first is a paper Formatting Pathology Reports: Applying Four Design Principles to Improve Communication and Patient Safety, by Valenstein in which he describes four design principles for pathology reports.

1. Using headlines to emphasise key findings
2. Maintain layout continuity
3. Optimise information density
4. Reduce extraneous data or clutter

A structured report lends itself to all of these principles however the study makes no other specific recommendations on style.

The second investigation into surgical pathology report formats is by Powsner et al in 2000. They found that there is a significant difference in clinicians understanding of the content of a pathological report based on the physical presentation of the information in the report. The participants compared existing formatted reports with a new streamlined report format with changes including spacing, highlighting, positioning of information, and font selection. They found that these items, did not in and of themselves, contribute to the content of the report; however, they did appear to contribute substantially to the comprehension of that report, despite nearly identical wording.

Powsner et al also found that changing the style (text layout) actually increased the error-rate which they postulate as “a negative short-term effect on .. readers (ie report disorientation).” In effect this finding demonstrating the impact of style (text layout) on the comprehension of documents. However the ‘streamlined’ format used in this study was based largely on aesthetic qualities preferred by the authors rather than any validated recommendations as “essentially no empirical literature exists on pathology report comprehension to provide a guide for improvement.”

The format in which pathology data is presented eg layout, font, tables, is often constructed solely for aesthetic and marketing (branding) purposes. Reports formats are generally perceived as simply a vehicle for information, inherently conveying a message as to the quality of work performed but with little regard to its impact on readability and comprehension by the clinician.
Ruby⁴ states that “reports are our product and effective communication is our connection with our clinical colleagues”. Have we missed a vital component in our communication?

**Review of Existing Pathology Reporting Standards**

The National Pathology Accreditation Advisory Council (NPAAC) which is managed by the Australian Government Department of Health and Ageing plays a key role in ensuring the quality of Australian pathology services and is responsible for the development and maintenance of standards and guidelines for pathology practices. In regard to reporting NPAAC have a single standard “Reports commensurate with good medical laboratory practice and patient care must be provided to the requesting practitioner.”⁵ NPAAC also includes standards for information communication but this is concerned with report messaging, identification, security and privacy rather than specific formatting concerns.

Neither the Australia Standard 4700.2-2007 “Implementation of Health Level Seven (HL7) Version 2.4 - Pathology and medical imaging (diagnostics)” nor the Handbook HB 262 “Guidelines for pathology messaging between pathology providers and health service providers” provide any information or recommendations on the style (text layout) of pathology reports.

**Readability**

Much of the early research on readability has focused on developing tools for objectively assessing the readability of texts in order to match them to the reading skills of readers. This effort has resulted in the development of many *readability formulas*, which have been widely used. The Flesch-Kincaid formula (1975) is one of the most popular and heavily tested formulas and it correlates 0.91 with comprehension as measured by reading tests.⁶ There are many individual articles evaluating the readability of specific documents using this formulae as well as the Flesch Reading Ease score (1948), FOG index (1952), Fry Readability Graph (1963), Simple Measure Of Gobbledygook (SMOG) (1969) formulae and several others (refer to examples ⁷⁻⁹). These formulae evaluate documents based on Natural Language Processing (NLP) and rely on measures such as sentence length, word length and number of polysyllabic words, which in a health environment may not necessarily be useful as the propensity for polysyllabic words in common use eg diabetes, may confuse the evaluation.

These formulae have been widely used to evaluate a variety of patient information pamphlets and instruction sheets which generally have concluded that the reading age of the material is inappropriate for the intended audience. However, almost none of these studies evaluate the effects of visual or design factors that could influence readability and comprehension. A study in 2008, by Kandula and Zeng-Treitler ¹⁰ in an attempt to create a ‘gold standard’ for the readability measurement of health texts, evaluated 324 documents noted that a document’s style - such as
On the whole very little has been written about the impact of layout of text on readability of pathology or healthcare related reports. This may be a result of the perception that style (text layout) is primarily about aesthetics and marketing (branding) with minimal impact on readability. In a study by Frost in 1999[^11] on the “importance of format and design in print patient information”, a survey of 44 oncology healthcare professionals identifying characteristics that they believed important for effective print educational materials was undertaken. Style (text layout) was initially perceived to be the least important factor impacting readability by the surveyed oncologists; however it in fact made a major impact on acceptance and a nine-fold increase in distribution of the redesigned material.

In the last decade there appears to be a slowly increasing recognition that style (text layout) may be a major contributor to the overall predictor of readability. A study by Arnold et al in 2006 in their Assessment of Newborn Screening Parent Education Materials developed the “User-Friendliness Tool (UFT)”[^12] which focussed on issues such as layout, illustrations, message, information, and cultural appropriateness. They viewed layout as “critical” to reading ease. Many very recent assessments of readability of educational and instructional material (examples[^13-14]) have used one of the abovementioned readability formulae’s in conjunction with the UFT. The UFT has taken the recommendations of several manuals and texts and distilled this into general layout recommendations such as font size ($\geq$ 12 point), type (avoiding all-capital letters, italics, and specialty fonts), white space (ample), paragraph size (<4-5 lines) and the presentation of manageable information (use of bullets and boxes to break up text).

In a 2009 MIT study[^15], Bufuka, recognising that current readability measures depend heavily on text analysis (NLP) but neglected style properties (text layout), found that style (text layout) is in fact an important predictor of a documents readability. His study developed a program to analyse images of a document and then review the style (text layout) and predict readability. This program, evaluating style (text layout) alone, provided a closer correlation to human expert assessments of readability than the Flesch-Kincaid and other NLP methods[^15].

One of the more recently developed readability formulae, the PMOSE/IKIRSCH formula[^16] in 1998 recognised the role of structure/design on readability. It was developed to evaluate overall document complexity based on structure (design) (e.g., simple list, combined list, and nested lists), density (e.g., number of labels and items) and dependency (whether or not any important information is to be found outside the document).

**Standardisation**

The surgical pathology report represents the focal point for determining the clinical management of cancer and it therefore is a highly technical document in which the details must be clearly articulated and logically presented. However there are wide variations in report formats, styles, and contents,
which are often designed to demonstrate the wealth of histological detail accumulated rather than a clear concise report aimed at patient care decision-making.

In the last decade studies have shown that a change from the traditional narrative style of report to a structured or synoptic style significantly enhances the completeness and quality of data provided to clinicians\(^\text{17-22}\). This change standardises the information within a pathology report (using approved data elements) as well as the way in which this information is structured (discrete question and answer). There are several publications which review the types of information to be included in these cancer specific report (examples\(^\text{22-23}\)) as well as national standards published by the US\(^\text{24}\), UK\(^\text{25}\), and Australia\(^\text{26}\). Some publications which favour structured or synoptic reporting while focussing on content also include broad recommendations in regard to placement of information and style such as “placing all demographic information in the top portion of the report”; making sure “printing is of sufficient quality to be read easily”\(^\text{27}\). However on the whole little has been done on the investigation of appropriate style (text layout) in which to display the recommended information.

The standardisation of style and layout has been recognised as a safety mechanism in health such as the development of the Australian National Inpatient Medication Chart (NIMC)\(^\text{28}\), the 2009 report by NEHTA on the pathology industry\(^\text{29}\), the RANZCR Radiology Written Report Guideline Project\(^\text{30}\), the National Health Service(UK) Common User Interface (CUI)\(^\text{31}\), as well as in other industries such as that described by Valenstein\(^\text{3}\) in the standardisation of aircraft instrumentation in the post war era. This is largely to do with consistency in positioning and style; having the same information presented in the same relative position and manner each time thereby avoiding delay in locating and recognising critical information. Consistency in presentation facilitates human perception, cognition, visual scanning, learning and remembering\(^\text{32}\). The facilitation of learning and remembering occurs by matching each visual stimuli to a stored representation\(^\text{33}\), allowing a user to quickly view and then understand the information being displayed. According to Ponton\(^\text{34}\) consistency should be applied to font, colours, terms, units, layouts , typography, margins amongst other characteristics.

Nygren et al\(^\text{35,36,37}\) in their analyses of the way in which doctors searched medical records, observed that doctors tended to skimmed rapidly over pages of text while continuously making assessments of relevance. Experienced physicians had developed perceptual skills that enabled them to use pattern-recognition as a complement to normal reading. So given the same content, document structure and format determined the time and effort needed for relevant information to be extracted. Consistency of format assisted in the speed of the ‘skimming’ process and enabled doctors to find the information they needed faster. They found that a standardised set of headings and subheadings in a predefined order enabled the doctors to locate themselves in lengthy documents. Indenting of subheadings created visual landmarks, making specific data items easier to find. This information grouping was also noted to improve decision making in ICU’s as well\(^\text{38}\). Similar findings were also noted with respect to abstracts by Hartley\(^\text{39}\) who recommended that texts be opened-up and clearly subdivided into their component parts, so that readers perceive their structure and the information should be sequenced in a consistent order under consistent subheadings to facilitate search and retrieval.
This consistency in style (layout) supports the development of a “cognitive map”. Cognitive maps enable a user to analyse and interpret the information available to them at speed and orientate themselves through an area/display. Cognitive maps are generally discussed in wayfinding literature, that is, car or pedestrian travel; and more recently in 3D virtual environments. As a user ‘interacts’ with a certain layout or display they acquire information and start to form an internal cognitive model of the structure, organisation, and relationships between the parts.

There is also evidence that the style (text layout) in which information is presented significantly influences clinical decisions in a range of different medical settings such as the accuracy of obstetric judgements, to the speed of interpreting laboratory tests, or the review of intensive care data. Elting’s study showed that doctors’ decisions can be manipulated by changing the format of words and data, for example, by presenting information in tables, graphs, or pictograms.

**Topography/Legibility**

While legibility is usually defined as the assessment of the fine details of typeface design, and in an operational context this usually means the ability to recognise individual letters or words, typography in relation to documents, is defined as the presentation/arrangement and layout of whole bodies of text. The two terms are often used interchangeably in the literature when reviewing aspects of style to improve readability.

The aim of any document is to communicate content to the reader as unambiguously as possible. To achieve that aim the content needs to be composed to create a readable, coherent, and visually satisfying whole that works, ideally without the awareness of the reader. The reader should be assisted in navigating around the information by optimal inter-letter, inter-word and particularly inter-line spacing, coupled with appropriate line length and position on the page, careful editorial “chunking” and choice of titles and reference links etc.

According to Rousseau, there are four steps of interaction between the viewer and the design of a document, for the design to effectively convey its meaning:

1. noticed
2. encoded [decoded]
3. comprehended
4. complied with by the viewer

Rousseau states that all four steps must be successfully completed to achieve successful communication. In relation to medical reports, step 1 that the document is ‘noticed’ is largely assumed through the act of sending it to a specific person. Step 2, that the information is compiled or ‘encoded’ is ultimately the responsibility of the author and to achieve step 2 the author will need to consider the individual elements of text and document design and how these elements interact to ensure a document’s legibility. Steps 3 and 4 are largely to do with the reader of the report, though
arguably comprehension is a matter of the author ensuring the language and content used can be understood by the reader as much as it is the reader’s responsibility to concentrate while reading and assimilate the information.

There are many published papers and books on legibility analysing the various individual aspects of text topography:

- Typeface /Font. Typeface consists of all characters, in all sizes, of a particular design. Font is all the characters (upper and lowercase, figures, fractions, reference marks, etc.) of one size of one particular typeface. There are investigations into various aspects of character design related to different fonts including x-height (the height of the lower case “x” in a typeface) Figure A; ascenders (vertical strokes which rise above the body of a character or x-height), descenders (strokes which fall below the baseline of the x-height) Figure B; counter forms (the “negative spaces” inside a character) Figure C; serifs (or lack of (sans) serifs), and stroke weight (how thick or thin the lines are).

Figure A

```
x-height    x-height
```

Arial, 51 point          Times New Roman, 51 point

Figure B

```
          Ascenders
            x
          Descenders
```

Figure C

```
a  b  g  e
```

Arial, 51 point

```
a  b  g  e
```

Arial Black, 51 point

- The legibility of seriffed fonts is of particular debate and there are many studies on this one area alone. Serifs are the small finishing strokes on the end of a character and one school of thought is that serifs are used to guide the horizontal “flow” of
the eyes, with the lack of serifs said to contribute to a vertical stress, which is supposed to compete with the horizontal flow of reading\textsuperscript{46}. Advocates of this theory say that serifs help in combining separate letters into word-wholes and letters with serifs are more easily differentiated by readers than letters without serifs\textsuperscript{47-49}. However others believe that this is not supported by research into eye movements which has established that eyes do not move along a line of text in one smooth sweep but in a series of quick jerks called saccadic movements\textsuperscript{50-51}. Other studies suggest that there is no difference between the legibility of serif and sans serif typefaces\textsuperscript{46,52,53-54,55}.

- Poulton\textsuperscript{56} tested several fonts for legibility and found that x height was a significant factor in legibility and that fonts with the same x height (rather than point size) had relatively the same level of legibility.
- A study conducted at the New England College of Optometry\textsuperscript{57} found that horizontal letter compression had a greater effect on readability than vertical letter height.
- Bernard et al\textsuperscript{58} identified 5 font types that were perceived as being the most legible: Courier, Comic, Verdana, Georgia, and Times. However, examining legibility of some of these commonly used fonts both serif and sans serif (Century Schoolbook Courier, Georgia, Times New Roman, Arial, Comic Sans MS , Tahoma and Verdana) by means of reading efficiency, resulted in no significant difference between the fonts\textsuperscript{59}. However, there were significant differences in reading time. Generally, Times and Arial were read faster than Courier, Schoolbook, and Georgia. Fonts at the 12-point size were read faster than fonts at the 10-point size. In this study Verdana appeared to be the best overall font choice. Besides being the most preferred, it was read fairly quickly and was perceived as being legible.
- Bernard et al\textsuperscript{58} also noted that different fonts were perceived in different ways eg ornate fonts like Bradley and Corsiva were perceived as having a great deal of personality and elegance; Courier and Times were perceived as being the most business-like, whereas Comic was perceived as being the most fun and youthful.
- Letter spacing relates to the amount of space used between letters which can be both positive and negative. Negative spacing is where space is removed between certain characters eg Yo, WA so that there is overlap between the letters. According to Bix\textsuperscript{60} letter spacing is widely recognised to impact legibility, but there is little documentation with regard to specific requirements for legible messages.

- **Leading or line spacing.** Leading is the space between the baseline characters line to line. 
  
  Eg.
  
  \begin{tabular}{l}
  \textbf{Alphabet} \\
  \textbf{Crumble} \\
  \textbf{Alphabet} \\
  \textbf{Crumble}
  \end{tabular}

  A small leading means that the ascenders and descenders, line to line are closer making it less legible. According to Bix\textsuperscript{60} the optimal amount of leading for maximum legibility is
dependent on the elements of both letter and message design eg bolder typefaces need more leading.

- Colour contrast. The majority of research findings are consistent in this area: dark text on a light background provides the best legibility\textsuperscript{61-63}. Bradley et al\textsuperscript{63} also suggest that these combinations avoid difficulties associated with red/green colour blindness.

- ALL CAPS are generally not recommended in the literature being regarded as less legible\textsuperscript{64} than sentence case.

- ‘Adequate’ or ‘ample’ white space is deemed an important feature of readability in several readability assessment tools such as The Suitability Assessment of Materials (SAM)\textsuperscript{65}, and the User Friendliness Tool (UFT)\textsuperscript{12} since it promotes visual breaks and as noted by Valenstein\textsuperscript{2}, reduces clutter.

Bix\textsuperscript{60} in 2002, in reviewing the many different aspects of legibility concluded that the individual elements of design and layout, do not determine legibility, but that sufficient legibility is the outcome of the sum of the parts eg font, font size, leading, colour etc.

Eye Tracking

Eye tracking is the technique of monitoring eye movement while a subject views a document or screen. Eye tracking evaluates specific eye movements and responses such as blinks, pupil dilation, fixations, saccades (a fast move or jerk of the eye) and regressions (right-to-left movements along the line or movements back to previously read lines), which can be indicative of fatigue, cognitive processing, difficulty processing information and areas of interest. Eye tracking is a useful tool in evaluating how people utilise an interface (paper or screen) optimally and efficiently in terms of assimilating information. Displays on paper or screen which are not laid out to exploit human spatial abilities can increase user orientation time drawing mental resources from other tasks and objectives\textsuperscript{66}. The efficiency with which a person is able to find and absorb information from a given page also contributes to their satisfaction.

Rayner\textsuperscript{67} in his review of eye tracking of people reading (from paper), explains that saccades (a fast move or jerk of the eye) correlate to about 7-9 letter spaces or characters and about 10-15% of saccades are regressions (right-to-left movements along the line or movements back to previously read lines). Many regressions tend to be only a few letters long and could be due to the reader making too long of a saccade forward, in which case a short saccade to the left may be necessary for reading to proceed efficiently. Short within-word regressive saccades may also be due to problems that the reader has processing the currently fixated word. Longer regressions (more than 10 letter spaces back along the line or to another line) occur because the reader did not understand the text. Rayner also notes that certain types of words such as function and content words impact on fixation time as does word length.

Eye movements are influenced by textual and typographical variables. For example, as text becomes conceptually more difficult, fixation duration increases, saccade length decreases, and the frequency
of regressions increases\textsuperscript{51,68} Factors such as the quality of the print (variations in fonts), line length, and letter spacing influence eye movements\textsuperscript{69}. Ponton\textsuperscript{34} explains that enhancing the usability of an interface (paper or screen) would result in the user fixating less, scanning / reading quickly\textsuperscript{70} and making fewer regressions to previously scanned areas\textsuperscript{67}.

The area which has used eye tracking as a means of exploring how people interact with an interface in more depth than any other is in the field of marketing and communications eg newsprint, advertising copy, website design. One of the dominant studies undertaken was the Eyetrack series\textsuperscript{71} by The Poynter Institute and its partners. The first, Eyetrack I, (1990-1991) investigated how news consumers interact with print editions of newspapers. Interestingly this investigation broke many myths about how people read newspapers such as “color photos do not automatically draw readers. Content, size, and placement are more important”. Eyetrack II, (1999-2000) reviewed first-generation news websites investigating reading patterns by age and gender. Eyetrack III (2003-2004) undertook further investigation into web design and multimedia components.

In the study of 46 participants in Eyetrack III, several key findings were recorded in reference to review of information onscreen including:

- Dominant headlines most often draw the eye first upon viewing a page, especially when they are in the upper left
- Smaller type encourages focused viewing behaviour (that is, reading the words), while larger type promotes lighter scanning
- Visual breaks, like a line, discouraged people from looking at items beyond the break
- Shorter paragraphs performed better than longer ones.
- People looked at text elements before their eyes landed on an accompanying photo

The Bridge\textsuperscript{72} is a website focused on marketing and it draws together some findings from eye tracking studies on websites as well:

- Type face - serif upper and lower case - enhances readability and comprehension
- Short paragraphs are more effective over longer paragraphs
- Short words or lines over long words or lines
- Using colour in body copy has a negative impact on readability and comprehension
- In the case of a single page the human eye tracks top left to bottom right
- A colour image is more effective than black and white

Ponton\textsuperscript{34} in her study for the \textit{Defence Science and Technology Organisation} on user interaction with Naval combat data systems, used eye tracking as one of the key investigative techniques but as with the Poynter Research and Bridge information and more recent investigations in eye tracking these are largely focused on human –computer interactions which involve the added complexities of navigation/ search and retrieval of information.
Online versus paper

Although there are several publications which show that people read more slowly onscreen than from paper, they attribute this to the quality of the image presented to the reader. Tinker\textsuperscript{52} reports dramatic interaction effects of image quality variables on paper and according to Gould et al\textsuperscript{73} (1986) it is likely that these occur on screen too.

Dillon\textsuperscript{74} noted that the reading process is affected by the medium of presentation though they found it extremely difficult to quantify and demonstrate such differences empirically. The major differences appeared to occur in manipulation which seems more awkward with electronic texts and navigation which seems to be more difficult with electronic and particularly hypertexts. Dillon found that eye movement patterns did not seem to be significantly altered by the presentation medium. However he noted that further process issues may emerge as our knowledge and conceptualisation of the reading process improves.

It appears caution should be applied when transferring a successful design for printed documents to the Web. The majority of eye tracking studies are largely devoted to analysis of viewing information onscreen. When Poynter\textsuperscript{75} studied newspaper readers it found that photos and graphics catch the eye first. However when they looked to news websites, they found that users read text before they looked at photos or other graphics. Boiarsky\textsuperscript{76} 2002 conducted an investigation into onscreen readability/useability and noted that additional research needed to be conducted concerning legibility in terms of font size and style and their relation to spacing between letters and words in reading onscreen as there were several contradictory findings between text and onscreen reading.

Therefore we can assume that any recommended format will need to be comprehensively analysed both on paper and on screen and that adjustments in format for the different media may be required.
Summary

There is a significant lack of investigation into style (text layout) on readability and comprehension of pathology reports in general and anatomical pathology reports specifically. This may be a result of a perception that style is not an important factor in communication of this type of information.

Medical documents including pathology reports have developed over a long period of time and are more focussed on content, aesthetics and branding that on readability and comprehension.

In recent studies there is recognition that style (text layout) impacts on readability, speed of assimilation of information and in reference to consistency in layout, impacts on patient safety.

Some of the key features noted are:

- Consistency in positioning
- Headings and subheadings to create visual landmarks
- Information grouping/chunking
- Topographical considerations – font, line spacing, white space

What is clearly lacking in the literature is any detailed study into a style of report for pathological information which supports quick and easy assimilation of often complex data. Of all the eye tracking and readability studies and tools, standardisation of format, investigations into effects of topography on reading and comprehension most are based on either non-health related documents or documents designed to be read by patients rather than complex documents design primarily to be read by medical personnel.

Wyatt suggested we tap into the design skills of psychology, graphical design, perception, and typography and as with the design of railway timetables to electricity bills use this type of information to create the ideal vehicle for communicating pathological information. We must ensure that the formats designed are based on empirical study, not merely by asking clinicians to indicate a preference, since preferences can mislead. The formats preferred by doctors are not necessarily the formats which lead to optimal comprehension. Good design usually goes unnoticed. In a study by Nygren, participants had no recall of the layout of a well-formatted document they had just read but good recall of one with text attributes that slowed reading or were likely to induce errors. So any proposed formatting must be confirmed by detailed analysis into its affect on speed to assimilate and comprehension of content, as visual analysis into what “looks better” may in fact lead to the propagation of a poor design.

Eye tracking has proven a useful tool in evaluating usability of interfaces most pre-dominantly on screen interfaces in recent years. It may be an important tool to provide empirical data to support and fine tune any proposed report design.
References


37 Nygren E (1997). From paper to computer screen. Human information processing and interfaces to patient data. IMIA WG6 Conference on Natural Language and Medical Concept Representation, Jacksonville, Florida, USA.


The Bridge Marketing that works. *Eye tracking produces effective copy.* Available at: [http://www.thebridgeeffect.co.nz/blog/marketing-research/eye-tracking-produces-effective-copy](http://www.thebridgeeffect.co.nz/blog/marketing-research/eye-tracking-produces-effective-copy).

Gould JD et al (1986). Why reading was slower from CRT displays than from paper. CHI '87 Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface New York, NY, USA.


RACP Project Report 1

Automatic population of colorectal structured reports

By
Jon Patrick

Health Language Laboratories
4th May 2012
QUPP Colorectal Project

The stages in this project are:

1. Acquire a corpus of 400 colorectal cancer reports.
2. Design a Tagset for annotating the reports.
3. Build a mapping between the Tagset and the Structured report fields.
4. Train staff in the use of the tagset.
5. Annotate the corpus using the tagset.
6. Build a machine learning model for computing the annotation tags;
7. Run a series of evaluations on different models;
8. Develop the programme code to extract the descriptive statistics;

An update on the stages of the project

Stage 1: A corpus of 312/400 colorectal reports was delivered on March 13, 2012. March 13, 2012 thus constitutes the beginning of the contract for this project. This original corpus has been cleaned and converted into plain text files with UTF 8 encoding. Non-supported characters were stripped from the files and images have not been included in any way (i.e. we have not replaced images with an image reference so unless there is an independent textual reference for that image it will not be recorded as having been present in the file).

A description of the corpus is provided at the bottom of this report. The development of the tag set is set up so that tagged files will automatically provide a measurable index of the structuredness of a report and details about adherence to guidelines and standards for colorectal reporting. The details provided under corpus description are intended to provide an initial general indication of the number of structured reports in the project and are also used for annotation purposes.

Stage 2: Design of a tagset. The design of the stable tagset took 6 draft versions. These draft versions have been recorded in the project wiki and in the repository. The final tagset is a mapping of the Colorectal Cancer Structured Report Sample provided in the protocol with the extractable items provided under chapter 6 of the protocol.

Stage 3: Build a mapping between the Tagset and the Structured report fields. This has been completed and is located on the wiki. The tag set has attribute values that map onto the extractable fields provided under chapter 6 of the protocol. We have not included the following standards and guidelines as these are not readily available from the corpus provided (please see corpus description below):

Clinical information and surgical handling

S1.01 Patient name; Date of birth; Sex; Identification and contact details of requesting doctor; Type of specimen; Date of surgical procedure; Clinical information relevant to the investigations requested
G1.01 Patient identifiers (e.g. MRN, UHI, NHI)
G1.02 Pathology accession number
S1.02 Principal clinician caring for the patient
S1.03 Operating surgeon: Contact address Phone (mobile) number
**Stage 4:** All staff on the project have been trained in the use of the tagset and have read the protocol document provided as well as numerous support references relevant to the project. The tagset and support notes are available on the project wiki.

**Stage 5:** Annotation of the corpus is approximately 1/3 complete and is expected to be completed by May 18, 2012., including the development of a gold standard.

*Any issues encountered and proposed action to resolve*

**Stage 1.** Delivery of files in various formats including pdf, xls, doc meant that files had to be extracted and converted to the correct format and encoding – delay impacted on stages 2-5 by 4 weeks.

**Proposed action:** simplified tagset to speed up annotation and implemented some changes to work flow.

**Stage 2 – 5.** Flow-on effect from the resolution action for stage 1 means that stages 2-5 have also been delayed.

**Proposed action:** given the anonymised nature of the reports and the low level security risk, it would be possible to bring one off-shore annotator on to complete files in a shorter time frame. Files will be annotated day and night.

*Any risks to the successful completion or impact on the timeline*

**Anticipated high risks to the completion of the remaining stages:** None

**Anticipated low-moderate risk elements:** staff availability; competing project deadlines; technical problems.

*A plan of action for the remaining stages*

Annotated Gold Standard to be passed to the computational team within the next two weeks (May 18 2012). Leaving 6 weeks for the remaining 4 stages (including writing the project report). Many of these tasks will happen concurrently.
Corpus description:

<table>
<thead>
<tr>
<th>Structuredness</th>
<th>No. of files</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured</td>
<td>136</td>
<td>43.58974359</td>
</tr>
<tr>
<td>Partially Structured</td>
<td>62</td>
<td>19.871794872</td>
</tr>
<tr>
<td>Unstructured</td>
<td>114</td>
<td>36.538461538</td>
</tr>
<tr>
<td><strong>Total Corpus</strong></td>
<td><strong>312</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>File Format</th>
<th>No. of files</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF</td>
<td>66</td>
<td>21.153846154</td>
</tr>
<tr>
<td>DOC</td>
<td>7</td>
<td>2.2435897436</td>
</tr>
<tr>
<td>XLS</td>
<td>239</td>
<td>76.602564103</td>
</tr>
<tr>
<td><strong>Total Corpus</strong></td>
<td><strong>312</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country of Origin</th>
<th>No. of files</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>17</td>
<td>5.4487179487</td>
</tr>
<tr>
<td>Australia (?)</td>
<td>275</td>
<td>88.141025641</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>4</td>
<td>1.2820512821</td>
</tr>
<tr>
<td>Malaysia</td>
<td>5</td>
<td>1.6025641026</td>
</tr>
<tr>
<td>Namibia</td>
<td>1</td>
<td>0.3205128205</td>
</tr>
<tr>
<td>New Zealand</td>
<td>3</td>
<td>0.9615384615</td>
</tr>
<tr>
<td>South Africa</td>
<td>2</td>
<td>0.641025641</td>
</tr>
<tr>
<td>UAE</td>
<td>1</td>
<td>0.3205128205</td>
</tr>
<tr>
<td>Unknown</td>
<td>4</td>
<td>1.2820512821</td>
</tr>
<tr>
<td><strong>Total Corpus</strong></td>
<td><strong>312</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient Sex</th>
<th>No. of files</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td>292</td>
<td>93.58974359</td>
</tr>
<tr>
<td>Female</td>
<td>11</td>
<td>3.5256410256</td>
</tr>
<tr>
<td>Male</td>
<td>9</td>
<td>2.8846153846</td>
</tr>
<tr>
<td><strong>Total Corpus</strong></td>
<td><strong>312</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient Age</th>
<th>No. of files</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td>295</td>
<td>94.551282051</td>
</tr>
<tr>
<td>Above 70</td>
<td>6</td>
<td>1.9230769231</td>
</tr>
<tr>
<td>50-70</td>
<td>8</td>
<td>2.5641025641</td>
</tr>
<tr>
<td>Below 50</td>
<td>3</td>
<td>0.9615384615</td>
</tr>
<tr>
<td><strong>Total corpus</strong></td>
<td><strong>312</strong></td>
<td></td>
</tr>
</tbody>
</table>