MASSACHUSETTS GENERAL HOSPITAL

PATHOLOGY

HARVARD MEDICAL SCHOOL

Data Quality/Information Quality

Veronica E. Klepeis, M.D., Ph.D.

Massachusetts General Hospital

Boston, MA

May 23, 2016

PATHOLOGY INFORMATICS SUMMIT 2016

May 23-26, 2016
Pittsburgh, PA

Brought to you by the Association for Pathology Informatics.

Notice of Faculty Disclosure

In accordance with ACCME guidelines, any individual in a position to influence and/or control the content of this ASCP CME activity has disclosed all relevant financial relationships within the past 12 months with commercial interests that provide products and/or services related to the content of this CME activity.

The individual below has disclosed the following financial relationship(s) with commercial interest(s):

Veronica Klepeis, MD, PhD

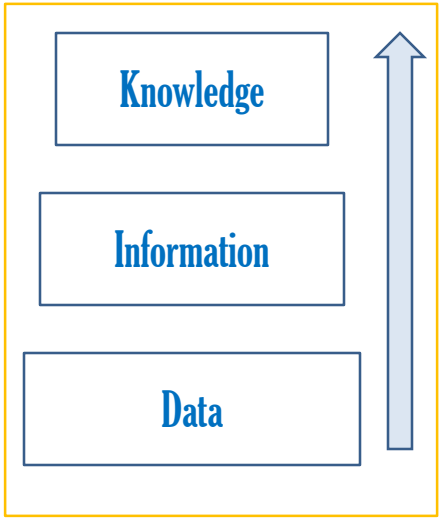
Philips
Consulting Fee
Consulting

Outline

- Introduction to Data Quality
  - Definitions
  - Data Quality in Healthcare
  - Data Quality Dimensions
  - Poor Quality Data
- Data Quality in (Surgical) Pathology
  - Variability in Free Text Pathology Reporting
  - Coding and Standards
  - Synoptic Reporting
  - Structured Data
  - Report formatting

Basic Definitions

- **Data** = raw facts (ex. blood glucose level)
  - May describe a particular event/specific observation
- **Information** = processed data (anything from explanations to formal analysis)
  - Cannot create information without data
- **Knowledge** = Information applied to rules, experiences and relationships with the result that it can be used for decision making (ex. heparin prevents clotting of blood)
- *Cannot pinpoint where data ends and information begins*
  - Same data may provide different information to different users
  - One person's data may be another person's information
  - More important to understand relationship between data and information



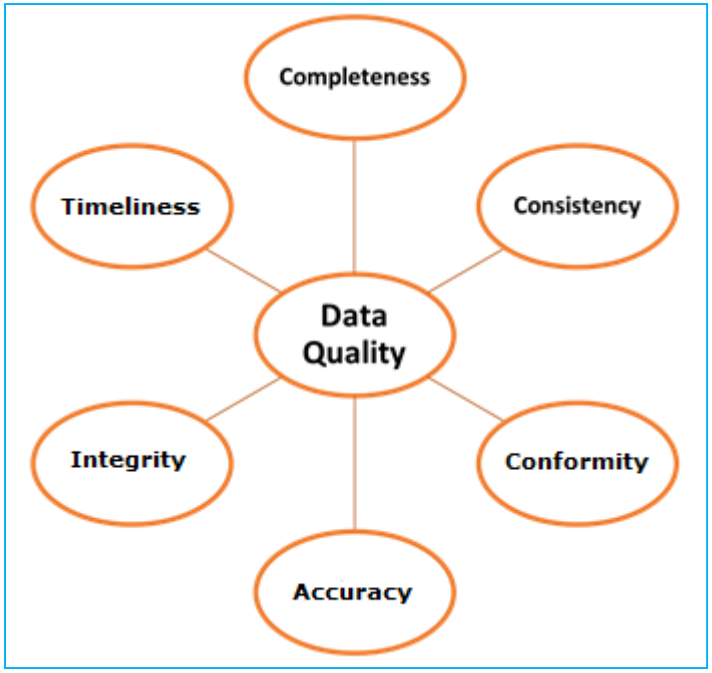
Wager KA, et al. Managing Healthcare Information Systems (2005).

Types of Data in Pathology

- Patient name and demographics
- Responsible physicians/staff/technicians
- Dates/Times
- Specimen Types/Procedure
- Clinical History
- Tests ordered
- Laboratory Results
- Molecular/Genetic Results
- Pathology Reports
  - Gross Description
  - Microscopic Description
  - Final Pathologic Diagnosis
- Billing information/codes
- Imaging
- etc…

Data Quality

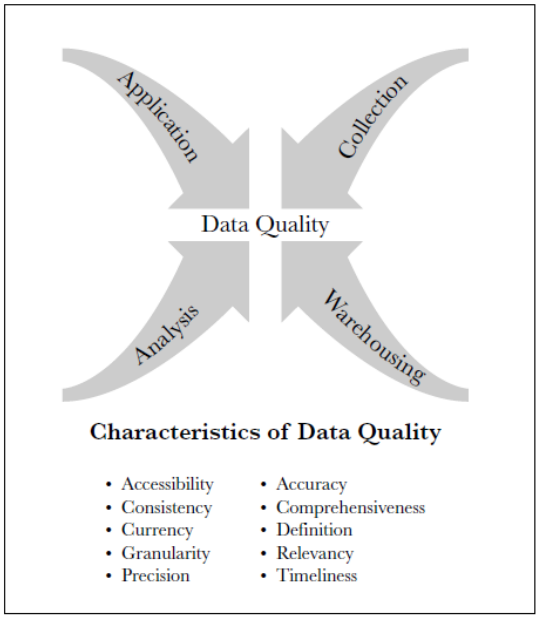
- More than just data accuracy
- Data quality cannot be assessed without data quality standards
- Dimensions of data quality
  - Method to measure data quality
  - Established for a variety of industries/businesses
- No universally recognized set of health care data quality standards
  - *Quality of data needed in any situation is driven by how data will be used*



Wager KA, et al. Managing Healthcare Information Systems (2005).

Data Quality in Healthcare

- American Health Information Management Association (AHIMA): Data Quality Management Tool
  - Guidance to assist health care organizations in establishing data quality standards



Characteristics of Data Quality

- Accessibility
- Consistency
- Currency
- Granularity
- Precision
- Accuracy
- Comprehensiveness
- Definition
- Relevancy
- Timeliness

Application – The purpose for which the data are collected.

Collection – The processes by which data elements are accumulated.

Warehousing – Processes and systems used to archive data and data journals.

Analysis – The process of translating data into information utilized for an application.

Defines characteristics of data quality that can be applied to application, collection, warehousing and analysis of data in healthcare

Wager KA, et al. Managing Healthcare Information Systems (2005).

Data Quality Characteristics

- **Data accuracy** = data are the correct values and are valid (ex. typographical error)
  - *Syntactic accuracy* – permissible value
  - *Semantic accuracy* – closeness to true value (correctness)
- **Data accessibility** = data items should be easily obtainable and legal to collect (can't access data it's of no use)
- **Data comprehensiveness** = all required data elements are included (data not useful if not complete)
- **Data consistency** = value of the data should be reliable and same across applications (ex. use of an abbreviation that has two different meanings)
- **Data currency** = data should be up-to-date (many data become obsolete after a period of time)

Wager KA, et al. Managing Healthcare Information Systems (2005).

Data Quality Characteristics

- **Data definition** = clear definitions should be provided so that current and future data users will know what the data mean (ex. use of data dictionaries)
- **Data granularity/atomicity** = attributes and values of data should be defined at the correct level of detail (ex. patient name recorded as three data elements: last name, first name, middle name)
- **Data precision** = how close to an actual size , weight or other standard a particular measurement is
- **Data relevancy** = data are meaningful to the performance of the process or application for which they are collected
- **Data timeliness** = defined by how data are being used and their context

Wager KA, et al. Managing Healthcare Information Systems (2005).

Data Errors

- **Data error** = failures of data to meet established quality standards
  - *Have negative impact on one or more of the characteristics of data quality*
- **Systematic errors** = flaw or discrepancy in adherence to standard operating procedures or systems
- **Random errors** = due to carelessness rather than lack of training (ex. transcription error)

Wager KA, et al. Managing Healthcare Information Systems (2005).

Potential Causes of Poor Quality Data

Systematic Errors

- Unclear data definitions
- Poor interface design
- Programming errors
- Guidelines not adhered to
- Lack of sufficient data checks
- No system for correcting detected errors
- Suboptimal data conversion
- System upgrades

Random Errors

- Illegible handwriting in data source
- Typing errors
- Frequent personnel turnover
- Calculation errors (not in system)

Wager KA, et al. Managing Healthcare Information Systems (2005).

Effect of Poor Quality Data

- Medical Records Institute (MRI) = professional organization dedicated to improvement of patient records through technology
- Health care documentation has two parts
  - Data capture
  - Report generation

Both must be considered in order to have high quality data

- 5 major areas negatively effected by poor quality data
  - Patient safety
  - Public safety
  - Continuity of patient care
  - Health care economics
  - Clinical research and outcomes analysis

Wager KA, et al. Managing Healthcare Information Systems (2005).

Methods for Improving Data Quality

Data Error Prevention

Compose a minimum set of necessary data items
Define data and data characteristics in a data dictionary
Develop a data collection protocol
Create user friendly data entry forms or interface
Compose data checks
Create a quality assurance plan
Train and motivate users

Data Error Detection

Perform automatic data checks
Perform data quality audits
Review data collection protocols and procedures
Check inter- and intraobserver varability (if appropriate)
Visually inspect completed forms (online or otherwise)
Routinely check completeness of data entry

Actions for Data Quality Improvement

Provided data quality reports to users
Correct inaccurate data and fill in incomplete data detected
Control user correction of data errors
Give feedback of data quality results and recommendations
Resolve identified causes of data errors
Implement identified system changes
Communicate with users

Wager KA, et al. Managing Healthcare Information Systems (2005).

Data Quality in Pathology

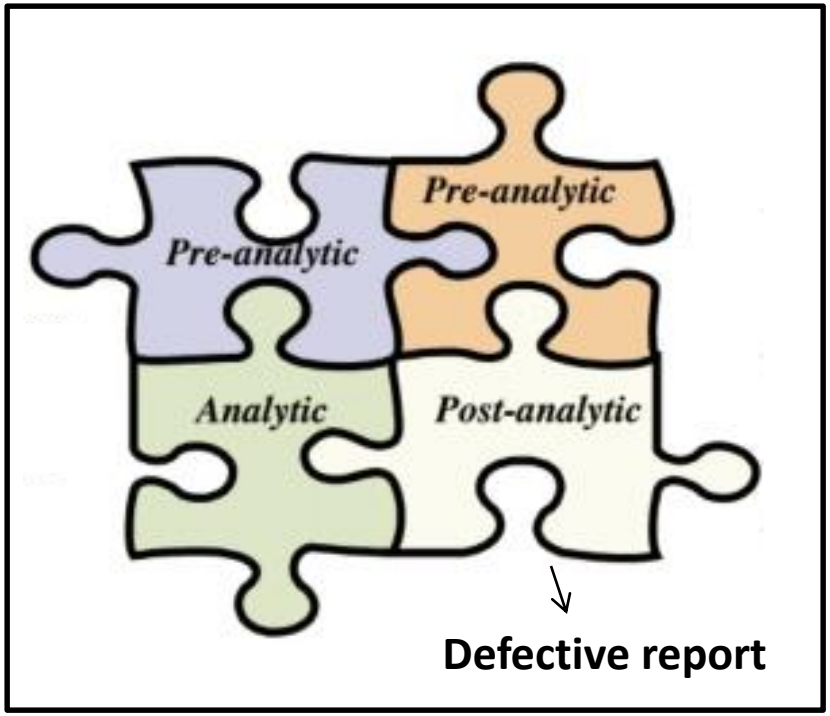
- Errors can occur across all phases of the process of pathology
  - *Preanalytic* = receiving and preparing specimen
  - *Analytic* = interpretation
  - *Postanalytic* = conveying results to clinician

Data quality issues are generally much more complex in surgical /anatomic pathology compared to clinical pathology, especially in post-analytic phaserelating to data capture and reporting

*For more on quality measures in anatomic pathology – archived preentations by Dr. Liron Pantanowitz

Data Quality in Surgical Pathology

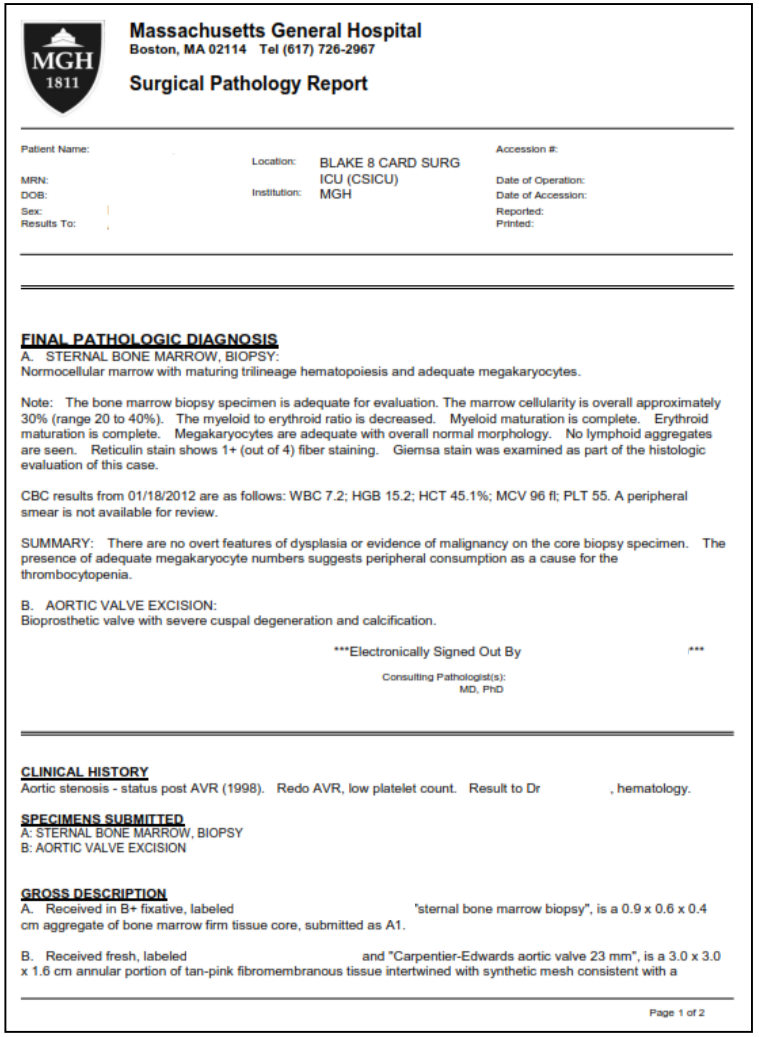
- Quality in surgical pathology is determined by
  - *Diagnostic accuracy*
  - *Report completeness*
  - *Timeliness*
- *Clear and comprehensible pathology reports*



Zarbo RJ, et al. Arch Pathol Lab Med. 2005;129:1237-45.
Nakhleh RE. Arch Pathol Lab Med. 2011;135:1394-97.

Reporting in Surgical Pathology

- Traditional pathology reports are written in **free text**
- Results in creation of *a few large text fields* correlating with specific parts of the surgical pathology report
  - Gross description
  - Clinical history
  - Final pathologic diagnosis
  - Etc.



Massachusetts General Hospital
Boston, MA 02114 Tel (617) 726-2967

Surgical Pathology Report

Patient Name:
MRN:
DOB:
Sex:
Results To:

Location: BLAKE 8 CARD SURG ICU (CSICU)
Institution: MGH

Accession #:
Date of Operation:
Date of Accession:
Reported:
Printed:

FINAL PATHOLOGIC DIAGNOSIS
A. STERNAL BONE MARROW, BIOPSY:
Normocellular marrow with maturing trilineage hematopoiesis and adequate megakaryocytes.

Note: The bone marrow biopsy specimen is adequate for evaluation. The marrow cellularity is overall approximately 30% (range 20 to 40%). The myeloid to erythroid ratio is decreased. Myeloid maturation is complete. Erythroid maturation is complete. Megakaryocytes are adequate with overall normal morphology. No lymphoid aggregates are seen. Reticulin stain shows 1+ (out of 4) fiber staining. Giemsa stain was examined as part of the histologic evaluation of this case.

CBC results from 01/18/2012 are as follows: WBC 7.2; HGB 15.2; HCT 45.1%; MCV 96 fl; PLT 55. A peripheral smear is not available for review.

SUMMARY: There are no overt features of dysplasia or evidence of malignancy on the core biopsy specimen. The presence of adequate megakaryocyte numbers suggests peripheral consumption as a cause for the thrombocytopenia.

B. AORTIC VALVE EXCISION:
Bioprosthetic valve with severe cuspal degeneration and calcification.

***Electronically Signed Out By ***

Consulting Pathologist(s):
MD, PhD

CLINICAL HISTORY
Aortic stenosis - status post AVR (1998). Redo AVR, low platelet count. Result to Dr , hematology.

SPECIMENS SUBMITTED
A: STERNAL BONE MARROW, BIOPSY
B: AORTIC VALVE EXCISION

GROSS DESCRIPTION
A. Received in B+ fixative, labeled "sternal bone marrow biopsy", is a 0.9 x 0.6 x 0.4 cm aggregate of bone marrow firm tissue core, submitted as A1.

B. Received fresh, labeled and "Carpentier-Edwards aortic valve 23 mm", is a 3.0 x 3.0 x 1.6 cm annular portion of tan-pink fibromembranous tissue intertwined with synthetic mesh consistent with a

Page 1 of 2

Problems with Free Text Reports

- *Variability* in diagnostic terminology
  - Makes it difficult to combine/compare datasets
- *Variability* in reporting *styles*
  - Adds complexity to understanding the reports
- *Variability* in report *content*
  - Increases likelihood of reports *missing important data* elements (eg., margins, lymphatic invasion)

Improving Data Quality of Free Text Reports

- Natural language processing
- Coding
- Secondary review and amendment

Variations in Wording of Breast Diagnoses

- Breast pathology reports were analyzed using natural language processing software to extract information on specific breast diagnoses
- 76,333 breast pathology reports from 3 institutions
- Widespread variation in how pathologists reported common diagnoses
  - 124 variations in wording for invasive ductal carcinoma
  - 95 variations in wording for invasive lobular carcinoma
  - 14 variations in wording for atypical ductal hyperplasia

⇨

Atypical ductal hyperplasia
Atypical duct hyperplasia
Atypical intraductal hyperplasia
Atypical ductal epithelial hyperplasia
Atypical hyperplasia with ductal and lobular features
Atypical hyperplasia with mixed ductal and lobular features
Atypical lobular and ductal hyperplasia
Atypical ductal and lobular hyperplasia
Atypical hyperplasia, both typical, and atypical
Atypical hyperplasia is present with pagetoid spread in to ducts
Papillary duct hyperplasia with atypia

Buckley JM, et al. J Pathol Inform. 2012; 3:23.

Classification of Free Text Reports by NLP

- Describes prototype natural language processing (NLP) system that automatically extracts lung cancer staging information from free text pathology reports
- Applied machine learning text categorization techniques to train machines using reports from 710 lung cancer patients
- Validated on a set of 179 patients
- Overall accuracy of 74% for tumor (T) stage, 87% for node (N) stage
- *Works pretty well, but definitely not perfect….data for clinical decision support requires higher quality….*

McCowan IA, et al. J Am Med Inform Assoc. 2007; 14:736-745.

Coding Systems

- Way to represent a set of related real world concepts using defined codes or terms that are readily adaptable for use by data processing systems
- Range from simple list of concepts to complex multiaxial hierarchical structures with defined relationships among underlying concepts
- Provides a controlled terminology = finite set of terms with an agreed-upon meaning
- Allows users to communicate information with confidence

Becich MJ, et al. Clin Lab Med. 2004; 24:913-43.

Coding: Key terms

- **Concept** = fundamental unit of meaning within a terminology or classification system (ex. ischemia)
- **Term** = word or phrase which names a particular concept
  - Multiple terms may convey the identical concept: ex. stroke and cerebrovascular accident
- **Classification** = system for organizing concepts in a particular area of knowledge into related groupings (ex. ICD-10-CM)
- **Terminology** or **nomenclature** = set of terms for concepts in a particular area of knowledge (ex. ICD-10-CM)
  - Definitions NOT required (unlike a vocabulary)
- **Ontology** = vocabulary that includes information about relationships among concepts (ex. SNOMED-CT)
- **Semantic relationships** = expressions of the connections between various concepts (ex. SNOMED-CT)

Coding Standards Important in Pathology

- CPT (Current Procedural Terminology)
  - Healthcare procedures (>8,500 terms)
  - Required for most billing transactions
- ICD-10-CM (International Classification of Disease, 10th Edition, Clinical Modification)
  - Disease and injury classification (68,000 codes)
  - Required for most billing transactions
- SNOMED-CT (Systematized Nomenclature of Medicine-Clinical Terms)
  - Broad-based medical nomenclature (>300,000 concepts)
  - Includes semantic relationships
  - Most complex coding system
- LOINC (Logical Observation Identifiers and Codes)
  - Set of numeric codes that identify a particular type of observation
  - Ensures that tests from two different labs are same

Beckwith, B. Coding in Pathology Informatics: Theory and Practice (2012)

Messaging Standards

- DICOM (Digital Imaging and Communications in Medicine) – image exchange standard
- HL7 (Health Level 7) - clinical and administrative messaging standard
- CDA/CCD (Clinical document architecture/continuitiy of care document) – clinical document standards

Beckwith, B. Coding in Pathology Informatics: Theory and Practice (2012)

Disadvantages of Coding Standards

- Complexity – difficult to know which code to use
- Incomplete standards – can't find specific code that you need

- Note on Autocoding: use NLP/text parsing to automatically code free text
  - Issues with negation, ambiguous findings (ex. most consistent with)

Data Quality in Pathology

- Healthcare documentation has two parts
  - **Data capture**
  - Report generation

Both must be considered in order to have high quality data

  - *Proactive/preventive: use more standardized terminology upfront when data is initially being captured*

Argument for Synoptic Reporting

- Incomplete pathology reports
- **Landmark retrospective study** of 15,940 pathology reports of colorectal cancer from 322 labs
- Essential elements (gross tumor size, depth of tumor invasion, status of resection margins and tumor grades) were omitted from a significant portion of surgical pathology reports
- *Use of a standardized report or checklist increased likelihood of complete report*

Synoptic Reports

- Terminology: “guidelines, protocols, templates, practice parameters, checklists”
- Provides uniform standardized data elements in the form of checklists to ensure that pathologists make note of these findings in their reports
- Data is summarized as a list of previously defined data elements

KIDNEY (LEFT): ADENOCARCINOMA
MACROSCOPIC
SPECIMEN TYPE: Radical Nephrectomy
LATERALITY: Left
TUMOR SITE: Upper pole
FOCALITY: Unifocal
TUMOR SIZE: Greatest dimension is 7.2 cm
MACROSCOPIC EXTENT OF TUMOR: Tumor extends into major veins
MICROSCOPIC
HISTOLOGIC TYPE: Clear cell (conventional) renal carcinoma
HISTOLOGIC GRADE: (Furhman Nuclear Grade): 2
PATHOLOGIC STAGING (pTN)
PRIMARY TUMOR (pT): pT3
REGIONAL LYMPH NODES (pN): Nx
Number of lymph nodes examined: 0
Number of lymph nodes involved: 0
MARGINS: Renal vein margin positive
ADRENAL GLAND: Univolved
VENOUS (LARGE VESSEL) INVASION (V)(excluding renal vein and inferior vena cava): Negative
LYMPHATIC (SMALL VESSEL) INVASION (L): present
ADDITIONAL PATHOLOGIC FINDINGS: Chronic glomerulonephritis present in non-involved renal parenchyma.

Benefits of Structured Synoptic Cancer Pathology Reports

- Significantly improves *completeness* of cancer reports across a broad range of tumor types
- Simplifies and *prioritizes* the recording of information
- Ensures that pathologists are kept abreast of the *latest* minimum reporting standards for all tumors
- Secondary users, such as *cancer registries*, can more efficiently extract meaningful staging and prognostic data than from narrative reports
- Improved information to support clinical *decision* making, i.e. increased clinician *satisfaction*

Lankshear S, et al. Arch Pathol Lab Med. 2013;137:1599-1602.

Some History….

- Early 1990's: a number of pathology professional societies began issuing recommendations specifying a minimum set of data elements that should be included in pathology reports for particular tissue types or pathologic diagnoses
- 2004: American College of Surgeons' Commission on Cancer required as a condition of cancer program accreditation that surgical pathology cancer reports contain validated or regularly used data elements in their reports for each site and specimen
- To facilitate this, College of American Pathologists (CAP) developed site-specific cancer protocols and checklists as a resource
- Today: CAP requires CAP-accredited laboratories to include all report elements specified in CAP cancer protocols in surgical pathology reports

Kang HP. Am J Clin Pathol. 2009;132:521-30.

CAP Cancer Protocols/Checklists

- Set of *standardized protocols* for the most commonly reported forms of cancer
- Goal: Improve quality and uniformity of information in pathology reports
- Developed by the CAP cancer committee
- Consist of data elements structured as a set of questions and prospective answers
- Staging is based on the AJCC Staging Manual
- Includes reference information and is updated periodically
- Available in doc and pdf versions
- Electronic cancer checklists (eCC) were released in early 2007 to advance use in computerized pathology reporting

CAP Electronic Cancer Checklists (eCC)

Implemented and overseen by the CAP PERT (pathology electronic reporting committee)

Benefits of the CAP eCC

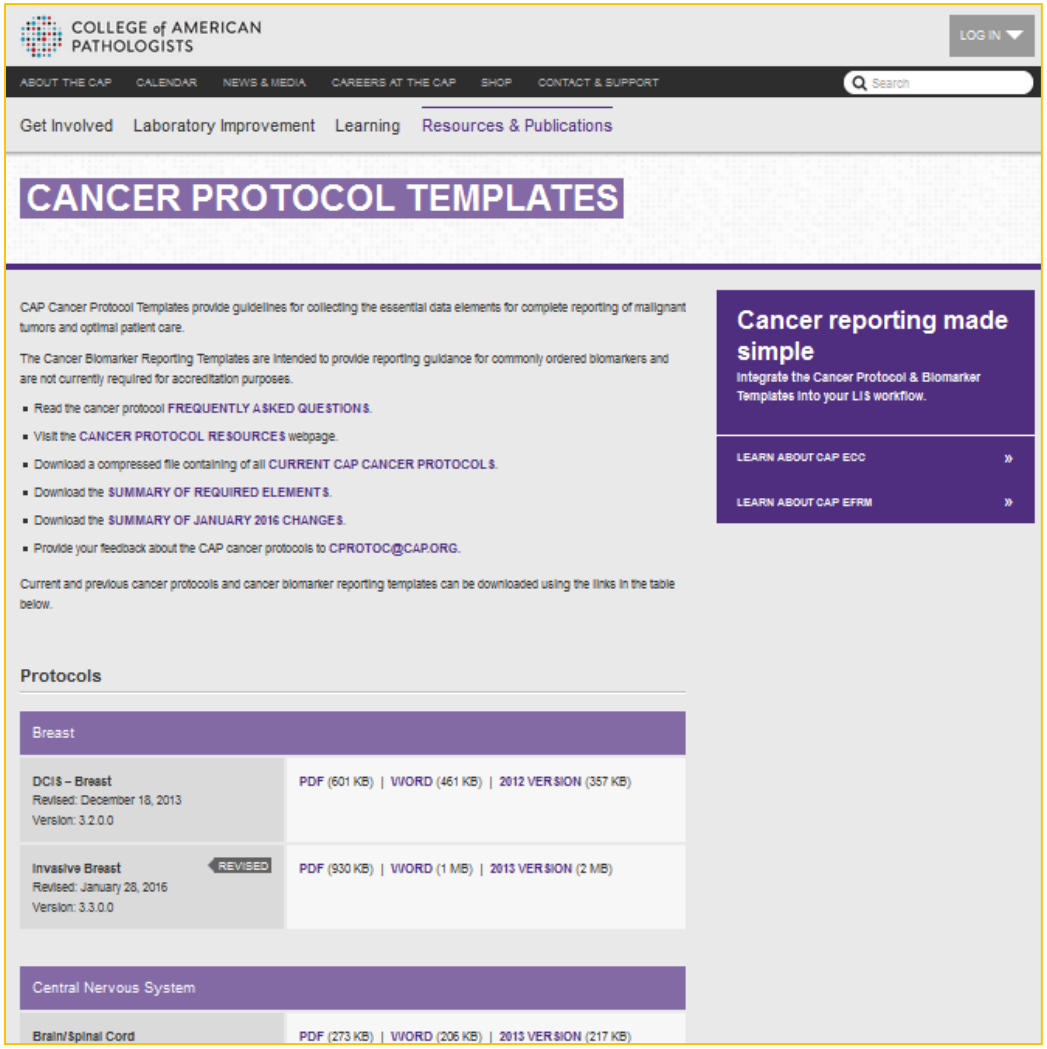
- Integrates into pathologist AP-LIS workflow
- Ensures each report is completed with the necessary required elements
- Improves and supports information exchange and data interoperability
  - Ckeys: provide unique codes for each data element in the CAP eCC
  - SNOMED CT
- Provides automated access to patient data through work with vendors

Vendors who integrate CAP eCC

- Cerner CoPathPlus
- Epic Beaker
- mTuitive (CAP's partner for CAP eFRM™) interfaces and/or integrations with:
  - Cerner Millennium
  - Cortex
  - Meditech
  - Sunquest Copath
  - Other LIS systems
- Novopath
- Psyche Systems
- Softworks Group
- Sunquest Powerpath
- Voicebrook

Note: eCC is available in a standardized software implementation using CAP eFRM

CAP Cancer Protocols



CAP Cancer Synoptic Reporting Format

- Data is displayed as required checklist item (required data element, RDE), followed by its answer (response), ex. “Tumor size: 5.5 cm”
- Each diagnostic parameter pair (checklist RDE: response) is listed on a separate line or in a tabular format, with a few exceptions:
  - Anatomic site or specimen, laterality and procedure
  - Pathologic Staging Tumor Node Metastasis (pTNM) staging elements
  - Negative margins, as long as all negative margins are specifically enumerated
- The synopsis can appear in the diagnosis section of the pathology report, at the end of the report or in a separate section, but all RDE and responses must be listed together in one location
- Additional items (not required for the CAP checklist) may be included in the synopsis but all required RDE must be present
- Narrative style comments are permitted in addition to, but are not a substitute for, the synoptic reporting.

Impact of Synoptic Reporting on Adequacy of Surgical Pathology Reporting in Cancer

- CAP Q-probes study: 2125 cancer reports from 86 institutions
- Each institution reviewed 25 consecutive surgical pathology reports (breast, colon, rectum, and prostate cancer)
- Recorded type and total number of missing required elements (deemed essential by the American College of Surgery) for each report
- *Only 68.8% of all surgical pathology cancer reports contained all the required elements*
- *Institutions in which checklists were routinely used reported all required elements at a higher rate than those that did not use checklists (88% vs. 34%)*
- The missing elements common to cancer reports of all tumor types were *extent of invasion* and *status of resection margins*

Idowu MO, et al. Arch Pathol Lab Med. 2010;134:969-74.

Impact of Standardized Synoptic Pathology Reporting on Physician Satisfaction

- 970 clinicians (pathologists and treating physicians) across 27 hospitals
- 11-item survey to obtain information regarding timeliness, completeness, clarity and usability (*5 point scale*); open-ended questions also employed
- 51% response rate
- ***The vast majority of physicians perceive synoptic reports as significantly better than narrative reports for all items*** (mean scores ranging from 3.84 to 4.77)
- ***Statistically significant difference in the overall satisfaction scores of oncologists and pathologists*** (mean 4.52, SD=0.991 vs mean 4.0, SD 1.34)
  - Pathologists reported that the time to produce the reports was more than that of narrative reports (mean 3.51, SD 1.43) (although end users did NOT perceive a difference in the time to obtain the report)
  - Comments revealed technology-related issues as the most frequent factor impacting timeliness of report completion

Lankshear S, et al. Arch Pathol Lab Med. 2013;137:1599-1602.c

Synoptic Reports vs. Structured Data

- *Not all synoptic reports contain structured data*
- Many **synoptic reports** are simply word processing documents that appear structured to humans
  - They provide visibly structured blocks of free text which is embedded in the pathology report
- Truly **structured data** is entered in many smaller specific text fields rather than a few large ones
  - Every single data element has its own predefined place in the database
  - Every discrete data element is directly linked to its inherent context

Synoptic reports clarify findings for clinicians while structured data clarifies findings for computers

Amin W, et al. Open Access Bioinform. 2010;2:105-112

Advantages of *Structured Data*

- Beyond the benefits of synoptic reporting, truly *structured data* in the form of discrete data elements allows for
  - Advanced data-querying capabilities
  - Automated analysis
  - Decision support
  - Predefined comment generation or staging

- *Required for the future of pathology informatics and computational pathology*

Ellis DW. Pathology. 2011;43:404-9.

Data Quality in Pathology

- Healthcare documentation has two parts
  - Data capture
  - **Report generation**

Both must be considered in order to have high quality data

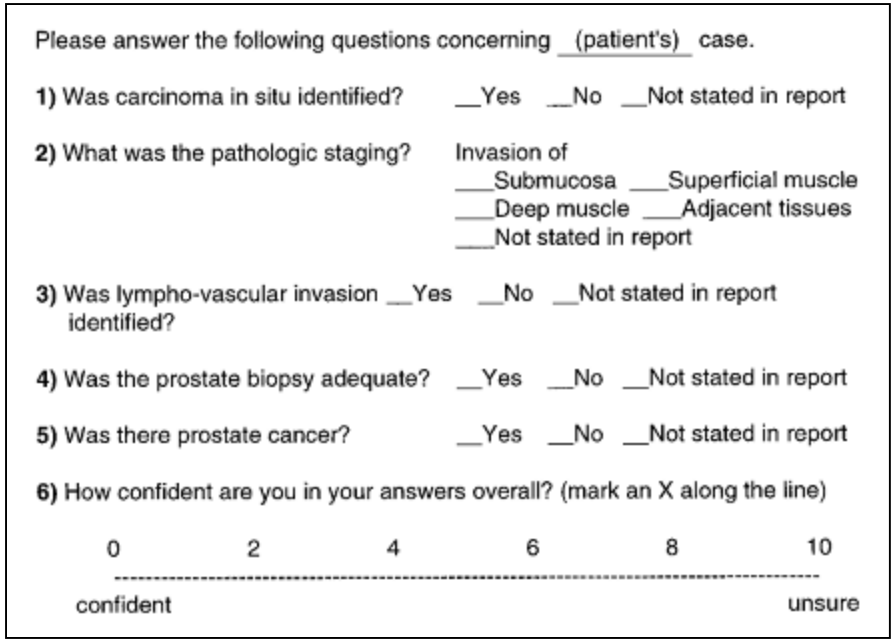
Structured data capture is important step towards improving data quality in surgical pathology (computer readable), but, if the report is of poor quality in the human readable format, data is considered poor quality

Clinician (Mis)interpretation of Pathology Reports

"Clinicians are from Mars and Pathologists are from Venus"

- Goal: Compare clinician comprehension with pathologist intent in written pathology reports
- Typical surgical pathology reports relevant to surgeons and covering a wide range of specimen complexity
- Questionnaires based on these cases administered open-book-examination style to surgical attendings physicians and trainees

- *Surgeons misunderstood pathology reports 30% of the time*
  - Surgical experience reduced but did not eliminate the problem
  - Familiarity with the report format helped reduce misinterpretation



Please answer the following questions concerning (patient's) case.

1) Was carcinoma in situ identified? __Yes __No __Not stated in report

2) What was the pathologic staging? Invasion of
___Submucosa ___Superficial muscle
___Deep muscle ___Adjacent tissues
___Not stated in report

3) Was lympho-vascular invasion identified? __Yes __No __Not stated in report

4) Was the prostate biopsy adequate? __Yes __No __Not stated in report

5) Was there prostate cancer? __Yes __No __Not stated in report

6) How confident are you in your answers overall? (mark an X along the line)

0 2 4 6 8 10

confident unsure

Powsner SM, et al.. Arch Pathol Lab Med. 2000;124:1040-1046.

What are elements that lead to successful reporting?

- Content is complete
- Reports are formatted well…

Nakhleh RE. Arch Pathol Lab Med. 2011;135:1394-97.

Design of the Pathology Report

It's not just about the content….

- Used principles from the publishing, commercial aviation and cognitive psychology literature to provide guidance to pathologists interested in formatting diagnostic reports to optimize information transfer

- Four principles are described that can help more effectively communicate information
  - Use headlines to emphasize key findings
  - Maintain layout continuity
  - Optimize information density for readers
  - Reduce clutter

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Use Headlines to Emphasize Key Findings

- Lessons from the newspaper industry
- Headline
  - Precedes story
  - Conveys key message in large bold typeface
  - Set off from the body of the story by visual “white space”
  - Often followed by a cascade of progressively smaller subheadings that convey important subpoints (inverted pyramid)

"All the News That's Fit to Print"

The New York Times

Late Edition

VOL. CL . . . No. 51,674 NEW YORK, WEDNESDAY, SEPTEMBER 12, 2001 75 CENTS

U.S. ATTACKED

HIJACKED JETS DESTROY TWIN TOWERS AND HIT PENTAGON IN DAY OF TERROR

A CREEPING HORROR

Buildings Burn and Fall as Onlookers Search for Elusive Safety

By N. R. KLEINFIELD

President Vows to Exact Punishment for 'Evil'

By SERGE SCHMEMANN

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Use Headlines to Emphasize Key Findings

- Traditional specimen-centered report
  - Each specimen receives its own diagnosis

A. PROSTATE, RIGHT BASE: ATROPHY AND CHRONIC INFLAMMATION. NO NEOPLASM IDENTIFIED.
B. PROSTATE, RIGHT MID: ACUTE AND CHRONIC INFLAMMATION. NO NEOPLASM IDENTIFIED.
C. PROSTATE, RIGHT APEX: ATROPHY. NO NEOPLASM IDENTIFIED.
D. PROSTATE, LEFT BASE: ADENOCARCINOMA, CONVENTIONAL TYPE, GLEASON 3+4=7, SIZE = 7 MM, PERINEURAL INVASION PRESENT.
E. PROSTATE, LEFT MID: SINGLE FOCUS OF ADENOCARCINOMA, CONVENTIONAL TYPE, GLEASON 3+3=6, SIZE = 2 MM.
F. PROSTATE, LEFT APEX: ATROPHY AND CHRONIC INFLAMMATION. NO NEOPLASM IDENTIFIED.

- Patient-centered diagnosis
  - Newspaper-style diagnostic headline
  - *Most important finding is set apart and above the other findings*

PROSTATE: ADENOCARCINOMA

Malignant locations: left base, left mid
Benign locations: left apex, right base, right mid, right apex
Gleason score: 3 + 4 = 7
Size: 7 mm (left base); 2 of 6 cores contain carcinoma
Histologic type: conventional prostatic adenocarcinoma
Perineural invasion is present

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

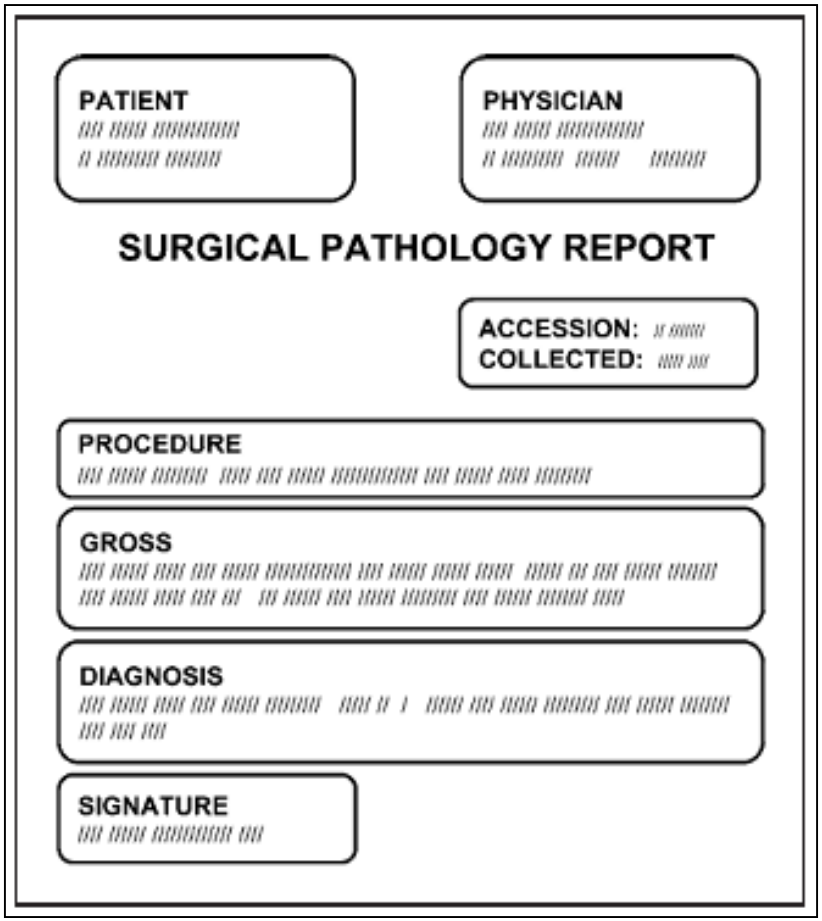
Maintain Layout Continuity

- Experience from the airline industry
  - *Uniform positioning of instrumentation eases pilot transition from one type of aircraft to another*

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Maintain Layout Continuity

- Pathology reports created by a single department usually position fields consistently
- There is less consistency in formatting text within an individual fields and little continuity with reports produced by other departments
- *Consistent positioning speeds up information transfer and reduces risk of confusion*
- **Caveat:** Change in report format to an arguably better layout results in a 17% - 54% increase in recall errors by clinicians
  - *Change creates confusion and should be carefully considered!*



Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Optimize Information Density for Readers

- How much information to include in a single report line
- 4 diagnoses of endometrial carcinoma with different *densities* of information
  - Progressively less information in the diagnostic headline and more information in the text that follows

ENDOMETRIUM: WELL DIFFERENTIATED ENDOMETRIOID ADENOCARCINOMA (FIGO 1), ARISING IN A BACKGROUND OF ATYPICAL HYPERPLASIA. ENDOCERVIX FREE.

ENDOMETRIUM: ENDOMETRIOID ADENOCARCINOMA (FIGO G1).

Atypical hyperplasia is present.
Endocervix not involved.

ENDOMETRIUM: ADENOCARCINOMA.

Histologic type: Endometrioid
Histologic grade: Well differentiated (FIGO 1)
Non-malignant endometrium: Atypical hyperplasia
Endocervix: Not involved

ENDOMETRIUM: CANCER.

Cancer type: Adenocarcinoma
Histologic subtype: Endometrioid
Histologic grade: Well differentiated (FIGO 1)
Non-malignant endometrium: Atypical hyperplasia
Endocervix: Not involved

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

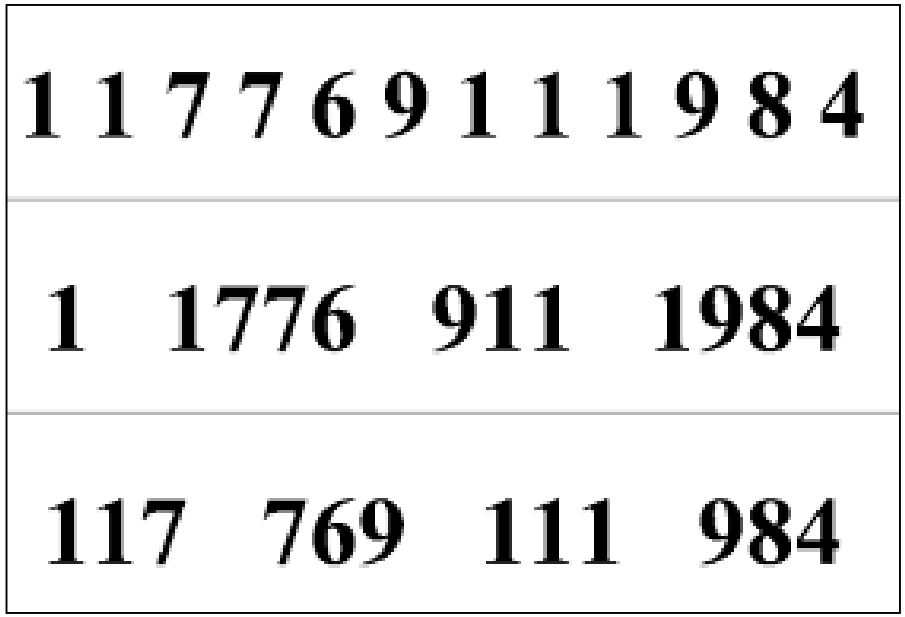
Optimize Information Density for Readers

- Classic example in the psychology literature is inability of an average person to store more than 7 bits of unrelated information in short term memory

- Average person will only be able to recall 7 of the 12 digits

- *When digits are grouped into 4 familiar numbers, average person will be able to recall all 12 digits*

- When digits are grouped into unfamiliar groups, recall is comparable to ungrouping



Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Reduce Clutter

- Inclusion of ***distractors*** in a visual field interferes with the person's ability to acquire information and interferes with short term memory
- ***Distractors*** include
  - Symbols
  - Text effects (font changes or colors) that draw attention away from the information
  - Unnecessary information
    - Additional information that does not influence patient management or have prognostic value
    - Histologic description (*arguable*)
    - Billing codes
    - Disclaimers required by regulation

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Preserving Formatting Over Electronic Interfaces

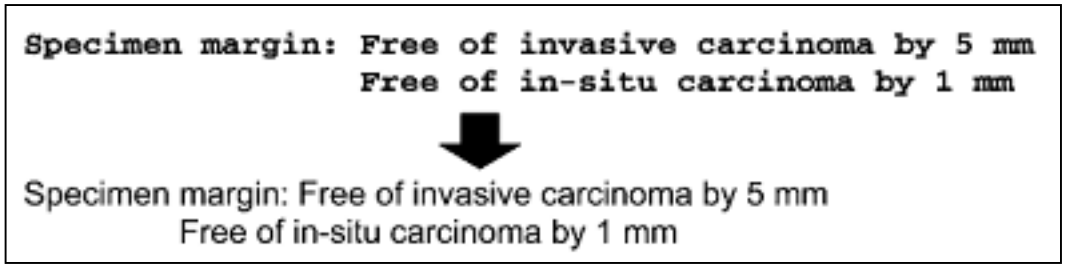
- *Over 75% of pathology reports for inpatients are first viewed on the hospital information system computer screen (vs. paper)*
- Degradation of formatting over interfaces depends on capabilities of the sending and receiving systems and the interface in use
  - HL7 is most commonly used
  - Supports lower and upper case text but not font type, bold, font size, color) or columns and tables

Therefore, reliance on text effects is discouraged where electronic interfaces are likely to be used

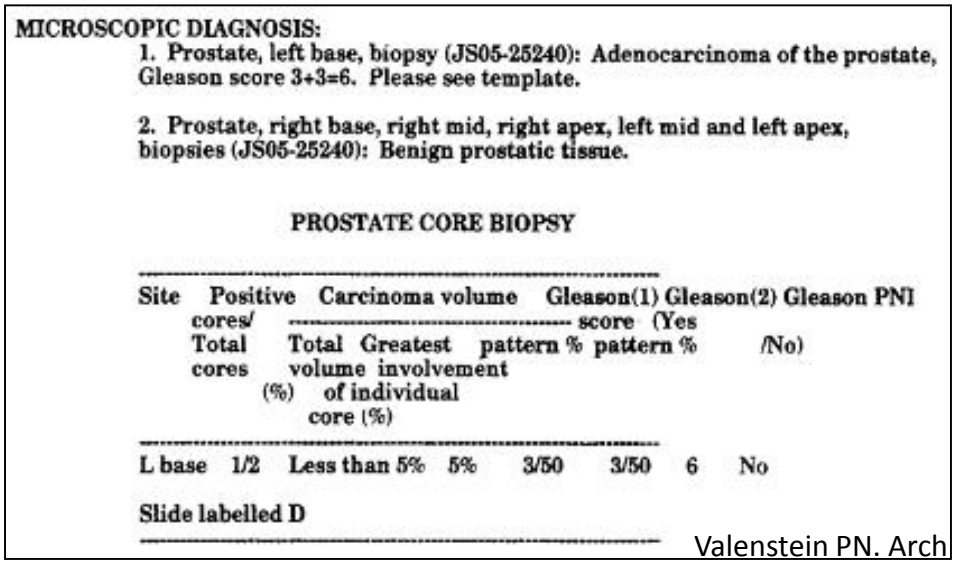
Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Common formatting problems

- Loss of alignment when text created in fixed-width font (Courier) is re-set in proportional font (Arial)



- Loss of column alignment when tables are constructed



MICROSCOPIC DIAGNOSIS:

1. Prostate, left base, biopsy (JS05-25240): Adenocarcinoma of the prostate, Gleason score 3+3=6. Please see template.

2. Prostate, right base, right mid, right apex, left mid and left apex, biopsies (JS05-25240): Benign prostatic tissue.

PROSTATE CORE BIOPSY

Site	Positive cores/ Total cores	Carcinoma volume Total volume (%)	Greatest involvement of individual core (%)	Gleason(1) pattern %	Gleason(2) pattern %	Gleason score	PNI (Yes /No)
L base	1/2	Less than 5%	5%	3/50	3/50	6	No

Slide labelled D

Valenstein PN. Arch Pathol Lab Med. 2008;132:84-94.

Summary

- Data quality cannot be measured without defining a set of standards. The dimensions of data quality used in other industries can be applied to healthcare.
- Ensuring data quality is more complex in AP due to the traditional narrative or free text nature of pathology reports.
- Natural language processing tools or coding standards can be used to try to improve quality of free text reports and data exchange, however these methods may currently be insufficient for clinical decision support application.
- Initially capturing data a structured synoptic format is the most effective method for ensuring data quality.
- In addition to data content, report formatting is an important aspect of data quality as it relates to communication of data.

Thank you!

How do you code for ambiguous terminology

What about the rest of pathology