The Integration of Whole Slide Imaging in the Clinical Anatomic Pathology – Limitations of Laboratory Information Systems, Image Capture Systems and Archives

Yukako Yagi, Drazen Jukic, Anil Parwani, Jon Ho, William Gross, Ellen Kokal, Tony Piccoli, Michael Kistler and John Gilbertson

University of Pittsburgh Medical Center (UPMC), Pittsburgh USA
This work was partially supported by funding from the U.S. Air Force administered by the U.S. Army Medical Research Acquisition Activity (USAMRAA), 820 Chandler Street, Fort Detrick MD 21702-5014, Contract No. DAMD17-03-2-0017. The content of the information does not necessarily reflect the position or policy of the U.S. Government and no official endorsement should be inferred.

I (Yukako Yagi) have been Scientific Advisor for Trestle, DMetrix and Aperio.
Background

The past five years has seen the emergence of whole slide imaging robots – devices that can automatically image entire microscope slides at high speed and high resolution. A typical device can capture a slide in 5 minutes at tissue sampling rates of 0.3-0.5 microns/pixel, resulting in an uncompressed image file of 5 to 10 GB, and a typical pathology case contains ten slides. Though these “high resolution whole slide images” provide diagnostic information similar to that obtained by direct examination of tissue under the microscope and are proving useful in a variety of clinical activities; their novelty and sheer volume has resulted in a number of image and data management challenges. One of these challenges is that Laboratory Information Systems, which drive workflow and data management in pathology departments, are not well equipped to manage image level information.
What is a whole slide image?
Microscope Imaging

If we try to send all information on a glass slide, it is more than 2.7GB/slide.

For Static Image Telepathology, a referring pathologist has to be able to select appropriate diagnostic fields. To select suitable fields for consultation requires experience.

Needed virtual slide
A Digital Slide

- A classical whole slide image is formed by imaging a entire physical (glass) slide, field by field, and then ‘knitting’ these fields together to form a seamless montage.

- With some display software, one can pan and zoom around the image set.
Pyramid File Structure

Low-magnification

Mid-level magnification

High-magnification
Whole Slide Imaging

- A Digital Slide is a massive data set
File Size

- Consider a WSI system:
  - 0.6 NA, 20x Primary Magnification
  - 8.8 x 6.6 mm CCD
  - 6.6 um pixels
  - 0.33 um/pixel
  - 900 million pixels / square cm of tissue
  - 3 bytes / pixel (24 bit color)
  - 2.7 GB / square centimeter of tissue per focal plane for the base image
Digital Slides have Issues

- 2.7 GB/cm is the base image
  - $2.7 + \frac{2.7}{4} + \left(\frac{2.7}{4}\right)^2 + \ldots$
  - $2.7 \text{ GB} \times 1.33 = 3.5 \text{ GB per square cm of tissue}$

- Assume 1.5 square cm per slide, 10 slides per case:
  - $\sim 52 \text{ GB per case!}$

- This is based on 20x optical magnification and one focal plan
- 40x magnification and 5 focal planes $\sim 50 \text{ GB} \times 20 = 1 \text{ TB}$
Digital Slides have Issues

- JPG2000 compression ~ 30:1
- ~ 115 MB per square cm per focal plane…
- 1725 MB per case
- UPMC → 80,000 cases/year
- 138 TB per year
- Is this unreasonable?

- This is based on 20x optical magnification and one focal plan
**Methods**

**Methods:** At the University of Pittsburgh, we have developed an infrastructure for the clinical use of whole slide imaging (WSI) including the implementation of different types of imaging robots, imaging quality assurance protocols, compression and storage mechanisms, mechanisms to serve whole slide images throughout the medical center, slide image viewers and a team of pathologists, imaging scientists and engineers dedicated to the evaluation of whole slide imaging systems in the clinical environment.

It was quickly realized that for the clinical evaluation of WSI to be realistic, images had to be managed (or at least accounted for) within the Laboratory Information System (LIS). The team, assisted by personnel from central IT and Radiology, examined 1) The image information needs of the pathologist, histologist, imager and image data manager and how these needs can be accommodated in with the LIS and Pathology Imaging Systems and 2) Mechanisms by which specific systems – the Copath C/S Laboratory Information System, the Aperio T2 Whole Slide Imager and UPMC’s DICOM Compliant Enterprise Image Archive - could share images and image information in support of clinical evaluations.
Why are we implementing a clinical whole slide image delivery system?
Histology Workflow
From Tissue Sample to Histological Examination

In pathology, imaging begins in histology!
Virtual Microscopy
Digital Slide Creation, Management & Analysis

Surgery Center
- Biopsy
- Frozen Section

Histology Lab
- Specimen Examination
- Histo-processing
- Sectioning
- Staining
- Scanning

Workflow Management
- Image Analysis
- Conferencing

Microscopic Examination
- Pathologist Office

Digital Slide(s)

Pathologist Office - Step that requires Pathologist
Available systems in US

Aperio

Hamamatsu

Dmetrix

Trestle
Aperio ScanScope T2/T3/CS

Digital Slide Creation, Management, and Analysis
Aperio's ScanScope Systems, comprised of award-winning ScanScope scanners and Digital Slide Information Management Software, deliver integrated digital slide creation, viewing, management, and analysis capabilities for virtual microscopy applications. ScanScope Systems are invaluable to pathologists for a multitude of applications, including education, tissue microarrays, toxicology pathology, telepathology, image analysis and workflow systems (PACS).
Hamamatsu

NanoZoomer Digital Pathology opens up new applications in pathology

With Hamamatsu Photonics scanning technology, rapid conversion from glass slides into high resolution digital slides become available. The digital slide technology will become indispensable for telepathology to link hospitals through networks and for the development of new diagnostic methods in pathology.

1. Pathological networks not only within a hospital but also among remote hospitals

By applying digital slide technology for telepathology through a network, it is expected to conduct frozen section diagnosis at hospitals with no resident pathologists. In addition, by establishing a network among pathologists, it is expected to conduct consultations quickly and conveniently with remote specialized pathologists for difficult cases. As the result, it is expected to improve the diagnosis accuracy. Furthermore, for pathological diagnosis contract businesses, it is expected to get higher efficiency and increasing turnover by eliminating the process of the slide delivery.

2. Ideal for medical students’ education

Since digital slides can be observed from multiple computers, there is no need to prepare individual slides sets for students. In addition, because of simple viewing software, even beginners can concentrate on the essential educational content without complexity of microscope operation.

3. Coordination of Pathological Laboratory Information system and NDP

Through the coordination of Pathological Laboratory Information system and NDP, it is expected to realize the improvement of efficiency and accuracy. Furthermore, digital slides will be shared with other departments, and it will be applicable for conferences.

4. Possibilities for new diagnostic methods using the benefits of new observation methods using PC and digital slide databases

Improvement of diagnosis accuracy is expected through the introduction of new observation methods that were not possible by using conventional microscopes, such as simultaneous observations of details and whole tissues, comparison of pre-treatment and post-treatment and simultaneous observations of multiple slides stained differently such as immunostains and DNA stains. In addition, database of digital slides enables easy search of slides. Furthermore, image analysis, such as the size of nuclei and stain concentration, as well as the creation of case libraries are expected.
DMetrix

Strategy for rapid throughput of pathology specimens: Glass slides are immediately scanned with an ultra-rapid virtual slide processor and read out by telepathologists.

* Final pathology report includes the input of specialists and other authorities as needed.
Trestle

We Enhance the Efficiency of Research Operations

Live review of slide over the internet or virtual slide creation

Immediately available for viewing, diagnosis and consults worldwide

Trestle provides an integrated, modular research solution
Work to be presented: In this paper we describe the nature of whole slide images and integration of whole slide imaging into the existing workflow of a pathology department. Using an image v glass slide equivalence study in anatomic pathology quality assurance as a context, we will discuss changes required in the Laboratory Information System (LIS) and Histology Laboratory to support image level information, departmental decisions surrounding the dissemination of images, the integration of imaging systems and the LIS and the development of a “DICOM wrapper” to communicate gross and histological images from Pathology to an Enterprise Image Archive.
Whole slide image clinical validation studies

- 3 pathologists
- 25 full cases, same workflow (3-24 slides)
- 200-500 whole slide images per study
- Integration of images with clinical information, histology information (ie staining) and workflow information (ie case status)
- Security issues
- Significant logistical effort
Goal

To create a whole slide image delivery system for pathologists for clinical sign-out responsibilities

*requires AP LIS integration, but workflow will reside in the AP LIS
Whole Slide Imaging at UPMC

The ability to digitize an entire histologic slide at high resolution and display the resulting image across a broadband network - is becoming an important technology for telepathology and Pathology Imaging.

Current Spec
Hardware Specification
Resolution: 0.47um/pixel (20x)
0.23um/pixel (40x)
Speed: 40mm2/min
File format: Tiff/Jpeg2000 (Pyramid)
Bar code: 2D
Autoloader: 120 slides
Scanning Process

- Place Slides (120) in Autoloader
- Push “Start” button
  - Change slide (30s)
  - Bar Code reading
  - Tissue finding
  - Auto focus
  - Scan (strip)
  - Compression (on board)
  - Stitch
  - Feeding Virtual slide image to the storage

Average time: 3-10 min/slide, about 7 hours for 120 slides

Image size: 1-8 GB
File size: 20-600 MB
Each vendor has its own formats, servers and clients
No real integration with LIS or Security

Vendor machine

WSI Robot
WSI Presentation Server
WSI DB
WSI Storage

LIS System
Client Pathologist
Security System
Better structure
Process 1:

Looks for new whole slide images
Pulls new images
Decodes the 2D barcode
Verifies slide/accession info with APLIS
Fetches specimen/patient info from APLIS
Constructs XML metadata wrapper
Reconstructs file
Sends file to WSI Image server

WSI Image Import Server

Barcode elements
Patient/specimen elements
Metadata wrapper

- **UID**: 1.2.840.152371.157.229.222.79.20050706.121311.30.1
- **XML**
- **Patient level** (last name, first name, med rec num, sex, birthdate…)
- **Accession level** (Accession number, date, time, pathologist…)
- **Study level** (Modality, date, time, manufacturer, IP address…)
- **Series level** *(whole slide image)*
  - Part
  - Block
  - Slide number
  - Stain
  - Components and descriptions (thumbnail, label, base image…)
- **Image level** (describes each component)
Example of metadata wrapper

- <SeriesLevel>
  #A Series is effectively one slide imaged one time.
  #A Series has multiple images usually in a TIFF container
  #Some images are "real" optical images, others are
  sampled as part of a Pyramid:
  - <S.1>
    #Series Identification:
    <UID>1.2.840.152371.157.229.221.31.20050322.112447.30.1</UID>
    <PartNumber>1</PartNumber>
    <BlockNumber>A</BlockNumber>
    <SlideNumber>1</SlideNumber>
    <PartDescription>Colon Resection</PartDescription>
    <BlockDescription>Proximal Margin</BlockDescription>
    <SeriesBeginTime>143456</SeriesBeginTime>
    <SeriesEndTime>144056</SeriesEndTime>
    <SeriesFileDescription>WSI-TIFF</SeriesFileDescription>
    <PixelSize>N/A</PixelSize>
    <SeriesComment>NA</SeriesComment>
    <Paths>
      <P.1>C:\Image Quality\Images\Focus Measure\Yukako Images</P.1>
    </Paths>
    <Images>
      <I.1>Thumbnail.jpg</I.1>
      <I.2>SlideLabel.tif</I.2>
      <I.3>BaseImage.jp2</I.3>
      <I.4>Level1Image.jp2</I.4>
      <I.5>Level2Image.jp2</I.5>
      <I.6>ScoutCameraImage.jpg</I.6>
    </Images>
  </S.1>
- <DICOMMetaTags>
  - <x00002002>1.2.840.10008.5.1.4.1.1.7</x00002002>
  - <x00002003>1.2.840.152371.157.229.221.31.20050322.112447.30</x00002003>
    - <x00020010>1.2.840.10008.1.2.4.50</x00020010>
    - <x00020016>SimpleDICOMWrap</x00020016>
    - <x00080005>ISO_IR 100</x00080005>
  - <x00080008>ORIGINAL\PRIMARY</x00080008>
  - <x00080016>1.2.840.10008.5.1.4.1.1.7</x00080016>
  - <x00080018>1.2.840.10008.1.2.4.50</x00080018>
  - <x00080020>20050322</x00080020>
  - <x00080023>20050322</x00080023>
  - <x00080030>112447</x00080030>
  - <x00080033>112447</x00080033>
  - <x00080034>1234</x00080034>
  - <x00080040>OT</x00080040>
  - <x00080050>N/A</x00080050>
  - <x00080060>Olympus</x00080060>
  - <x00080070>UPMC Presbyterian</x00080070>
  - <x00080090>Dr. John Kirkwood</x00080090>
  - <x00080100>Olympus CC12</x00080100>
  - <x00080120>N/A</x00080120>
  - <x00080130>N/A</x00080130>
  - <x00080140>UPMC Presbyterian</x00080140>
  - <x00080160>Number Two</x00080160>
**Process 2**

- Receives the image
- Parses metadata wrapper
- Writes database entry
- Stores the image
- Updates LIS on image status
- Fetches image’s context
- Serves the image
System communicates image status to LIS
Results & Discussion

The work revealed a number of structural and procedural issues in the LIS, Imaging System and Archive that hindered the implementation of large scale imaging in pathology. Some of these issues include:

- LIS systems did not support the concept of an whole slide image
- Current pathology imaging systems did not well support the “groups of associated images (i.e. “series”)”
- The Enterprise Image Archive had difficulties with the size of the WSI and the proprietary internal structure of some of the WSI image files.

The team implemented a series of work-arounds for these problems and tested them as part of clinical evaluations. On the basis of their results, we are working to develop long term solutions.
Patient Name: Madison, Dolley
Accession #: S01-00104

Accession Date: 10/19/2000
Procedure Date: 10/18/2000
Signout Date: Not Signed Out
Sex: Female
Attending MD: Walter Brown, DR

MRN: 999820372
DOB: 6/19/1948

Patient History:
Polyps

Gross Description:
1. Esophagogastric Junction, Bx. A formalin container is received labeled with the name "D. Madison" and "bx E-O Junction". It contains three 0.1 cm diameter items of tan soft tissue that are submitted in toto as #1.

2. Stomach, Not otherwise specified, Bx. A formalin container is received labeled with the name "D. Madison" and "gastric bx". It contains a 0.1 cm diameter item of tan soft tissue that is submitted in toto as #2.

3. Colon, Sigmoid, Polypectomy. A formalin container is received labeled with the name "D. Madison" and "sigmoid colon polyp". It contains multiple fragments of tan soft tissue that in aggregate are 0.4 cm in diameter. They are submitted in toto as #3.
Old cases on this patient
Is this patient on a conference list etc
Does case need QA, etc
Final thoughts

• The real world
  – Multiple vendors
  – Multiple information systems to interface with

• Difficulties we encountered
  – LIS limitations (Unique slide problem)
  – Lack of standard for modality output format
  – Lack of DICOM standards for whole slide images
Thank you!