Virtual Microscopy:
A Tipping Point in Tissue Based Research and Education

James Diamond
Peter Hamilton
Queen’s University of Belfast
Pathology
Skilled interpretation of tissue morphology
Pathology diagnostics is central to patient care
Pathology diagnostics is central to translational research and biomarker discovery.

(Courtesy of Imgenex)

AE1AE3  E-Cadherin  Ki67
The Tipping Point?
Turning glass slides into bits and bytes
Hardware – a competitive market
Image tiling

Single objective

Line scanning

The array microscope

The array microscope
Virtual Fluorescence Slides
Z-axis scanning and virtual focus

Cervical cytology

17 Gigabyte : TIFF JPEG compressed

25 slides = 1 Terabyte
These are large images!

At x40 objective magnification the required resolution is 0.24 microns per pixel = 104,000 x 85,000 linear pixels

25 Gigabytes

JPEG2000 compression 25:1

RVH Belfast: 300 cases a day > 100 Terabytes of data a year
Technical challenges

♦ Defining standards ♦

♦ Storage of Virtual Slides ♦
  ♦ Enhancing on-line delivery of gigapixel imagery ♦
  ♦ Tools for viewing and interacting with virtual slides ♦
  ♦ Algorithms for automated machine vision of gigapixel images ♦
  ♦ High performance computing for tissue imaging ♦

It’s not the virtual slide that counts – it’s what we do with it.
Application opportunities

Making virtual microscopy work

Education and Training

Quality Assurance

Tissue archiving

Tissue Research
intranet

Virtual slide server

internet
Slide Annotations for Training and Education

**Infiltration**
- Tissue infiltration by tumor.

**Cohesive clusters**
- Cohesive clusters of epithelial cells forming relatively uniform cell groups with evenly spaced nuclei.

**Koilocytes**
- Cells with enlarged nuclei and granular cytoplasm; a hallmark of HPV infection.

http://www.path.md.eub.ac.uk/BZ_PathVZ_Virtual_Slide_Viewe_microscopy Firefox

Transforming data from PathVZ eub.ac.uk...
Virtual Slides for Examination and Proficiency Testing

Setting virtual slide questions on-line
Recording responses centrally on-line
Setting examinations on-line
Case 9

Breast Histopathology Case 9

View in PathXL slide viewer
View in standard slide viewer
View instructional video with PathStream
Unique approach to training in pathology

Stepping through the diagnostic clues

Visual comparison with template images

Diagnostic probability

Diagnostic map
Bayesian belief network engine with fuzzy logic representation of language

Allows us to model and record diagnostic knowledge from experts
A probability decision map

Diagnostic clues:
- Bare nuclei
- Cellularity
- Cohesion
- Pleomorphism
- Cell arrangement
- Nuclear size
- Nucleoli
- Intracytoplasmic lumina
- Apocrine cells
- Mucinous background

Diagnostic probability

Benign to Malignant
InView®
“Learning by diagnostic simulation”

Endometrial Neoplasia Grading in Prostate Cancer
Soft Tissue Tumours
Salivary Gland Tumours
Melanocytic Skin Lesions
Virtual EQA
External Quality Control and Proficiency Testing
Virtual Slides for EQA

Melanocytic Skin Lesion Diagnosis in UK and Japan
The need for worldwide VM service network of data centers with rapid access to image data globally
Proficiency Testing
Worldwide server architecture for delivery of high resolution imagery
Virtual Tissue Banking

Glasgow Biobank
onCore UK
Tayside Tissue Bank
Wales Cancer Bank
Northern Ireland Tissue Bank
Virtual slides for Tissue Archiving and Research
On-line scoring of TMAs and tissue samples
Enhancing biomarker evaluation

Leeds University (Gastric TMAs)
Cambridge University (Bladder BOXIT trial)
Queen’s University (Mesothelioma)
Queen’s University (Prostate response TMA)
Fusion Antibodies Ltd

Controlled vocabularies
Compatibility with Data Exchange Standards
(Berman et 2003, CaBIG, CancerGrid, Dublin Core Header)
Virtual slides for Tissue Archiving and Research
On-line evaluation of TMA

TMA generation

TMA context mapping

After

Before
Shortfalls in visual interpretation

Making the diagnosis

Subjective criteria
Language – associated with uncertainty
Disease “patterns” are based on predefined mind sets
Not good at detecting quantitative differences
Open to interpretation
Poor reproducibility
Inaccuracy

Scoring the biomarker
Disagreement in pathological diagnosis

<table>
<thead>
<tr>
<th>Condition</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-invasive lesions of the bronchus</td>
<td>0.55</td>
</tr>
<tr>
<td>(Nicholson – Histopathology 2001;38:202-208)</td>
<td></td>
</tr>
<tr>
<td>Cervical cytology</td>
<td>0.46</td>
</tr>
<tr>
<td>(Stoler – JAMA 2001;285:1500-1505)</td>
<td></td>
</tr>
<tr>
<td>Cervical Histology</td>
<td>0.15 – 0.62</td>
</tr>
<tr>
<td>Prostate Cancer</td>
<td>0.58</td>
</tr>
<tr>
<td>(Egevad – Urology 2001;57:291-295)</td>
<td></td>
</tr>
<tr>
<td>Oral Dysplasia</td>
<td>0.27 – 0.45</td>
</tr>
<tr>
<td>(Warnakulasuriya – J Pathol 2001;194:294-297)</td>
<td></td>
</tr>
</tbody>
</table>

Variation in interpretation of renal transplant biopsies  
Furness et al. 2001

Aberrant diagnoses by surgical pathologists  
Wakely 1998

Dysplasia classification: pathology in disgrace  
Bosman. 2001

“Individuality” in the specialty of surgical pathology  
Ackerman 2001
Enormous opportunities to explore improved methods of tissue diagnostics that are objective, reproducible and reliable

Image analytics  Machine vision  Quantitative evaluation  Tissue classification

Locating Nuclei

Cervical preneoplasia

Colorectal adenomas
Progression of cervical intraepithelial neoplasia

Increasing dysplasia
Augmented Visualisation in Pathology

BPH

Prostatic neoplasia

PIN

Prostate Cancer
Machine vision on virtual slides
Computer vision analysis of Tissue Microarrays

IOD = 1.75 = Score 3

Cumulative survival

\[ P = 0.0003 \]

- p53-negative \((n = 204)\)
- p53-positive \((n = 136)\)
High Performance Computing (HPC) using Clusters

Significant speed enhancements for densitometric analysis of gastric TMA

David McCleary
Augmented Visualisation in TMA assessment
High Throughput Analysis of TMAs

Speeding analysis
Providing objectivity and reproducibility

For algorithm development and evaluation

For pre-selection of markers for manual analysis

For more sensitive scoring of markers for clinical evaluation
Multiresolution texture analysis, CBIR and CBR

Tissue archive

Similarity metric (Hamming Distance)

Test case

Case-based reasoning

### Table 1: Texture Feature Descriptions

<table>
<thead>
<tr>
<th>Feature</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average intensity</td>
<td>$I = \frac{1}{Z} \sum_{i} p(i, j)$</td>
</tr>
<tr>
<td>Contrast</td>
<td>$\sigma = \sqrt{\mu_2 - \mu_1^2}$</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>$\rho_{ij} = \frac{\sum (i - \mu_1)(j - \mu_2)p(i, j)}{\sigma_1 \sigma_2}$</td>
</tr>
</tbody>
</table>

where $Z$ is the number of intensity levels in the image, $L$ is the number of possible intensity levels, $\mu_1$ is the mean, $\mu_2$ is the variance, $p(i, j)$ is the pixel intensity at location $(i, j)$, and $\sigma_1$ and $\sigma_2$ are the standard deviations.

Hamming Distance:

$$d_H = \sum_{i,j} |p(i, j) - q(i, j)|$$
Growing numbers of virtual tissue archives

Storage and Search?
Image Search and content-based image retrieval
Acknowledgements

Jim Diamond
David McCleary
Andrena McCavigan
Yin Hai Li

Sid Trewin
Alan Lyons
Adam Askew
Stewart Church